THE PRACTICE OF MEDICINAL CHEMISTRY

THIRD EDITION

EDITED BY

CAMILLE GEORGES WERMUTH

Biography

Camille-Georges Wermuth PhD, Prof. and Founder of Prestwick Chemical, was Professor of Organic Chemistry and Medicinal Chemistry at the Faculty of Pharmacy, Louis Pasteur University, Strasbourg, France from 1969 to 2002. He became interested in Medicinal Chemistry during his two years of military service in the French Navy at the "Centre d'Etudes Physio biologiques Appliquées à la Marine" in Toulon. During this time he worked under the supervision of Dr Henri Laborit, the scientist who invented artificial hibernation and discovered chlorpromazine.

Professor Wermuths' main research themes focus on the chemistry and the pharmacology of pyridazine derivatives. The 3-aminopyridazine pharmacophore, in particular, allowed him to accede to an impressive variety of biological activities, including antidepressant and anticonvulsant molecules; inhibitors of enzymes such as mono-amine-oxidases, phosphodiesterases and acetylcholinesterase; ligands for neuro-receptors: GABA-A receptor antagonists, serotonine 5-HT_3 receptor antagonists, dopaminergic and muscarinic agonists. More recently, in collaboration with the scientists of the Sanofi Company, he developed potent antagonists of the 41 amino-acid neuropeptide CRF (corticotrophin-releasing factor) which regulates the release of ACTH and thus the synthesis of corticoids in the adrenal glands. Professor Wermuth has also, in collaboration with Professor Jean-Charles Schwartz and Doctor Pierre Sokoloff (INSERM, Paris), developed selective ligands of the newly discovered dopamine D_3 receptor. After a three-year exploratory phase, this research has led to nanomolar partial agonists which may prove useful in the treatment of the cocaine-withdrawal syndrome.

Besides about 300 scientific papers and about 80 patents, Professor Wermuth is co-author or editor of several books including; *Pharmacologie Moléculaire*, Masson & Cie, Paris; *Médicaments Organiques de Synthèse*, Masson & Cie, Paris; *Medicinal Chemistry for the Twenty-first Century*, Blackwell Scientific Publications, Oxford; *Trends in QSAR and Molecular Modeling*, ESCOM, Leyden, two editions of *The Practice of Medicinal Chemistry*, Academic Press, London and *The Handbook of Pharmaceutical Salts, Properties Selection and Use*, Wiley-VCH.

Professor Wermuth was awarded the Charles Mentzer Prize of the Société Française de Chimie Thérapeutique in 1984, the Léon Velluz Prize of the French Academy of Science in 1995, the Prix de l'Ordre des Pharmaciens 1998 by the French Academy of Pharmacy and the Carl Mannich Prize of the German Pharmaceutical Society in 2000. He is Corresponding Member of the German Pharmaceutical Society and was nominated Commandeur des Palmes Académiques in 1995. He has been President of the Medicinal Chemistry Section of the International Union of Pure and Applied Chemistry (IUPAC) from 1988 to 1992 and from January 1998 to January 2000 was President of the IUPAC Division on Chemistry and Human Health.

Section Editors

Michael J. Bowker studied chemistry and received his doctorate in Organic Chemistry from the University of Leeds, UK. After 5 years working for a multinational polymer company, he moved to May & Baker Ltd., a UK subsidiary of Rhône-Poulenc Santé (now Sanofi-Aventis). He was a Director of Analytical Chemistry for about 15 years and, more recently, Director of Preformulation at Aventis Pharma Ltd. He has been intimately involved in preformulation and solid-state activities, on a worldwide basis for more than 15 years. He has published several research papers and one chapter for a book on pharmaceutical salts and is currently a Director of M. J. Bowker Consulting Limited, a small company undertaking consultancy in salt selection, polymorph selection and pharmaceutical preformulation.

Hugo Kubinyi is a Medicinal Chemist with 35 years of industrial experience in drug design, molecular modeling, protein crystallography and combinatorial chemistry, in Knoll and BASF AG, Ludwigshafen. He is a Professor of Pharmaceutical Chemistry at the University of Heidelberg, former Chair of The QSAR and Modelling Society and IUPAC Fellow. From his scientific work resulted more than 100 publications and seven books on QSAR, drug design, chemogenomics, and drug discovery technologies.

John R. Proudfoot received his Ph.D. from University College Dublin, Ireland in 1981 working with Professor Dervilla Donnelly. He completed post doctoral studies with Professor Carl Djerassi at Stanford University and Professor John Cashman at the University of California San Francisco. In 1987, he joined Boehringer Ingelheim and is presently a Distinguished Scientist in the medicinal chemistry department.

Bryan G. Reuben is Professor Emeritus of Chemical Technology at London South Bank University. He has written widely on the technology and economics of the chemical and pharmaceutical industries. His most recent experimental work was on hydrogen–deuterium exchange in protonated peptides and on the downstream processing of nisin.

Richard B. Silverman is the John Evans Professor of Chemistry at Northwestern University. He has published 240 research articles, holds 38 domestic and foreign patents, has written four books, and is the inventor of Lyrica™ (pregabalin), marketed worldwide by Pfizer for refractory epilepsy, neuropathic pain, fibromyalgia, and (in Europe) for generalized anxiety disorder.

David J. Triggle is a SUNY Disinguished Professor and the University Professor State University of New York at Buffalo. Educated in United Kingdom and Canada in physical and organic chemistry he has served a variety roles at Buffalo including Dean of the School of Pharmacy and University Provost. His work has been principally in the area of the chemical pharmacology of drug–receptor and drug–ion channel interactions. He is the author and editor of some 30 books and several hundred publications.

Han van de Waterbeemd studied organic and medicinal chemistry and got his PhD at the University of Leiden. After his academic years at the University of Lausanne with Bernard Testa he worked for 20 years in the pharmaceutical industry for Roche, Pfizer and AstraZeneca. His research interests are in optimizing compound quality using measured and predicted physicochemical and DMPK properties. He contributed to 145 research papers and book chapters, and (co-)edited 13 books.

Contributors

Raffaella. G. Balocco Mattavelli
Manager of the International Nonproprietary
Names Programme
Quality Assurance & Safety: Medicines
World Health Organization
20, av. Appia
CH-1211, Geneva 27

Paul L. Bartel
Myriad Genetics, Inc.
320 Wakara Way
Salt Lake City, UT 84108
USA

Patrick Bazzini
Prestwick Chemical Inc.
Boulevard Gonthier
d'Andernach
67400 Illkirch
France

Frans M. Belpaire
Heymans Institute for Pharmacology
Jeroom Duquesnoylaan 37
9051 Gent
Belgium

Koen Boussery
Laboratory of Medical Biochemistry and Clinical Analysis
Faculty of Pharmaceutical Sciences
Gent University
Harelbekestraat 72
9000 Gent
Belgium

Michael J. Bowker
M.J. Bowker Consulting Ltd.
36, Burses Way
Hutton, Brentwood
Essex CM13 2PS
UK

Sharon D. Bryant
Medicinal Chemistry Group
Laboratory of Pharmacology and Chemistry
National Institute of Environmental Health Sciences
P.O. Box 12233, MD: B3-05
Research Triangle Park, NC 27709
USA

David Cavalla
Arachnova
St. John's Innovation Centre
Cambridge CB4 4WS
UK

François Chast
Pharmacy, Pharmacology, Toxicology Department
Hôtel-Dieu
1, Place du Parvis Notre-Dame
75004 Paris
France

Paola Ciapetti
Head of Medicinal Chemistry
Novalyst Discovery
Boulevard Sébastien Brant BP 30170
F-67405 Illkirch Cedex
France

Jean-Marie Contreras
Prestwick Chemical Inc.
Boulevard Gonthier d'Andernach
67400 Illkirch
France

Gordon M. Cragg
Natural Products Branch
National Cancer Institute
1003 W 7th Street, Suite 206
Frederick, MD 21701
USA

Patrick M. Dansette
Laboratoire de Chimie et Biochimie
Pharmacologiques et Toxicologiques
Université PARIS Descartes
UMR 8601 – CNRS
45, Rue des Saints Pères
F-75270 Paris Cedex 06
France

Ji-Cui Dong
International Nonproprietary Names Programme
Quality Assurance & Safety: Medicines
World Health Organization
20, av. Appia
CH-1211, Geneva 27

Bernard Faller
Novartis Pharma AG
Werk Klybeck
Klybeckstrasse 141
WKL-122.P.33
CH-4057 Basel
Switzerland

Bennett T. Farmer
Boehringer Ingelheim Pharmaceuticals, Inc.
900 Ridgebury Road
P.O. Box 368
Ridgefield, CT 06877
USA

Bruno Galli
Novartis Pharma AG
TRD-PTM WSJ-340-451
Lichtstrasse 35
CH-4056 Basel
Switzerland

Jean-Pierre Gies
Université Louis Pasteur
Faculté de Pharmacie
Equipe de Signalisation Cellulaire
74, Route du Rhin
67401 Illkirch-Cedex,
France

Bruno Giethlen
Prestwick Chemical Inc.
Boulevard Gonthier d'Andernach
67400 Illkirch
France

Fumitoshi Hirayama
Faculty of Pharmaceutical Sciences
Sojo University
4-22-1 Ikeda
Kumamoto 860-0082
Japan

Adrian N. Hobden
Myriad Genetics, Inc.
320 Wakara Way
Salt Lake City, UT 84108
USA

Andrew L. Hopkins
Division of Biological Chemistry and Drug Discovery
College of Life Sciences
University of Dundee
Dundee
Scotland DD1 5EH
UK

Peter Imming
Institut für Pharmazie
Martin-Luther-Universitaet Halle-wittenberg Wolfgang-Langenbeck-Str. 4
06120 Halle (Saale) Germany

Paul F. Jackson
Johnson & Johnson
Pharmaceutical R&D, L.L.C.
Welsh McKean Roads
P.O. Box 776
Spring House, PA 19477
USA

David G. I. Kingston
Virginia Polytechnic Institute & State University
Department of Chemistry, M/C 0212
3111 Hahn Hall
West Campus Drive
Blacksburg, VA 24061
USA

Sabine Kopp
Medicines Quality Assurance Programme
Quality Assurance & Safety: Medicines
World Health Organization
20, av. Appia
CH-1211 Geneva 27

Hugo Kubinyi
Donnersbergstrasse 9
67256 Weisenheim am Sand
Germany

Kamal Kumar
Max Planck Institute of Molecular Physiology
Otto-Hahn-Str. 11
D-44227 Dortmund
Germany

Yves Landry
Université Louis Pasteur
Faculté de Pharmacie
Equipe de Signalisation Cellulaire
74, Route du Rhin
67401 Illkirch-Cedex,
France

Thierry Langer
Inte:Ligand GmbH
Clemens Maria Hofbauer-G.6
2344 Maria Enzersdorf
Austria

Institute of Pharmacy
University of Innsbruck
Innrain 52
6020 Innsbruck
Austria

Sophie Lasseur
International Nonproprietary Names Programme
Quality Assurance & Safety: Medicines
World Health Organization
20, av. Appia
CH-1211, Geneva 27

Christopher A. Lipinski
Melior Discovery
10 Conshire Drive
Waterford, CT 06385-4122
USA

Anne-Christine Macherey
Unité de Prévention du Risque Chimique
UPS 831–Bat.11
CNRS
Avenue de la Terrasse
F-91198 Gif sur Yvette Cedex
France

André Mann
Département de Pharmacochimie de la Communication Cellulaire
UMR 7175 LC 1 ULP/CNRS
Faculté de Pharmacie
74 route du Rhin
67401 Illkirch
France

Christophe Morice
Prestwick Chemical Inc.
Boulevard Gonthier
d'Andernach
67400 Illkirch
France

Richard Morphy
Organon Laboratories Ltd.
A part of the Schering Plough Corporation
Newhouse
Lanarkshire
Scotland ML1 5SH
UK

David J. Newman
Natural Products Branch
National Cancer Institute
1003 W 7th Street, Suite 206
Frederick, MD 21701
USA

Jean-Pierre Nowicki
Sanofi-Aventis RD
31, Avenue Paul Vaillant-Couturier
92220 Bagneux
France

Alex Polinsky
Research Technologies
Pfizer Global Research and Development
620 Memorial Drive
Cambridge, MA 02138
USA

John R. Proudfoot
Boehringer Ingelheim
Pharmaceuticals Inc.
900 Ridgebury Road
P.O. Box 368
Ridgefield, CT 06877
USA

Z. Rankovic
Organon Laboratories Ltd.
A part of the Schering Plough Corporation
Newhouse
Lanarkshire
Scotland ML1 5SH
UK

Allen B. Reitz
Johnson & Johnson
Pharmaceutical Research and Development, LLC
Welsh McKean Rds.
Spring House, PA 19477
USA

Bryan G. Reuben
London South Bank University
24 Claverley Grove
London N3 2DH
UK

Jean-Michel Rondeau
Novartis Pharma AG
Novartis Institutes for BioMedical Research
WSJ-88.8.08A
CH-4056 Basel
Switzerland

Sally Rose
Cresset BioMolecular Discovery Ltd
BioPark Hertfordshire
Broadwater Road
Welwyn Garden City
Herts., AL7 3AX
UK

Bernard Scatton
Sanofi-Aventis RD
31, Avenue Paul Vaillant-Couturier
92220 Bagneux
France

Laurent Schaeffer
Prestwick Chemical Inc.
Boulevard Gonthier
d'Andernach
67400 Illkirch
France

Jean-Michel Scherrmann
INSERM U 705; CNRS 7157
University Paris Descartes and Paris Diderot
Department of Pharmacokinetics Faculty of Pharmacy
4, avenue de l'Observatoire
75006 Paris
France

Herman Schreuder
Aventis Pharma Deutschland GmbH
Building G 6865A
D-65926 Frankfurt am Main
Germany

Brian C. Shook
Johnson & Johnson
Pharmaceutical R&D, L.L.C.
Welsh McKean Roads
P.O. Box 776
Spring House, PA 19477
USA

Richard B. Silverman
Department of Chemistry
Northwestern University
2145, Sheridon Road
Evanston, IL 60208-3113
USA

Wolfgang Sippl
Department of Pharmaceutical Chemistry
Martin-Luther-Universität Halle-Wittenberg
Wolfgang-Langenbeck-Str. 4
06120 Halle (Saale)
Germany

Maria Souleau
Sanofi-Aventis
20, Rue Raymond Aron
92160 Antony
France

P. Heinrich Stahl
Lerchenstrasse 28
79104 Freiburg im Breisgau
Germany

Bernard Testa
Service de Pharmacie, CHUV
Centre Hospitalier Universitaire Vaudois
Rue du Bugnon 46
CH-1011 Lausanne
Switzerland

David J Triggle
SUNY at Buffalo
School of Pharmaceutical Sciences
126 Cooke Hall
Buffalo, NY 14260
USA

Kaneto Uekama
Faculty of Pharmaceutical Sciences
Sojo University
4-22-1 Ikeda
Kumamoto 860-0082
Japan

Johan Van de Voorde
Ghent University
Vascular Research Unit
De Pintelaan 185 – Blok B
9000 Gent
Belgium

Han van de Waterbeemd
AstraZeneca
LG DECS, Global Compound Sciences
Alderley Park, 50S39
Macclesfield
Cheshire SK10 4TG
UK

Herbert Waldmann
Max Planck Institute of Molecular Physiology
Otto-Hahn-Str. 11
D-44227 Dortmund
Germany

Camille G. Wermuth
Prestwick Chemical Inc.
Boulevard Gonthier d'Andernach
67400 Illkirch
France

Stefan Wetzel
Max Planck Institute of Molecular Physiology
Otto-Hahn-Str. 11
D-44227 Dortmund
Germany

Kenton H. Zavitz
Myriad Genetics, Inc.
320 Wakara Way
Salt Lake City, UT 84108
USA

Preface to the First Edition

The role of chemistry in the manufacture of new drugs, and also of cosmetics and agrochemicals, is essential. It is doubtful, however, whether chemists have been properly trained to design and synthesize new drugs or other bioactive compounds. The majority of medicinal chemists working in the pharmaceutical industry are organic synthetic chemists with little or no background in medicinal chemistry who have to acquire the specific aspects of medicinal chemistry during their early years in the pharmaceutical industry. This book is precisely aimed to be their 'bedside book' at the beginning of their career.

After a concise introduction covering background subject matter, such as the definition and history of medicinal chemistry, the measurement of biological activities and the three main phases of drug activity, the second part of the book discusses the most appropriate approach to *finding a new lead compound or an original working hypothesis*. This most uncertain stage in the development of a new drug is nowadays characterized by high-throughput screening methods, synthesis of combinatorial libraries, data base mining and a return to natural product screening. The core of the book (Parts III to V) considers the *optimization of the lead in terms of potency, selectivity, and safety*. In 'Primary Exploration of Structure-Activity Relationships', the most common operational stratagems are discussed, allowing identification of the portions of the molecule that are important for potency. 'Substituents and functions' deals with the rapid and systematic optimization of the lead compound. 'Spatial Organization, Receptor Mapping and Molecular Modelling' considers the three-dimensional aspects of drug-receptor interactions, giving particular emphasis to the design of peptidomimetic drugs and to the control of the agonist- antagonist transition. Parts VI and VII concentrate on the definition of satisfactory drug-delivery conditions, i.e. means to ensure that the molecule reaches its target organ. Pharmacokinetic properties are improved through adequate chemical modifications, notably prodrug design, obtaining suitable water solubility (of utmost importance in medical practice) and improving organoleptic properties (and thus rendering the drug administration acceptable to the patient). Part VIII, 'Development of New Drugs: Legal and Economic Aspects', constitutes an important area in which chemists are almost wholly self taught following their entry into industry.

This book fills a gap in the available bibliography of medicinal chemistry texts. There is not, to the author-editor's knowledge, any other current work in print which deals with the practical aspects of medicinal chemistry, from conception of molecules to their marketing. In this single volume, all the disparate bits of information which medicinal chemists gather over a career, and generally share by word-of-mouth with their colleagues, but which have never been organized and presented in coherent form in print, are brought together. Traditional approaches are not neglected and are illustrated by modern examples and, conversely, the most recent discovery and development technologies are presented and discussed by specialists. Therefore, *The Practice of Medicinal Chemistry* is exactly the type of book to be recommended as a text or as first reading to a synthetic chemist beginning a career in medicinal chemistry. And, even if primarily aimed at organic chemists entering into pharmaceutical research, all medicinal chemists will derive a great deal from reading the book.

The involvement of a large number of authors presents the risk of a certain lack of cohesiveness and of some overlaps, especially as each chapter is written as an autonomic piece of information. Such a situation was anticipated and accepted, especially for a first edition. It can be defended because each contributor is an expert in his/her field and many of them are 'heavyweights' in medicinal chemistry. In editing the book I have tried to ensure a balanced content and a more-or-less consistent style. However, the temptation to influence the personal views of the authors has been resisted. On the contrary, my objective was to combine a plurality of opinions, and to present and discuss a given topic from different angles. Such as it is, this first edition can still be improved and I am grateful in advance to all colleagues for comments and suggestions for future editions.

Special care has been taken to give complete references and, in general, each compound described has been identified by at least one reference. *For compounds for which no specific literature indication is given, the reader is referred to the Merck Index*.

The cover picture of the book is a reproduction of a copperplate engraving designed for me by the late Charles Gutknecht, who was my secondary school chemistry teacher in Mulhouse. It represents an extract of Brueghel's engraving *The alchemist ruining his family in pursuing his chimera*, surmounted by the aquarius symbol. Represented on the left-hand side is my lucky charm caster oil plant (*Ricinus communis L., Euphorbiaceae*), which was the starting point of the pyridazine chemistry in my laboratory. The historical cascade of events was as follows: cracking of caster oil produces n-heptanal and aldolization of

n-heptanal – and, more generally, of any enolisable aldehyde or ketone – with pyruvic acid leads to a-hydroxy-γ-ketonic acids. Finally, the condensation of these keto acids with hydrazine yields pyrodazones. Thus, all our present research on pyridazine derivatives originates from my schoolboy chemistry, when I prepared in my home in Mulhouse n-heptanal and undecylenic acid by cracking caster oil!

Preparing this book was a collective adventure and I am most grateful to all authors for their cooperation and for the time and the effort they spent to write their respective contributions. I appreciate also their patience, especially as the editing process took much more time than initially expected.

I am very grateful to Brad Anderson (University of Utah, Salt Lake city), Jean-Jacques André (Marion Merrell Dow, Strasbourg), Richard Baker (Eli Lilly, Erl Wood, UK), Thomas C. Jones (Sandoz, Basle), Isabelle Morin (Servier, Paris), Bryan Reuben (London South Bank University) and John Topliss (University of Michigan, Ann Arbor) for their invaluable assistance, comments and contributions.

My thanks go also to the editorial staff of Academic Press in London, Particularly to Susan Lord, Nicola Linton and Fran Kingston, to the two copy editors Len Cegielka and Peter Cross, and finally, to the two secretaries of our laboratory, Franqois Herth and Marylse Wernert.

Last but not least, I want to thank my wife Renée for all her encouragement and for sacrificing evenings an Saturday family life over the past year and a half, to allow me to sit before my computer for about 2500 hours!

Camille G. Wermuth

Preface to the Second Edition

Like the first edition of *The Practice of Medicinal Chemistry* (nicknamed 'The Bible' by medicinal chemists) the second edition is intended primarily for organic chemists beginning a career in drug research. Furthermore, it is a valuable reference source for academic, as well as industrial, medicinal chemists. The general philosophy of the book is to complete the biological progress – Intellectualization at the level of function using the chemical progress Intellectualization at the level of structure (Professor Samuel J. Danishevsky, *Studies in the chemistry and biology of the epothilones and eleutherobins*, Conference given at the XXXIVémes Rencontres Internationales de Chimie Th6rapeutique, Facult6 de Pharmacie, Nantes, 8–10 July, 1998).

The recent results from genomic research have allowed for the identification of a great number of new targets, corresponding to hitherto unknown receptors or to new subtypes of already existing receptors. The massive use of combinatorial chemistry, associated with high throughput screening technologies, has identified thousands of hits for these targets. The present challenge is to develop these hits into usable and useful drug candidates. This book is, therefore, particularly timely as it covers abundantly the subject of drug optimization.

The new edition of the book has been updated, expanded and refocused to reflect developments over the nine years since the first edition was published. Experts in the field have provided personal accounts of both traditional methodologies, and the newest discovery and development technologies, giving us an insight into diverse aspects of medicinal chemistry, usually only gained from years of practical experience.

Like the previous edition, this edition includes a concise introduction covering the definition and history of medicinal chemistry, the measurement of biological activities and the three main phases of drug activity. This is followed by detailed discussions on the discovery of new lead compounds including automated, high throughput screening techniques, combinatorial chemistry and the use of the internet, all of which serve to reduce pre-clinical development times and, thus, the cost of drugs. Further chapters discuss the optimization of lead compounds in terms of potency, selectivity, and safety; the contribution of genomics; molecular biology and X-ray crystallization to drug discovery and development, including the design of peptidomimetic drugs; and the development of drug-delivery systems, including organ targeting and the preparation of pharmaceutically acceptable salts. The final section covers legal and economic aspects of drug discovery and production, including drug sources, good manufacturing practices, drug nomenclature, patent protection, social-economic implications and the future of the pharmaceutical industry.

I am deeply indebted to all co-authors for their cooperation, for the time they spent writing their respective contributions and for their patience during the editing process. I am very grateful to Didier Rognan, Paola Ciapetti, Bruno Giethlen, Annie Marcincal, Marie-Louise Jung, Jean-Marie Contreras and Patrick Bazzini for their helpful comments.

My thanks go also to the editorial staff of *Academic Press* in London, particularly to Margaret Macdonald and Jacqueline Read. Last but not least, I want to express my gratitude to my wife Renée for all her encouragements and for her comprehensiveness.

Camille G. Wermuth

Preface to the Third Edition

Like the preceding editions of this book, this third edition treats of the essential elements of medicinal chemistry in a unique volume. It provides a practical overview of the daily problems facing medicinal chemists, from the conception of new molecules through to the production of new drugs and their legal/economic implications. This edition has been updated, expanded and refocused to reflect developments in the past 5 years, including 11 new chapters on topics such as hit identification methodologies and cheminformatics. More than 50 experts in the field from eight different countries, who have benefited from years of practical experience, give personal accounts of both traditional methodologies and the newest discovery and development technologies, providing readers with an insight into medicinal chemistry.

A major change in comparison to the previous editions was the decision to alleviate my editorial burden in sharing it with seven section editors, each being responsible for one of the eight sections of the book. I highly appreciated their positive and efficacious collaboration and express them my warmest thanks (in the alphabetical order) to Michael Bowker, Hugo Kubinyi, John Proudfood, Bryan Reuben, Richard Silverman, David Triggle and Han van de Waterbeemd.

Another change was the decision taken by Elsevier/Academic Press to publish the book in full colors thus rendering it more pleasant and user-friendly. I take this occasion to thank Keri Witman, Pat Gonzales, Kirsten Funk and Renske van Dijk for having successively ensured the editorial development of the book. Taking into account that we had to work with a cohort of about 50 authors, each of them having his personality, his original approach and his main busy professional live, this was not an easy task. I am deeply indebted to my assistant Odile Blin for the way she had mastered, efficiently and with friendliness, all the secretarial work and particularly the contacts with the different authors and with the Elsevier development editors. As for the earlier editions, I also want to express my gratitude to my wife Renée and my daughters Delphine, Joëlle and Séverine for all their encouragements and for sacrificing many hours of family life in order to leave me enough free time to edit this new version of the "Medicinal Chemist's Bible."

My final thoughts go to the future readers of the book, and especially to the newcomers in Medicinal Chemistry having the curiosity to read the preface. I cannot resist giving them some advice for doing good science.

First of all, be open-minded and original. As Schopenhauer noted, the task of the creative mind is "not so much to see what no one has seen yet; but to think what nobody has thought yet, about what everyone sees." A wonderful illustration is found in Peter Hesse's cartoon below.

" IT'S CALLED FIRE... IT RECYCLES WOOD."

Second, always keep in mind that the object of Medicinal Chemistry is to synthetize new drugs useful for suffering patients. Like many scientists, medicinal chemists, have to navigate between two tempting reefs. On one side they should avoid doing "NAAR": non-applicable applied research, on the other side they may be attracted by "NFBR": non-fundamental basic search."

Third, convinced as they may be that the neighbors grass is always greener, they may be attracted to start their research in using as a hit a recently published competitor's product. In fact, the published compound may exhibit only a weak activity, therefore be very careful when starting a new program and never forget that the worst thing a medicinal chemist can do is to prepare a me-too of an inactive compound!

Camille G. Wermuth

Part I

General Aspects of Medicinal Chemistry

Hugo Kubinyi

Section Editor

Chapter 1

A History of Drug Discovery

From first steps of chemistry to achievements in molecular pharmacology

François Chast

I. INTRODUCTION
A. The renewal of chemistry
B. The dawn of the organic chemistry crosses the birth of biology
II. TWO HUNDRED YEARS OF DRUG DISCOVERIES
A. Pain killers: best-sellers and controversies
B. Giving back the heart its youth
C. Fight against microbes and viruses
D. Drugs for immunosuppression
E. Contribution of chemists to the fight against cancer
F. Drugs for endocrine disorders
G. Anti-acid drugs
H. Lipid lowering drugs
I. From neurotransmitters to receptors
J. Drugs of the mind
III. CONSIDERATIONS ON RECENT TRENDS IN DRUG DISCOVERY
A. From genetics to DNA technology
B. Hopes and limits for drug hunting
REFERENCES

Le médicament place l'organisme dans des conditions particulières qui en modifient heureusement les procédés physiques et chimiques lorsqu'ils ont été troublés.

Claude Bernard*

During more than 2,000 years, Hippocratic medical tradition weighed on the development of a modern medicine and a renewed approach of the treatment of diseases. The basis for the use of drugs remained founded on empirical theories linked to the equilibrium of body's "humors" consisting in sanguine, melancholic, phlegmatic and choleric. Health and disease were seen as a question of balance or imbalance with foods and herbs classified according to their ability to affect natural homeostasis. Later, during the Middle Ages, Muslim world made significant contributions to medicine and a major medical advance was the founding of many hospitals and university medical schools.

Before the 1800s, pharmacy remained an empiric science, guided by traditional medicine, inherited from "Ancients." Numerous drugs, most of them being prepared with plant extracts, (Figure 1.1) sometimes efficacious, were available. But none of them could respond to a chemical definition of what we call today a drug, except drugs coming from mineral reign.

The technology of making drugs was crude at best: tinctures, poultices, soups, and infusions were made with water- or alcohol-based extracts of freshly ground or dried herbs or animal products such as bone, fat, or even pearls, and sometimes from minerals best left in the ground.[1]

The objective of this first chapter is to offer a presentation of the fabulous history of drug discoveries, from traditional pharmacy emerged from ethnopharmacy, till the recent

*Leçons sur les Effets de Substances Médicamenteuses et Toxiques (1857) deuxième leçon (5 mars 1856), p.38: "Drugs place the body in particular conditions which modify fortunately the physical and chemical processes when they have been disturbed."

Copyright © 2008, Elsevier Ltd
All rights reserved.

FIGURE 1.1 Opium latex flowing out of poppy.

concepts of drug design, production and development, born from molecular genetics and molecular pharmacology.

Of course, it is not possible to describe exhaustively, in such a short chapter, such a complex and diversified history. We made the choice to describe the evolution of few families of drugs as examples of mankind ingenuity and intelligence to make pharmaceutical progress more and more successful in treating or preventing diseases.

I. INTRODUCTION

A. The renewal of chemistry

The 18th century concluded its progress in chemistry with an enthusiastic environment. Joseph Priestley in the United Kingdom, Carl Wilhelm Scheele in Sweden, Antoine Laurent de Lavoisier in France,[2] gave a precise signification to the chemical reactivity and promoted a large number of substances to the statute of chemical reagents. Scheele and Priestley prepared and studied oxygen. Both of them discovered nitrogen as a constituent of air, carbon monoxide, ammonia, and several other gases ; manganese, barium and chlorine; isolated glycerin and many acids, including tartaric, lactic, uric, prussic, citric, and gallic. Lavoisier is generally considered as the founder of modern chemistry as creating the oxygen theory of combustion.[3] He should be known as one of the most astonishing 18th century "men of the Enlightenment," the founder of modern scientific experimental methodology. By formulating the principle of the conservation of mass, he gave a clear differentiation between elements and compounds, something so important for pharmaceutical chemistry. Few years later, Antoine François de Fourcroy, Louis Nicolas Vauquelin, Joseph Louis Proust, Jöns Jakob Berzelius, Louis-Joseph Gay-Lussac, and Humphrey Davy introduced new concepts in chemistry. Those scientists integrated the practical advancements of a new generation of experimenters. All these industrial innovations would have their own impact on other developments in industrial and then medicinal chemistry.[4] At the turn of the 19th century, as the result of a scientific approach, drugs are becoming an industrial item. Claude Louis Berthollet began the industrial exploitation of chlorine (1785). Nicolas Leblanc prepared sodium hydroxide (1789) and then, bleach (1796). Davy performed electrolysis and distinguished between acids and anhydrides. Louis Jacques Thénard prepared hydrogen peroxide and Antoine Jérôme Balard discovered bromide (1826). The growing of therapeutic resources was mainly due to the mastery of chemical or physico-chemical principles proposed by Gay-Lussac and Justus Von Liebig.[5] This chemists' generation, by realizing all these discoveries, established the compost of the therapeutic discoveries of the 19th century. The constitution of chemistry as a scientific discipline found a new turn few decades later by crossing the road of biology which included revolutionary works of Claude Bernard,[6] Rudolph Virchow,[7] and Louis Pasteur.[8] Besides these fundamental sciences, physiology, biochemistry, or microbiology were becoming natural tributaries of the outbreak of pharmacology. Thus, rational treatments were about to be designed on the purpose of new knowledge in various clinical or fundamental fields. After a period characterized by extraction and purification from natural materials (mainly plants), drugs would be synthesized in chemical factories or prepared through biotechnology (fermentation or gene technology) after a rational research, design and development in research laboratories. Whereas the purpose was to isolate active molecules from plants during the first half of the 19th century, the birth of organic chemistry following charcoal and oil industries, progressively led chemists and pharmacists toward organic synthesis performed in what would be called "laboratory" a new concept created by this generation of scientists. Even when those laboratories hosted discoveries like active principles extracted from plants, progresses in drug compounding and packaging made irreversible industrialization processes. At the same time, the economical dimension of growing pharmaceutical industry transformed drugs as strategic items, mainly when it could interfere with military processes, for instance during colonial expeditions.

The "modern" word "*pharmacology*" became more and more often used by physicians after the works of François Magendie (Figure 1.2) in France or Oscar Schmiedeberg in Germany. Progressively a clear dichotomy took place between those two entities. *Materia Medica* considered drugs with a static and conservative view as for their

FIGURE 1.2 François Magendie.

FIGURE 1.3 Friedrich Wöhler.

production and the compounding of medicines. It was somewhere considered as the natural history of drugs. At the contrary, pharmacology was embracing the creation of drugs through a more dynamic point of view, studying drugs with respect of their site and mechanism of action. At the same time, medicinal chemistry was becoming the application of chemical research techniques to the synthesis of new pharmaceuticals. During the early stages of medicinal chemistry development, chemists were primarily concerned with the isolation of medicinal agents found in plants. Today, in this field they are also equally concerned with the creation of new synthetic drug compounds. As a constant, medicinal chemistry is almost always geared toward drug discovery and development.

B. The dawn of the organic chemistry crosses the birth of biology

A radical turn in the development of new chemicals occurred when charcoal and then oil distillation offered so many opportunities. After the extract of paraffin, carbon derivatives chemistry knew considerable developments with a lot of industrial consequences during the second third of the century. The first organic molecules used for their therapeutic properties had acyclic structures: chloroform was discovered in 1831 by three independently working chemists: Eugene Soubeiran of France (1831),[9] Justus Von Liebig of Germany,[10] and Samuel Guthrie of the United States (1832).[11] Von Liebig taught chemistry through books like *Organic Chemistry and its Application to Agriculture and Physiology* (1840), and *Organic Chemistry in its Application to Physiology and Pathology* (1842)[12] and editing the journal that was to become the preeminent chemistry publication in Europe: *Annalen der Chemie und Pharmazie*. Liebig and Friedrich Wöhler (Figure 1.3) began in 1825 various studies over two substances that had apparently the same composition – cyanic acid and fulminic acid – but very different characteristics. The silver compound of fulminic acid, investigated by Liebig was explosive; whereas Wöhler's silver cyanate was not. These substances, called "isomers" by Berzelius, lead chemists to suspect that substances were defined not simply by the number and kind of atoms in the molecule but also by the arrangement of those atoms.

The most famous creation of an isomeric compound was Wöhler's "accidental" synthesis of urea (1828), when failing to prepare ammonium cyanate. For the first time someone prepared an organic compound by the means of inorganic ones.[13] That "incident" made Wöhler saying: "I can no longer, so to speak, hold my chemical water and must tell you that I can make urea without needing a kidney, whether of man or dog; the ammonium salt of cyanic acid is urea".[14] Liebig and Wöhler's original objective was to interpret radicals as organic chemical equivalents of inorganic atoms. It was an early step along the path to structural chemistry. Organic chemistry precipitously entered the medicinal arena in 1856 when the youngster William Perkin, in an unsuccessful attempt to synthesize quinine, stumbled upon mauveine, the first synthetic dye, leading to the development of many other synthetic dyes, which will

give birth few decades later to the first antiseptic and anti-infectious drugs. Indeed, industrial world understood that some of these dyes could have therapeutic effects.

Synthetic dyes, and especially their medical "side effects," helped to put Germany and Switzerland in the forefront of both organic chemistry and synthesized drugs. The dye–drug connection began to be a very prolific way to discover drugs. After the first developments in organic chemistry during the first half of the 19th century, the question of the chemical origin of life was clearly put in the forefront of the scientific debate. Since Wöhler's works, it was clear that chemistry was a unique science, with the same rules governing reactions kinetics and atomic, radical, or molecular arrangements. A characteristic of the way to continue on discovery pathway was a beginning of scientific cooperation meaning as well muldisciplinary approaches as more curiosity from scientists taking here and there the knowledge necessary to understand natural or experimental phenomena. As an example, Louis Pasteur, the French emblematic physicist and chemist after beginning his career as a specialist in crystallography, studied the impact of bacteria on stereochemical properties of tartaric acid crystals, and after productive research on alcoholic and acetic fermentations, put the concept of spontaneous generation to pieces. As bacteria could react on organic substances, he presumed that they also could be active on living beings.

BOX 1.1 Alkaloids

The first alkaloid ever isolated, emetine, was found by Pierre Joseph Pelletier, the pharmacist, and François Magendie, the physician, in the traditional Ipecacuanha (1817).[16] The same year, Joseph Pelletier and Joseph Bienaymé Caventou extracted strychnine, a powerful neurostimulating agent, from Strychnos. Three years later (1820) they extracted quinine from various Cinchona species.[17] Pelletier and Caventou began an industrialization of quinine production, the drug being more and more popular as a tonic and antifever drug (before being recognized as a treatment of choice for malaria). An impressive cohort of alkaloids would be extracted in the following years. Brucine (1819), caffeine (1819), colchicine (1820), codeine (1832)[18], atropine (1833)[19], papaverine (1848)[20] were subsequently obtained. Coniine, extracted in 1826, was the first alkaloid to have its structure established (Schiff, 1870) and to be synthesized (Ladenburg, 1889)[21], but for others, such as colchicine, it was well over a century before the structures were finally elucidated. Between the years 1817 and 1850, a new generation of scientists gave rise to a new relationship between medicine and these new therapeutic tools. Nevertheless, in the first two-thirds of the 19th century, pure alkaloids were seldom used, even if the first medical textbook presenting alkaloids source of drugs the "Formulaire des médicaments" by François Magendie, where he tries to make more popular the use of morphine, and fought against old formulas[22] was published in 1822.

II. TWO HUNDRED YEARS OF DRUG DISCOVERIES

Besides conceptual progresses, the formal evolution in the concept of medicines was based on the radical transformation of the nature of medicines. One of the theorists of this trend, Charles Louis Cadet de Gassicourt,[15] reported in the inaugural issue of the *Bulletin de Pharmacie* (1809) that the use of complex preparations had to be withdrawn in favor of pure substances. Pharmacist and physicians had, first, to classify drugs and their use. This trend was much more convenient with pure substances. Between 1815 and 1820, the first active principles were isolated from plants. At that time, a new era in pharmaceutical chemistry opened. Hereafter, drug activity would not depend on the quality of extracts or tinctures and their inherent variability in active principles. The only variability acceptable in therapeutics would be the patient himself.

A. Pain killers: best-sellers and controversies (Box 1.1)

1. *Poppy extracts led to brain receptors*

The first controversy is to know who discovered morphine. Jean-Francois Derosne,[23] in Paris, prepared a crude extract of opium (with alcohol and water), and obtained, after potassium carbonate precipitation, what he called "sel de Derosne." Derosne's alkaloidal fraction lacked narcotic properties and was probably largely made of narcotine (also known as noscapine), perhaps mixed with meconic acid. This work, has been presented at the Institute of France in 1804, but only published in 1814.[24] It describes the isolation of a compound, but did not report any animal or human experiment. A young German apothecary from Paderborn (Germany), Friedrich Sertürner did, in fact, begin publishing on opium in 1805,[25] and claimed to have begun work before a paper on opium by Derosne had appeared in 1804. This claim has been interpreted to mean that Sertürner began work in 1803. However, Sertürner's earlier work fixated on acid constituents of opium. Thus, his 1806 paper[26] is mainly concerned with the constituent we now know as meconic acid. It was only in 1817 that he unequivocally reported the isolation of pure morphine.[27] He prepared it by extracting opium with hot water and precipitating morphine with ammonia. He obtained colorless crystals, poorly soluble in water, but soluble in acids and alcohol. He then established that the crystals carried the pharmacological activity of opium. The name "morphine" has been coined later. The discovery was received by great perplexity: morphine had an alkaline reaction toward litmus paper. The scientific world was doubtful and Pierre Jean Robiquet performed new experiments in order to check Sertürner results. For the first time a substance extracted from a plant was not an acid!

BAYER
PHARMACEUTICAL PRODUCTS.
We are now sending to Physicians throughout the United States literature and samples of
ASPIRIN
The substitute for the Salicylates, agreeable of taste, free from unpleasant after effects.
HEROIN
The Sedative for Coughs,
HEROIN HYDROCHLORIDE
Its water-soluble salt.
You will have call for them. Order a supply from your jobber.
Write for literature to
FARBENFABRIKEN OF ELBERFELD CO.
40 Stone Street, New York,

FIGURE 1.4 Aspirin and Heroin "co-advertising".

Gay-Lussac finally accepted the revolutionary idea that alkaline drugs could be found in plants. All alkaline substances isolated in plants would be given a name with the suffix "-ine" (Wilhelm Meissner, 1818) in order to remind the basic reaction of all these drugs. Morphine gained wide medical use in the beginning of the 1860s during the American Civil War, but many injured soldiers returned from the war as morphine addicts, victims of the "soldiers' disease." In 1874, English researcher, C. R. Alder Wright (Saint Mary's Hospital, London) first synthesized (diacetylmorphine) by boiling morphine acetate over a stove.

Twenty years later, Heinrich Dreser working for the Bayer Company of Elberfeld, Germany, found (erroneously) that diluting morphine with acetyls produced a drug without the common morphine side effects. In 1895, Bayer began the production of diacetylmorphine and coined the name "heroin" and introduced it, commercially, after another three years (Figure 1.4). At the beginning of the 20th century, heroin addiction rose to alarming rates driving United Kingdom, United States and France to ban opium and opiate drugs. During next 70 years, morphine will be almost completely withdrawn from medical use, before its "rehabilitation" that came through the so-called *Hospice movement*, founded in the United Kingdom in order to alleviate suffering of dying patients within hospitals.

Candace Pert, together with Solomon Snyder (Johns Hopkins, Baltimore, USA), first identified opioid receptors in the brain in 1972[28] (Figure 1.5). In 1975 Hans Kosterlitz and John Hughes (Aberdeen, UK) reported the existence of an endogenous morphine-like substance[29] and named it *enkephalin* (for "in the head"). Enkephalins, endorphins, and dynorphins bind to specific receptor sites in the brain.

FIGURE 1.5 Soloman Snyder (left) and Candace Pert (right).

Scientific studies of opioid neurotransmitters during the 1970s have uncovered a complex and subtle system that exhibited impressive diversity in terms of endogenous ligands for only three major receptors. The opioid peptide precursors were subject to complex post-translational modifications resulting in the synthesis of multiple active peptides all of them sharing the common N-terminal sequence of Tyr-Gly-Gly-Phe-(Met or Leu), which has been termed the *opioid motif*. Based on the results of theses studies, the endogenous opioids have been implicated in circuits involved in the control of sensation, emotion, and affect and a role has been ascribed to them in addiction, not only to opiates such as morphine or heroin, but also to alcohol.[30]

2. *Aspirin and* NSAIDs

Another active principle soon extracted from plants was salicylic acid. Salicin, extracted from the willow tree, has been launched in 1876 by a Scottish physician, Thomas John McLogan[31]. It was in extensive competition with Cinchona bark and quinine and never became a very popular treatment for fever or rheumatic symptoms. The Italian chemist Raffaele Piria, after having isolated salicylaldehyde (1839)[32] in *Spireae species*, prepared salicylic acid from salicin in Dumas' laboratory in the Sorbonne, Paris. This acid was easier to use and was an ideal step before future syntheses. Its structure was closely related to benzoic acid, an effective preservative useful as an intestinal antiseptic for instance in typhoid fever. Acetylsalicylic acid has been first synthesized by Charles Frederic Gerhardt in 1853[33] and then, in a purer form, by Johann Kraut (1869). Acetylsalicylic acid synthesis with carbolic acid and carbon dioxide was improved by Hermann Kolbe in1874, but in fact nobody noticed its pharmacological interest.

During the 1880s and 1890s, physicians became intensely interested in the possible adverse effects of fever on the human body and the use of antipyretics became one of the hottest fields in therapeutic research. The name of Arthur Eichengrün, who performed the research and development-based pharmaceutical division where Felix Hoffmann worked, and Heinrich Dreser (Figure 1.6) in charge of testing the drug with Kurt Witthauer and Julius Wohlgemuth are to be memorized for this historical discovery (1897).

FIGURE 1.6 Heinrich Dreser.

It is likely that acetylsalicylic acid was synthesized under Arthur Eichengrün's direction and that it would not have been introduced in 1899 without his intervention.[34] Dreser carried out comparative studies of aspirin and other salicylates to demonstrate that the former was less noxious and more beneficial than the latter.[35] Bayer built his fortune upon this drug which received the name of "Aspirin," the most familiar drug name. For the first time, an industrial group illustrated the close relationship between chemistry and practical therapeutics. It was not until the late 1970s that aspirin's ability to inhibit prostaglandins production by the cyclo-oxygenase enzymes was identified as the basis of its therapeutic activity. Prostaglandins are known as end-products of the so-called arachidonic acid cascade. Arachidonic acid is normally stored in membrane-bound phospholipids and released by the action of phospholipases. Enzymatic conversion of released arachidonic acid into biologically active derivatives proceeds through several routes. First, cyclo-oxygenase converts arachidonic acid to unstable cyclic endoperoxides from which prostaglandins, prostacyclin and thromboxanes are derived.[36] Second, the production of the leukotrienes from arachidonic acid is initiated by the action of 5-lipoxygenase producing leukotrienes which are also believed to play an important pathophysiological role in allergic broncho-constriction of asthma. Through pharmacological intervention in the arachidonic acid cascade various anti-inflammatory agents have been developed. These include aspirin-like drugs, which inhibit cyclo-oxygenase. Corticosteroids appear to indirectly inhibit phospholipases thus preventing release of arachidonic acid. Future progress in this field is likely to produce drugs which antagonize arachidonic acid derivatives or inhibit the enzymes involved in their synthesis with greater specificity.[37] Using an ingenious "real time" biological assay of bloodstream hormones irrigating an isolated organ, called the "blood-bathed organ cascade," John Vane

FIGURE 1.7 John Vane.

(Figure 1.7) developed a system for highly sensitive monitoring of several mediators like angiotensin, bradykinin and prostaglandins and discovered prostacyclin, a potent platelet aggregation inhibitor. John Vane explained anti-inflammatory drugs effects (among which aspirin remains a true leader) through their activity on cyclo-oxygenase and inhibition of prostacyclin and thromboxane production. The impact of aspirin administration at low dose for the prevention of stroke or coronary attack resulted from its effect on enzymes regulating the production of prostaglandins.

Vane then assigned a major physiological function to the vascular endothelium which became a pharmacological target for new drugs. He won Albert Lasker Prize in 1977 and Nobel Prize in medicine and physiology (with Sune Bergström and Bengt I. Samuelson) in 1982.[38]

3. *Controversies over "coxibs"*

Another cyclo-oxygenase isoform, so-called type 2 (COX-2) has been discovered in the early 1990s by Daniel Simmons and W. L. Xie,[39] chemists at Brigham Young University in Provo, Utah. Simmons immediately understood the importance of his discovery. The same day the enzyme was sequenced,[40] and he kept his notebook notarized as proof of his discovery. Subsequently, a new class of drugs, COX-2 inhibitors was developed after researchers at the University of Rochester discovered the gene in humans that is responsible for producing the COX-2 and revealed the enzyme's role in causing inflammation within individual cells. The team, lead by Donald Young (University of Rochester Medical Centre), provided the basic understanding of the role of COX-2 in disease showing that selectively blocking the activity of the enzyme would be beneficial in treating

inflammation.[41] Besides the constitutive COX-1, participating to stomach protection and renal artery vasodilatation, this COX-2 enzyme, induced by inflammatory phenomena and cytokines stimulation, allowed to design specific inhibitors, "coxibs," playing an increasing but controversial role in the struggle against inflammation. This discovery set in motion a worldwide race among pharmaceutical companies to identify drugs that would restrain the action of the enzyme and, in turn, reduce inflammation and pain. There may be other forms of COX that could account for some of the remaining discrepancies in action amongst non-steroidal anti-inflammatory drugs (NSAIDs).[42] COX-2 inhibitors were apparently safer from a digestive point of view but questionable for their cardiovascular effects. Selective inhibitors of COX-2 cause less endoscopically visualized gastric ulceration in arthritis patients than equi-efficacious doses of traditional NSAIDs, which coincidentally inhibit COX-1 and COX-2. COX-2 inhibitors suppress substantially platelet inhibitory, vasodilator prostaglandins, such as prostacyclin (PGI_2), without coincidental inhibition of the platelet agonist vasoconstrictor thromboxane (TxA_2). As PGI_2 counters the cardiovascular effects of TxA_2 and augments the response to thrombotic stimuli *in vivo*, this affords a plausible mechanism by which COX-2 inhibitors might enhance the risk of thrombosis in otherwise predisposed individuals. After being marketed in 1999 rofecoxib (Vioxx®) has been withdrawn in 2004, because of an excess risk of myocardial infarctions and strokes. Despite the withdrawal, controversies remain. Although the non-selective NSAIDs can cause life-threatening gastric toxicity, the risk for any single patient is fairly low when COX-2 inhibitors are compared with two non-selective NSAIDs.[43] Among those controversies, the question whether selective COX-2 inhibitors are prothrombotic, or not, is not theoretical. Whereas aspirin and traditional NSAIDs inhibit both thromboxane A_2 and prostaglandin I_2, the coxibs leave thromboxane A_2 generation unaffected, reflecting the absence of COX-2 in platelets. Thus, this single mechanism might be expected to elevate blood pressure, accelerate atherogenesis, and predispose patients receiving coxibs to an exaggerated thrombotic response to the rupture of an atherosclerotic plaque.[44] Clinical observations and studies found that taking common NSAIDs was linked to a lower risk of certain cancers. When celecoxib was approved for familial adenomatous polyposis in 1999, there was hope that other COX-2 inhibitors would also prove to be safe and powerful anticancer treatments. This is not the case. Structural differences between celecoxib and rofecoxib could explain this discrepancy. A systematic chemical approach allowed to produce 50 compounds tested for their ability to induce apoptosis in human prostate cancer cells, confirmed that the structural requirements for the induction of apoptosis are distinct from those that mediate COX-2 inhibition. Apoptosis induction requires a bulky terminal ring, a heterocyclic system with negative electrostatic potential and a benzenesulfonamide or benzenecarbonamide moiety. Ching Shih Chen *et al.* (Columbus, USA) modified the structure of rofecoxib to create compounds that mimicked the surface electrostatic potential of celecoxib, one of which showed a substantial increase in apoptotic activity.[45] What a challenge for the future!

FIGURE 1.8 Marc Feldmann and Ravinder Maini.

4. *New strategies for rheumatoid arthritis*

Drug therapy for rheumatoid arthritis (RA), a chronic inflammatory and destructive joint disease, rests on two bases: symptomatic treatment with NSAIDs, not interfering with the underlying immuno-inflammatory and disease-modifying antirheumatic drugs (DMARDs), "modifying" the disease process. DMARDs are divided into small-molecule drugs and biological therapies. The initial approach to understanding the pathogenesis of RA and defining a novel therapeutic target was to investigate the role of cytokines by blocking their action with antibodies on cultured synovial-derived mononuclear cells *in vitro*. In a series of experiments using tissue taken from joints, Marc Feldmann and Ravinder Maini (Kennedy Institute, London) investigated the role of cytokines (Figure 1.8), protein messenger molecules that drive inflammation, and found that a number of pro-inflammatory cytokines were indeed present in the inflamed joints.

These investigations suggested that neutralization of tumor necrosis factor-alpha (TNF-α) with antibodies significantly inhibited the generation of other pro-inflammatory cytokines. Their first clinical trial was performed in 1992 at Charing Cross Hospital and revealed rapid and dramatic improvement of rheumatoid disease activity with anti-TNF therapy. The blockade of a single cytokine, TNF-α, had far-reaching effects on multiple cytokines and thereby exerted significant anti-inflammatory and protective effects on cartilage and bone of joints. A chimeric anti-TNF-α high-affinity antibody was initially tested, with substantial and universal benefit. Then, a randomized placebo-controlled double-blind trial supported the proposition that TNF-α was implicated in the pathogenesis of RA and was thus a

FIGURE 1.9 Docteur Gachet (with a digital flower). Painting by Vincent Van Gogh.

key therapeutic target.[46] Three TNF inhibitors have been approved since 1998 for the treatment of RA. First was infliximab (Remicade®), a chimeric (human-murine) IgG1 anti-TNF-α antibody, administered intravenously. It binds with high affinity to soluble and membrane-bound TNF-α thus inhibiting it. The two others are Etanercept (Enbrel®) and Adalimumab (Humira®) a recombinant humanized monoclonal anti-TNF-α antibody administered subcutaneously.[47] Feldmann and Maini received the Albert Lasker award for their discovery in 2003.

B. Giving back the heart its youth

1. *Digitalis*

In the second half of 18th century, William Withering, an English physician, heard that the local population was able to cure dropsy using a complex plant decoction. After having tested the various herbs on dropsy, digitalis leaf remained the most active and probably contained a substance increasing the ability of the weakened heart to improve pumping blood (Figure 1.9). In 1775, Withering published a pamphlet in which he reported his discovery, meticulously describing how the extract of the digitalis should be prepared, and giving precise instructions on dosage, including warnings about side effects and overdose from the experience learnt from 163 patients.[48]

The only but not least problem was a dreadful continuous vomiting and diarrhea during the treatment that was caused by the fact that the boundary between the therapeutic dose and poisoning was exceedingly narrow. It was therefore evident and absolutely necessary to purify the active substance in order to fix the effective and non-toxic dosage. After decades of works, Augustin Eugène Homolle and Théodore Quevenne, two Parisian pharmacists obtained from foxglove leaves an amorphous substance they called "digitaline," keeping the "ine" terminology, as they were sure that it was an alkaloid.

In fact it was a complex substance containing a specific sugar. It is not until 1867 that another French pharmacist, Claude Adolphe Nativelle was able to purify foxglove leaves and to produce the effective substance in the form of white crystals[49] that he called "crystallized digitalin." Just a few years, later the German, Oswald Schmiedeberg, managed to produce digitoxin (1875).[50] Shortly thereafter reports began to come in about other medicinal herbs which had the same effect on the heart as the foxglove products. Ethnopharmacy gave birth to ouabain, extracted by Albert Arnaud from *Acocanthera* roots and bark, and strophantin, extracted from *Strophantus*. Both of these drugs had previously been used by arrow hunters in Equatorial Africa. One hundred years later, explanation for the cardiotonic properties of digitalis, ouabain and strophantin were given through molecular pharmacology experiments. The story began when Jens Christian Skou (Aarhus, Denmark) (Figure 1.54) studied in the early 1950s the action of local anesthetics. He thought that membrane protein might be affected by local anesthetics. He therefore had the idea of looking at an enzyme which was embedded in the membrane: ATPase, discovering that it was most active when exposed to the right combination of sodium, potassium and magnesium ions.[51] Only then did he realize that this enzyme might have something to do with the active transport of sodium and potassium across the plasma membrane. Skou left out the term "sodium-potassium pump" from the title of his publication, continuing his studies on local anesthetics. In 1958, Skou met Robert L. Post (Nashville, USA), who had been studying the pumping of sodium and potassium in red blood cells[52] recently discovered that three sodium ions were pumped out of the cell for every two potassium ions pumped in,[53] his research being made by the use of a substance called ouabain which had recently been shown to inhibit the pump. Conversations between Post and Skou about ATPase drove Skou to verify if ouabain inhibited the pump. Indeed, it did inhibit the enzyme, thus establishing a link between the enzyme and the sodium–potassium pump. Skou received a Nobel Prize in Chemistry (1997). Julius C. Allen and Arnold Schwartz (Houston, USA) then studied digitalis effect on cardiac contractility (the positive inotropic effect), caused by the drug's highly specific interaction with Na^+/K^+-ATPase. It has been established that partial inhibition of the ion pumping function of cardiac

Na^+/K^+-ATPase by digitalis glycosides led to a modest increase in intracellular Na^+, which in turn, affected the cardiac sarcolemmal Na^+/Ca^{2+} exchanger, causing a significant increase in intracellular Ca^{2+} and in the force contraction.[54]

2. N*itroglycerin*

Nitroglycerin synthesis has been performed in 1844 by Antoine Jérôme Balard (Montpellier, France) who observed the collapse of animals few minutes after the administration of the drug. The vasodilatating effect of the drug was exploited by Ascani Sobrero (Torino, Italy) following work with Theophile-Jules Pelouze (1847) in Paris. Two years later, Konstantin Hering and Johann Friedrich Albers, developing the sublingual administration of nitroglycerin, observed the violent headache caused by the drug. Alfred Nobel, later founder of Nobel Prize, joined Pelouze in 1851 and recognized the potential of this yellow liquid with explosive interest.[55] He began manufacturing nitroglycerin in Sweden, overcoming handling problems with his patent detonator. Nobel suffered acutely from angina but refused what he considered as a chemical for a treatment. When the English physician Thomas Lauder Brunton succeeded to relieve severe recurrent angina pain in refractive patients except by bleeding, he realized that phlebotomy provided relief by lowering arterial blood pressure. This gave birth to the concept that reduced cardiac after load and work were beneficial. When administering amyl nitrite, a potent vasodilatator, by inhalation, Brunton noticed, in 1867, that coronary pain was transiently relieved within 30–60 s.[56] In 1876, William Murrell (Westminster Hospital London) proved that the action of nitroglycerin mimicked that of amyl nitrite, and he established the use of sublingual nitroglycerin for relief of the acute angina attack and as a prophylactic agent to be taken prior physical exercise. Almost a century later, research in the nitric oxide (NO) field explaining the mode of action of nitroglycerin, has dramatically extended. In 1977, Ferid Murad (Houston, USA) discovered the release of NO from nitroglycerin and its action on vascular smooth muscle.

Robert Furchgott (Figure 1.11) and John Zawadski (New York, USA) recognized the importance of the endothelium in acetylcholine-induced vaso-relaxation (in 1980) and Louis Ignarro (Figure 1.11) and Salvador Moncada (London, UK) (Figure 1.10) identified endothelial-derived relaxing factor (EDRF) as NO (in 1987).[57] Today, glycerol trinitrate remains the treatment of choice for relieving angina; other organic esters and inorganic nitrates are also used, but the rapid action of nitroglycerin and its established efficacy make it the mainstay of angina pectoris relief. The role of NO in cellular signaling has become one of the most rapidly growing areas in biology during the past two decades. As a gas and free radical with an unshared electron, NO participates in various biological processes. NO is formed from the amino acid L-arginine by a family of enzymes, the NO synthases, and plays a role in many physiological functions.

FIGURE 1.10 Salvadore Moncada.

FIGURE 1.11 Robert Furchgott and Louis Ignarro.

Its formation in vascular endothelial cells, in response to chemical stimuli and to physical stimuli such as shear stress, maintains a vasodilator tone that is essential for the regulation of blood flow and pressure. NO also inhibits platelet aggregation and adhesion, inhibits leukocyte adhesion and modulates smooth muscle cell proliferation. NO is also synthesized in neurons of the central nervous system (CNS), where it acts as a neuromediator with many physiological functions, including the formation of memory, coordination between neuronal activity and blood flow, and modulation of pain. In the peripheral nervous system, NO is now known to be the mediator released by a widespread network of nerves.[58]

3. *Antihypertensive drugs*

Scipione Riva-Rocci (University of Pavia) introduced his mercury sphygmomanometer, easy to use and giving reliable results for measuring blood pressure, in 1896. This device led to many developments in the therapy of hypertension disease.[59] A fundamental role in spreading the use of the instrument was played by Harvey Cushing. But the importance of arterial blood pressure monitoring was not understood before 1913, when researchers reported a clear association of hypertension with heart failure, stroke, and kidney impairment.[60] Few years later, in 1925, the American Society of Actuaries published an epidemiological analysis concerning 560,000 men and demonstrating the link between cardiovascular risks and an elevated blood pressure.[61] In the 1930s and 1940s, no pharmacological antihypertensive treatment being available, physicians could choose between sympathectomy,[62] very-low-sodium diets,[63] and pyrogen therapy.[64] Those treatments had life-threatening complications or unpleasant side effects. The first successful drug treatment for hypertension was introduced in 1946. Blocking the sympathetic nervous system by the means of tetraethylammonium, a drug known for 30 years to block nerve impulses was introduced few years before hexamethonium, another ammonium derivative, available as a treatment for hypertension by 1951.[65] Another effective blood pressure-lowering drug, hydralazine,[66] resulted from the search for antimalarial compounds. It was diverted to the treatment of hypertension when it was found to have no antimalarial activity but to lower blood pressure and increase kidney blood flow. Unfortunately, both hexamethonium and hydralazine often caused severe side effects. The final drug developed in those early days, reserpine,[67] was the product of more than two decades of research into compounds derived from *Rauwolfia serpentina*, a plant used for centuries by physicians and herbalists on the Indian subcontinent.[68] Antihypertensive therapy gave birth to one of the most demonstrative clinical trial in the history of drug discovery. Men with elevated blood pressure were randomly divided into two groups. Hydralazine, hydrochlorothiazide, and reserpine, were given to the first group, the other group receiving a placebo. Previously planned to last for 5 years, the study stopped after 18 months: patients who had received the placebo were dying at a greater rate than those who had received the antihypertensive drugs.[69] The clinical interest of treating hypertension was definitively proven (Table 1.1).

TABLE 1.1 Main Steps in Antihypertension Drugs Discovery

1890	Veratrum alkaloids (protoveratrin)[70]
1949	Pyrogens
1936	Thiocyanates[71]
1950	Ganglion blocking agents[72]
1954	Catecholamine depletors (*Rauwolfia* derivatives)
1953	Vasodilators (hydralazine)
1960	Peripheral sympathetic inhibitors (guanethidine)
1952	Monoamine oxidase inhibitors (iproniazide)
1957	Diuretics (chlorothiazide)
1963	Calcium channel blockers (verapamil)
1964	β-Adrenergic inhibitors (propranolol)[73]
1969	Central α_2-agonists (clonidine)
1969	α-Adrenergic inhibitors (prazosin)
1976	α-β-Blockers (labetalol, carvedilol)
1978	ACE inhibitors (captopril)[74]
1991	Angiotensin II (ATII) receptor antagonists

4. *Diuretics*

A major progress was the use of diuretics, effective through an increase of urine flow and sodium excretion. They act directly on nephrons acting on various targets, including tubules and glomerules. The first thiazidic drug, chlorothiazide, became available in 1958, it was a real breakthrough,[75] and remains, 50 years later the basic diuretic used in a majority of antihypertensive regimens. Discovery of diuretic drugs followed two unrelated endeavors in the 1930s: the development of sulfanilamide[76] which had an unexpected diuretic action and the identification of carbonic anhydrase, the enzyme responsible for transport of carbon dioxide by the blood and its excretion in the lungs,[77] enzyme on which, sulfanilamide is active. By 1938, physiologists demonstrated that sulfanilamide was an inhibitor of carbonic anhydrase. Various compounds with no antibacterial effect were synthesized and, among them, acetazolamide, a potent carbonic anhydrase inhibitor, increased urination and resulted in weight loss and clinical improvement of patients with heart failure and edema. It is still used to treat glaucoma and cranial hypertension. Diuretic therapy had a dramatic effect of on hypertension. It was possible to obtain a normalized blood pressure and to lower fluid accumulation, with few side effects. Diuretics rapidly get extension of their indication in heart failure and other conditions caused by the inability of the kidneys to regulate the salt and water balance. The importance of diuretics discovery was recognized in 1975, when the Albert Lasker Award was given to Karl H. Beyer, James M. Sprague, John E. Baer, and Frederick C. Novello (Merck Sharp and Dohme Research Laboratories),

structural chemists and kidney physiologists responsible for chlorothiazide development for the control of high blood pressure and of edema associated with cardiac failure. Thiazidic compounds and furosemide, another "sulfonylurea-derived" diuretic, are now universally accepted as a primary treatment for hypertension.

5. β-Blockers

George W. Oliver and Edward A. Schafer (University College, London) demonstrated, in 1895, that an injection of adrenal gland extract could raise blood pressure in experimental animals, mimicking the stimulation of the sympathetic nervous system.[78] After epinephrine discovery by Jokichi Takamine (1901),[79] Friedrich Stolz performed in 1904 the synthesis of norepinephrine (noradrenalin) and epinephrine (adrenalin). The same year, it was suggested that sympathetic nerves produce epinephrine.[80] Physiologists demonstrated that a stimulation of sympathetic nerve increased the heart rate and the force of the heart's contractions in mammalians. The conditions of the release of the active substance remained unknown till the 1940s, first when Ulf Von Euler (Karolinska Institute, Stockholm), 1970 Nobel Prize-winning, demonstrated that norepinephrine was produced by sympathetic nerves[81] and released from them when they are stimulated. Second, with Raymond Alhquist's (Augusta, USA) 1948 discovery, result of pure serendipity, of two types of receptors in tissues: α- and β-receptors, explaining why the same chemical, norepinephrine, could have various effects on tissular functions.[82] This work first refused for publication and then ignored several years, revolutionized pharmacological concepts. Fifteen years later, in 1963, James Black (Imperial Chemical Industries, UK) (Figure 1.34) discovered propranolol. It was developed from the early β-adrenergic antagonists dichloroisoprenaline and pronethalol. The key structural modification, which was carried through to essentially all subsequent β-blockers, was the insertion of an oxymethylene bridge into the arylethanolamine structure of pronethalol thus greatly increasing the potency of the compound. By binding to the β-receptors, β-blockers are limiting the rise in heart rate, decreasing the force of contraction, and reducing the oxygen requirements of heart muscle with a significant lowering of blood pressure.[83] Black received Nobel Prize in 1988. Since 40 years, many β-blockers have been synthesized. Some have slightly different effects than propranolol, but none has been shown to be superior to propranolol in controlling hypertension or angina.

6. Calcium antagonists

The discovery of Ca^{++} antagonism as a new principle of action of coronary drugs reaches back to 1964, when Albrecht Fleckenstein (Freiburg, Germany) observed two new compounds (verapamil and prenylamine), mimicking the cardiac effects of simple Ca^{++} withdrawal, in that they diminished Ca^{++}-dependent high energy phosphate utilization, contractile force, and oxygen requirement of the beating heart without impairing the Na^{+}-dependent action potential parameters (Box 1.2).

Verapamil demonstrates an action clearly distinguishable from beta-receptor blockade, as it could promptly be neutralized with elevated Ca^{++}, β-adrenergic catecholamines, or cardiac glycosides, in order to restore the Ca^{++} supply to the contractile system.[84] The term "Calcium antagonist" was introduced in 1969 as a novel drug designation. In an extensive search for other Ca^{++} antagonists, a considerable number of substances were identified: dihydropyridines (nifedipine, amlodipine, nicardipine, nitrendipine and others); verapamil, which is structurally similar to papaverine; bepridil, a non-specific calcium blocker. In 1975, Fumio Ariyuki, (Osaka, Japan) contributed diltiazem, a benzothiazepine derivative to this group. All specific calcium antagonists interfere with the uptake of Ca^{++} into the myocardium and prevent myocardial necrosis arising from deleterious intracellular Ca^{++} overload; they also

BOX 1.2 Ion Channels

Erwin Neher and Bert Sakmann (awarded for Nobel Prize in Medicine in 1991 for their discoveries concerning "the function of single ion channels in cells") developed the "patch clamp method" proving the existence of ion molecular channels by measuring the ionic current on a tiny membrane patch to which a pre-determined voltage-clamp is applied. Ion channels are not specific of myocardic cells, but ubiquitous. Cell membranes of the nervous system contain a number of specific transport systems, which bring different agents in and out of the cell, transporting ions. The interior of the cell has a high concentration of K^{+}, whereas Na^{+} dominates on the outside. This leads to a difference in electric potential between the two sides of the cell membrane, which can amount to as much as a tenth of a volt. This membrane potential is used for a number of different tasks: the nerve cells to send rapid electrical signals, a large variety of cells to communicate with each other, etc. Every single ion channel (specific to one type of cation like Na^{+} or K^{+}, or anion Cl^{-}) consists of one protein molecule or a molecular complex, which forms the walls of a thin channel, connecting the interior of the cell with its exterior, with such a small diameter that it corresponds to the width of only one single ion. When one of the channels is opened, a very small current will flow, which can be measured through the patch-clamp technique indicating when a single ion channel is opened or closed, that is, when a single molecule changes its shape. This technique has considerably changed pharmacological studies. A number of diseases are either influenced or caused by a modified ion channel function: anxiety, cardiovascular disease, epilepsy, diabetes, and even reproduction.

block excitation-contraction coupling of vascular smooth muscle and, in this manner, lower Ca^{++}-dependent coronary vascular tone and neutralize all types of experimental coronary spasms.[85]

7. ACE *inhibitors and sartans*

Among recent discoveries, the research poined out the crucial role of converting enzyme. Angiotensin-converting enzyme (ACE) inhibitors are unique in the history of antihypertensive drug development. Those antihypertensive drugs history began in 1898, when R. Tigerstedt and P. G. Bergman Swedish scientists discovered a substance in kidneys (the reason why they called it renin), raising blood pressure when injected into animals.[86] Harry Goldblatt (Cleveland, USA) (Figure 1.12) revealed, in 1934, that the constriction of the renal arteries causes a chemical chain reaction leading to hypertension.[87] If Goldblatt demonstrated that hypertension could be related to a reduced blood flow in kidneys, the renin story remained buried more few years till experiments demonstrating in 1940, that renin was that enzyme which could act to produce angiotensin, the protein which narrows small blood vessels[88] and thus raises blood pressure.[89]

By the early 1950s, research had shown that the cornerstone of this system was an enzyme, the ACE, active in a two-steps cascade. The first step is the production of angiotensin I (decapeptide) from angiotensinogen consisting of 453 amino acid residues, a blood circulating protein, through the catalytic action of renin. This reaction occurs not only in plasma but also in the kidneys, brain, adrenal glands, ovaries, and possibly other tissues. The second step is the conversion of angiotensin I (with no effect on blood pressure) to angiotensin II (octapeptide) which elevates blood pressure (Figure 1.13).

After having discovered the amino acid sequence of angiotensin II, angiotensinogen and angiotensin I were synthesized[90] and, in a logical process, antagonists of angiotensin II were sought as treatment for hypertension. The only fruitful research came to the study of ACE inhibition itself.[91] In the 1960s, John R. Vane (London, UK) was actively investigating the cause of hypertension. During this time, a Brazilian post-doc, Sergio Ferreira, joined Vane's group and brought with him an extract bradykinin potentiating factor (BPF) of the venom of the Brazilian viper *Bothrops jararaca*. This venom was found to contain compounds increasing the potency of bradykinin by blocking the enzyme, kininase II that destroys it.[92] It is a history where chance, serendipity and clear scientific reasoning weaved together the work of several scientists. It is also a classical example of drug development for which the initial basic research was made at the university, but the useful product was achieved by industry.[93] BPF tested on ACE was found to be a potent inhibitor thereof. This led to Vane's strong interest in ACE and its inhibition as a means of treating hypertension. ACE was the same enzyme as kininase II.[94] Vane advocated the Squibb's concept consisting to "test the hypothesis" by investigating the snake venom peptide by injection and to tackle the problem of making an orally available form of the drug. It was also necessary to demonstrate that the peptide would block the conversion of angiotensin I to II, the biochemical reaction mediated by ACE. Knowing the enzyme and its naturally occurring inhibitors from snake venom, chemists performed the synthesis of the

FIGURE 1.12 Harry Goldblatt.

FIGURE 1.13 The renin–angiotensin cycle.

first ACE inhibitor, called teprotide, only active when given intravenously (Table 1.2).

Information learned from the synthesis of teprotide and an increasing knowledge about ACE, another compound, captopril, was synthesized[95] then followed by a number of other ACE inhibitors. US Food and Drug Administration (FDA) approval came in the early 1980s. Captopril was Squibb's first billion dollar drug and it opened a new approach for the treatment of hypertension.[96] ACE inhibitors discovery resulted from the systematic exploration of a major system that controls blood pressure and the targeted synthesis of compounds that block the system. Since angiotensin II is the effector molecule of the renin–angiotensin system, the most direct approach to block this system is to antagonize angiotensin II at the level of its receptor. A new hypothesis emerged at Du Pont Merck when it was possible to identify metabolically stable and orally effective angiotensin II-receptor antagonists. They could constitute a new and superior class of agents useful in treating hypertension and congestive heart failure: some simple *N*-benzylimidazoles originally described by Takeda (Osaka, Japan) were in the "pipe-line." Potent and orally effective non-peptidic antagonists were found. The first major breakthrough in order to increase the potency of the compounds (sartans) came with the development of a series of *N*-benzylimidazoles, phthalamic acid derivatives, and the discovery of losartan, a highly potent selective receptor antagonist with a long duration of action[97] (Table 1.3).

TABLE 1.2 Actions of angiotensin II

Tissue affected	Action
Artery	Stimulates contraction growth
Adrenal zona glomerulosa	Stimulates secretion of aldosterone
Kidney	Inhibits release of renin Increases tubular reabsorption of sodium Stimulates vasoconstriction Releases prostaglandins Affects embryogenesis
Brain	Stimulates thirst and the release of vasopressin
Sympathetic nervous system	Increases central sympathetic outflow Facilitates peripheral sympathetic transmission Increases adrenal release of epinephrine
Heart	Increases contractility and ventricular hypertrophy

C. Fight against microbes and viruses

1. *Identifying the role of germs*

Among the scientific advances of 19th century, the emergence of microbial theory of infectious diseases and the discovery of first vaccines to prevent those diseases have to be considered as main milestones. The brilliance of European lens makers and microscopists, coupled with the tinkering of laboratory scientists who developed the technologies

TABLE 1.3 Ion Channels and Drugs that Affect Them

Type of channel	Drug family	Drugs
Ca^{++}	Antiangina drugs	Amlodipine, diltiazem, felodipine, nifedipine, verapamil
	Antihypertensive drugs	Amlodipine, diltiazem, felodipine, isradipine, nifedipine, verapamil
	Class IV antiarrhythmics	Diltiazem, verapamil
Na^{+}	Anticonvulsant drugs	Carbamazepine, phenytoin, valproic acid
	Class I antiarrhythmics	IA Disopyramide, procainamide, quinidine IB Lidocaine, mexiletine, phenytoin, tocainide IC Encainide, flecainide, propafenone
	Diuretic drugs	Amiloride
	Local anesthetic drugs	Bupivacaine, cocaine, lidocaine, mepivacaine, tetracaine
Cl^{-}	Anticonvulsant drugs	Clonazepam, phenobarbital
	Hypnotic or anxiolytic drugs	Clonazepam, diazepam, lorazepam
	Muscle-relaxant drugs	Diazepam
K^{+}	Antidiabetic drugs	Glipizide, glyburide, tolazamide
	Antihypertensive drugs	Diazoxide, minoxidil
	Class III antiarrhythmics	Amiodarone, clofilium, dofetilide, N-acetylprocainamide, sotalol
	Drugs opening K^{+} channels	Adenosine, aprikalim, levcromakalim, nicorandil, pinacidil

TABLE 1.4 Pre-antibiotic Era Discoveries in the Field of Infectious Diseases

1859	Louis Pasteur suggests that microorganisms may cause many human and animal diseases
1867	Joseph Lister publishes *On the Antiseptic Principle in the Practice of Surgery* showing that disinfection reduces postoperative infections
1879	Louis Pasteur demonstrates value of vaccine to protect sheep against anthrax
1882	Robert Koch isolates microorganism responsible for tuberculosis (TB), then leading cause of death
1883	Robert Koch isolates microorganism responsible for cholera, major epidemic disease in 19th century
1885	Louis Pasteur develops first rabies vaccine
1890	Emil von Behring and Shibasaburo Kitasato develop effective diphtheria antitoxin
1897	George Nuttall demonstrates that flies can spread plague bacilli
1906	August von Wasserman introduces diagnostic test for syphilis
1907	Clemens Von Pirquet introduces skin test for TB
1911	Paul Erhlich tests salvarsan, first treatment effective against syphilis; regarded as birth of modern chemotherapy
1916	Polio epidemics break out in New York and Boston; polio outbreaks continue sporadically in summers for decades to come
1918–1919	An influenza pandemic kills nearly 40–50 million people worldwide
1928	Alexander Fleming discovers penicillin although it does not become available in a therapeutically usable form until 1940

of sterilization and the media and methods for growing and staining microbes, provided the foundation of the new medical science: microbiology that would explode in the 20th century. The practical use of disinfectant fumigations inaugurated by Carl Wilhelm Scheele, just came before "Guytonian's fumigations," based on chlorine activity (Table 1.4).

As soon as in 1785, a solution of chlorine gas in water was used to bleach textiles. Potassium hypochlorite (Eau de Javel) was prepared by Berthollet in 1789. In 1820, Antoine G. Labarraque replaced potash liquor by the cheaper caustic soda liquor and thus was born sodium hypochlorite. At the end of the 1820s, Robert Collins, then Oliver Wendell Holmes showed that puerperal (childbed) fever frequency decreased when midwives wash their hands in chlorinated water.[98] Few decades later (1861), Ignaz Philip Semmelweis published his research on the transmissible nature of puerperal fever. But he failed to convince physicians either in Vienna or in Budapest that they were at the origin of the contamination of pregnant women.[99] Eventually, with the works of Robert Koch (Berlin, Germany), Joseph Lister and Louis Pasteur adding proof of the existence and disease-causing abilities of microorganisms, a worldwide search for the microbial diseases began. Koch demonstrated, in 1881, the lethal effect of hypochlorites on pure cultures of bacteria. Few years after, in 1894, Isidor Traube established the purifying and disinfecting properties of hypochlorites in water treatment. During the World War I, Dakin's hypochlorite solution has been extensively used for disinfection of open and infected wounds. Milton® fluid (containing 1% sodium hypochlorite and 16.5% sodium chloride) was marketed in the United Kingdom, in 1916, as a general disinfectant and antiseptic in pediatrics and child care. Another halogen, iodine, had been discovered by Bernard Courtois (Dijon, France) in 1811, who extracted the element from wracks at seashore. Iodine "tincture," proposed in 1835 by William Wallace (Dublin, Ireland) to disinfect wounds, was contested by iodoform, invented by Georges Simon Serullas (Paris, France). Structurally, it was very comparable to chloroform, the chlorine atom being substituted by an iodine one. In 1829, Jean Lugol, a French physician researching the medical uses of iodine in infectious diseases, observed that the presence of potassium iodide in water increases markedly the aqueous solubility of iodine. He used his preparation for dermatological treatments at Saint-Louis Hospital in Paris. The antiseptic properties of iodine were widely used from the discovery of iodine until today. In 1873, the French bacteriologist Davaine used tincture of iodine as an agent to treat anthrax.[100] The revolutionary change in hospital hygiene was introduced when Friedlieb Runge (Breslau, Germany) prepared carbolic acid (phenol) by distillation of coal tar (1834).[101] Joseph Lister (Glasgow, UK) proposed to use "phenolic" surgical ligatures and dressings. Phenol sprays were used by the French surgeon, Just Lucas-Championniere (Paris) in operating rooms around 1860. At that time, Joseph Lister not only reduced the incidence of wound infection by the introduction of antiseptic surgery, he also showed that urine could be kept sterile after boiling in swan-necked flasks. He was the first person to isolate bacteria in pure culture (*Bacillus lactis*) and so, can be considered a co-founder of medical microbiology with Koch, who later isolated bacteria on solid media.[102] More and more the experimental proof confirmed the empirical behavior. In this environment, the microbial theory of a lot of diseases constituted the hallmark of 19th century medicine. The idea that infectious diseases were caused by invisible agents, gave an opportunity for many progresses. The laboratory took its entire place when microscopes, staining of preparations and sterilization were available. As an

example, *Escherichia coli*, discovered in 1879, became the perfect example of easily grown, "safe" bacteria for laboratory practice. Working with pure cultures of the diphtheria bacillus in Pasteur's laboratory in 1888, Emile Roux and Alexandre Yersin first isolated the deadly toxin that causes most of diphtheria's lethal effects.[103] One by one over the next several decades, various diseases revealed their microbial cause, including digestive ulcers.

2. *Sulfonamides*

Till the beginning of the 20th century, struggle against microbes remained devoted to the disinfection of wounds and sanitization of drinking water. Pasteur's objective to treat infectious diseases as cholera, tuberculosis, and diphtheria, remained a dream although some vaccines (smallpox or rabies) were already available. The breakthrough would come from an unexpected side of the scientific field: dyes industry.[104] In 1865, Friedrich Engelhorn founded BASF (Mannheim, Germany) to produce coal tar dyes and precursors. In 1871, the company marketed the red aniline-dye alizarin. Other new dyestuffs followed: eosin, auramine, and methylene blue, together with the azo dyes, which would eventually develop into the largest group of synthetic dyestuffs. Around the 1880s, German chemists, following in Paul Ehrlich's wake, discovered the fact that living cells absorbed dyes in a different way as dead cells. He expected healing, probably more than he obtained. Nevertheless, he noticed some improvements. Thereafter, Ehrlich (Berlin, Germany) (looked for a cure or treatment for "sleeping sickness" or African trypanosomiasis and found that a chemical called Atoxyl® worked well but was a fairly strong arsenical compound and thus poisonous. Ehrlich began an exhaustive search for other arsenicals that could be a "magic bullet" able to kill the microbe but not the person when killing the disease. In 1909, after testing over 900 different compounds on mice, Ehrlich's Japanese colleague Sahachiro Hata went back to no. 606 or dihydroxy-diamino-arsenobenzol-dihydrochlorid (Figure 1.14). It did not do much for the sleeping sickness microbe, but it seemed to kill another (recently discovered) microbe, the one which caused syphilis: *Treponema pallidum*. If a microorganism could be colored, vital properties of the bacteria or the parasite could also be transformed.[105] Which conclusion could be drawn from these informations about the viability of colored bacteria? Ehrlich refined the use of methylene blue in bacteriological staining and used it to stain the tubercle bacillus, showing the dye bound to the bacterium and resisted discoloration with an acid alcohol wash.[106] Following this hypothesis, Ehrlich administrated methylene blue to patients suffering from malaria.

At that time, syphilis was a disabling and prevalent disease. Ehrlich and Hata tested 606 over and over on mice, guinea pigs, and then rabbits with syphilis. They achieved complete cures within 3 weeks, with no dead animals[107].

FIGURE 1.14 Paul Ehrlich and Sahachiro Hata.

A production of the first large batch of Salvarsan® took place at Hoechst (Frankfurt) on July 1910. It was an almost immediate success and was sold all over the world. It spurred Germany to become a leader in chemical and drug production and it made syphilis a curable disease. The concept of the "magic bullet" was born simultaneously to the concept of chemotherapy. Arsenicals, unlike vaccines, were not tightly controlled and were far more subject to misprescription and misuse as far as they had to be administered intravenously, which was very hazardous at that time. In Julius Morgenroth's laboratory (Berlin, Germany), the following year, other works were performed on *Pneumococcus* and particularly on the nature of the external capsule of the microorganism and the power of biliary salts to dissolve *Trypanosome's or Pneumococcus* external structures. Another concept concerning unspecific targets for drugs in infectious diseases was built, also explaining the activity of various isoquinoline derivatives for treating different infectious diseases.[108] Most clinicians thought the future would be in immunotherapy rather than in chemotherapy, and it was not until the antibiotic revolution of the 1940s that the balance would shift. But the influenza pandemic (so-called Spanish Flu) of 1918–1920 clearly demonstrated the inability of medical science to stand up against disease. Forty to Fifty million people worldwide were killed. Chemotherapy research had to be improved and continued. In the year 1927, Gerhardt Domagk (Figure 1.15), who got a promotion in Bayer's research department (Wuppertal, Germany), aimed to find a drug capable to destroy microorganisms after oral administration. The experimental model he used

FIGURE 1.15 Gerhardt Domagk.

FIGURE 1.16 René Dubos.

was the streptococcal infection of mice, allowing experiments on the effect of the intake of a large amount of drugs. Domagk turned his attention to azo dyes, so-called because the two major parts of the molecule are linked by a double bond between two nitrogen atoms. Some of these dyes attach strongly to protein in wool fibers or leather, so that they hold fast against fading or cleaning. Domagk reasoned that they might also attach themselves to the protein in bacteria, inhibiting if not killing microorganisms. Chrysoidine, which was a marvelous deep-red dye, had to be grafted to a sulfonylurea derivative (sulfamidochrysoidine) to be active. Fritz Mietzsch and Josef Klarer tested the new dye in 1932 on laboratory rats and rabbits infected with streptococci bacteria. Domagk found that it was highly antibacterial but not toxic and named the substance Streptozan®, soon changed to Prontosil® giving birth to the new era of antimicrobial chemotherapy. The first human cure occurred in 1932.

At least two versions of the same story are coexisting. It is not still clear whether it was administered in an act of desperation, in a 10-month-old boy who was dying of staphylococcal septicemia; the baby made an unexpectedly rapid recovery. Another account is that Domagk used Prontosil® to treat his own daughter, who was deadly ill from a streptococcal infection following a pin prick. Domagk did not immediately publish his results. His landmark paper (February 1935) published shortly after that it took a patent, attracted the curiosity of a great number of researchers in Europe. Domagk was awarded the Nobel Prize in Physiology or Medicine in 1939, but due to the Nazi veto, received his medal after World War II, in 1947. Even if Domagk discovered sulfonamides, he did not discover the way they were active. The work was done by a French team at the Pasteur Institute in Paris by Ernest Fourneau, Jacques and Thérèse Trefouël, Federico Nitti and Daniel Bovet (the last received the Nobel Prize in Physiology or Medicine in 1957). Prontosil® was inactive on bacteria cultures because it needed presence of a reductase to split the molecule. The active part was the sulfonamide (amino-4-benzene sulfonamide) itself and not the dye! Doctors in whole Europe in 1936 had stunning results using the new drug to treat childbed fever and meningitis. Sulfanilamide was brought to the United States by Perrin H. Long and Eleanor A. Bliss, who used it in clinical applications at Johns Hopkins University (Baltimore) in 1936. Prontosil® won wide publicity in the United States in 1936 when it was used to treat President Franklin Delano Roosevelt's son Franklin, Jr., who was severely ill from a streptococcic infection. More than 5,000 sulfa drugs were prepared in the late 1930s and early 1940s. Among them, sulfapyridine was used against pneumonia (it was used to treat Winston Churchill when he came down with pneumonia during World War II); sulfathiazole was used against both pneumonia and staphylococcal infections; sulfadiazine was used against pneumococcal, streptococcal, and staphylococcal bacteria; and sulfaguanadine against dysentery.

3. *Antibiotics*

The 1930s were also the period for a new era, the birth of antibiotic treatments.[109] René Dubos (Figure 1.16) had been recruited by Oswald Avery, to the Rockefeller Institute

FIGURE 1.17 Alexander Fleming.

(New York) and challenged to find a soil microbe that could destroy a bacteria.[110] In 1939, he discovered a substance extracted from a soil bacillus. Tyrothricin (later showed composition of two substances, gramicidin (20%) and tyrocidine (80%), cured mice infected with pneumococci. It was the first natural antibiotic extracted from soil bacteria, able to arrest the growth of staphylococcus, but proved highly toxic.

(a) Penicillins

It was the desire to find an internal antiseptic that drove Alexander Fleming (Figure 1.17) in his pioneering work in London in the 1920s after the amazing observation that the human teardrop contained a chemical capable of destroying bacteria – and at an alarming rate. However, the excitement at this discovery was soon dashed. While the new discovery – which Fleming called lysozyme – was effective at dissolving harmless microbes, it proved ineffective at negating those that caused disease. Fleming, however, did not give up. In 1928, his diligence was rewarded.

In his laboratory (Saint Mary's Hospital, London) Fleming was in the process of developing *staphylococci*. Removing the lid from one of these cultures, Fleming was surprised to see that around the mould, the colonies of staphylococci had been dissolved. Something (penicillin) produced by the mould was dissolving the bacteria. After further testing, Fleming was able to isolate the juice of the mould and it was then that he named it penicillin.[111] Penicillin was seen by the medical community as a non-event. The overwhelming casualties on the battlefield during the World War II led two medical researchers, Howard

FIGURE 1.18 Boris Ernest Chain Howard Florey.

Florey and Boris Ernst Chain (Figure 1.18), to look at resurrecting Fleming's work with penicillin. The experimental job was performed and published in July 1940, without raising any (positive or negative) reaction among pharmaceutical world.[112] After much refinement they were able to develop a powdered form of penicillin. In 1941, the first human was successfully treated. Before long, penicillin was in full production. Fleming, Florey and Chain were awarded the Nobel Prize for Medicine in 1945.[113] As early as 1945, in an interview with The *New York Times*, Fleming warned that the misuse of penicillin could lead to selection of resistant forms of bacteria.[114] Fourteen years elapsed between discovery of penicillin (in 1928) and its full-scale production for therapeutic use (in 1942). A great number of factors were responsible for such a delay: initial difficulty of other bacteriologists in reproducing Fleming's discovery, identifying the chemical makeup of penicillin, search for other penicillin-producing organisms to enhance production of the drug, its purification and crystallization, experiments on animals (chiefly mice) to determine toxicity, hesitancy to administer the drug to humans, standardization of an effective dosage for humans, and search for equipment and financial resources to enhance full-scale production. The adjunctive role of serendipity in overcoming these obstacles and in contributing to the successful conclusion of the penicillin project constituted an amazing story.[115]

The agricultural industry contributed, through its fermentation facilities and corn steep liquor used for the medium of culture, to penicillin development. The production of penicillin increased by more than 10-fold. In fact, by 1944, there was sufficient penicillin to treat all of the severe battle wounds incurred on D-day at Normandy. Also, diseases like syphilis and gonorrhea could suddenly be treated more easily than with earlier treatments.

Dorothy Crowfoot Hodgkin (Oxford), Nobel Prize in Chemistry in 1964, determined the chemical structure of penicillin by crystallography, in the early 1940s, enabling synthetic production of derivatives. John Sheehan at MIT

(Boston, USA), completed the first total synthesis of penicillin and some of its analogs in the early 1950s, but his methods were not efficient for mass production. The narrow spectrum of activity of the penicillins, along with the selective activity of the orally active phenoxymethylpenicillin, led to the search for derivatives of penicillin which could treat a wider range of infections. The first major development was ampicillin, which offered a broader spectrum of activity than either of the original penicillins.[116] Ampicillin (1961) was produced in a batch process by enzymatic acylation of 6-aminopenicillanic acid (6-APA) with the aid of a phenylglycine derivative such as D-phenylglycine amide. Before even ampicillin, however, in 1960, two companies: Beecham and Bristol brought out methicillin more resistant to the beta-lactamase enzyme produced by *Staphylococcus aureus*. But almost immediately, strains of methicillin-resistant *S. aureus* were discovered, although for many years their number was low. Meanwhile, the new drug seemed to bring a solution to the threat of *S. aureus*. Famously, the life of the actress Elizabeth Taylor was rescued after she was treated for *Staphylococcus pneumoniae* during the shooting of the film "Cleopatra". Further development with great commercial impact yielded beta-lactamase-resistant penicillins including flucloxacillin, dicloxacillin, and oxacicillin.[117]

(b) Streptomycin and antituberculosis drugs

Since 1914, Selman A. Waksman (Figure 1.19) screened systematically soil bacteria and fungi and, at the University of California, in 1939 he discovered the marked inhibitory effect of certain fungi, especially actinomycetes, on bacterial growth. In 1940, he and his team were able to isolate actinomycin from *Actinomyces griseus* (later named *Streptomyces griseus*), an antibiotic effective against Koch's bacillus, but too toxic for use in humans. Few months later, Waksman, with Albert Schatz and Elizabeth Bugie, isolated the first aminoglycoside, streptomycin, from *S. griseus*.[118] In 1942, several hundred thousand deaths resulted from tuberculosis in Europe, and another 5–10 million people suffered from the disease worldwide. Sulfonamides and penicillins being ineffective against *Mycobacterium tuberculosis,* Waksman studied the value of streptomycin in treating that disease. Merck immediately started manufacturing streptomycin. Simultaneously, studies by William H. Feldman and H. Corwin Hinshaw at the Mayo Clinic (Rochester, USA) confirmed streptomycin's efficacy and relatively low toxicity against tuberculosis in guinea pigs. On November 20, 1944, doctors administered streptomycin for the first time to a seriously ill tuberculosis patient and observed a rapid, impressive recovery. His advanced disease was visibly arrested, the bacteria disappeared from his sputum, and he made a rapid recovery. The only problem was that the new drug made the patient deaf: streptomycin was particularly toxic on the inner ear. In 1952, Waksman was awarded the Nobel Prize in Physiology or Medicine for his discovery of streptomycin (and 17 other antibiotics discovered under his guidance). A succession of tuberculicid drugs appeared during following years. These were important because with streptomycin monotherapy, resistant mutants began to appear. In 1950, British physician Austin Bradford Hill demonstrated that a combination of streptomycin and *p*-aminosalicylic acid (PAS) could better cure the disease, although the toxicity of streptomycin was still a problem. By 1951, an even more potent antituberculosis drug was developed simultaneously and independently by the Squibb Co. and Hoffmann-LaRoche. Purportedly, after an experiment on more than 50,000 mice and the examination of more than 5,000 compounds, this drug, isonicotinic acid hydrazide was proved to be able to protect against a lethal inoculum of tubercle bacteria. It was marketed ultimately as isoniazid and proved especially effective in mixed dosage with streptomycin or PAS. In two decades, after PAS acid (1949) and isoniazid (1952), pyrazinamide (1954), cycloserine (1955), ethambutol (1962) and rifampin (1963) were introduced as other weapons against tuberculosis. Aminoglycosides such as capreomycin, viomycin, kanamycin, and amikacin, and recently, the newer quinolones (ofloxacin and ciprofloxacin) are only used in drug resistance situations. With acquired immune deficiency syndrome (AIDS) pandemic, tuberculosis in particular, experienced a dreadful come-back.

FIGURE 1.19 Selman A. Waksman.

(c) Chloramphenicol

Originally isolated by David Gottlieb (University of Illinois, USA) from the soil organism *Streptomyces*

venezuelae in 1947, has been introduced into clinical practice in 1949. It was the first antibiotic to be manufactured synthetically on a large scale.[119] Usually chloramphenicol was a bacteriostatic, but at higher concentrations or against some very susceptible organisms it could be bactericidal. Manufacture of oral chloramphenicol in the United States stopped in 1991, because the hematological toxicity of the drug, which appeared very early in the 1950s.[120]

(d) Tetracyclines

Closely congeneric derivatives of the polycyclic napthacene-carboxamide were discovered as natural products by Benjamin Minge Duggar, consultant in mycological research to Lederle Laboratories (American Cyanamid), in 1948. The discovery of the tetracycline ring system also enabled further development in antibiotics.[121] Since that time more than one hundred of various molecules of the tetracycline family (Figure 1.20), active against a wide range of bacteria, were discovered. The first of these compounds, chlortetracycline was isolated from *Streptomyces aureofaciens*. At Pfizer, Lloyd Conover joined a team which was exploring the molecular architecture of the broad-spectrum antibiotics oxytetracycline (Terramycin®) and chlortetracycline (Aureomycin®). With his team and working with Robert B. Woodward (1965 Nobel laureate in chemistry), in Harvard University, Conover realized it was possible to chemically alter an antibiotic to produce other antibiotics that could be very effective.

In 1952, Conover developed tetracycline (Figure 1.20) from chlortetracycline by removal of its chlorine atom by catalytic hydrogenation, and then oxytetracycline.[122] The discovery prompted an industry-wide search for superior structurally modified antibiotics, which has provided most of the important antibiotic discoveries made since then.

FIGURE 1.20 The tetracycline molecular model.

Tetracyclines including semi-synthetic derivatives like doxycycline and minocycline are offering a wide range of antimicrobial activity against Gram-positive, Gram-negative bacteria and even some protozoa infections.

(e) Erythromycin

In 1949, Abelardo Aguilar, a Filipino scientist, sent some soil samples to his employer Eli Lilly. Eli Lilly's research team, led by J. M. McGuire, managed to isolate erythromycin from the metabolic products of a strain of *Streptomyces erythreus* (designation changed to "*Saccharopolyspora erythraea*") found in the samples. The product was launched commercially in 1952 under the brand name Ilosone® (after the Philippines region of Iloilo, where it was originally collected from) after formerly being also called Ilotycin®. Even if the drug has earned American drug giant Eli Lilly billions of dollars, neither Aguilar nor the Philippine government ever received royalties. In 1981, Robert B. Woodward and a large team of researchers reported the first stereo-controlled asymmetric chemical synthesis of Erythromycin. To overcome the acid instability of erythromycin, Taisho Pharmaceutical (Tokyo, Japan) found a new antibiotic, with a structure close to macrolides: clarithromycin.[123]

(f) Vancomycin

In 1952, a missionary in Borneo (Indonesia) sent a sample of dirt to his friend, E. C. Kornfield, an organic chemist at Eli Lilly. An organism isolated from that sample *(Streptomyces orientalis)* produced a substance ("compound 05865") that was active against most Gram-positive organisms, including penicillin-resistant *staphylococci*,[124] *clostridia*, and *Neisseria gonorrhea. In vitro* experiments were initiated to determine whether the activity of compound 05865 would be preserved despite attempts to induce resistance. After 20 serial passages of *staphylococci*, resistance to penicillin increased 100,000-fold, compared with only a 4- to 8-fold increase in resistance to compound 05865.[125] Isolates from other laboratories were also tested, and the results were similar. Subsequent animal experiments suggested that compound 05865 might be safe and effective in humans. Before clinical trials, the compound, dubbed "Mississippi mud" because of its brown color, was purified and the resulting drug, which was named "vancomycin" (from the word "vanquish"), was made available.[126] Vancomycin kept its major interest for Gram-positive infections in which bacteria had proved resistant to methicillin-group antibiotics.

(g) Cephalosporins

In the early 1960s, came the emergence of the cephalosporins (related to the penicillins) first isolated from cultures of *Cephalosporium acremonium* from a sewer

in Sardinia in 1948 by Italian scientist Guiseppe Brotzu (University of Cagliari). He noticed that these cultures produced substances that were effective against *Salmonella typhi*, the cause of typhoid fever. Researchers at the William Dunn School of Pathology (Oxford, UK) isolated cephalosporin C, which had stability to β-lactamases but was not sufficiently potent for clinical use. The cephalosporin nucleus, 7-aminocephalosporanic acid (7-ACA), was derived from cephalosporin C and proved to be analogous to the penicillin nucleus 6-aminopenicillanic acid. Modification of the 7-ACA side-chains resulted in the development of useful antibiotic agents. The first agent cephalothin was launched by Eli Lilly in 1964. Since 40 years, four "generations" of cephalosporins had been marketed. First-generation (cephalexin, cefazolin, cefadroxil) are moderate spectrum agents, with a spectrum that includes penicillinase-producing methicillin-susceptible *cocci*, though they are not the drugs of choice for such infections. They also have activity against some *E. coli, K. Pneumoniae*, but have no activity against *B. fragilis, enterocci*, methicillin-resistant *staphylococci, Ps. aeruginosa*, etc. The second-generation cephalosporins (cefuroxime cefoxitin) have a larger Gram-negative spectrum while retaining some activity against *streptococci* or *staphylococci*. They are also more resistant to β-lactamase. Third generation cephalosporins (cefotaxime, cefoperazone, ceftriaxone, ceftazidime) have a broad spectrum of activity and further increased activity against Gram-negative organisms. Some members of this group (particularly those available in an oral formulation, and those with antipseudomonas activity) have decreased activity against Gram-positive organisms. They may be particularly useful in treating hospital-acquired infections, although increasing levels of extended-spectrum β-lactamases are reducing the clinical utility of this class of antibiotics. Fourth generation cephalosporins (cefepime) are extended-spectrum agents with similar activity against Gram-positive organisms as first-generation cephalosporins. They also have a greater resistance to β-lactamases.

(h) Quinolones

The prolific development of the quinolones began in 1962, when George Y. Lesher (Sterling Research, Albany, USA) (Figure 1.21) made the accidental discovery of nalidixic acid as a by-product, 1-8-naphthyridine, during an attempt of the synthesis of the antimalarial compound chloroquine.[127] Since that discovery, the utility of nalidixic acid was largely limited to the treatment of Gram-negative urinary tract infections. Thus, the quinolones have evolved to become important and effective agents in the treatment of bacterial infection. The molecular structures of the quinolones have been adapted over time in association with clinical need (Table 1.5).

The addition of specifically selected substituents at key positions on the quinolone nucleus made it possible to target

FIGURE 1.21 George Y. Lesher.

TABLE 1.5 Quinolones: Structure–Activity Relationship

One fluorine atom at position C-6	Increased DNA gyrase Inhibitory activity
Second fluorine atom at position C-8	Increased absorption Longer elimination half-life Increased phototoxicity
Piperazine group at position C-7	Greatest activity : aerobic Gram-negative bacteria Increased activity: both *staphylococci* and *Pseudomonas* species
Alkylation of the C-7 ring	Improved activity: aerobic gram-positive bacteria
Methyl to the distal nitrogen of the C-7 piperazine ring	Increased elimination half-life and improved bioavailability

specific groups of bacteria (topoisomerases II and IV being the lethal targets) and to improve the pharmacokinetics of the earlier quinolone compounds. The first fluoroquinolone, flumequine, was used transiently until ocular toxicity was reported. Shortly afterwards, second-generation agents were developed, epitomized by ciprofloxacin produced after addition of a cyclopropyl group at position N-1. This agent has a wider spectrum of *in vitro* antibacterial activity, in particular against Gram-negative bacteria, and is effective in the treatment of many types of infection. Despite excellent results in many respiratory infections, reports of failure in pneumococcal infection have limited its use in this area.[128] This drug got celebrity when it later famously resorted to in the 2001 anthrax scare.

4. *The problem of resistance*

Over time, some bacteria have developed ways to circumvent the effects of antibiotics. René Dubos had the foresight

TABLE 1.6 Agent for Which Resistance was Observed

Drug agents	Introduction	First resistance described
Penicillin G	1943	1943
Streptomycin	1947	1947
Tetracycline	1952	1952
Methicillin	1960	1961
Nalidixic acid	1964	1966
Gentamicin	1967	1969
Cefotaxime	1981	1981
Linezolid	2000	1999

to understand the unfortunate potential of antibiotic-resistant bacteria and encouraged prudent use of antibiotics. As a result of this fear, Dubos stopped searching for naturally occurring compounds with antibacterial properties. The widespread use of antibiotics enhanced evolutionarily adaptations that enable bacteria, viruses, fungi, and parasites to survive powerful drugs. Antimicrobial resistance provides a survival benefit to microbes and makes it harder to eliminate infections from the body. Ultimately, the increasing difficulty in fighting off microbes leads to an increased risk of acquiring infections within hospitals. There is a great worry that while many variants of older drugs had been produced, new families of antibiotics were not being found. The promise of molecular biology through the sequencing of bacterial genomes and the design of chemicals designed to attack their weak points had not yet borne fruit by the early 21st century[129] (Table 1.6).

Antimicrobial resistance is a growing threat worldwide, especially within hospitals harboring critically ill patients who are less able to fight off infections without the help of antibiotics. Resistance mechanisms have been found for every class of antibiotic agent and the search for new antibiotics effective against various multiresistant germs is probably one of the most difficult challenges of medicinal chemistry for the next decade. Development of new classes of antibiotics or more robust versions of old classes will be essential for the future.[130]

5. *Anti-*HIV *drugs*

On June 5, 1981, the Centers for Disease Control and Prevention (CDC, Atlanta, USA) published an unusual notice in its *Morbidity and Mortality Weekly Report*: the occurrence of *Pneumocystis carinii* pneumonia among gay men. At the same time, in New York, a dermatologist encountered cases of a rare cancer, Kaposi's sarcoma. By the end of 1981,

FIGURE 1.22 Samuel Broder.

those symptoms were recognized as harbingers of a new and deadly disease later named AIDS. Twenty five years after, more than 40 millions cases were estimated worldwide. In 1984, Luc Montagnier of the Pasteur Institute[131] and Robert Gallo of the National Cancer Institute (NCI) proved that AIDS was caused by a retrovirus (whose replication is linked to a key enzyme, reverse transcriptase). Since the beginning of the epidemics, many of the therapeutic strategies have yielded positive results.[132] In 1985, nucleoside analogs called dideoxynucleosides were discovered to be potent inhibitors of human immunodeficiency virus (HIV) replication *in vitro*. The first drug introduced to treat the disease was AZT (3′-azido-3′-deoxythymidine, azidothymidine, zidovudine) a thymidine analog previously developed in 1964 as an anticancer drug by Jerome Horowitz of the Michigan Cancer Foundation (Detroit, USA). But AZT, being ineffective against cancer, Horowitz did not register a patent.[133] In 1987, after 3 years of intensive research,[134] the ultimate approval of AZT as an antiviral treatment for AIDS was the result of pharmacological technology and the personal determination of Samuel Broder (Figure 1.22), a physician and researcher at the NCI.[135] A 6-week clinical trial of AZT was sufficient to prove its potent antiviral activity against HTLV-III in patients with AIDS or AIDS-related complex.[136] Dideoxynucleosides selectively inhibit HIV reverse transcriptase after they are phosphorylated within the cell to 5′-triphosphates. AZT was only a first step in developing new therapy for AIDS. Its use has been associated with toxicities, particularly bone marrow suppression and several groups have reported the development of AZT-resistant strains of HIV. Other dideoxynucleosides with toxicity profiles different from that of AZT had

also shown activity against HIV in early clinical studies. In September 1995, the results of the "Delta trial" showed that combining AZT with ddI (didanosine) or ddC (zalcitabine) did provide a major improvement in treatment compared with AZT on its own.[137]

Other nucleoside analogs including stavudine, lamivudine, abacavir, and tenofovir, began to be used in 1994 while non-nucleoside reverse transcriptase inhibitors (non-NRTI), nevirapine, delavirdine, and efavirenz came to the market in 1996. Studies have shown that the binding of HIV to lymphocytic CD4 receptor may be blocked by genetically engineered forms of CD4 protein. It has been proved that HIV protease could be inhibited by substrate analogs. Protease inhibitors (PIs) were thus an excellent area for drug design, development, and production. They constitute the third sub-class of antiretroviral (ARV) drugs. The discovery of PIs has been hampered by a number of significant obstacles. Foremost among these, is the identification of inhibitors which simultaneously embody potent anti-HIV activity and high oral bioavailability accompanied by a long elimination half-life to yield sustained virus-suppressive drug levels in the blood and infected tissues. HIV PIs represent a critical milestone on the path to therapeutic efficacy.[138] Saquinavir was the first PI approved by the FDA in 1995 as Invirase®, a poorly absorbed hard gel capsule which quickly led to viral resistance. It was approved again on November 1997 as Fortovase®, a soft gel capsule reformulated with improved bioavailability. Last step, in 2006, owing to reduction in demand, Fortovase® ceased being in favor of Invirase® boosted with another potent PI ritonavir. Already, in 2000, the co-formulation of lopinavir and low-dose ritonavir (Kaletra®) had benefited from the fact that a sub-therapeutic dose of ritonavir, a potent cytochrome P_{450}3A4 inhibitor, inhibited the metabolism of lopinavir, resulting in higher lopinavir concentrations than when lopinavir is administered alone. This pharmacokinetic interaction is associated with a high lopinavir trough level and good general tolerability when compared with other PIs. The concept of pharmacokinetic enhancement – boosting – was not new as ritonavir has previously been used in this context with other PIs. Even if the relationship between plasma and intracellular drug levels has not been clarified, ARVtrough plasma concentrations are correlated with virological outcome. On the other hand, Kaletra® has reduced pill-burden and aids compliance.[139] Since 1996, highly active ARV therapies usually consisting of two NRTIs plus a PI (saquinavir, ritonavir, indinavir, nelfinavir, amprenavir or lopinavir) have been widely used. They produce durable suppression of viral replication with undetectable plasma levels of HIV-RNA in more than half of patients. Immunity recovers, and since 1995, AIDS morbidity and mortality fall by more than 80%. Besides these successes, however, ARV therapies also produce numerous side effects. These challenges prompt the search for new drugs and new therapeutic strategies to control chronic viral replication;[140] among them, antisense oligonucleotide therapy could target the regulatory genes of HIV.[141] Nevertheless, it is admitted that the preparation of an effective vaccine is probably the only way to eradicate the disease[142] (Table 1.7).

TABLE 1.7 Milestones in the Fight Against HIV

1981	Centers for Disease Control and Prevention (CDC Atlanta) report an alarming recrudescence of Kaposi's sarcoma in healthy gay men
1982	The term AIDS is used for the first time on July 27th
1983	CDC (USA) warns blood banks of a possible problem with the blood supply Major outbreak of AIDS in central Africa Luc Montagnier in Pasteur Institute, Paris, France (and later, Robert Gallo in the United States) isolates a retrovirus, later known as human immunodeficiency virus or HIV
1985	FDA approves the first enzyme linked immunosorbent assay (ELISA) test kit to screen for antibodies to HIV
1987	AZT (zidovudine) is the first anti-HIV drug approved. It has to be taken by one 100 mg capsule every 4 h around the clock
1991	DDI (didanosine): NRTI
1992	DDC (zalcitabine): NRTI The first clinical trial of multiple drugs is held
1994	D4T (stavudine): NRTI
1995	Saquinavir, the first anti-HIV drug as aPI 3TC (lamivudine): NRTI The FDA approves Saquinavir in a record 97 days
1996	• The FDA approves an HIV viral load test • Nevirapine, the first anti-HIV drug of the non-nucleoside reverse transcriptase inhibitors (NNRTI) • Ritonavir: PI • Indinavir: PIs

D. Drugs for immunosuppression

In the years following World War II, Sidney Farber, a cancer scientist at Boston's Children's Hospital, was testing the effects of folic acid on cancer. Some of his results, which now look dubious, suggested that folic acid worsened cancer conditions, inspiring chemists at Lederle to make antimetabolites – structural mimics of essential metabolites that interfere with any biosynthetic reaction involving the intermediates – resembling folic acid to block its action. These events led to the 1948 development of methotrexate, one of the earliest anticancer agents and the mainstay of leukemia chemotherapy.[143] At the time, George Hitchings

FIGURE 1.23 George Hitchings and Gertrude Elion.

and Gertrude Elion (Wellcome Research, Tuckahoe, USA) (Figure 1.23) pioneered the design of immunosuppressants demonstrating also anticancer activity.[144] The corticosteroids were used in 1948 by Kendall and Hench and gave the spectacular demonstration of their successful use in patients with Rheumatoid Arthritis (RA). They were subsequently used as immunosuppressants in various clinical situations. Among immunosuppressants described, a nitrogen mustard-like[145], cyclophosphamide, alkylates DNA bases and preferentially suppresses immune responses mediated by β-lymphocytes.

Methotrexate and its polyglutamate derivatives suppress inflammatory responses through release of adenosine; they suppress immune responses by inducing the apoptosis of activated T-lymphocytes and inhibiting the synthesis of both purines and pyrimidines.[146] Azathioprine, studied by Roy Calne (Cambridge, UK), inhibits several enzymes of purine synthesis.[147] Often, their mechanisms of action were established long after their discovery. For instance, glucocorticoids inhibit the expression of genes coding for interleukin-2 (IL-2) and other mediators.[148] After the "heroic period" of the 1950s and 1960s when corticosteroids were proposed to prevent organ rejection in renal transplantation, pharmacology helped surgery to enter a new era of optimism, characterized by improving allograft survival rates. Revolutionary methods of rejection treatment have been responsible for this new era[149] few years after the first heart transplant performed in Cape Town in 1967 by Christiaan

FIGURE 1.24 Jean-François Borel.

Barnard. A few determined individuals in the medical and research community spent the next two decades attempting to solve the organ rejection puzzle. One of these scientists was Jean-Francois Borel, (Figure 1.24) who worked for Sandoz Pharmaceuticals (Basel, Switzerland). He discovered the immunosuppressant agent that ultimately moved transplantation from the realm of curiosity into routine therapy. Both J. Borel and H. Stähelin markedly contributed to the discovery and characterization of the biological profile of what will become a revolutionary drug.[150] In its subsequent exploitation, Borel played the leading role[151]. He chose to examine a compound that was isolated from the soil fungus *Tolypocladium inflatum* Gams.

Borel discovered that, unlike immunosuppressants that acted indiscriminately, this compound selectively suppressed the T-cells of the immune system. Excited by these characteristics, Borel continued his study and, in 1973, purified the compound called cyclosporine.[152] Cyclosporine is active in two ways. First, it impedes the production and release of IL-2 by T-helper white blood cells. Secondly, it inhibits IL-2 receptor expression on both T-helper and T-cytotoxic white blood cells. Tacrolimus and cyclosporine, known as calcineurin inhibitors, act on the IL-2 by inhibiting its production thus leading to a decrease in the proliferation of the activated lymphocyte. In addition, these compounds have recently been found to block signaling pathways triggered by antigen recognition in T-cells.[153] In contrast, rapamycin inhibits kinases required for cell cycling and responds to IL-2. Rapamycin also induces apoptosis of activated T-lymphocytes. Mycophenolate mofetil reduces the proliferation of T-cell by inhibiting purine synthesis and by its action on inosine monophosphate dehydrogenase, thereby depleting guanosine nucleotides and inducing apoptosis of activated T-lymphocytes.[154] Antilymphocyte antibodies (globulines) deplete circulating lymphocytes while selective

monoclonal antibodies are directed against IL-2 receptor thus reducing the rate of proliferation of activated T-cells. With the availability of such potent and diverse agents it is now possible to develop multi drug regimens that can depress the immune system at the different steps of the activation cascade, with minimal side effects, thus improving graft and patient survival rates.[155]

E. Contribution of chemists to the fight against cancer

1. *Drugs against the cancer cell*

Chemistry has had a major role in the discovery and development of most new anticancer drugs, and some of these are still relevant to drug discovery today. The dawning of cancer chemotherapy is generally accepted to have been the serendipitous discovery of the mustard family of agents in the first half of the 20th century. The application of medicinal chemistry to sulfur mustard gas led to agents that are still clinically useful today. The change of a sulfur atom in favor of a nitrogen one transformed the worse war weapon into a beneficial drug active against cancer.

(a) Nitrogen mustards

Thus, the first agents were nitrogen mustards (halogenated alkyl amine hydrochlorides) among which 2-2′-2″-trichlotriethylamine was the prototype first studied by pharmacologists from Yale University (USA). Louis S. Goodman, Maxwell M. Wintrobe, William Dameshek, Morton Goodman, Alfred Gilman, and Margaret McLennan[156] performed in 1943 but only presented in 1946, the salutary results obtained in patients treated for Hodgkin's disease, lymphosarcoma, and leukemia by this first nitrogen mustard. Indeed, in the first two disorders, dramatic improvement has been observed in an impressive proportion of terminal and so-called radiation resistant cases. First constant success in hematological malignancies where obtained when Vincent T. De Vita, Arthur A. Serpick, and Paul P. Carbone (NCI) increased in 1970 the response rate with "MOPP" therapy combining mechlorethamine, vincristine, procarbazine, and prednisone. This protocol was superior to that previously reported with the use of single drugs with 35 of 43, or 81% of the patients achieving a complete remission, defined as the complete disappearance of all tumor and return to normal performance status.[157] Cyclophosphamide and the related alkylating agent ifosfamide were further developed by Norbert Brock for ASTA pharmaceuticals (Bielefeld, Germany). Brock and his team synthesized and screened more than 1,000 candidate oxazaphosphorine compounds.[158] The main effect of cyclophosphamide is due to its metabolite phosphoramide mustard which is only formed in cells with low levels of aldehyde deshydrogenase. Phosphoramide mustard forms DNA cross links between and within DNA strands at guanine N-7 positions, leading to cell death.

FIGURE 1.25 Catharanthus roseus.

(b) Vinca alkaloids, taxans and other plant anticancer drugs

Although "serendipity" is not a reliable source of new anticancer-drug leads, more molecules with interesting anticancer properties might still appear through chance in the future, especially from natural products. Second, synthetic chemistry has been used to modify drug leads discovered in plant material – the so-called semi-synthetic approach.

A lot of anticancer drugs have been extracted from plants.[159] Eighteen centuries ago, Galen proposed the juice expressed from woody nightshade (*Solanum dulcamara*) to treat tumors and warts, which has been demonstrated to exert anti-inflammatory properties.[160] In the recent decades, more than 1,600 genera have been examined.[161] Extracts of the leaves of the subtropical plant *Catharanthus roseus* (L.), (Figure 1.25) Madagascar periwinkle, were reputed among ethnobotanists to be useful in the treatment of diabetes. The attempt to verify the antidiabetic properties of the extracts led instead to the discovery and isolation of two complex indole alkaloids, vinblastine and vincristine, which are used in the clinical treatment of a variety of lymphomas, leukemias, and cancers as small cell lung or cervical and breast cancer.[162] The two alkaloids, although structurally almost identical, nevertheless differ markedly in the type of tumors that they affect and in their toxic properties. From the use of *Podophyllum* in ancient China, a lot of plant derivatives are being used in cancer chemotherapy: two glycosides were extracted to prepare podophyllotoxin, and subsequently two semi-synthetic derivatives, etoposide and

teniposide.[163] As the 1970s opened, US President Richard Nixon established the National Cancer Program, popularly known as the "war on cancer," with an initial half-billion dollars of new funding. This may explain why, in the following years, many new compounds with antineoplastic properties were isolated in plants. Among them, the pyridocarbazole alkaloids ellipticine and 9-methoxyellipticine from *Ochrosia elliptica*, intercalate between the base pairs of DNA.[164] Camptothecin and its derivatives, alkaloids from Chinese tree *Camptotheca acuminata*, showed a broad-spectrum anticancer activity. 9-Aminocamptothecin gave birth to topotecan and irinotecan.[165] Alkaloids from *Cephalotaxus* species were isolated for experimental and clinical studies. If the parent alkaloid is inactive, the esters harringtonine and homoharringtonine obtained from *Cephalotaxus harringtonia*, according to Hagop M. Kantarjian and Moshe Talpaz, (M. D. Anderson Cancer Center, Houston, USA), give new hopes in the cure of solid tumors or leukemias.[166]

The most enthusiastic reports concern the diterpenoids paclitaxel, Taxol® (from *Taxus brevifolia*) and docetaxel, Taxotere® (from *Taxus baccata*) having unique tri- or tetracyclic 20 carbon skeletons extracted from the bark of yew. This tree was known as a toxic plant for animals and humans for centuries.[167] Monroe E. Wall and Mansukh C. Wani, at the Research Triangle Park (Chapel Hill, USA), identified the active principle of the yew tree in 1971.[168] In 1979, Susan Horwitz of the Department of Molecular Pharmacology, Albert Einstein College of Medicine (New York) suggested that paclitaxel's mechanism of action was different from that of any previously known cytotoxic agent. She observed an increase in the mitotic index of P388 cells and an inhibition of human HeLa and mouse fibroblast cells in the G2 and M phases of the cell cycle.[169] It has been suggested that Taxol exerted its activity by preventing depolymerization of the microtubule skeleton. Clinical use of paclitaxel includes a lot of solid tumors with best results in ovarian and breast cancers. Extraction of paclitaxel from the yew bark was quite difficult: three trees for 1 g of drug (one cure of chemotherapy). This difficulty encouraged the pursuit of semi-synthetic production.

The strategy included immediately increasing the amount of paclitaxel derived from yew bark and establishing a broad research program to evaluate alternative sourcing options and their commercial feasibility.[170] The prospects for finding a solution to the paclitaxel supply problem through semi-synthesis using a naturally occurring taxan as a starting material, were considerable. This approach was pioneered by Pierre Potier (Figure 1.26 in the Institut de Chimie des Substances Naturelles (Gif-sur-Yvette, France).[171] He found in the early 1980s that a naturally occurring taxan containing the paclitaxel core, 10-deacetylbaccatin III, was 20 times more abundant than paclitaxel and was primarily contained in the needles of the abundant English Yew (*Taxus baccata*). Potier succeeded in the difficult conversion of 10-DAB into paclitaxel, in 1988, with only four steps with an overall yield being 35%, still significantly less than needed for an efficient commercial process.[172] The discovery and development of the taxans class of antitumor compounds, involved the discovery of a paclitaxel semi-synthetic analog, docetaxel (Taxotere®), by Pierre Potier *et al.*, represent significant advances in the treatment of patients with a variety of malignancies. Although paclitaxel and docetaxel have a similar chemical root, extensive research and clinical experience indicate that important biological and clinical differences exist between the two compounds. Although the mechanism by which they disrupt mitosis and cell replication is unique, there are small but important differences in the formation of the stable, non-functional microtubule bundles and in the affinity of the two compounds for binding sites.[173] These differences may explain the lack of complete cross-resistance observed between docetaxel and paclitaxel in clinical studies.[174] Besides natural products, synthetic anticancer drugs flourished in various directions.

FIGURE 1.26 Pierre Potier.

(c) Antimetabolites

A third way in which chemistry has generated anticancer-drug was conducted through screening programs using cancer cell lines *in vitro*. In the 1950s, the NCI set up a series of screening programs that invited chemists from around the world to submit their novel compounds for screening against a range of *in vitro* tumor cell lines. Antimetabolites interest in cancer treatment had been discovered by George Hitchings, head of the department of biochemistry at Burroughs Wellcome, and Gertrude Elion,

utilizing what today is termed "rational drug design." They methodically investigated areas where they could see cellular and molecular targets for the development of useful drugs. During their long collaboration, they produced a number of effective drugs to treat a variety of illnesses, including leukemia, malaria, herpes, and gout. They began examining the nucleic acids, particularly purines, including adenine and guanine, two of DNA's building blocks. They discovered that bacteria could not produce nucleic acids without the presence of certain purines, and set to work on antimetabolite compounds which locked up enzymes necessary for the incorporation of these purines into nucleic acids.[175] They synthesized two substances, diaminopurine and thioguanine, which the enzymes apparently latched onto instead of adenine and guanine. These new substances proved to be effective treatments for leukemia, a blood malignancy characterized by a great increase in white blood cells count, due to the activity of oncogenes. Later, Elion substituted an oxygen atom with a sulfur atom on a purine molecule, thereby creating 6-mercaptopurine used to treat leukemia. After this success, Elion and Hitchings developed a number of additional drugs using the same principle. Later, these related drugs were found to not only interfere with the multiplication of white blood cells, but also suppress the immune system. This latter discovery led to a new drug, azathioprine (Imuran®), and a new application – organ transplants.[176] The team also developed allopurinol, a drug successful in reducing the body's production of uric acid, thereby treating gout, pyrimethamine, used to treat malaria, and trimethoprim used to treat bacterial infections. With Howard Schaeffer, Elion was also at the origin of acyclovir, marketed as Zovirax® which interferes with the replication process of the herpes virus,[177] drug characterized by a radical antiviral efficiency, a fair inocuity and, despite an extensive use for 30 years, very few cases of viral resistance. Hitchings and Elion won the Nobel Prize in Physiology or Medicine in 1988 for their discoveries of "important principles for drug treatment," which constituted the groundwork for rational drug design.

(d) Anticancer antibiotics

Anthracyclines may be listed among the main anticancer drugs. Daunomycin (also called daunorubicin) was isolated from *Streptomyces peucetius* in 1962 by Aurelio Di Marco from Farmitalia (Milan, Italy).[178] With adriamycin it is the prototypical member in the anthracyclines antitumor antibiotic family. Adriamycin (a 14-hydroxy derivative of daunorubicin) was isolated from the same microorganism, in 1967. Despite their severe cardiotoxicity and other side effects, these drugs have been widely used as dose-limited chemotherapeutic agents for the treatment of human solid cancers or leukemias since their discovery.[179] These antibiotics contain a quinone chromophore and an aminoglycoside sugar. Their antineoplastic activity has been mainly attributed to a strong interaction with DNA in the target cells.[180] While anthracyclines can be very effective against breast, lung and other cancers, they pose a risk of cardiotoxicity and therefore, they are typically used in limited doses. Doxorubicin and Epirubicin are commonly used in combination with other chemotherapy drugs to help decrease the risk of side effects.[181]

FIGURE 1.27 Barnett Rosenberg.

(e) Platinum anticancer drugs

Cisplatin was discovered serendipitously in 1965 while Barnett Rosenberg (Figure 1.27) and Loretta Van Camp (Michigan State University, USA) were studying the effect of an electric current on *E. coli*.

It was found that cell division was inhibited by the production of *cis*-diammine-dichloroplatinum from the platinum electrodes rather than by the method expected[182] (Figure 1.28). Further studies on the drug indicated that it possessed antitumor activity.

In 1972, the NCI introduced cisplatin into clinical trials. It now has a major role in the treatment of testicular, ovarian, head and neck, bladder, esophageal, and small cell lung cancers. Cisplatin is a square planar compound containing a central platinum atom surrounded by two chloride atoms and two molecules of ammonia moieties. The antitumor activity has been shown to be much greater when the chloride and ammonia moieties are in the *cis* configuration as opposed to the *trans* configuration. The cytotoxicity of cisplatin is due to its ability to form DNA adducts.[183] The drug is able to enter the cell freely in its neutral form, yet once in the cell the chloride ions are displaced to allow the formation of a more reactive, aquated compound. In 1975, Memorial Sloan-Kettering Cancer Center (New York) initiated trials

FIGURE 1.28 Effect of the electric current transmitted through platinum electrodes on *E. coli* culture.

of cisplatin alone and later in combination with cyclophosphamide and/or adriamycin in patients with urothelial tract cancer. The results were not as positive as those seen in the testicular cancer studies, but they were favorable.[184] Studies using cisplatin alone and in combination with adriamycin to fight ovarian cancer done by Holland gave substantial improvements.[185] Due to the extreme toxicity of cisplatin, as well as resistance against it, there has been a need for the development of analogs which are just as potent, but not as toxic. Several cisplatin analogs were considered as viable alternatives to the parent drug in terms of their toxicities, antitumor properties, and potential biochemical selectivity but it has been concluded that diamine(1, 1-cyclobutane-dicarboxylato)platinum (II), or carboplatin, was the most interesting.[186] From that time a lot of comparisons were performed between cisplatin and carboplatin.[187] Triplatin tetranitrate was discovered by Nicholas Farrell (Virginia Commonwealth University)[188] and John Broomhead (Camberra, Australia)[189] and developed by Swiss pharmaceutical company Roche. It is a new platinum anticancer drug, result of a trinuclear platinum coordination complex with chloride and amine ligands. It is active through covalent adducts with DNA inducing apoptosis.

2. A *new generation of anticancer drugs*

The fourth, and most recent, application of chemistry in anticancer drugs conception has been to generate drug leads following the discovery of a new target; for example, the specific identification of a cancer-related gene from the sequence of the human genome. Such targets can be proteins, enzymes, or nucleic acids. Knowledge of the three-dimensional structure of a target, obtained using X-ray crystallography, can lead to the rational design of specific inhibitor molecules that target functionally important parts of the structure. It is to be considered that most drugs that have been discovered through screening are highly toxic agents, whereas most efforts are now to discover anticancer drugs targeting molecular aberrations that are specific to tumor cells. Whether this goal will be attained for common human solid cancers that have become established is still unclear, in view of the widespread misregulation of signaling pathways. The emergence of tumor-specific, molecularly targeted agents signifies a paradigm shift in cancer therapy, with less reliance on drugs that non-discriminately kill tumor and host cells. Two examples may illustrate this purpose.

(a) Protein tyrosine kinases (TKs)

TKs are enzymes that catalyze the transfer of phosphate from adenosine triphosphate (ATP) to tyrosine residues in proteins. The human genome contains about 90 TKs regulating cellular proliferation, survival, differentiation, function, and motility. TKs were largely ignored in drug development because of a paucity of evidence for a causative role in human cancer and concerns about drug specificity and toxicity. The knowledge of their importance was evident. Imatinib mesylate, an inhibitor of the BCR–ABL was successfully used in chronic myeloid leukemia.[190] First results with imatinib were obtained by Ciba Company in 1996.[191] TKs are now regarded as excellent targets for cancer chemotherapy, but reality lies somewhere between the extremes of triumph and tribulation.[192]

(b) Antiangiogenic agents

Antiangiogenic agents may target angiogenesis. During cancer cell proliferation, tumor growth is accompanied by the formation of new blood capillaries from preexisting vessels. This neovascularization plays both beneficial and damaging roles in the organism. Vascular endothelial growth factor (VEGF) identified in the 1980s, is one of the most important pro-angiogenic factors involved in tumor angiogenesis. VEGF increases vascular permeability, which might

FIGURE 1.29 Judah Folkman.

facilitate tumor dissemination via the circulation causing a greater delivery of oxygen and nutrients; it recruits circulating endothelial precursor cells, and acts as a survival factor for immature tumor blood vessels. VEGF and its receptors play a central role in tumor angiogenesis, and therefore the blockade of this pathway is a promising therapeutic strategy for inhibiting angiogenesis and tumor growth.

A number of different strategies to inhibit VEGF signal transduction were developed including anti-VEGF monoclonal antibodies, receptor antagonists, soluble receptors, antagonistic VEGF mutants, and inhibitors of VEGF receptor function. These agents can be divided in two broad classes, namely agents designed to target the VEGF activity and agents designed to target the surface receptor function.[193] The idea that inhibiting tumor angiogenesis might be an effective anticancer strategy was proposed by Judah Folkman (Boston, USA) (Figure 1.29) over 30 years ago.[194] In 1993, it was shown that a monoclonal antibody that targeted VEGF resulted in a dramatic suppression of tumor growth *in vivo*, which led to the development of bevacizumab (Avastin®, Genentech), a humanized variant of this anti-VEGF antibody, as an anticancer agent. Approval of bevacizumab in 2004 as a first-line therapy for metastatic colorectal cancer validated the ideas that VEGF was a key mediator of tumor angiogenesis and that blocking angiogenesis was an effective strategy to treat cancer.[195] Although the diversity of targets giving rise to this new generation of anticancer drugs has expanded, many challenges persist in the design of effective treatment regimens. The complex interplay of signal-transduction pathways further complicates the customization of cancer treatments to target single mechanisms. However, despite uncertainty over precise or dominant mechanisms of action, especially for compounds targeting multiple gene products, emerging agents are producing significant therapeutic advances against a broad range of cancers[196] (Table 1.8).

TABLE 1.8 Main Steps of Anticancer Chemotherapy[197]

1942	Louis Goodman and Alfred Gilman use nitrogen mustard to treat a patient with non-Hodgkin's lymphoma, inducing tumor regression
1948	Sydney Farber uses antifolates successfully to induce remissions in children with acute lymphoblastic leukemia
1948	George Hitchings and Gertrude Elion synthesize the purine analog 6-mercaptopurine
1958	Roy Hertz and Min Chu Li demonstrate that methotrexate as a single agent can cure choriocarcinoma (first solid tumor cured)
1959	Approval of cyclophosphamide, alkylating agent
1965	Treatment of acute lymphoid leukemia with combination chemotherapy including vincristine (Bernard)
1970	Treatment of lymphoma with combination chemotherapy (De Vita)
1972	Emil Frei demonstrates that chemotherapy given after surgical removal of a tumor increases cure rates (adjuvant chemotherapy)
1975	Cyclophosphamide methotrexate fluorouracil combination is effective for treatment of node-positive breast cancer
1978	Cisplatin is proved to be effective in ovarian cancer
1989	Pierre Potier succeeds in docetaxel synthesis
1992	FDA approves paclitaxel (Taxol) which is the first "blockbuster" in oncology
2001	Brian Druker studies imatinib, first tyrosine kinase inhibitor for treatment of chronic myelogenous leukemia
2004	Approval of bevacizumab (first anti-VEGF) in the treatment of colorectal cancer

F. Drugs for endocrine disorders

At the beginning of the 20th century, two chemical discoveries gave a new turn in the research of endocrine disorders. First, in 1906, Mikhail Tswett (Warsaw, Poland) developed the all-important technique of column chromatography allowing separating chemical entities in complex media. Almost immediately after, Svante August Arrhenius (Stockholm, Sweden) and Soren Sorensen (Copenhagen, Denmark) demonstrated in 1909 that pH could be measured; Sorensen pointing out that pH could affect enzymes activity. This discovery was a critical step in the development of a biochemical model of metabolism and kinetics. Some critical breakthroughs in metabolic medicine had

FIGURE 1.30 Charles Best and Frederick Banting.

been made in the 1890s, but they were exceptions rather than regular occurrences. In 1891, myxedema was treated with sheep thyroid injections. This was the first proof that animal gland extracts could benefit to patients. In 1896, Addison's disease was treated with chopped up adrenal glands from a pig. These test treatments provided the starting point for all hormone research. From the 1920s to the 1940s new major treatments for physiological disorders were discovered and mainly among them, insulin for diabetes mellitus and cortisone for inflammatory diseases.

1. *Antidiabetic drugs*

(a) Insulin

For most of human history, diabetes mellitus meant certain death. Since the late 19th century, scientists attempted to isolate the essential hormone and inject it to patients to control the disease. Using dogs, numerous researchers had tried and failed, but in the late spring of 1921, Frederick Banting worked on his project in Toronto University (Canada) with his young medical student, Charles Best (Figure 1.30). After many failures, one of the dogs whose pancreas had been tied off showed signs of diabetes. Banting and Best removed the pancreas, ground it up, and dissolved it in a salt solution to create the long-sought extract. They injected the extract into the diabetic dog, and within a few hours the canine's elevated blood sugar returned to normal. The scientists had created the first effective treatment for diabetes.

FIGURE 1.31 Billy Leroy, first patient having received insulin therapy – before treatment (left) and after (right).

John MacLeod, physiologist at the same department, provided facilities for Banting's work, biochemists James Collip and E. C. Noble joining the research team to help purify and standardize the hormone, which was renamed insulin.[198] Only Banting and MacLeod will be awarded with Nobel Prize in 1923. Connaught and Lilly in Northern America and Novo in Europe (Denmark) performed technical developments that enabled large-scale collection of raw material, extraction and purification of insulin, and supplying of the drug in a suitable state for clinical use (Figure 1.31).

During the 1960s new developments in peptides engineering led to synthetic insulin and in the 1970s, biotechnology developments gave birth to genetically manufactured insulin. Since the end of the 1990s, two new short-acting semi-synthetic analogs were marketed. Insulin lispro and insulin aspart can be administered as pre-prandial bolus injections, thereby synchronizing insulin administration and food absorption. In clinical trials, blood glucose levels were significantly less after treatment with insulin lispro or insulin aspart than with regular insulin.[199] Because of their short duration of action, a slightly greater basal insulin supply may be needed when those new insulins are used. Insulin glargine is a long-acting human insulin analog also prepared by recombinant DNA technology. Modification of the human insulin molecule at position A_{21} and at the C-terminus of the B-chain results in the formation of a more stable compound. The plasma concentration versus time profile of insulin glargine is therefore relatively constant in relation to conventional human insulins, with no pronounced peak over 24 h. This allows once-daily administration as basal therapy. Early randomized trials with insulin glargine generally showed greater reductions in fasting blood or plasma glucose levels and a reduced frequency of nocturnal hypoglycemia relative to neutral protamine Hagedorn (NPH) insulin in patients with type 1 diabetes mellitus.[200] Inhalation devices for aerosolized regular insulin offer an alternative to pre-meal subcutaneous bolus injections. It is absorbed more rapidly than subcutaneous insulin and may

therefore be given closer to mealtime; this theoretical interest remains to be confirmed.[201] During the last 50 years, diabetes became an increasing problem in healthcare systems, spreading among developed as well as in the poor countries, with severe consequences for heart, vascular, kidney, retina, and nerve diseases.[202] The prevalence of all types of diabetes is on the rise in the world's population, increasing by 4–5% per year with an estimated 40–45% of individual's over the age of 65 years having either type II diabetes or impaired glucose tolerance. This is why the search for orally active antidiabetic agents became so exciting for drug industry. Oral antidiabetic agents differ with regard to mechanisms of action, hemoglobin A_{1c}-lowering efficacy and safety. Traditional agents consist of those that enhance insulin secretion (sulfonylurea and glinides), those that enhance insulin sensitivity (metformin and thiazolidinediones) and those that inhibit intestinal carbohydrate absorption (α-glucosidase inhibitors).

(b) Biguanides

Biguanides are prescribed for many decades. *Galega officinalis* was used for diabetes treatment in traditional medicine for centuries and, in the 1920s, guanidine compounds (galegine) were discovered in *Galega officinalis* extracts. It was showed in animals that they lower blood glucose, leading to their therapeutic use. But they were withdrawn in 1932 due to their hepatotoxic effects. Some less toxic derivatives, synthalin A and B were used for diabetes treatment but after the discovery of insulin they were forgotten for several decades. Biguanides were reintroduced by two German diabetologists, Hellmut Mehnert and Walter Seitz (Munich) into type 2 diabetes treatment in the late 1950s.[203] The first to be marketed, phenformin, has been widely used but its potential for fatal lactic acidosis in the elderly patient or in case of renal impairment resulted in its withdrawal at the end of the 1970s. Metformin had a much better safety profile and remains the only biguanide drug used in pharmacotherapy worldwide. It is now extensively used and was recently postulated that it could work as an "antidote" of glyoxal and methylglyoxal as advanced glycation endproducts when glyoxalases enzyme systems are inefficient.[204]

(c) Sulfonylureas

Sulfonylureas were discovered in the 1940s by the French pharmacologist, Marcel Janbon (Montpellier) who tried to find an effective antityphoid molecule, tested few sulfonylureas in animals, one of them, sulfamidothiodiazol (carbutamide), induced a rapid drop in the animals' blood sugar. Then, Janbon convinced August Loubatieres, a brilliant clinician to try the drug on diabetic patients. It triggered a fall in these patients' blood sugars, inducing a severe hypoglycemia.[205] Later, these products were found to stimulate insulin secretion by the endocrine pancreas. *In vitro* studies have also shown that they bind specifically to an ATP-dependent K^+ channel of the β-cell membrane. This binding closes the channel so that K^+ outflow ceases, the β-cell membrane depolarizes and voltage-dependent Ca^{++} channels open to allow an influx of extracellular calcium. The result is migration and extrusion of insulin granules.[206]

(d) Thiazolidinediones

Thiazolidinediones act by binding to peroxisome proliferator-activated receptors (PPARs), a group of nuclear receptors, and specifically PPARγ. The usual ligands for these receptors are free fatty acids (FFA) and prostaglandins. When activated, the receptor migrates to the DNA, activating specific gene transcription. In the early 1980s, ciglitazone, the prototypical 2,4-thiazolidinedione, was discovered by Hiroshi Imoto (Takeda, Osaka, Japan) and Takashi Sohda (Fukuoka University, Japan). It had antihyperglycemic activity in insulin-resistant animal models, but no effect in insulin-deficient animal models of diabetes. During structure–activity relationship studies on 2,4-thiazolidinediones and related compounds, highly potent compounds, such as pioglitazone were discovered.[207] Pioglitazone, rosiglitazone, and troglitazone were then synthesized, the last being withdrawn from market because of liver toxicity.[208]

(e) Meglitinides

Meglitinides help the pancreas produce insulin and are often called "short-acting secretagogues." Their mode of action is original. By closing the potassium channels of the pancreas β-cells, they open the calcium channels, hence enhancing insulin exocytosis.[209] Repaglinide or nateglinide are taken with meals to boost the insulin response to each meal. The action of agonists on various PPARs results in improved glucose, lipid, and weight management, with effects dependent on full or partial agonist activity at single or multiple receptors. Although the dual PPAR compounds have been associated with unacceptable toxicities, new PPAR agonist medications continue to be developed and investigated to discover a safe drug with benefits in multiple disease states.[210] New oral agents recently included the dipeptidyl peptidase-4 (DPP-4) inhibitors, which potentiate the activity of the incretin glucagon-like peptide-1 (GLP-1) and enhance glucose-dependent insulin secretion.[211] Exenatide is the first of this new class of medications. The peptide exendin-4, derived from the saliva of the Gila Monster (*Heloderma suspectum*), a venomous lizard (with neurotoxic bites) discovered by Hans Christoph Fehmann, Rüdiger and Burkhard Göke (Marburg, Germany),[212] is a 39 amino acid peptide that mimics the GLP-1 incretin, an insulin secretagogue with glucoregulatory effects but is more potent and longer acting in humans than GLP-1. While it may lower blood glucose levels on its own, it can also be

combined with other medications such as pioglitazone, metformin, sulfa drugs, or insulin to improve glucose control. The medication has to be injected twice per day. New oral DPP-4 inhibitors vildagliptin or sitagliptin give some hope in glycemic control by increasing GLP-1 as its catabolism is inhibited by dipeptidyl peptidase.

2. *Corticosteroids*

If insulin revolutionized diabetes mellitus treatment, cortisone discovery was another revolution in inflammation and arthritis management. The discovery of corticosteroids as therapeutic agents can be linked to Thomas Addison (Guy's Hospital, London), who made the connection between the adrenal glands and the rare Addison's disease in 1855.[213] But the turn came when Edward Calvin Kendall[214] at the Mayo Clinic (Rochester, USA) and Tadeusz Reichstein[215] at the University of Basel (Switzerland) independently isolated several hormones from the adrenal cortex. In 1948, Kendall and Philip S. Hench demonstrated the successful treatment of patients with RA using cortisone.[216] Kendall, Reichstein, and Hench were awarded the 1950 Nobel Prize in Physiology or Medicine (Figure 1.32). Corticosteroids are being used in various clinical conditions. Their ability to modulate the immune response and to diminish inflammation make them useful in rheumatology, respiratory diseases, allergies, endocrine and metabolic disorders, blood disorders, gastrointestinal diseases, neurological and muscular diseases, renal diseases, cardiovascular disorders, and skin diseases. They have been widely tried empirically and, sometimes, they have proved unequivocally effective. Among the most current use of corticosteroids is chronic obstructive pulmonary disease, including asthma. Findings of several large randomized clinical trials have shown benefits for this population of regular treatment with low doses of inhaled corticosteroids. Additional drugs are rarely needed, and although leukotrienes modifiers are effective, they are less so than inhaled corticosteroids.[217]

3. *The contraceptive "pill"*

The birth of steroids chemistry gave the idea that the female hormonal cycle was being controllable.[218] The modern knowledge of the menstrual cycle began when Edgar Allen and Edward Doisy (St. Louis University, USA) showed that uterine bleeding occurs as a withdrawal effect when estrogen ceases to act on the endometrium.[219] At the same time, the chemistry of steroids became clearer with the works of Adolf Butenandt.[220] (Göttingen, Germany), John Browne[221] (McGill University, Montréal, Canada), Leopold Ruzicka[222] (Zürich, Switzerland), etc. Perhaps no contribution of chemistry in the second half of the 20th century had a greater impact on social customs than the development of oral contraceptives. Several people were important in its development – among them Margaret Sanger, Katherine Mc Cormick, advocates of birth control as the means to solving the world's overpopulation,[223] Russell Marker, Carl Djerassi (Figure 1.33) and George Rosenkrantz (Syntex, Mexico) and Gregory Pincus (Worcester Foundation for Experimental Biology, Shrewsbury), as scientists to make this idealistic project. Pincus agreed with the project when he had been asked by the feminist leaders to produce a physiological contraceptive. The key of the problem was

FIGURE 1.32 (From left to right) Charles Slocomb, Howard Polley, Edward C. Kendall, Philip S. Hench.

FIGURE 1.33 Carl Djerassi.

the use of a female sex hormone such as progesterone. This hormone prevented physiological ovulation and could be imagined as a pregnancy-preventing hormone. But the first difficulty to solve was to find a suitable, inexpensive source of the scarce compound to do the necessary research.[224] The job has been done by Russell Marker who converted sapogenin steroids extracted from dioscoreas into progesterone.

Until 1970, this source for the sapogenins remained a yam grown in Mexico. The period from late 1949 through 1951 was an extraordinarily productive one in steroid chemistry[225] with the synthesis of 19-nor-17α-ethynyltestosterone (norethindrone) and preparation of cortisone from diosgenin. Carl Djerassi synthesized an "improved" progesterone, one that could be taken orally after a minimal change in the carbone 19 of the steroid, the withdrawal of a methyl group.[226] Those derivatives were called "19–nor." In 1951, his group developed a progesterone-like compound called norethindrone. Enovid®, the first "contraceptive pill" was a combination of the progestin norethynodrel and the estrogen mestranol. It was first approved in 1957 for the treatment of a variety of disorders associated with the menstrual cycle. The era of oral contraception began in May 1960, when Enovid® was approved by the FDA for ovulation inhibition, and was immediately thereafter introduced for such use.[227] The pill offered women the ability to decide on their own, in private, whether or when to become pregnant, thus undermining the historical dominance of men in all matters relating to sex and reproduction. The consequences range from cultural to economic, professional, and educational aspects, most of them positive. The effort to discover better steroid drugs than those available at that time was remarkably successful and resulted in the introduction of several important pioneering drugs. These included norethandrolone, marketed in 1956 as Nilevar®, the first anabolic agent with a favorable separation between protein building and virilization, and spironolactone, introduced in 1959 as Aldactone®, the first steroid antialdosterone antihypertensive agent.

G. Anti-acid drugs

1. *Anti-H_2 drugs*

Rapid progress in gastroenterological research was initiated by the discovery by William Prout (Guy's and St. Thomas' hospitals London), in 1823, of the presence of inorganic, hydrochloric acid in the stomach and by Ivan P. Pavlov (Saint-Petersburg) in 1890, of neuro-reflex stimulation of secretion of this acid that was awarded Nobel Prize in 1904. Then, James W. Black (Figure 1.34), at that time pharmacologist at Smith Kline and French, who followed L. Popielski's concept of histamine involvement in the stimulation of this secretion, was awarded second Nobel Prize (1988) in gastrology within the same century for the identification of histamine 2-receptor antagonists.

FIGURE 1.34 James W. Black.

There was still controversy regarding the physiology of acid secretion in 1964 when a team at Smith Kline and French Laboratories in England started a project to prove the existence of more than one receptor for histamine and to find a substance capable of blocking the effects not blocked by the commonly used antihistamines. The team was convinced that histamine was the final mediator of acid secretion. In 1972, James Black *et al.* published evidence of the first H2-receptor antagonist, burimamide.[228] As this substance was not suitable for oral therapy, the research continued. Metiamide was synthesized with promising clinical effects but questionable safety. The final answer was cimetidine (Tagamet®), approved in 1976. Cimetidine was a breakthrough in the treatment of peptic ulcers.[229] The concept of H2-receptor interaction with other receptors such as muscarinic receptors (M_3-R), mediating the action of acetylcholine released from local cholinergic nerves, and those mediating the action of gastrin (CCK_2-R) on parietal cells, has been confirmed both *in vivo* studies and *in vitro* isolated parietal cells.

2. *Proton pump inhibitors*

Another target gave another opportunity for new treatments in the control of gastric secretion of acid. In 1968, George Sachs and his collaborators at SmithKline and

French began work that established an H^+,K^+ ATPase as the proton pump that moves acid across the gastric mucosa and gastric parietal cells.[230] For Astra Pharmaceuticals, in Sweden, the search for drugs that might improve upon the emerging H2-receptor blockers for control of acid secretion began in the mid 1970s. In 1974, the French Xavier Pascaud (Servier Laboratories, France) discovered the antisecretory activity of pyridyl-2-thioacetamide compounds,[231] but it was too toxic and other sulfur derivatives had to be examined. As a substituted imidazole was important for acid control by H2-receptors, Per Lennard Lindberg, at the Astra group (Sweden) added a benzimidazole moiety to pyridine-2-thioacetamide.[232] Chemically, the resulting derivatives were sufficiently novel so that an argument could be made in favor of their patentability. In addition, the weakly basic nature of the benzimidazoles might contribute to drug effectiveness by permitting their accumulation in acidic environments at the canalicular spaces of gastric parietal cells, precisely where acid control was needed. Substituted benzimidazoles with pK_a values around 4.0 accumulate about 1,000-fold in these low-pH spaces and thus result in great organ selectivity. Moreover, benzimidazoles behave as prodrugs, undergoing an acid-catalyzed rearrangement that provides a reactive sulfenamide species inhibiting the H^+,K^+-ATPase in the gastric fluid.[233] Specifically, early success in the suppression of acid secretion by the substituted benzimidazoles (timoprazole and picoprazole) synthesized by Astra scientists was mechanistically complemented by Sachs's work on the gastric H^+,K^+-ATPase. From discussions at a scientific meeting in Sweden in 1977, Sachs and Hässle Gastrointestinal Research scientists began the collaboration that yielded omeprazole in 1978 and then tested in humans in 1983 and 1984.[234] During the past two decades, enormous changes occurred in the management of gastric acid-related diseases. After the H2-receptor antagonists offering patients the first single-agent therapy that effectively reduced gastric acid secretion, proton pump inhibitors became widely available in the early 1990s, and they generally appeared to be superior to the previous drugs in symptom control and healing. Most physicians now use proton pump inhibitors as first-line treatment for many patients with acid-peptic disorders, including erosive or non-erosive gastro-esophageal reflux disease and duodenal and gastric ulcers. In the 1980s, *Helicobacter pylori*,[235] a spiral bacteria, has been discovered in the stomach and recognized as an important factor in the pathogenesis of gastritis and peptic ulcer by two Australian clinical researchers, R. J. Warren and B. J. Marshall (Perth Hospital, Australia) (Figure 1.35) who received the Nobel Prize in Physiology or Medicine (2005), the third, after Pavlov and Black, related to ulcers healing; three achievements appreciated by millions of ulcer patients all over the world.[236] It has been clearly demonstrated that *H. pylori* eradication could dramatically reduce chronic gastric and duodenal ulcers, and widely accepted therapeutic regimens for *H. pylori* eradication now include proton pump inhibitors and two or more antibiotics.[237]

FIGURE 1.35 Barry J. Marshall and J. Robin Warren.

H. Lipid lowering drugs

1. A *better knowledge of lipoproteins*

During the 1960s, hypercholesterolemia had been considered the hallmark of atherosclerosis, a condition understood as a bland lipid-storage disease. The acute complications of atherosclerosis were attributed to high-grade coronary stenosis. This concept was reassessed during the 1990s. Inflammation, rather than plaque size, was established as the fundamental determinant of plaque instability and thrombosis. In addition, low-grade chronic inflammation has been found to predict cardiovascular events, independent of the severity of the atherosclerotic burden. This shift from the degenerative to the inflammatory paradigm has favored a rediscovery of an 1858 description of "atheromatous affections of arteries" by Rudolf Virchow.[238] The idea of a fat transport system in the plasma of mammals evolved slowly over three centuries. The high density and the low density lipoproteins (HDL and LDL) were, respectively, isolated from horse serum in 1929 and 1950. Very low and intermediate density lipoproteins (VLDL and IDL) were then revealed. Subsequently, it was discovered that the FFA in plasma were bound to albumin and varied with feeding and fasting. The protein components of the lipoproteins (apopeptides) were characterized in the period from 1960 to 1970 and the LDL-receptor was identified in 1974. Fat transport was then established as a receptor-mediated delivery system of lipoproteins to targeted tissues. Genetic defects in this transport and receptor protein system explained dyslipidemias, which promoted atherosclerosis and other diseases.[239] An elevated level of lipoproteins, except for HDL, is the basis of all hyperlipidemias. LDL and remnant particles are potential risk factors for atherogenesis and subsequent cardiovascular disease. This is the reason why pharmacological agents capable to increase

FIGURE 1.36 Feodor Lynen.

FIGURE 1.37 Akria Endo.

the breakdown and reduce the synthesis of LDL and remnant factors were searched through the second half of the 20th century. These include nicotinic acid and its analogs, fibric acid derivatives (clofibrate, gemfibrozil, bezafibrate) then described as PPARα agonists, biliary acids and resins (cholestyramine), β-hydroxy-β-methylglutaryl-CoA (HMG-CoA) reductase inhibitors (lovastatin, simvastatin, pravastatin) and probucol. Lipid lowering drugs of different classes have a synergistic effect on lipid metabolism and combination therapy is often used. They are prescribed as long-term preventive therapy in apparently asymptomatic people. Several studies indicate that secondary prevention with lipid lowering drugs is cost-effective, particularly in patients with symptomatic coronary artery disease.[240] Among those drugs, statins appear to be the most spectacularly active.

2. *Statins*

Various experiments on animals and humans had shown that cholesterol could either be absorbed from the diet, or if the diet was lacking sufficient cholesterol to meet the body's needs, then it could be synthesized. Cholesterol production within the body is controlled by a feedback mechanism in which cholesterol inhibited the enzyme HMG CoA reductase, an enzyme discovered in 1959 by Feodor Lynen *et al.* (Figure 1.36) at the Max Planck Institute (Munich).[241]

By inhibiting this enzyme, the conversion of HMG-CoA to mevalonic acid is stopped.[242] The most active drugs, statins, were discovered in Tokyo (Japan), in 1971. Knowing that many microorganisms require cholesterol for growth, Akira Endo (Figure 1.37) and Masao Kuroda were hoping to identify novel factors that would inhibit HMG-CoA-reductase, shown by basic researchers to be critically important for cholesterol synthesis. They looked to microorganisms, especially fungi, as a source for these factors, hoping to find a microorganism that produced an HMG-CoA reductase inhibitor as a defense mechanism against attack by other microbes which relied on sterols as part of their biochemical make up.[243]

The search for a suitable compound took 2 years and involved more than 6,000 microbes. The second mould shown to inhibit lipid synthesis was *Penicillium citrinum*. The active compound from *P. citrinum* was ML-236B (Mevastatin) capable of inhibiting lipid synthesis from either ^{14}C-HMG CoA, or ^{14}C-acetate. However, there was no inhibitory effect on lipid production from ^{3}H-labeled mevalonate. From this, it was possible to deduce that mevastatin did, in fact, inhibit the enzyme HMG CoA reductase. Two moulds were found to meet the requirements. Firstly *Pythium ultimum* was found to produce a substance called citrinin that was shown to irreversibly inhibit HMG CoA reductase.

The critical step was the discovery by Michael Brown and Joseph Goldstein (University of Dallas, USA) (Figure 1.38), of how the use of a statin could dramatically reduce the level of LDL or "bad" cholesterol in the blood, by causing liver cells to increase the amount of LDL they would snatch up and use for themselves. They were awarded the Nobel Prize in 1985.[244] By 1976 Carl Hoffman (Merck & Co, USA) successfully repeated the experiments of Endo and Kuroda an isolated lovastatin from a strain of the

FIGURE 1.38 Michael Brown and Joseph Goldstein.

FIGURE 1.39 Paul Ehrlich.

fungus *Aspergillus terreus*. The new compound was slightly more effective than mevastatin. In 1979, while developing and researching lovastatin, Merck scientists synthetically derived a more potent HMG-CoA reductase inhibitor from a fermentation product of *Aspergillus terreus*, which was designated MK-733 (later to be named simvastatin).[245] Development of other drugs based on mevastatin and lovastatin has continued all over the world. Three main approaches have been utilized. Firstly, synthetic compounds, such as fluvastatin, were produced. Research in this area concentrated on replacing the decalin ring of the fungal compounds with an aromatic ring. Secondly, chemical alteration of fungal products created drugs such as Simvastatin. In this drug, modifications were made to the acyl group. Finally, microbial alteration of fungal compounds has lead to drugs such as pravastatin. By altering the basic chemical composition of the mevastatin molecule, potency of the drug can be increased. Simvastatin is approximately twice as potent as pravastatin and lovastatin, whilst mevastatin is the least powerful. However, in changing the shape of the active molecule, the chances and severity of side effects was also altered. For example, there is an increased risk of muscle toxicity with lovastatin in comparison to pravastatin. Atorvastatin was designed based in part on molecular modeling comparisons of the structures of the fungal metabolites and other synthetically derived inhibitors. In addition to development of the structure–activity relationship which led to atorvastatin, another critical aspect of the development of this area was the parallel improvement in the chemistry required to prepare compounds of the increased synthetic complexity needed to potently inhibit the target enzyme. Ultimately, the development of several chiral syntheses of enantiomerically pure atorvastatin calcium was accomplished.[246]

At approximately this same time, population-based studies such as the well-known Framingham Heart Study, demonstrated that high cholesterol was a major risk factor for the development of coronary artery disease. As a result of these studies, statins, first approved in 1987, are now some of the most widely prescribed drug. Several studies have identified a link between heart attacks and systemic markers of inflammation, and suggest that statins might help by decreasing the degree of inflammation. Recent studies not only suggest that statins can help heart disease patients by decreasing coronary artery inflammation, but also offer hope that statins might also be useful in organ transplantation, as well as in the treatment of autoimmune diseases such as multiple sclerosis, RA, and psoriasis.

I. From neurotransmitters to receptors

As for giving a symbolic landmark to drugs history at the beginning of the century, Paul Ehrlich (Institut für experimentelle Therapie, Frankfurt) (Figure 1.39) introduced, in 1900, the term "receptor." The receptor concept as such, was in fact developed in the context of immunology. The drug receptor theory, in turn, would be later developed in Ehrlich's chemotherapy.

Previously, in animal experiments in the 1870s, John Newport Langley, Professor of physiology in Cambridge (UK) had shown that jaborandi extract, containing pilocarpine, modified the heart rate, the effect being reversed by atropine. Similarly, pilocarpine stimulated the secretion of saliva and atropine inhibited it. In both heart and salivary gland experiments, the effect depended on the amount of

each drug present. Langley concluded that pilocarpine and atropine formed "chemical compounds" with tissue components, the end result depending on "their relative chemical affinity to the tissue and the mass of each present." As a conclusion of these observations, Langley described in 1905 the concept of "receptive substances" as mediators of drug action[247] to the site of action of drugs. He was the one who, first, proposed a receptor theory of drug action. Later, Alfred Joseph Clark (University College, London), in his book, *The Mode of Action of Drugs on Cells* (1933), had a considerable impact on the discipline of pharmacology by showing that for many drugs the relationship between drug concentration and biological effect corresponded to a hyperbolic curve expressing the equilibrium between a drug interacting with a specific number of receptors on the cell, and that the pharmacological action produced by the drug was "directly proportional to the number of receptors occupied".[248] However, it was not until after the World War II that the work of E. J. Ariëns, in Utrecht[249] and R. P. Stephenson, in Edinburgh,[250] modified Clark's occupancy theory to explain the affinity (i.e. the attraction between a compound and a receptor) and the efficacy, introducing the concept of partial agonist, and making an important distinction between the affinity of a drug for a receptor, and its potency, a concept that later became important in beta-blockade. At the same time, the conceptualization of the reactions between drugs and tissues in terms of receptors was gradually becoming more acceptable, and more deeply ingrained in laboratory practice.[251] Thus, receptor theory emerged progressively from immunology, metabolism, pharmacokinetics, but mainly with the physiology of the autonomic (sympathetic and parasympathetic) nervous system in connection with mediators. Some drugs, like adrenaline, which produce an effect similar to electrical stimulation of the sympathetic nerves, bound to receptive substances in cells and that there were two types of such substances, "motor" (excitatory) and "inhibitory."

Henry Dale (University College, London) (Figure 1.40) showed that while the excitatory actions (i.e. vasoconstriction and contraction of smooth muscle) of adrenaline and other structurally similar compounds in most tissues were blocked by ergot alkaloids, their inhibitory effects (i.e. vasodilation and relaxation of bronchial muscle in the lungs) were not.[252] However, Dale, who became a hugely influential figure in pharmacology, like other researchers remained skeptical about Langley's idea that drugs combined with specific receptive substances or "side-chains," and subsequently failed to give his full support to the concept of receptor. The concept of β-blockade eventually made the most circumspect physiologists move their mind. This happened to Ahlquist who published his landmark paper in 1948[253] establishing that there were two different β-receptors. Since that time, neuropsychopharmacology is organized primarily according to the neurotransmitters involved for synaptic transmission.

FIGURE 1.40 Henry Dale.

1. *Dopamine receptors*

L-Dopa (l-3,4-dihydroxyphenylalanine) was first isolated from seedlings of *Vicia faba* (broad bean) by Marcus Guggenheim in 1913 (Hoffmann LaRoche, Basel, Switzerland) who suspected that L-dopa was a precursor to adrenaline. In 1938, Peter Holz (Pharmakologischen Institut, Greifswald, Deutschland) found that L-dopa decarboxylizes into dopamine in mammalian tissue. During the 1950s, L-dopa was found in many tissues, especially in the brain. By 1959, it was found that dopamine was present in most parts of the CNS but in particular in the nigrostriatal pathway as shown by the discovery that degeneration of this pathway occurs in the brains of patients afflicted with Parkinson's disease. The depletion of dopamine resulting from the degeneration of the specific neurons led George C. Cotzias (Brookhaven, New York) (Figure 1.41) to develop dopamine-replacement therapies (L-dopa associated with or without dopa decarboxylase inhibitors) for alleviating Parkinson's disease.

In 1967 after 6 years of studies, trials of L-dopa in patients with Parkinson's disease showed dramatic improvements in all motor deficits. The hypothesis that dopamine is involved in the pathogenesis of psychosis, in particular schizophrenia, rests on the finding that most antipsychotic drugs are dopamine-receptor antagonists and that agents which cause excessive release of dopamine mimic schizophrenia-like states. In 1979, John Kebabian and Donald Calne (NIH, Bethesda, USA) found that dopamine exerts its effects by binding to two subtypes of receptors,

FIGURE 1.41 George C. Cotzias.

known as the D_1 and D_2 receptors.[254] These receptors could be differentiated pharmacologically, biologically, physiologically, and by their anatomical distribution. D_1 receptor is to bind the benzazepine antagonist SCH 23390, while that of the D_2 receptor is to recognize with high affinity various butyrophenones as spiperone and haloperidol. For 10 years, this two-subtype classification has accounted for most of the activities attributed to the dopaminergic system. The existence of other dopamine receptors has been proposed but had been refuted as well. Since that time, most neuroleptics were developed as D_2 receptor antagonists and thus were expected to bind to this receptor with higher affinity than to the D_3 and D_4 receptors. Clozapine, an "atypical" neuroleptic, shows a higher selectivity for the D_4 receptor than for any other D_2-like receptors. The D_4 receptor binds clozapine with a 10-fold higher affinity than does the D_2 receptor.[255] Therefore, the D_4 receptor may be the specific target of clozapine. The discovery of the "unexpected" dopamine receptors has and will continue to impact the understanding of the dopaminergic system.

2. *Serotonin receptors*

Serotonin was isolated and named in 1948 by Maurice M. Rapport, Arda Green, and Irvine Page (The Cleveland Clinic, USA).[256] Within a decade there were indications for its existence in the CNS of animals and for a neurotransmitter function. Serotonin, or 5-hydroxytryptamine (5-HT), has been implicated in almost every conceivable physiologic or behavioral function: aggression, appetite, cognition, emesis, migraine, neurotrophism, sex, sleep, vascular, endocrine, gastrointestinal, motor and sensory functions. Moreover, most drugs that are currently used for the treatment of psychiatric disorders (e.g., depression, mania, schizophrenia, autism, obsessive compulsive, or anxiety disorders) are thought to act, at least partially, through serotoninergic mechanisms. It is possible for 5-HT to be involved in so many different processes because of the anatomy of the serotoninergic system, in which serotoninergic neurons influence all regions of the neuraxis.[257] Another answer lies in the molecular diversity and differential cellular distribution of the many 5-HT receptor subtypes that are expressed in brain and other tissues. By the late 1950s, evidence for 5-HT receptor heterogeneity was found in the periphery, and in 1979 two distinct populations of 5-HT binding sites were identified in rat brain: 5-HT_1 and 5-HT_2. A class of drugs which specifically antagonizes the 5-HT type 3 receptor (5-HT_3) now occupies a major place in the supportive care of cancer patients since these drugs allow the use of high-dose cytotoxic treatment by blocking the nausea and vomiting triggered by cancer chemotherapeutic agents and/or radiotherapy.[258] After ondansetron, granisteron and tropisetron were also useful as prophylactic agents in preventing postoperative nausea and vomiting. The 5-HT_3 receptor antagonists (with or without other antiemetic drugs as neuroleptics or corticosteroids) have become the agents of choice in controlling emesis because of higher efficacy and relatively lower adverse effect profile as compared to the conventional antiemetic agents. The major site of action of these drugs appears to be the central 5-HT_3 receptors, although inhibition of peripheral receptors may also play a role in the control of vomiting. The new era in antimigraine drugs began in 1973 with efforts to synthesize a selective serotonin agonist, following up on numerous observations implicating serotonin in the generation of a migraine attack. In 1991, Glaxo Wellcome introduced the first of the new serotonin agonists, sumatriptan effective at two serotonin (5-HT) receptors, 5-HT_{1B} and $5\text{-HT}_{1D,}$ with weaker effects at other 5-HT_1 receptors.

3. *Acetylcholine receptors*

By 1914, Henry Dale had isolated a compound from ergot that produced effects on organs similar to those produced by nerves. He called the compound acetylcholine. When Dale heard of Loewi's discovery of "*vagustoffe*" 7 years later (Figure 1.42), he suggested that it was identical to the acetylcholine he had discovered earlier (for their discoveries, both shared the 1936 Nobel Prize for physiology or medicine). They made acetylcholine the first known neurotransmitter. It can be found in the brain, neuromuscular junctions, spinal cord, and ganglia of the autonomic nervous system. It is synthesized from acetyl-CoA and choline. Acetylcholine receptor sites can be ionotropic (nicotinic receptor) or metabotropic (muscarinic receptor), making possible various responses to a stimulus by acetylcholine. When an ionotropic receptor is activated, it opens a channel that allows ions such as Na^+, K^+, or Cl^- to flow. In contrast, when a metabotropic receptor is activated, a series

FIGURE 1.42 Otto Loewi and his students.

of intracellular events are triggered that also results in ion channel opening but must involve a range of second messenger chemicals. A numerous cholinergic projections together make up a wide source of acetylcholine in the brain. Nicotinic acetylcholine receptor and five different subtypes of the muscarinic receptor have been cloned to date, and a majority of those are known to be expressed in the brain. It is known that acetylcholine contributes to cognitive processes and, dramatically, in case of Alzheimer's disease (AD), a disease first discovered in 1906 by Alois Alzheimer. It is a progressive, degenerative, and irreversible neurological with no cure. One of the characteristic changes that occur in this disease is the loss of memory and the loss of acetylcholine from both cholinergic and non-cholinergic neurons of the brain. However, acetylcholinesterase activity is increased around amyloid plaques.[259] This increase in acetylcholinesterase has been of significance for therapeutic strategies using acetylcholinesterase inhibitors. Amyloid β-protein, the major component of amyloid plaques, acts on the expression of acetylcholinesterase inhibited by tetrahydroaminoacridine (tacrine, Cognex®), the first anti-AD drug. Its development began with its synthesis as an antiseptic in 1940 by Adrian Albert in Australia. In the 1970s, William Summers (Figure 1.43) began using tacrine in treating drug overdose coma and delirium. He felt it might have application in Alzheimer's based on work done in England by Peter Davies. In 1981, Summers *et al.*, giving intravenous tacrine to Alzheimer's patients, showed measurable improvement. Between 1981 and 1986, Summers worked with Art Kling and his group at UCLA to demonstrate usefulness of oral tacrine in treatment of Alzheimer's patients.[260]

The average length of tacrine use in 14 completing patients was 12.6 months and improvement was quite robust, but this sparked controversy in the field. In 1993, after larger studies replicated the positive effect of tacrine, it was approved by the FDA for treatment of AD.[261] The first step toward the best possible long-term management is early diagnosis of AD, thereby facilitating early initiation

FIGURE 1.43 William Summers.

of cholinesterase inhibitor treatment, which may stabilize/reduce the rate of symptomatic cognitive and functional decline. Cholinesterase inhibitor therapy with rivastigmine,[262] donepezil,[263] or galantamine[264] is endorsed as standard first-line therapy in patients with mild-to-moderate AD. The *N*-methyl-D-aspartate receptor antagonist, memantine, may be used as monotherapy or in combination with a cholinesterase inhibitor for patients with moderate AD, and as monotherapy for patients with severe AD.[265] Despite the slight variations in the mode of action of the three cholinesterase inhibitors there is no evidence of any differences between them with respect to efficacy. It may be that galantamine and rivastigmine match donepezil in tolerability if a careful and gradual titration routine over more than 3 months is used. Titration with donepezil is more straightforward and the lower dose may be worth consideration.[266] Unfortunately, AD eradication seems to be a long way to go for pharmacologists as for psychiatrists or neurologists. Future anti-AD therapies will likely be multi-modal and individually tailored depending on the patient's immune status, genetic background and their amyloid burden, as determined by imaging studies using specific labeling ligands.[267]

4. *Cannabinoid receptors*

The main active principle of Cannabis, Δ^9-tetrahydrocannabinol, has been isolated and characterized by Raphaël Mechoulam (Jerusalem, Israel) (Figure 1.44) in 1964.[268] The term "endocannabinoid" coined after the discovery of membrane receptors for this psychoactive principle, indicates

FIGURE 1.44 Raphaël Mechoulam.

a whole signaling system that comprises cannabinoid receptors, endogenous ligands, and enzymes for ligand biosynthesis and inactivation.

Analogous to the discovery of endogenous opiates, isolation of cannabinoid receptors provided the appropriate tool to isolate an endogenous cannabimimetic eicosanoid, anandamide.[269] Cannabinoids were shown to bind to selective cannabinoid receptor subtypes, CB_1, CB_2, and CB_{1A} all of which belong to the superfamily of G-protein-coupled plasma membrane receptors. Recent studies indicate that anandamide is a member of a family of fatty acid ethanolamides that may represent a novel class of lipid neurotransmitters.[270] Many of the enzymes involved in endocannabinoid synthesis and degradation have now been characterized and are currently being pursued as therapeutic targets. Inhibitors of endocannabinoid re-uptake include *N*-(4-hydroxyphenyl)-arachidonylamide (AM404), which extensively blocks anandamide transport. AM404 is supposed as a possible active metabolite of acetaminophen (paracetamol), which, following deacetylation to its primary amine, is conjugated with arachidonic acid in the brain and the spinal cord to form the potent transient receptor potential vanilloid (TRPV1) agonist AM404. AM404 also inhibits purified COX-1 and COX-2 and prostaglandin synthesis. These findings identify fatty acid conjugation as a novel pathway for drug metabolism and provide a molecular mechanism for the occurrence of the analgesic *N*-acylphenolamine AM404 in the nervous system following treatment with acetaminophen.[271] AM404 activity is prevented by the CB_1 cannabinoid antagonist SR 141716A, better known as rimonabant. A more thorough characterization of the roles of endocannabinoids in health and disease will be necessary to define the significance of endocannabinoid inactivation mechanisms as targets for therapeutic drugs.[272] Therapeutic strategies include small-molecule cannabinoid receptor agonists and antagonists, and the use of non-psychotropic plant cannabinoids. A CB_1 receptor antagonist looks promising against obesity, metabolic syndrome, and nicotine dependence. Rimonabant is the first selective blocker of the CB_1 being developed for the treatment of multiple cardiometabolic risk factors, including abdominal obesity and tobacco addiction. Gérard Le Fur *et al.* (Sanofi-Aventis, France) discovered rimonabant as an agent with a novel mechanism of action and a potential to be a useful adjunct to lifestyle and behavior modification in treatment of multiple cardiometabolic risk factors,[273] dyslipidemia,[274] including abdominal obesity[275] and smoking. Clinical applications of rimonabant are still to be specified, for instance for the treatment of obesity-associated liver diseases and related features of metabolic syndrome.[276] Clinical trials carried out with oral THC and plant cannabinoids for the treatment of multiple sclerosis and Parkinson's disease have shown some efficacy and few side effects.[277] Curiously, just when the first CB_1 agonists are being studied and probably soon introduced for pain treatment, it comes out that an indirect cannabino-mimetic (paracetamol) had been extensively used for more than a century.

J. Drugs of the mind

1. Psychotropic drugs

The first modern reflection upon pharmacological treatments of psychiatric disorders has been performed by Louis Lewin (Berlin, Germany), who succeeded, in 1924, in presenting a classification of drugs and plants founded on their psychoactive properties. "*Inebriantia*" included alcohol and ether, "*excitantia*" included amphetamine, "*euphorica*" included morphine or heroin, "*hypnotica*" included kava and "*phantastica*" included peyotl or ayahuasca. The field of psychiatry is so complex that till the middle of the 20th century, it was clear that the only behaviorist approach could represent the final way to explore and treat mental disorders. At that time, drugs used in mental disorders were almost exclusively extracted from plants, except barbiturates, bromine salts or amphetamine. These drugs could provoke depression or psychosis but were unable to demonstrate any activity to treat psychiatric disorders. This may explain the reluctance of psychiatrists toward "biological" explanations of schizophrenia, depression or other mental illnesses including an imbalance in the chemical constituents of the brain. Early treatments for depression involved dosing patients with barbiturates, keeping them unconscious for several days, in the hope that sleep would restore them to a healthier frame of mind.[278] Convulsive therapy was introduced in 1934 by Hungarian neuropsychiatrist Ladislas Meduna (Budapest) who, believing mistakenly that schizophrenia and epilepsy were antagonistic disorders, induced seizures first with the injection of camphor oil and then with pentetrazol (Cardiazol®). In Italy, Ugo Cerletti (La Sapienza University,

Rome), who had been using electric shocks to produce seizures in animal experiments, and his colleague Lucio Bini, developed the idea of using electricity as a substitute for metrazol in convulsive therapy and, in 1937, experimented for the first time on a person. Psychiatrists found they could lessen the effects of depression. Electro-convulsive therapy (ECT) is still used as a treatment for severe depression. The understanding of depression depended on the understanding of the brain itself. This took a leap forward in 1928, when Otto Loewi discovered acetylcholine.[279] It was another 24 years before scientists would discover the presence of other neurotransmitters in the brain, such as serotonin, noradrenalin, and dopamine. By the 1980s, 40 different neurotransmitters had been isolated in the brain.[280]

2. *Lithium*

Modern psychiatric treatments were introduced in 1949, when lithium carbonate was discovered as treatment for mania by Australian psychiatrist John F. Cade (Figure 1.45). After Cade's initial report, lithium therapy was principally developed in 1954 by Mogens Schou (Aarhus University, Denmark).[281] In 1969, 20 years after its discovery by John Cade and after a decade of trials, the Psychiatric Association and the Lithium Task Force recommended lithium to the FDA for therapy of mania. A breakthrough had been achieved in the treatment of manic depression, and the genetically related forms of recurrent depression. Bipolar disorders, which afflict about 1% of adults, are now treated with drugs called mood stabilizers, especially lithium and valproic acid, both discovered decades earlier, but nothing better has yet emerged.[282]

Lithium is thought to act by blocking the function of an enzyme called glycogen synthase kinase-3 (GSK-3β) in the brain. Among other substances found to block GSK-3β elsewhere in the body, lithium salts are the only capable to get into the brain and then to combat bipolar disorder. Alan Kozikowski (University of Illinois in Chicago) took a directed, rational approach showing that newly discovered compounds called (3-(benzofuran-3-yl)-4-indol-3-yl) maleimides) are potent and relatively selective GSK-3β inhibitors.[283]

3. *Neuroleptics*

In 1937, Daniel Bovet (Nobel Prize for Physiology or Medicine, in 1957) and Anne-Marie Staub discovered the first antihistamine (anti-H_1),[284] and then Bernard Halpern discovered Antergan in Rhône-Poulenc (Lyon, 1942) and Promethazine (1946), a phenothiazine antihistamine. It was the compost of the adventure of psychotropic drug discovery. Henri Laborit, (Figure 1.46) a naval surgeon in Paris, had begun experimenting with antihistamines in 1949 to make patients to recover quicker from the anesthesia. He tried first promethazine, usually used to fight allergies and he noted how some patients became obviously calmer. In 1950, the famous 4560 RP, later called chlorpromazine, was synthesized by Paul Charpentier, a chemist from Laboratoires Rhône-Poulenc (Vitry, France). Laborit tried to administer 4560 RP, hoping it would enhance autonomic blockade. He noticed when he gave a strong dose to his patients, a change in their mental state : they did not seem anxious, in fact, they were rather indifferent. Laborit was able to operate using much less anesthetic. At that time, no

FIGURE 1.45 John Cade.

FIGURE 1.46 Henri Laborit.

one in psychiatry was working with drugs. Shock or various psychotherapies were used. Laborit kept pressing his point, however, and he persuaded several apparently skeptical psychiatric colleagues to try it out. Interest of Jean Delay and Pierre Deniker, at Sainte-Anne Hospital in Paris, was piqued and they tried chlorpromazine on their most agitated, uncontrollable patients. On January 19, 1952, a 24-year old man who had mania was successfully injected with chlorpromazine. They first published the results obtained by treating, in May and June 1952, 38 psychotic patients with injections of 75–150 mg a day of chlorpromazine reporting.

It was stunning: patients who had stood in one spot without moving for weeks, patients who had to be restrained because of violent behavior, could now make contact with others and be left without supervision. Another psychiatrist reported, "For the first time we could see that they were sick individuals to whom we could now talk".[285] Severe mental illnesses had been increasing since the beginning of the century. In 1904, 0.2% of people were hospitalized in mental hospitals; by 1955, this feature had doubled. Psychiatrists argued whether it was a result of biology or of experience, but there was nothing to help the chronic mentally ill, usually warehoused in mental institutions. Meanwhile, Smith Kline purchased the rights to chlorpromazine from the leading French company Rhône-Poulenc in 1952, putting the drug on the US market as an antivomiting treatment. Pierre Deniker succeeded in convincing resistant US psychiatry practitioners to try the drug. In 1954, chlorpromazine was approved throughout the world. It had a calming effect without sedating patients, allowing them to live a nearly normal life. By 1964, some 50 million people around the world had taken the drug, and Smith Kline revenues doubled three times in 15 years. As a fact of a modern society the demand for sedatives was profound, and the drug marketplace responded rapidly. Chemists built a lot of derivatives, in order to find another magic bullet: more activity and better tolerated. By replacing the chlorine group of chlorpromazine with a trifluoromethyl group, one obtains trifluoperazine and by adding a terminal ethyl alcohol group to trifluoperazine one obtains fluphenazine. Instead of the phenothiazine heterocyclic ring, it is possible to substitute a thioxanthine heterocyclic ring. With the rest of the "tail amines" of the substituted phenothiazines, a whole new series of substituted thioxanthenes such as thiothixene could be synthesized. Another possibility was the use of other heterocyclics to obtain clozapine or one of the newer "atypical" antipsychotics chemically and pharmacologically similar to clozapine such as the thienobenzodiazepine-derivative, olanzapine (Zyprexa®).[286] Another early major development in antipsychotic drugs was haloperidol. The "Haldol story" began in a small Belgian company that Paul Janssen inherited from his parents. About 1953, Janssen decided that the company could only survive if it had exclusive patented drugs. They decided early to first work with anticholinergic derivatives. His success in all aspects of pharmaceutical development is truly remarkable. In 1957, one of Janssen's lead chemicals, a butyrophenone derivative of normeperidine, had a mixture of narcotic and neuroleptic effects in animals. Subsequent molecular modifications led to the development of the potent antipsychotic, haloperidol.[287]

4. *Anti-anxiety drugs*

The development of drugs for the treatment of anxiety has gradually evolved from less selective agents, such as alcohol, opiates, and the bromides, to progressively more specific drugs, leading ultimately to the development of the benzodiazepine anxiolytics. But first anti-anxiety drug, meprobamate, discovered in the early 1950s by Frank Berger (Wallace laboratories, Cranbury New Jersey) was the first of the major "tranks." The pharmacological properties of this carbamate derivative were described as producing reversible flaccid paralysis of skeletal muscles without significantly affecting the heart, respiration, and other autonomic functions with muscular relaxation and sedation.[288] Even if Miltown® had been called the "Wonder Drug of 1954," sedatives were not widely used until 1961, when Librium® (chlordiazepoxide) was discovered and marketed. In 1954, Leo Sternbach (Hoffmann LaRoche, Nutley, USA) (Figure 1.47) began to study a class of unexplored compounds, the benzheptoxdiazines, It was subsequently proved that the series prepared by Sternbach was not the expected benzheptoxdiazines, it was found that the actual

FIGURE 1.47 Leo Sternbach.

substance he obtained was quinazoline 3-oxides. None of these compounds gave any interesting results. The program was abandoned in 1955 in order for Sternbach to work on a different project. In 1957, during a general laboratory cleanup a vial shelved was submitted, as a last effort, for pharmacological testing. Unlike all of the other compounds, this one gave very promising results in six different tests used for preliminary screening of tranquilizers.

Further investigation revealed that this compound was not a quinazoline 3-oxide, but was instead the benzodiazepine 4-oxide. Roche's head of pharmacology, Lowell Randall, stated that "the substance has hypnotic, sedative, and antistrychnine effects in mice similar to meprobamate." In 1958, the compound Ro 5-0690 became chlordiazepoxide. Librium® proved a phenomenal success.[289] Then Valium® (diazepam), discovered in 1960, was marketed by Roche 1963 and rapidly became between 1969 and 1982 the most prescribed drug in America. Sternbach has to be credited not only with the invention of chlordiazepoxide (Librium®), diazepam (Valium®), but also flunitrazepam (Rohypnol®), nitrazepam (Mogadon®), and clonazepam (Rivotril®). Benzodiazepines affect the γ-aminobutyric-acid (GABA) receptors at the level of the subcortical nuclei. This implies that they have a tranquilizing function with only minor influence on the cognitive functions and level of consciousness. Physical and psychological addiction followed for many patients. Benzodiazepines, 50 years after their discovery, remain precious for various clinical situations: anxiety disorders, insomnia, panic disorders, restless legs syndrome, and addiction. Benzodiazepines provide a relatively safe means of providing sedation in a variety of clinical situations. Midazolam, which is shorter acting than other benzodiazepines, is the drug of choice for sedation in ambulatory patients. Last, flumazenil is a highly effective specific competitive benzodiazepines antagonist which provides a safe means of rapidly attenuating or terminating benzodiazepines sedation.[290]

5. *Antidepressants*

Iproniazid, the first modern antidepressant, was originally developed as an antituberculosis drug in the early 1950s. In addition to its ability to treat tuberculosis, iproniazid was observed to elevate mood and stimulate activity in many patients. After having pioneered the introduction and use of the *Rauwolfia* to treat psychiatric disorders, Nathan Kline (Rockland State Hospital, Orangeburg, New York) investigated the ability of iproniazid to treat the symptoms of depression. After promising preliminary findings reported in 1957, iproniazid was prescribed widely to patients with major depression. Thus, monoamine oxidase inhibitors (MAOIs) were developed at the end of the 1950s. By blocking the action of oxidases which break down neurotransmitters in the brain, it is possible to "bath" brain in large quantities of neurotransmitters, and to fight off the depression.

Kline received the Albert Lasker Award in 1964. The first tricyclic antidepressant, imipramine, was originally developed in a search for drugs useful in the treatment of schizophrenia. The therapeutic and commercial success of substituted phenothiazines such as promethazine, promazine, and chlorpromazine, initiated an enormous effort in the molecular modification of the polycyclic phenothiazine ring structure and its *N*-aminoalkyl side chain. After a while, a substance was unearthed, iminodibenzyl. It was not a new drug. Iminodibenzyl had been discovered in 1898 and used briefly as an intermediate, in the preparation of Sky Blue, a dye stuff. Iminodibenzyl, however, had a tricyclic ring structure, similar in appearance to the phenothiazines. Robert Domenjoz, head of the pharmacological laboratories at Geigy (Basel, Switzerland), asked two organic chemists, Walter Schindler and Ernst Häfliger, to prepare derivatives of iminodibenzyl where the sulfur bridge of the phenothiazine ring of promethazine is replaced with an ethylene bridge. They produced 42 separate basic alkylated derivatives, each being distinguished only by slight differences in their side-chains. Among the new molecules, they synthesized *N*-(γ-dimethylaminopropyl)-iminodibenzyl, which came to be called imipramine, a weak antihistaminic and mild anticholinergic with sedative properties in normal human volunteers. Although clinical trials demonstrated a lack of effect in treating schizophrenia, Roland Kuhn (Figure 1.48) at the Psychiatric Clinic in Munsterlingen, Switzerland, decided to examine its effectiveness in depressed patients. He discovered that of some 500 patients with various psychiatric disorders that were treated, only those with endogenous depression with mental

FIGURE 1.48 Roland Kuhn.

and motor retardation showed a remarkable improvement after about 1–6 weeks of daily imipramine therapy. "They again become interested in things, are able to enjoy themselves, despondency gives way to a desire to undertake something, despair gives place to renewed hope in the future," Kuhn wrote. These effects led to the idea that imipramine was selectively reversing the depression, rather than simply producing a general activating effect.[291] Subsequent biochemical studies on imipramine demonstrated that this drug increased the activity of the monoamine neurotransmitters, norepinephrine, and serotonin, by preventing the re-uptake.[292] It did not take long for the diamine structure of an additional secondary amine group in imipramine to be substituted with an ethylene group in amitriptyline (Laroxyl®). It was another tricyclic antidepressant subsequently widely used. Understanding of the activities of these drugs, in combination with other observations, provided the foundation for the monoamine hypothesis of depression, which proposes that depression results from a central deficiency of monoamine function. Most of the early antidepressants worked by affecting several different neurotransmitter chemicals at the same time, but scientists began to work on drugs that would target one specific neurotransmitter, while leaving others unaffected (Table 1.9).

In the early 1970s, evidence of the role of serotonin (5-hydroxytryptamine or 5-HT) in depression began to emerge and the hypothesis that enhancing 5-HT neurotransmission would be a viable mechanism to mediate antidepressant response was put forward. In 1968, Arvid Carlsson (Göteborg, Sweden) had already found that, when an electrical impulse passed from one neuron to another, serotonin was released into the space between the neurons – the synapse – to help the "message" to be transmitted and its re-uptake, inhibited, which, in clinical terms resulted in helping the patient to recover from depression.[293] The first selective serotonin re-uptake inhibitor was zimelidine developed by Carlsson, but rapidly withdrawn due to its adverse effects. At the same time, in the 1970s, at Eli Lilly (Indianapolis, USA), David Wong, Bryan Molloy and Robert Rathburn were also looking for an antidepressant that could emerge from a molecular design close to the 3-phenoxy-3-phenylpropylamine, but the result was only a compound that was active on norepinephrine re-uptake: nisoxetine. In further research, Wong re-tested other molecules and tests carried by Jong-Sir Horng showed a compound later named fluoxetine to be the most potent and selective inhibitor of serotonin re-uptake of the series. In fact, fluoxetine was the third specific serotonin re-uptake inhibitor (SSRI) on the market. The first had been fluvoxamine (1983). Fluoxetine, named Prozac® made its appearance in Europe in 1986 just before the United States in December 1987,[294] the term "SSRI" being specially coined for it. Fluoxetine provided rapid relief from the symptoms of depression, without any of the unpleasant side effects associated with the "older" tricyclic antidepressants (dry mouth, constipation, blurred vision, sweating, and weight gain) or the dietary restrictions that were necessary with MAOI drugs. By 1994, it was the number two best-selling drug in the world.[295]

6. *General and local anesthetics*

One of the most important therapeutic revolutions during the 19th century was the introduction of general anesthesia in the practice of surgery. In 1776, Joseph Priestley discovered the laughing gas (nitrous oxide), but analgesia seemed to be unreachable. Priestley and Humphrey

TABLE 1.9 Antidepressant Medication Classes

Tricyclic antidepressants (nonselective NE and/or 5-HT inhibitors reuptake inhibitor)	Imipramine Amitriptyline Nortriptyline Clomipramine Desipramine Amoxapine Doxepine Protriptyline Trimipramine	Selective serotonin-reuptake Aminoketone (NE and DA uptake inhibitor) Phenethylamine (5-HT, NE, and DA uptake inhibitor) Other (5-HT re-uptake inhibitor and antagonist)	Fluoxetine Fluvoxamine Paroxetine Sertraline Citalopram Bupropion Venlafaxine Nefazodone Trazodone
Tetracyclic antidepressants	Maprotiline (presynaptic NE re-uptake inhibitor)		Mirtazapine (plus a2-antagonist)
Monoamine oxidase inhibitors	Phenelzine tranylcypromine	Selective norepinephrine re-uptake inhibitor	Reboxetine

DA = dopamine; 5-HT = serotonin; NE = norepinephrine.

FIGURE 1.49 "The Ether Dome," Massachussets General Hospital, Boston.

Davy commented in 1796: "it may probably be used with advantage during surgical operations in which no great effusion of blood takes place." Michael Faraday proposed the use of diethyl ether to induce similar action. However, "hilarant gas" inhalation was proposed during exhibitions for shows named "ether frolics." Neither diethyl ether (sulfuric ether) nor nitrous oxide were clinically used before 1846. Surgery was so difficult before that and it was very uncommon till the midst of the century: pain and infection risk resulting from surgery were discouraging.

Dentists set the pace in the field of analgesia. They become familiar with both diethyl ether and nitrous oxide. They are in permanent contact with pain in complaining patients. They also produced pain through unfair or badly controlled operations. Horace Wells, a dentist, asked a colleague to extract his own teeth while under the influence of nitrous oxide.[296] This trial, held in 1844, was successful and painless. Shortly thereafter, in 1845, he attempted to demonstrate his discovery at the Massachusetts General Hospital in Boston (Figure 1.49). His first try has been a total failure. Another Bostonian dentist, William T. G. Morton (Figure 1.50), familiar with the use of nitrous oxide from his friendship with Wells, asked the surgeons of the Massachusetts General Hospital to demonstrate his technique after many tries on animal as on himself or friends. The first patient, Gilbert Abbott was to be operated by the chief surgeon Dr. John Collins Warren. Morton came with a special apparatus with which to administer the ether and only a few minutes of ether inhalation were necessary to make the patient unconscious.[297]

The eminent surgeon Henry J. Bigelow noted: "I have seen something today that will go around the world," a new era is to begin in the history of medicine.[298] Techniques and safety of anesthesia will not stop to improve. Even if ether was a very interesting agent, other drugs were rapidly tested among which was chloroform, introduced in surgery by the Scottish obstetrician James Simpson in 1847. Ether was flammable, but if chloroform was safer from this point of view,[299] it was a severe hepatotoxic drug and cardiovascular depressant. Despite the relatively high incidence of death associated with the use of chloroform, it became

FIGURE 1.50 William T. G. Morton.

the anesthetic of choice for nearly 100 years. Many other halogenoalcanes had been synthesized, among which ethylene chloride and recently, halothane, a non-flammable anesthetic that was introduced into clinical practice in 1956, after its preparation at Imperial Chemical Industries. It revolutionized anesthesia. In the 1860s, the introduction of the hypodermic syringe give new opportunities for the use of drugs for anesthesia. Injectable anesthetics were introduced after the works of Eugene Baumann and Alfred Kast who introduced, in 1887, a major advance with sulfones, mainly Sulfonal, a long-acting sedative drug.[300] In the 1940s and early 1950s, muscle relaxants were introduced, firstly with curare (derived from the original South American Indian poison studied by Claude Bernard 100 years before) and then over subsequent decades a whole series of other agents.[301] Curare in the form of tubocurarine was first used in clinical anesthesia in Montreal in 1943 by Harold Griffith and Enid Johnson[302] and first used in the United Kingdom in 1946 by Gray in Liverpool: "The road lies open before us and ... we venture to say we have passed yet another milestone, and the distance to our goal is considerably shortened".[303]

Local anesthesia began in Vienna (1884) when Carl Koller administered cocaine, locally, over cornea, in order to anaesthetize the eye before cataract surgery.[304] He noticed that the drug was able to prevent the oculomotor reflex in frogs. Cocaine had been previously isolated from coca leaves by Albert Nieman in 1860.[305] Before this founding step, local sensitivity could be abolished by the dermal administration of organic derivatives like diethyl ether or ethylene chloride on the skin. Few years later, William Halsted in the United States used cocaine

FIGURE 1.51 Alfred Einhorn.

FIGURE 1.52 Emil Fischer.

FIGURE 1.53 Joseph Von Mering.

to block nerves. Paul Reclus (Paris, France) and August Bier (Berlin, Germany) used it for loco-regional anesthesia.[306] Unfortunately, cocaine is an addicting drug and between the years 1890s to 1910s, it became a pillar of drug addiction.[307] Cocaine will be completely eradicated from clinical use in the years 1914–1916 with restrictive law in the United States as well as in Europe. Fortunately, synthetic local anesthetics will appear due to the works of Alfred Einhorn[308] (Figure 1.51) and Wilehm Filehne,[309] in Germany, and Ernest Fourneau[310] in France.

7. *Antiepileptic drugs*

Historically, agents introduced for the treatment of epilepsy have also been turned to almost simultaneously for psychiatric indications. The original first-generation antiepileptic drug, a bromide salt, which appeared in 1857[311] was also known for its tranquilizing properties. The discovery of barbituric acid by the German chemist Adolf Von Baeyer (Nobel Prize in Chemistry in 1905) by condensation of malonic acid and urea, took place in Ghent (Belgium) and led to a series of other derivatives of similar structure that opened avenues to drugs that were significant both therapeutically and socially. In 1903, Emil Fischer (Figure 1.52) and Josef Von Mering (Figure 1.53), working at Bayer, were the first to synthesize a therapeutically active "barbiturate" by substituting two ethyl groups for two hydrogens attached to carbon. Diethyl barbituric acid (Veronal®) also called Barbital became a very popular drug.[312]

It allowed sleep at night, and even caused drowsiness and relaxation when taken during the day. Problem was it was slow to take effect and was very slow to wear off due to slow metabolism. So those who took it may wind up sleeping a day and a half! Chemists tried to find how onset of action could be made faster, and the duration of action not quite as long with still a hypnotic effect. One century after those first discoveries, the problem remains incompletely resolved. In 1912, two independent teams of chemists synthesized phenobarbital what became marketed as Luminal®.

Subsequent research on barbiturates began to understand that the lack of drug activity in barbituric acid and the slow acting effect of Veronal were caused by a negligible or slow passage into the circulatory system after passage through the gastrointestinal tract. The problem was the insolubility of barbituric acid in fat and Veronal is only slightly more soluble. Chemists needed to develop molecules containing larger hydrocarbon groups that resembled the fatty components of the body's barriers. This idea led to modification in the chemicals to yield very lipophilic compounds that crossed the blood-brain barrier quickly and those that could be administered intravenously for pre-surgical anesthesia. Manipulations of the side chain at position 5 have resulted in amobarbital (Amytal®), pentobarbital (Nembutal®), and secobarbital (Seconal®) which became drugs of abuse. Changes in position 2 have resulted in short-acting barbiturates: hexobarbital (Evipal®), thiopental (Pentothal®) and methohexital (Brevital®). In addition to an excellent hypnotic action, Alfred Hauptmann (Halle University, Germany) discovered phenobarbital to have potent antiepileptic activity in 1912.[313] The compound, given twice daily, kept seizures under control and its hypnotic effects could be counteracted by administering amphetamines without affecting the anticonvulsant properties. Barbiturates mainly used as antiepileptic drugs or sleep-inducers were also very useful in the operating room, especially when the interest of Thiopental, a very rapid and short-acting derivative was launched in 1935, after the works of the American anesthetist, John S. Lundy (Mayo Clinic, Rochester, USA). Thiopental has been enthusiastically accepted as an agent for the rapid induction of general anesthesia.[314] Barbiturates enable the patient to go off to sleep quickly, smoothly and pleasantly contrary to inhalated agents. In the early 1920s, Parke-Davis laboratories began to develop an experimental model for studying the anticonvulsivant interest of various substances in animals. Diphenylhydanytoin (phenytoin) which had been synthesized by the German chemist Heinrich Blitz in 1908 had been shelved for decades before to be the first item on the list of compounds sent to Houston Merritt and Tracy Putnam for experiment with their new apparatus allowing to demonstrate that the threshold current at which convulsions occurred in cats remained relatively constant over several days; eventually this model was used to characterize the anticonvulsant activity of over 600 compounds. Phenytoin was found to have anticonvulsivant properties in animals late in 1936 and its clinical efficacy was established in 1937. Dilantin sodium capsules were prepared by Parke, Davis & Co. and were ready for marketing the same year.[315] Phenytoin was introduced as an antiepileptic drug in 1938.[316] It is the most widely used anticonvulsant drug, but has many side effects. Although its chemical mode of action is unknown, phenytoin is believed to function primarily by interference with the transport of sodium ions across the neuronal membrane. In the early 1960s, there was near-simultaneous introduction of carbamazepine and valproic acid and its derivatives, as new treatments for epilepsy. Although valproic acid (chemically near from valeric acid) had been first prepared by Beverly S. Burton, an American working in Europe, in 1882,[317] its antiepileptic utility was not appreciated until this was serendipitously discovered 80 years later by Pierre Eymard, as he was preparing his doctorate in Georges Carraz (Grenoble, France). In 1962, he used valproic acid (*n*-dipropylacetic acid) as a vehicle to dissolve novel compounds being tested for anticonvulsant activity. He found anticonvulsant activity in each compound tested and eventually in valproate itself.[318] The first clinical trials of sodium valproate were published in 1964 and its marketing in France began in 1967.[319] Carbamazepine was not first synthesized until 1960, in the United States, by Walter Schindler (Geigy, Basel, Switzerland) who had a decade earlier patented the structurally closely related imipramine and it was found to have antiepileptic properties.[320] When concurrent remedial effects on mood and behavior were noted with both carbamazepine and valproic acid in the very early epilepsy trials, both drugs were soon appropriated by psychiatrists, first by Lambert[321] (Chambéry, France), in 1966 using the amide derivative of valproic acid (Figure 1.54).

FIGURE 1.54 Jens Christian Skou.

It has only been since the mid-1990s that a series of novel antiepileptic drug has been approved. Five of these agents are currently available, which might then be termed

the "third generation". These are felbamate, lamotrigine,[322] gabapentin,[323] topiramate and tiagabine.[324] The search for new antiepileptic drugs focuses now toward precise targets like the α_2-δ-protein, an auxiliary subunit of voltage-gated calcium channels. This is the case of Pregabalin, a drug structurally related to the antiepileptic drug gabapentin. Their site of action is similar. Pregabalin reduces the synaptic release of several neurotransmitters, apparently by binding to calcium channel α_2-δ-subunits, and possibly accounting for its actions *in vivo* to reduce neuronal excitability and seizures.[325] Pregabalin is also approved in United States and Europe for adjunctive therapy of partial seizures in adults, and also has been approved for the treatment of pain from diabetic neuropathy or post-herpetic neuralgia in adults.

III. CONSIDERATIONS ON RECENT TRENDS IN DRUG DISCOVERY

A. From genetics to DNA technology

In 1935, George Wells Beadle before collaborating with Edward Lawrie Tatum (Columbia, New York) began studying the development of eye pigment in *Drosophila* with Boris Ephrussi.

After producing mutants of *Neurospora crassa*, a bread mold by irradiation and searching for interesting phenotypes, they concluded in a 1940 report that each gene produced a single enzyme, also called the "single gene–single enzyme" concept, leading the two scientists to share the Nobel Prize in Physiology or Medicine in 1958.[326] Scientific key stones were collected by Joshua Lederberg (plasmid concept), John Franklin Enders, Thomas H. Weller, Frederick Chapman Robbins (virus cultures), Salvador Luria, and Alfred Day Hershey (bacteriophage), but radical turn came in 1950. This is the time when Watson joined the Cavendish laboratories at King's College (London) at a time when Francis Crick, Maurice Wilkins, Rosalind E. Franklin and Linus Pauling attempted to determine the structure of DNA. The X-ray crystallography experiments of Franklin and Wilkins provided much information about DNA. Crick and Watson made the intuitive leap: in 1953, proposing the Watson–Crick Model of the DNA double helix which provided enormous impetus for research in the emerging fields of molecular genetics and biochemistry when published in *Nature* in 1953[327] (Figure 1.55).

In the following years, the enzymes puzzle found its place through Severo Ochoa at New York University School of Medicine discovering in 1955, polynucleotide phosphorylase, an RNA-degrading enzyme, then, Arthur Kornberg (Washington University, St. Louis) (Figure 1.56) discovering DNA polymerase and Mahlon Bush Hoagland and Paul Zamenick (Harvard Medical School) discovering transfer RNA (tRNA). All of the pieces were in place for Francis Crick to postulate in 1958 the "central dogma" of DNA – that genetic information is maintained and transferred in a one-way process, moving from nucleic acids to proteins.

FIGURE 1.55 James Watson and Francis Crick.

FIGURE 1.56 Arthur Kornberg.

In 1960, François Jacob and Jacques Monod and André Lwoff (Institut Pasteur, Paris) (Figure 1.57) proposed their operon model and were awarded the Nobel Prize in 1965. This was the birth of gene regulation models, which launched a continuing quest for gene promoters and triggering agents. In 1952, Frederick Sanger (Cambridge, UK) (Figure 1.58) had used paper chromatography to sequence

FIGURE 1.57 François Jacob, Jacques Monod and André Lwoff.

FIGURE 1.58 Frederick Sanger.

the insulin amino acids. In 1945, Sanger made an important technological breakthrough that made possible his later sequencing work on amino acids discovering that dinitrophenol (DNP) could bind tightly to one end of an amino acid and that this bond was stronger than the one formed by two amino acids bonding with one another.

This fact made it possible for Sanger to use DNP to take apart the insulin molecule one amino acid at a time. Each amino acid could then be identified by the newly discovered process of paper chromatography. The technique resulted

FIGURE 1.59 Bruce Merrifield.

in the eventual identification of all amino acid groups in the insulin molecule.[328] Sanger received the Nobel Prize in Chemistry, twice, in 1958 and 1980! After fundamental breakthroughs in the analysis of protein structure and the elucidation of protein functions, the following step was the manufacturing of therapeutic proteins. In 1954, Vincent Du Vigneaud at Cornell University (New York) synthesized oxytocin. The same year, ribosomes were identified as the site of protein synthesis. In 1956, the three-dimensional structure of proteins was linked to the sequence of its amino acids, so that by 1957, John Kendrew (Cambridge, UK) was able to solve the first three-dimensional structure of a protein (myoglobin); this was followed in 1959 with Max Perutz's determination of the three-dimensional structure of hemoglobin. In 1964, Bruce Merrifield (Rockefeller Institute, New York) (Figure 1.59) invented a simplified technique for protein and nucleic acid synthesis. Sixty years before, Emil Fischer's process involved blocking the carboxyl group of one amino acid and the amino group of the second amino acid. Then, by activation of the free carboxyl group, the peptide bond could be formed, and selective removal of the two protecting groups would lead to the free dipeptide.

The peptide was then separated from the by-products and unreacted starting material and the process was repeated. Merrifield succeeded in assembling a peptide chain in a stepwise manner while it was attached at one end to a solid support that could easily be removed by the proper solvents. It soon became apparent to Merrifield that the solid phase technique should be applicable to units other than amino acids. He extended it to the synthesis of

FIGURE 1.60 Kary B. Mullis.

depsipeptides while other laboratories succeeded in synthesizing polynucleotides and polysaccharides.[329]

These discoveries were made outside of the pharmaceutical industry, but gave enormous contributions to understanding the mechanisms of diseases and therapeutic drugs. These developments proved critical to functional analysis in basic physiological research and to drug discovery, through specific targeting. A significant advance in protein manufacturing was performed in when Kary B. Mullis (Cetus Company, Emeryville, USA) (Figure 1.60) coined a tool so ingenious that it revolutionized the field of therapeutic protein synthesis but also of DNA and protein characterization and synthesis, in general. *Polymerase chain reaction* (PCR) affected many aspects of biology. This technique amplifies DNA, enabling scientists to make billions of copies of a DNA molecule in a very short time. PCR has been used to detect DNA sequences, to diagnose genetic diseases, to carry out DNA fingerprinting, to detect bacteria or viruses (AIDS diagnostics), and to research human evolution (Box 1.3).

Another major advance in medicinal chemistry was imagined by Cesar Milstein (Cambridge, UK) and Georges Köhler (Max Planck Institute for Immunobiology, Freiburg), (Figure 1.61) who were awarded the Nobel Prize for Physiology or Medicine in 1984. They conducted a groundbreaking work into the synthesis of antibodies, proteins that are produced by the cells of the immune system in response to attacks by antigens. Their work was instrumental in the development of monoclonal antibody technology. By fusing antibody-producing B lymphocyte cells with tumor cells that are "immortal," they produced a "hybridoma," which could continuously synthesize antibodies. All of the antibodies produced by this type of hybridoma cell were identical, and came from a single clone of hybridoma cells, they were called monoclonal antibodies. This technique, developed in 1975, has been used extensively in the commercial development of new drugs and diagnostic tests. Initial use of monoclonal antibodies to treat disease in humans was limited, because they were produced in mice and induced an immune response in the human host. Later, Gregory Winter and Michael Neuberger at the Laboratory for Molecular Biology (Cambridge, UK) discovered how

BOX 1.3 Recombinant DNA

In 1970, Robert J. C. Harris, a cytogeneticist, proposed the term "genetic engineering." But DNA recombinant technology needed the discovery, in 1970, of restriction enzymes which cut DNA in the middle of a specific symmetrical sequence, by Werner Arber (Basel, Switzerland). The same year, Werner Arber (Basel, Switzerland) discovered restriction enzymes. Hamilton O. Smith and Daniel Nathans (Baltimore) verified Arber's hypothesis, showing, first that this enzyme cuts DNA in the middle of a specific symmetrical sequence, second the usefulness of restriction enzymes in the construction of genetic maps. Since that time, restriction enzymes are one of the main tools allowing solving various problems in genetics. The three scientists shared the 1978 Nobel Prize in Physiology or Medicine for their work in producing the first genetic map (of the SV40 virus). In 1972, was the date of birth for DNA recombinant technology, when Stanley Cohen and Herbert Boyer, (San Francisco), combined plasmid isolation with DNA splicing. They had the idea to combine the use of the restriction enzyme EcoR1 with DNA ligase to form engineered plasmids capable of producing foreign proteins in bacteria – the basis for the modern biotechnology industry. By 1973, Cohen and Boyer had produced their first recombinant plasmids. They received a patent on this technology for Stanford and UCSF that would become one of the biggest money-makers in pharmaceutical history. The year 1975 was the year of DNA sequencing. Walter Gilbert and Allan Maxam (Harvard, Boston) and Fred Sanger (Cambridge, UK) simultaneously developed different methods for determining the sequence of bases in DNA. Gilbert and Sanger shared the 1980 Nobel Prize in Physiology or Medicine. Genetic tools and principles being discovered, the link between these scientific discoveries and drug manufacturing could easily take place. By 1976, Robert Swanson teamed up with Herbert Boyer to form Genentech Inc., the harbinger of a wild proliferation of biotechnology companies over the next decades. Genentech's goal of cloning human insulin in E. *coli* was achieved in 1978, and the technology was licensed to Eli Lilly. It has been the first drug genetically engineered, but recombinant DNA era grew from these beginnings and had a major impact on pharmaceutical production and research in the 1980s and 1990s. Since that time, dozens of protein drugs had been marketed: growth hormone, colony stimulating factors, erythropoietin, tissue plasminogen activator, antihemophilic factors, interferons, monoclonal antibodies, etc.

FIGURE 1.61 Cesar Milstein and Georg Köhler.

to engineer the combining site of the mouse monoclonal antibody into the human immunoglobulin gene resulting in chimeric antibodies allowing a convenient administration to humans with little or no antimouse response.

The benefits of humanized monoclonal antibodies in the treatment of otherwise intractable diseases have been dramatic as for monoclonal antibodies to TNF-α, or to CD20, respectively, administered to patients with RA or lymphomas.

1. *Gene therapy and cell therapy*

If genetic sciences gave birth to a revolutionary rebound in drug discovery from insulin genetically engineered to the last monoclonal antibodies largely used in cancer treatments, gene therapy by itself remains more promising than successful and clinical trials on humans remain disappointing. In September 1990, the first human gene therapy was started by W. French Anderson at NIH (Bethesda, USA) in an attempt to cure adenosine deaminase (ADA) deficiency by inserting the correct gene for ADA into an afflicted 4-year-old girl.[330] Although the treatment did not provide a complete cure, it did allow the young patient to live a more normal life with supplemental ADA injections. One objective of gene therapy is tissue engineering: production of functional, biocompatible tissues involving the genetic modification of cells that are seeded onto (or into) scaffolds prior to implantation. The genetic modification is achieved through gene delivery, which can utilize viral transduction or non-viral transfection systems. Although novel non-viral systems have continued to emerge as innovative vehicles for controlled gene delivery, retrovirus, adenovirus, adeno-associated virus, and herpes virus remain the most efficient means by which exogenous genes can be introduced into and expressed by mammalian cells.[331] Among problems linked to gene therapy to treat certain diseases (mainly monogenic) at their origins include the problem of insertional oncogenesis. For the future, it seems necessary to minimize the genotoxic risk of gene therapy protocols, which will be important if gene therapy is to fulfill its conceptual promise.[332] Perhaps more hopeful, stem cell therapy held great promise since the beginning of the 1990s. This treatment uses human cells to repair and ameliorate inborn or acquired medical conditions, from Parkinson's disease to heart, lung, kidney diseases and from diabetes to traumatic spinal paralysis. Many studies have tried to manipulate the growth and differentiation conditions with varied success.[333]

B. Hopes and limits for drug hunting

At the end of the 1980s, progress in physiology, biochemistry and molecular biology, gave rise to new approaches in drug discovery. More than the knowing of the intimacy of cell structure and metabolism, new molecular techniques allowed aiming hundreds of proteins as pharmacological targets Thus, an improved understanding of the biochemical mechanisms of diseases allowed the identification of new drug targets and the development of disease models. Combinatorial chemistry for the synthesis of large series of compounds, high-throughput screening for the rapid identification of new leads, X-ray crystallography and NMR for the determination of protein 3D structures and the identification of ligands and computer-aided drug design for the search for new leads and their rational optimization helped to fill the gap in the knowledge of the set of molecular targets.[334] Nevertheless, a selective optimization of side activities of drug molecules (the "SOSA" approach) remains an intelligent and potentially efficient strategy for the generation of new biological activities.[335] But the question remains for pharmaceutical scientists to know whether all possible targets for drugs to act are currently described or if new discoveries are needed in order to find new pharmacological families. In 1997, 483 drug targets were identified.[336] In 2002, only 120 underlying molecular targets were proposed;[337] the following year, the number of 273 molecular targets[338] was proposed and recently, a consensus number of 324 drug targets for all classes of approved therapeutic drugs, appeared, reconciling earlier reports into a current and comprehensive survey.[339] One of the ultimate goals of the drug discovery process is to provide an understanding of the complete set of molecular mechanisms describing an organism. Although this goal is a long way off, many useful insights can already come from currently available information and technology. In the 1990s, dorzolamide, a carbonic anhydrase inhibitor used for the topical treatment of glaucoma, was the first drug developed on the basis of a known target structure.[340] Other structure-based developments have followed, notably in the field of PIs who opened a new era for treatment of HIV disease. The objective of numerous research teams is to elucidate three-dimensional structures of pharmacological targets to leave guesses of hypothetical binding sites and permit to study the interaction of drugs

and their sites of action. In the "real life", the reality is often more complex than the theoretical point of view. In some cases, drugs act through a single target ; for example, the histamine H1 receptor is believed to be the major mechanistic target for cetirizine and hydroxyzine, and acebutolol acts through an adrenoceptor, although all these drugs show binding to other G-protein-coupled receptors. In other cases, the drug can act through multiple distinct mechanisms, and therefore unrelated targets. A complicating feature of any such analysis is that many drugs have complex and relatively poorly understood pharmacology, and often limited selectivity against related proteins, and some targets are actually complex multimeric proteins with variable subunit compositions and so on. The identification of therapeutic targets requires knowledge of a disease's etiology and the biological systems associated with it outlining the need to integrate chemistry into biological research. Modern medicinal chemistry is more and more focused on the interactions of small molecules with proteins than with genes, which code for the synthesis of those proteins. Among the characteristics of modern evolution to drug innovation, the use of molecule extracted from human body, tissues or cells has to be pointed out. At the beginning of the 20th century, knowledge of human hormones drove to the extraction of sexual hormones from gonads, or insulin from pancreas. During the last years, the discovery of the nature of genes and the tools for genetic engineering permitted to biosynthesize proteins involved in genetic or non-genetic diseases, including, cancer, infectious diseases, endocrines or immune disorders, etc. or to discover small molecules interacting with proteins. In few years, mammalian cells became the main source of drug discovery and molecular biology has revolutionized the process of drug discovery. An increased interaction between chemistry and biology will perhaps help to fight one of the biggest challenges in drug discovery today: the high attrition rate. Many promising candidates prove ineffective or toxic owing to a poor understanding of the molecular mechanisms of biological systems they target.[341] During the last 150 years, the average life expectancy of occidentals has almost doubled. Medical and pharmaceutical progresses take a large part of this performance. Old drugs, still useful, and recent discoveries add their powerful actions to provide considerable advances in healthcare and longevity. In the past, the natural world, mineral, plant and animal reigns were the source of most medicines. With the dawn of magic bullets in the 20th century, complex organic chemistry opened up a world of drugs designed in pharmaceutical industry by modifying old molecules, or by *de novo* creation. But this is not the end of the story. Even if a wide understanding has been achieved for a drug acting on a perfectly known pharmacological target, new features may be discovered. For the pharmaceutical industry, the discovery of a new drug presents an enormous scientific (and financial) challenge and consists essentially in the identification of new molecules or compounds. Ideally, the latter will become drugs that act in new ways upon biological targets specific to the diseases requiring new therapeutic approaches. In the last decades, we have witnessed a decline in the number of new drugs that were introduced into therapy and, for some drug families such as for antibiotics, it is almost a total lack of innovation. Sometimes this fact is discussed as an argument against the contribution of modern drug design strategies. However, reasons for this decline are complex. Many scientists tell that it could be worse. The tools for drug discovery, today, do not leave the same chance for serendipity in drug discovery today as it was in the past. Genomics (DNA), transcriptomics (RNA), proteomics (peptides and proteins) allow for a much more rapid and precise discovery in the etiology and then treatment of much of diseases. The results obtained from this considerable progress are quite deceiving: among new molecular entities and biologics approved by the FDA in 2006, 19 were small molecules, 2 were enzymes, and 2 were monoclonal antibodies.[342] Drug discovery remains an uncertain, hazardous, and unpredictable adventure! This is probably why most of drug "hunters" largely share the opinion of Nobel laureate James Black who famously declared: "The most fruitful basis for the discovery of a new drug is to start with an old drug." (Tables 1.10 and 1.11).

TABLE 1.10 Nobel Prizes in Medicine in Relation with the Discovery of Drugs or in Biological Systems Directly Connected with Drug Discovery

Year	Name of laureates	Discovery as it is cited par the Nobel Committee
2005	Barry J. Marshall, J. Robin Warren	*Helicobacter pylori* and its role in gastritis and peptic ulcer disease
1998	Robert F. Furchgott, Louis J. Ignarro, Ferid Murad	Nitric oxide as a signaling molecule in the cardiovascular system
1991	Erwin Neher, Bert Sakmann	Function of single ion channels in cells
1988	James W. Black, Gertrude B. Elion, George H. Hitchings	Important principles for drug treatment
1986	Stanley Cohen, Rita Levi-Montalcini	Growth factors

(Continued)

TABLE 1.10 (Continued)

Year	Name of laureates	Discovery as it is cited par the Nobel Committee
1985	Michael S. Brown, Joseph L. Goldstein	Regulation of cholesterol metabolism
1984	Niels K. Jerne, Georges J. F. Köhler, Cesar Milstein	Development and control of the immune system and the principle for production of monoclonal antibodies
1982	Sune K. Bergström, Bengt I. Samuelsson, John R. Vane	Prostaglandins and related biologically active substances
1978	Werner Arber, Daniel Nathans, Hamilton O. Smith	Restriction enzymes and their application to problems of molecular genetics
1977	Roger Guillemin, Andrew V. Schally, Rosalyn Yalow	Peptide hormone production of the brain and development of radioimmunoassays of peptide hormones
1971	Earl W. Jr. Sutherland	Mechanisms of the action of hormones
1970	Sir Bernard Katz, Ulf Von Euler, Julius Axelrod	Humoral transmitters in the nerve terminals and the mechanism for their storage, release and inactivation
1957	Daniel Bovet	Synthetic compounds that inhibit the action of certain body substances, and especially their action on the vascular system and the skeletal muscles
1952	Selman Abraham Waksman	Streptomycin, the first antibiotic effective against tuberculosis
1950	Edward Calvin Kendall, Tadeus Reichstein, Philip Showalter Hench	Hormones of the adrenal cortex, their structure and biological effects
1945	Alexander Fleming, Ernst Boris Chain, Howard Walter Florey	Penicillin and its curative effect in various infectious diseases
1943	Henrik Carl Peter Dam Edward Adelbert Doisy	Vitamin K and its chemical nature
1939	Gerhard Domagk	Antibacterial effects of Prontosil
1937	Albert Szent-Györgyi Von Nagyrapolt	Biological combustion processes, with special reference to vitamin C and the catalysis of fumaric acid
1936	Henry Hallett Dale, Otto Loewi	Chemical transmission of nerve impulses
1929	Christiaan Eijkman, Frederick Gowland Hopkins	Antineuritic vitamin and the growth-stimulating vitamins
1923	Frederick Grant Banting, John James Richard MacLeod	Insulin
1913	Charles Robert Richet	Work on anaphylaxis
1908	Ilya Ilyich Mechnikov, Paul Ehrlich	Work on immunity
1905	Robert Koch	Investigations and discoveries in relation to tuberculosis
1901	Emil Adolf Von Behring	Work on serum therapy, especially its application against diphtheria, by which he has opened a new road in the domain of medical science and thereby placed in the hands of the physician a victorious weapon against illness and deaths

TABLE 1.11 Nobel Prizes in Chemistry in Relation with the Discovery of Drugs or in Chemical Systems Directly Connected with Drug Discovery

Year	Name of laureates	Discovery as it is cited par the Nobel Committee
2003	Peter Agre, Roderick Mackinnon	Water channels and structural and mechanistic studies of ion channels.
1997	Paul D. Boyer, John E. Walker	Enzymatic mechanism underlying the synthesis of ATP.
1997	Jens C. Skou	First ion-transporting enzyme, Na^+, K^+-ATPase
1984	Robert Bruce Merrifield	Methodology for chemical synthesis on a solid matrix.
1980	Paul Berg	Biochemistry of nucleic acids, with particular regard to recombinant-DNA
1980	Walter Gilbert, Frederick Sanger	Determination of base sequences in nucleic acids
1969	Derek H. R. Barton, Odd Hassel	Development of the concept of conformation and its application in chemistry
1965	Robert Burns Woodward	The art of organic synthesis
1964	Dorothy Crowfoot Hodgkin	Determinations by X-ray techniques of the structures of important biochemical substances
1962	Max Ferdinand Perutz, John Cowdery Kendrew	Structures of globular proteins
1958	Frederick Sanger	Structure of proteins, especially that of insulin
1955	Vincent Du Vigneaud	Biochemically important sulphur compounds, especially for the first synthesis of a polypeptide hormone
1947	Robert Robinson	Plant products of biological importance, especially the alkaloids
1939	Adolf Friedrich Johann Butenandt	Sex hormones
1938	Richard Kuhn	Carotenoids and vitamins
1937	Sir Walter Norman Haworth	Carbohydrates and vitamin C
1937	Paul Karrer	Carotenoids, flavins and vitamins A and B2
1928	Adolf Otto Reinhold Windaus	Constitution of the sterols and their connection with the vitamins
1918	Fritz Haber	Synthesis of ammonia from its elements
1915	Richard Martin Willstätter	Plant pigments, especially chlorophyll
1905	Johann Friedrich Wilhelm Adolf Von Baeyer	Advancement of organic chemistry and the chemical industry, through his work on organic dyes and hydroaromatic compounds
1902	Hermann Emil Fischer	Work on sugar and purine syntheses

REFERENCES

1. Grmek, M. D., Gourevitch, D. Les expériences pharmacologiques dans l'antiquité. *Arch. Int. Hist. Sci.* 1985, *35*, 3–27.
2. Lavoisier, A. A. *Traité Elémentaire de Chimie présenté dans un ordre nouveau et d'après les découvertes récentes, avec figures*. Cuchet: Paris, 1789.
3. Perrin, C. E. The Chemical revolution. In *Companion to the History of Modern Science* (Colby, R. C., Cantor, G. N., Christie, J. R. R., Hodge, M. J. S., Eds). Routeledge: London and New York, 1990.
4. Aftalion, F. *Histoire de la Chimie*. Masson: Paris, 1988.
5. Bensaude-Vincent, B., Stengers, I. *Histoire de la Chimie*. La Découverte: Paris, 1993.
6. Bernard, C. *Leçons de physiologie expérimentale appliquée à la medicine*. J.B. Baillière: Paris, 1855–1856.
7. Virchow, R. L. *Die Cellularpathologie in ihrer Begründung auf physiologische und pathologische Gewebelehre*. A. Hirschwald: Berlin, 1858.
8. Pasteur, L.*Œuvres complètes réunies par Pasteur Vallery-Radot*, Vol. 7. Masson: Paris, 1922–1939.
9. Soubeiran, E. Recherches sur quelques combinaisons du chlore. *Ann. Chim.* 1831, *48*, 113–157.
10. Von Liebig, J. Ueber die Verbindungen, welche durch die Einwirkung des Chlors auf Alkohol, Aether, ölbildenes Gas und Essiggeist entstehen. *Ann. Pharm.* 1832, *1*, 182–230.
11. Guthrie, S. New mode of preparing a spirituous solution of chloric ether. *Am. J. Sci. Arts.* 1832, *21*, 64–65. *22*, 105–106.

12. Von Liebig, J. *Die organische Chemie in ihrer Anwendung auf Physiologie und Pathologie*. F. Braunschweig: Vieweg, 1842.
13. Wöhler, F. Üeber künstliche Bildung des Harnstoffe. *Poggendorfs Ann. Phys. Chem.* 1828. Leipzig
14. Müller-Jahncke, W. D., Friedrich, C. *Geschichte der Arzneimitteltherapie*. Deutscher Apotheker Verlag: Stuttgart, 1996.
15. Cadet de Gassicourt, C. L. Considérations sur l'Etat actuel de la Pharmacie. *Bull. Pharm.* 1809, *1*, 5–12.
16. Pelletier, P. J., Magendie, F. Recherches chimiques et physiologiques sur l'ipecacuanha. *Ann. Chim. Phys.* 1817, *4*, 172–185.
17. Pelletier, P. J., Caventou, J. B. Recherches chimiques sur les quinquinas. *Ann. Chim. Phys.* 1820, *15*, 289–318. 337–365
18. Robiquet, P. J. Nouvelles observations sur les principaux produits de l'opium. *Ann. Chim. Phys.* 1832, *51*, 225–267.
19. Mein, H. F. Üeber die Darstellung des Atropins in weissen Krystallen. *Ann. Chem. Pharm.* 1833, *6*, 67–72.
20. Merck, G. F. Vorläufige Notiz über eine neue organische Base im Opium. *Ann. Phys. Chem.* 1848, *66*, 125–128.
21. Ladenburg, A. Die Natürlich vorkommenden mydriatische wirkenden Alkaloïde. *Ann. Chem. Pharm.* 1881, *206*, 274–307.
22. Magendie, F. *Formulaire pour la préparation et l'emploi de plusieurs nouveaux médicamens, tels que la noix vomique, la morphine, etc*. Méquignon-Marvis: Paris, 1822.
23. Derosne, J. F. Mémoire sur l'opium. *Ann. Chim.* 1803, *45*, 257–285.
24. Seguin, M. A. Premier mémoire sur l'opium. *Ann. Chim.* 1814, *92*, 225–245.
25. Sertürner, F. Auszuge aus briefen an den Herausgeber. (a) Säure im opium. (b) Ein deres Schreiben von Ebendenselben. Nachtrag zur Charakteristik der Saüre im Opium,. *J. Pharmazie für Artze, Apotheker, und Chemisten von D.J.B. Trommsdorff* 1805, *13*, 29–30.
26. Sertürner, F. Darstellung der reinen Mohnsäure (Opiumsäure) nebst einer Chemischen Untersuchung des Opiums mit vorzüglicher Hinsicht auf einen darin neu entdeckten Stoff und die dahin gehörigen Bemerkungen. *J. der Pharmacie fuer Aerzte und Apotheker* 1806, *14*, 47–93.
27. Sertürner, F. Ueber das Morphium, eine neue salzfähige Grundlage, und die Mekonsäure, als Hauptbestandtheile des Opiums. *Annalen der Physik* 1817, *55*, 56–89.
28. Pert, C. B., Snyder, S. H. Opiate receptor: demonstration in nervous tissue. *Science* 1973, *179*, 1011–1014.
29. Kosterlitz, H. W., Hughes, J. Some thoughts on the significance of enkephalin, the endogenous ligand. *Life Sci.* 1975, *17*, 91–96.
30. Mitchell, J. M., Tavares, V. C., Fields, H. L., D'Esposito, M., Boettiger, C. A. Endogenous opioid blockade and impulsive responding in alcoholics and healthy controls. *Neuropsychopharmacology* 2007, *32*, 439–449.
31. McLagan, T. J. The treatment of acute rheumatism with salicin. *Lancet* 1876, *1*, 342–343. 383–384
32. Piria, R. Recherches sur la salicine et les produits qui en dérivent. *C. R. Acad. Sci.* 1839, *8*, 479–485.
33. Gerhardt, C. F. Untersuchungen über die wasserfrei organischen Saüren. *Ann. Chem. Pharm* 1853, *87*, 149–179.
34. Sneader, W. The discovery of aspirin: a reappraisal. *Br. Med. J.* 2000, *321*, 1591–1594.
35. Dreser, H. Pharmakologisches über Aspirin (Acetylsalicylsaüre). *Pflüg. Arch. Ges. Physiol.* 1899, *76*, 306–318.
36. Vane, J. R. Nobel Lecture, 8th December 1982. Adventures and excursions in bioassay: the stepping stones to prostacyclin. *Br. J. Pharmacol.* 1983, *79*, 821–838.
37. Samuelsson, B. An elucidation of the arachidonic acid cascade. Discovery of prostaglandins, thromboxane and leukotrienes. *Drugs* 1987, *33*(Suppl. 1), 2–9.
38. Chast, F. John Vane, 1927–2004, pharmacologue de l'endothelium vasculaire. *Ann. Pharm. Fr.* 2006, *64*, 416–424.
39. Simmons, D. L., Xie, W., Evett, G., Merrill, J., Robertson, D. L., Bradshaw, W. S. Drug inhibition and cellular regulation of prostaglandin G/H synthase isoenzyme 2. *J. Lipid. Mediat.* 1993, *6*, 113–117.
40. Xie, W., Chipman, J. G., Robertson, D. L., Erikson, R. L., Simmons, D. L. Expression of a mitogen-responsive gene encoding prostaglandin synthase is regulated by mRNA splicing. *Proc. Natl. Acad. Sci. USA* 1991, *88*, 2692–2696.
41. O'Banion, M. K., Sadowski, H. B., Winn, V., Young, D. A. A serum- and glucocorticoid-regulated 4-kilobase mRNA encodes a cyclooxygenase-related protein. *J Biol. Chem.* 1991, *266*, 23261–23267.
42. Flower, R. J. The development of Cox2 inhibitors. *Nat. Rev. Drug Discov.* 2003, *2*, 179–191.
43. Topol, E. J. Failing the Public Health — Rofecoxib, Merck, and the FDA. *New Engl. J. Med.* 2004, *351*, 1707–1709.
44. Fitzgerald, G. A. COX-2 and beyond: approaches to prostaglandin inhibition in human disease. *Nat. Rev. Drug Discov.* 2003, *2*, 879–890.
45. Zhu, J., Song, X., Lin, H. P., Young, D. C., Yan, S., Marquez, V. E., Chen, C. S. Using cyclooxygenase-2 inhibitors as molecular platforms to develop a new class of apoptosis-inducing agents.. *J. Natl. Cancer Inst.* 2002, *94*, 1745–1757.
46. Maini, R. N., Elliott, M. J., Brennan, F. M., Williams, R. O., Chu, C. Q., Paleolog, E., Charles, P. J., Taylor, P. C., Feldmann, M. Monoclonal anti-TNF alpha antibody as a probe of pathogenesis and therapy of rheumatoid disease. *Immunol. Rev.* 1995, *144*, 195–223.
47. Scott, D.l., Kingsley, G. H. Tumor necrosis factor inhibitors for rheumatoid arthritis. *New Engl. J. Med.* 2006, *355*, 704.
48. Withering, W. *An account of the foxglove, and some of its medical uses*. G.G.J. & J. Robinson: Birmingham, AL, 1785.
49. Nativelle, C. A. De la digitaline et de la digitale. *Bull. Acad. Méd.* 1872, *2*, 201.
50. Schmiedeberg, J. E. O. Untersuchungen über die pharmakologische wirksamen Bestandteile der Digitalis purpurea. *L. Arch. Exp. Path. Pharmak.* 1875, *3*, 16–43.
51. Skou, J. C. The influence of some cations on an adenosine triphosphatase from peripheral nerves. *Biochim. Biophys. Acta.* 1957, *23*, 394–401.
52. Post, R. L., Merritt, C. R., Kinsolving, C. R., Albright, C. D. Membrane adenosine triphosphatase as a participant in the active transport of sodium and potassium in the human erythrocyte. *J. Biol. Chem.* 1960, *235*, 1796–1802.
53. Sen, A. K., Post, R. L. Stoichiometry and localization of adenosine triphosphate-dependent sodium and potassium transport in the erythrocyte. *J. Biol. Chem.* 1964, *239*, 345–352.
54. Schwartz, A., Lindenmayer, G. E., Allen, J. C. The sodium-potassium adenosine triphosphatase: pharmacological, physiological and biochemical aspects. *Pharmacol. Rev.* 1975, *27*, 3–134.
55. Albers, J. F. H. Die Physiologische und Therapeutische wirkung. *Dtsch. Klin.* 1864, *42*, 405.
56. Brunton, T. L. On the use of nitrite of amyl in angina pectoris. *Lancet* 1867, *2*, 97–98.
57. Sorelle, R. Nobel prize awarded to scientists for nitric oxide discoveries. *Circulation* 1998, *98*(22), 2365–2366.
58. Murad, F. Shattuck lecture. Nitric oxide and cyclic GMP in cell signaling and drug development. *New Engl. J. Med.* 2006, *355*(19), 2003–2011.
59. Brown, W. C., O'Brien, E. T., Semple, P. F. The sphygmomanometer of Riva-Rocci 1896–1996. *J. Hum. Hypertens.* 1996, *10*(11), 723–724.
60. Janeway, T. C. A clinical study of hypertensive cardiovascular disease. *Arch. Intern. Med.* 1913, *12*, 755–798.
61. Society of Actuaries. *Blood Pressure: Report of the Joint Committee on Mortality of the Association of Life Insurance Medical Directors and the Actuarial Society of America*, New York, 1925.
62. Smithwick, R. H. Surgery of the autonomic nervous system. *New Engl. J. Med.* 1947, *236*, 662–669.
63. Kempner, W. Treatment of hypertensive vascular disease with rice diet. *Am. J. Med.* 1948, *4*, 545–577.
64. Page, I. H., Corcoran, A. C. *Arterial Hypertension: Its Diagnosis and Treatment.*. Year Book Medical Publishers Inc.: Chicago, IL, 1949.

65. Freis, E. D., Finnerty, F. A., Schnaper, H. W., Johnson, R. L. The treatment of hypertension with hexamethonium. *Circulation* 1952, *5*, 20–27.
66. Schroeder, H. A. The effect of 1-hydrazinophthalazine in hypertension. *Circulation* 1952, *5*, 28–37.
67. Wilkins, R. W., Judson, W. E. Use of *Rauwolfia serpentina* in hypertensive patients. *New Engl. J. Med.* 1953, *248*, 48–53.
68. Vakil, R. J. A clinical trial of *Rauwolfia serpentina* in essential hypertension. *Br. Heart J.* 1949, *11*, 350–355.
69. Veterans Administration Cooperative Study Group on Antihypertensive Agents. Effects of treatment on morbidity in hypertension, I: results in patients with diastolic blood pressure averaging 115 through 129mmHg. *J. Am. Med. Assoc.* 1967, *202*, 1028–1034.
70. Moser, M. Seven decades of progress. *J. Clin. Hypertens.* 2007, *9*, 310–313.
71. Baker, M. H. The blood cyanates in the treatment of hypertension. *J. Am. Med. Assoc.* 1936, *106*, 762–767.
72. Page, I. H., Dustan, H. P. Current status of bretylium and guanethidine as antihypertensive drugs. *Circulation* 1960, *22*, 181–183.
73. Black, J. W., Crowther, A. F., Shanks, R. G., Smith, L. H., Dornhorst, A. C. A new adrenergic betareceptor antagonist. *Lancet* 1964, *1*, 1080–1081.
74. Case, D. B., Atlas, S. A., Laragh, J. H., Sealey, J. E., Sullivan, P. A., McKinstry, D. N. Clinical experience with blockade of the renin-angiotensin-aldosterone system by an oral converting-enzyme inhibitor (SQ 14,225, captopril) in hypertensive patients. *Prog. Cardiovasc. Dis.* 1978, *21*, 195–206.
75. Tapia, F., Dustan, H. P., Schneckloth, R. E., Corcoran, A. C., Page, I. H. Enhanced effectiveness of ganglion blocking agents in hypertensive patients during administration of a saluretic agent (chlorothiazide). *Lancet* 1957, *2*, 831–833.
76. Strauss, M. B., Southworth, H. Urinary changes due to sulfanilamide administration. *Bull. Johns Hopkins Hosp.* 1938, *63*, 41–45.
77. Mann, T., Keilin, D. Sulphanilamide as a specific inhibitor of carbonic anhydrase. *Nature* 1940, *146*, 164–165.
78. Oliver, G., Schafer, E. A. The physiological effects of extracts of suprarenal capsules. *J. Physiol.* 1895, *18*, 230–276.
79. Takamine, J. The blood pressure raising principle of supra-renal glands. *Am. J. Pharm.* 1901, *73*, 523–531.
80. Elliott, T. R. On the action of adrenalin. *J. Physiol.* 1904, *31*, 20–21.
81. Von Euler, U. S. A specific sympathomimetic ergone in adrenergic nerve fibres (sympathin) and its relations to adrenaline and nor-adrenaline. *Acta. Physiol. Scand.* 1946, *12*, 73–97.
82. Ahlquist, R. P. A study of the adrenotropic receptors. *Am. J. Physiol.* 1948, *153*, 586–600.
83. Gillam, P. M. S., Prichard, B. N. C. Discovery of the hypotensive effect of propranolol. *Postgrad. Med. J.* 1976, *52*, 70–75.
84. Tritthart, H., Volkmann, R., Weiss, R., Fleckenstein, A. Inhibition of calcium-dependent action potentials in mammalian myocardium by specific inhibitors of the transmembrane calcium conductivity (verapamil, D 600).. *Recent. Adv. Stud. Cardiac Struct. Metab.* 1975, *5*, 27–33.
85. Fleckenstein, A. History of calcium antagonists. *Circ. Res.* 1983, *52*, 13–16.
86. Tigerstedt, R., Bergman, P. G. Niere und Kreislauf. *Scand. Arch. Physiol.* 1897–1898, *8*, 233–271.
87. Goldblatt, H., Lynch, J., Hanzal, R. F., Summerville, W. W. Studies on experimental hypertension. I. The production of persistent elevation of systolic blood pressure by means of renal ischemia. *J. Exp. Med.* 1934, *59*, 347.
88. Page, I. H., Helmer, O. M. A crystalline pressor substance (angiotonin) resulting from the reaction between renin and renin activator. *J. Exp. Med.* 1940, *71*, 29–42.
89. Braun-Menendez, E., Fasciolo, C., Leloir, L. F., Munoz, M. The substance causing renal hypertension. *J. Physiol.* 1940, *98*, 283–298.
90. Skeggs, L. T., Kahn, J. R., Shumway, N. P. The preparation and function of the hypertensin-converting enzyme. *J. Exp. Med.* 1956, *103*, 295–299.
91. Davis, J. O., Freeman, R. H. Historical perspectives on the renin-angiotensin-aldosterone system and angiotensin blockade. *Am. J. Cardiol.* 1982, *49*, 1385–1389.
92. Ferreira, S. H. A bradykinin-potentiating factor (BPF) present in the venom of *Bothrops jararaca*. *Br. J. Pharmacol. Chemother.* 1965, *24*, 163–169.
93. Ferreira, S. H. History of the development of inhibitors of angiotensin I conversion. *Drugs* 1985, *30*(Suppl. 1), 1–5.
94. Yang, H. Y. T., Erdos, E. G., Levin, Y. A dipeptidyl carboxypeptidase that converts angiotensin I and inactivates bradykinin. *Biochim. Biophys. Acta.* 1970, *214*, 374–376.
95. Ondetti, M. A., Rubin, B., Cushman, D. W. Design of specific inhibitors of angiotensin converting enzyme: new class of orally active antihypertensive agents. *Science* 1977, *196*, 441–444.
96. Smith, C. G., Vane, J. R. The discovery of captopril. *FASEB J.* 2003, *17*, 788–789.
97. Wexler, R. R., Carini, D. J., Duncia, J. V., Johnson, A. L., Wells, G. J., Chiu, A. T., Wong, P. C., Timmermans, P. B. Rationale for the chemical development of angiotensin II receptor antagonists. *Am. J. Hypertens.* 1992, *5*, 209S–220S.
98. Dunn, P. M. Oliver Wendell Holmes (1809–1894) and his essay on puerperal fever. *Arch. Dis. Child Fetal Neonatal.* 2007, *92*, F325–F327.
99. Dunn, P. M. Ignac Semmelweis (1818–1865) of Budapest and the prevention of puerperal fever. *Arch. Dis. Child Fetal Neonatal.* 2005, *90*, F345–F348.
100. Abraham, G. E. The Safe and Effective Implementation of Orthoiodosupplementation In Medical Practice. *The Original Internist* 2004, *3*, 17–36.
101. Runge, F. F. Üeber einige Producte den Steinkohlendestillation. *Ann. Phys. Chem.* 1834, *31*, 65–77. *31*, 513–524, *32*, 308–332.
102. Newsom, S. W. Pioneers in infection control – Joseph Lister. *J. Hosp. Infect.* 2003, *55*, 246–253.
103. Opitz, B. Die Problemgeschichte der Giftstoff-Antitoxin Theorie in der medizinischen Mikrobiologie. *Z. Gesamte Hyg.* 1990, *36*, 378–380.
104. Chast, F. Les colorants, outils indispensables de la révolution biologique et thérapeutique du 19 siècle. *Rev. Hist. Pharm. (Paris)* 2005, *348*, 487–504.
105. Ehrlich, P. Beitrag zur Kenntnis der Anilinfärbungen und ihrer Verwendung in der mikroskopischen Technik. *Arch. mikr. Anat.* 1877, *13*, 263–277.
106. Ehrlich, P. Ueber das Methyleneblau und seine klinisch-bakterioskopische Verwerthung. *Ztschr. F. Klin. Med.* 1881, *II*, 710–713.
107. Ehrlich, P. Ueber Laboratoriumsversuche und Klinische Erprobung von Heistoffen. *Chem. Ztg.* 1912, *36*, 637–638.
108. Parascandola, J. Alkaloids to arsenicals: systematic drug discovery before the first world war. *Publ. Am. Inst. Hist. Pharm.* 1997, *16*, 77–91.
109. The American Chemical Society. *The Pharmaceutical Century. Ten decades of drug discoveries*. The American Chemical Society: Washington, DC. ACS Web Site, 2000.
110. Van Epps, H. L. Rene Dubos: unearthing antibiotics. *J. Exp. Med.* 2006, *203*, 259.
111. Fleming, A. On the antibacterial action of cultures of a penicillium with special reference to their use in the isolation of *B. influenzae*. *Br. J. Exp. Pathol.* 1929, *10*, 226–236.
112. Chain, E., Florey, H. W., Gardner, A. D., Heatley, N. G., Jennings, M. A., Ewing, J. O., Sanders, A. G. Penicillin as a chemotherapeutic agent. *Lancet* 1940, *2*, 226–228.
113. Wainwright, M. *Miracle Cure: The Story of Penicillin and the Golden Age of Antibiotics*. Basil Blackwell: Oxford, 1990.
114. Levy, S. B. *The Antibiotic Paradox*. Plenum Press: New York, 1992.
115. Henderson, J. W. The yellow brick road to penicillin: a story of serendipity. *Mayo Clin. Proc.* 1997, *72*, 683–687.
116. Doyle, F. P., Nayler, J. H., Smith, H., Stove, E. R. Some novel acid-stable penicillins. *Nature* 1961, *191*, 1091–1092.

117. Sheehan, J. C. The chemistry of synthetic and semisynthetic penicillins. *Ann. N. Y. Acad. Sci.* 1967, *145*, 216–223.
118. Schatz, A., Bugie, E., Waksman, S. A. Streptomycin: a substance exhibiting antibiotic activity against Gram-positive and Gram-negative bacteria. *Proc. Soc. Exp. Biol. Med.* 1944, *55*, 66–69.
119. O'Neill, W. P., Nystrom, R. F., Rinehart,, K.I., Jr., Gottlieb, D. Biosynthesis of chloramphenicol. Origin and degradation of the aromatic ring. *Biochemistry* 1973, *12*, 4775–4784.
120. Rich, M. I., Ritterhoff, R. J., Hoffmann, R. J. A fatal case of aplastic anemia following chloramphenicol (chloromycetin) therapy. *Ann. Intern. Med.* 1950, *33*, 1459–1467.
121. Finland, M. Twenty-fifth anniversary of the discovery of Aureomycin: the place of the tetracyclines in antimicrobial therapy. *Clin. Pharmacol. Ther.* 1974, *15*, 3–8.
122. Schach von Wittenau, M., Blackwood, R. K., Conover, L. H., Glauert, R. H., Woodward, R. B. The stereochemistry at C-5 in oxytetracycline. *J. Am. Chem. Soc.* 1965, *87*, 134–135.
123. Morimoto, S., Takahashi, Y., Watanabe, Y., Omura, S. Chemical modification of erythromycins. I. Synthesis and antibacterial activity of 6-*O*-methylerythromycins A. *J. Antibiot. (Tokyo)* 1984, *37*, 187–189.
124. Anderson, R. C. G. R., Higgins, H. M., Pettinga, C. D. Symposium: how a drug is born. *Cincinnati J. Med.* 1961, *42*, 49–60.
125. Geraci, J. E., Heilman, F. R., Nichols, D. R., Ross, G. T., Wellman, W. E. Some laboratory and clinical experiences with a new antibiotic, vancomycin. *Mayo Clin. Proc.* 1956, *31*, 564–582.
126. Levine, D. P. Vancomycin: a history. *Clin. Infect. Dis.* 2006, *42*, S5–S12.
127. Lesher, G. Y., Froelich, E. D., Gruet, M. D., Bailey, J. H., Brundage, R. P. 1, 8-Naphthyridine derivatives. A new class of chemotherapeutic agents. *J. Med. Pharm. Chem.* 1962, *5*, 1063–1068.
128. Ball, P. The quinolones: history and overview. In *The Quinolones* (Andriole, V. T., Ed.), 2nd Edition. Academic Press: London, 1998, pp. 1–28.
129. McDevitt, D., Rosenberg, M. Exploiting genomics to discover new antibiotics. *Trends Microbiol.* 2001, *9*, 611–617.
130. Woods, C. R. Antimicrobial resistance: mechanisms and strategies. *Paediatr. Respir. Rev.* 2006, *7*(Suppl. 1), S128–S129.
131. Barre-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W., Montagnier, L. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 1983, *220*, 868–871.
132. Mitsuya, H., Broder, S. Strategies for antiviral therapy in AIDS. *Nature* 1987, *325*, 773–778.
133. Horwitz, J. P., Chua, J., Curby, R. J., Tomson, A. J., Darooge, M. A., Fisher, B. E., Mauricio, J., Klundt, I. Substrates for cytochemical demonstration of enzyme activity. I. Some substituted 3-indolyl-beta-D-glycopyranosides. *J. Med. Chem.* 1964, *7*, 574–575.
134. Mitsuya, H., Weinhold, K. J., Furman, P. A., St. Clair, M. H., Lehrman, S. N., Gallo, R. C., Bolognesi, D., Barry, D. W. S. Broder 3'-Azido-3'-deoxythymidine (BW A509U): An antiviral agent that inhibits the infectivity and cytopathic effect of human T-lymphotropic virus type III/lymphadenopathy associated virus *in vitro*, *Proc. Natl. Acad. Sci.* 1985, *82*, 7096–7100.
135. Yarchoan, R., Broder, S. Development of antiretroviral therapy for the acquired immunodeficiency syndrome and related disorders. A progress report. *New Engl. J. Med.* 1987, *316*, 557–564.
136. Yarchoan, R., Klecker, R. W., Weinhold, K. J., Markham, P. D., Lyerly, H. K., Durack, D. T., Gelmann, E., Lehrman, S. N., Blum, R. M., Barry, D. W. et al Administration of 3′-azido-3′-deoxythymidine, an inhibitor of HTLV-III/LAV replication, to patients with AIDS or AIDS-related complex. *Lancet* 1986, *8481*, 575–580.
137. Delta Coordinating Committee. Delta: a randomised double-blind controlled trial comparing combinations of zidovudine plus didanosine or zalcitabine with zidovudine alone in HIV-infected individuals. *Lancet* 1996, *348*, 283–291.
138. Kempf, D. J., Marsh, K. C., Denissen, J. F., McDonald, E., Vasavanonda, S., Flentge, C. A., Green, B. E., Fino, L., Park, C. H., Kong, X. P., Wideburg, N. E., Saldivar, A., Ruiz, L., Kati, W. M., Sham, H. L., Robins, T., Stewart, K. D., Hsu, A., Plattner, J. J., Leonard, J. M., Norbeck, D. W. ABT-538 is a potent inhibitor of human immunodeficiency virus protease and has high oral bioavailability in humans. *Proc. Natl. Acad. Sci. USA* 1995, *92*, 2484–2488.
139. Qazi, N. A., Morlese, J. F., Pozniak, A. L. Lopinavir/ritonavir (ABT-378/r). *Expert. Opin. Pharmacother.* 2002, *3*, 315–327.
140. Richman, D. D. HIV chemotherapy. *Nature* 2001, *410*(6831), 995–1001.
141. Broder, S., Mitsuya, H., Yarchoan, R., Pavlakis, G. N. NIH conference. Antiretroviral therapy in AIDS. *Ann. Intern. Med.* 1990, *113*(8), 604–618.
142. Moore, J. P. AIDS vaccines: on the trail of two trials. *Nature* 2002, *415*(6870), 365–366.
143. Farber, S., Diamond, L. K., Mercer, R. D., Sylvester, R. F., Wolff, J. A. Temporary remissions in acute leukemia in children produced by folic acid antagonist, 4-aminopteroyl-glutamic acid (aminopterin). *New Engl. J. Med.* 1948, *238*, 787.
144. Elion, G. B., Hitchings, G. H. Antagonists of nucleic acid derivatives. IV. Reversal studies with 2-aminopurine and 2, 6-diaminopurine. *J. Biol. Chem.* 1950, *187*, 511–522.
145. Gilman, A., Philips, F. S. The biological actions and therapeutic applications of the chlorethylamines and sulfides. *Science* 1946, *103*, 409–415.
146. Jukes, T. H. The history of methotrexate. *Cutis* 1978, *21*, 396–398.
147. Calne, R. Y., Alexander, G. P. J., Murray, J. E. A study of drugs in prolonging survival of homologous renal transplant in dogs. *Ann. N. Y. Acad. Sci.* 1962, *99*, 743–761.
148. Vacca, A., Martinotti, S., Screpanti, I., Maroder, M., Felli, M. P., Farina, A. R., Gismondi, A., Santoni, A., Frati, L., Gulino, A. Transcriptional regulation of the interleukin-2 gene by glucocorticoid hormones. Role of steroid receptor and antigen-responsive 5'-flanking sequences,. *J. Biol. Chem.* 1990, *265*, 8075–8080.
149. Helderman, J. H., Gailiunas,, P., Jr. *Transplant. Am. J. Kidney Dis.* 1983, *3*, 194–198.
150. Stähelin, H. The history of cyclosporin A (Sandimmune) revisited: another point of view. *Experientia* 1996, *52*, 5–13.
151. Heusler, K., Pletscher, A. The controversial early history of cyclosporine. *Swiss Med. Wkly.* 2001, *131*, 299–302.
152. Borel, J. F., Kis, Z. L. The discovery and development of cyclosporine (Sandimmune). *Transplant. Proc.* 1991, *23*, 1867–1874.
153. Borel, J. F., Feurer, C., Gubler, H. U., Stähelin, H. Biological effects of cyclosporin A: a new antilymphocytic agent. *Agents Actions* 1976, *6*, 63–83.
154. Taylor, D. O., Ensley, R. D., Olsen, S. L., Dunn, D., Renlund, D. G. Mycophenolate mofetil (RS-61443): preclinical, clinical, and three-year experience in heart transplantation. *J. Heart Lung Transplant.* 1994, *13*, 571–582.
155. Masri, M. A. The mosaic of immunosuppressive drugs. *Mol. Immunol.* 2003, *39*, 1073–1077.
156. Goodman, L. S., Wintrobe, M. M., Dameshek, W., Goodman, M. J., Gilman, A., McLennan, M. T. Nitrogen Mustard Therapy. Use of methyl-bis (beta-chlorethylamine hydrochloride and tris (beta-chloroethyl)amine hydrochloride for Hodgkin's disease, lymphosarcoma, leukemia and certain allied miscellaneous disorders. *J. Am. Med. Ass.* 1946, *132*, 126–132.
157. De Vita, V. T., Serpick, A. A., Carbone, P. P. Combination chemotherapy in the treatment of advanced Hodgkin's disease. *Ann. Intern. Med.* 1970, *73*, 881–895.
158. Brock, N. The history of the oxazaphosphorine cytostatics. *Cancer* 1996, *78*, 542–547.
159. Abelson, P. H. Medicine from plants. *Science* 1990, *247*, 513.

160. Tunon, H., Olavsdotter, C., Bohlin, L. Evaluation of anti-inflammatory activity of some Swedish medicinal plants. Inhibition of prostaglandin biosynthesis and PAF-induced exocytosis. *J. Ethnopharmacol.* 1995, *48*, 61–76.
161. Dewick, P. M. Tumor inhibitors from plants. In *Trease and Evans' Pharmacognosy* (Evans, W. C., Ed.), 14th Edition. Saunders Company: London, 1996.
162. Noble, R. L. The discovery of the vinca alkaloids – chemotherapeutic agents against cancer. *Biochem. Cell Biol.* 1990, *68b*, 1b344–1b351.
163. Stahelin, H. F., Von Wartburg, A. The chemical and biological route from podophyllotoxin glucoside to etoposide: 9th Cain Memorial Award Lecture. *Cancer Res.* 1991, *51*, 5–15.
164. Kouadio, K., Chenieux, J. C., Rideau, M., Viel, C. Antitumor alkaloids in callus cultures of *Ochrosia elliptica*. *J. Nat. Prod.* 1984, *47*, 872–874.
165. Wall, M. E., Wani, M. C. Camptothecin. Discovery to clinic. *Ann. N. Y. Acad. Sci.* 1996, *13*(803), 1–12.
166. Kantarjian, H. M., Talpaz, M., Santini, V., Murgo, A., Cheson, B., O'Brien, S. M. Homoharringtonine: history, current research, and future direction. *Cancer* 2001, *92*, 1591–1605.
167. Langreth, R. Whom do you trust?. *Science* 1991, *252*, 1780.
168. Wani, M. C., Taylor, H.l., Wall, M. E., Coggon, P., McPhail, A. T. Plant antitumor agents. VI. The isolation and structure of taxol, a novel antileukemic and antitumor agent from *Taxus brevifolia*. *J. Am. Chem. Soc.* 1971, *93*, 2325–2327.
169. Schiff, P. B., Fant, J., Horwitz, S. B. Promotion of microtubule assembly *in vitro* by taxol. *Nature* 1979, *277*, 665–667.
170. Defuria, M. D., Horovitz, Z. Taxol commercial supply strategy. *J. Natl. Cancer Inst. Monogr.* 1993, *15*, 195–198.
171. Chast, F. Obituary: Pierre Potier (1934–2006). *Nature* 2006, *440*, 291.
172. Gueritte-Voegelein, F., Guenard, D., Lavelle, F., Le Goff, M. T., Mangatal, L., Potier, P. Relationships between the structure of taxol analogues and their antimitotic activity. *J. Med. Chem.* 1991, *34*, 992–998.
173. Guenard, D., Gueritte-Voegelein, F., Dubois, J., Potier, P. Structure–activity relationships of Taxol and Taxotere analogues. *J. Natl. Cancer Inst. Monogr.* 1993, *15*, 79–82.
174. Von Hoff, D. D. The taxoids: same roots, different drugs. *Semin. Oncol.* 1997, *24*, S13–S13. S13–10.
175. Aarbakke, J., Janka-Schaub, G., Elion, G. B. Thiopurine biology and pharmacology. *Trends Pharmacol. Sci.* 1997, *18*, 3–7.
176. Medzihradsky, J. L., Hollowell, R. P., Elion, G. B. Differential inhibition by azathioprine and 6-mercaptopurine of specific suppressor T cell generation in mice. *J. Immunopharmacol.* 1981, *3*, 1–16.
177. Bowden, M. E. *Pharmaceutical Achievers*. Chemical Heritage Foundation: Philadelphia, PA, 2001.
178. Di Marco, A., Cassinelli, G., Arcamone, F. The discovery of daunorubicin. *Cancer Treat. Rep.* 1981, *65*(Suppl. 4), 3–8.
179. Weiss, R. B., Sarosy, G., Clagett-Carr, K., Russo, M., Leyland-Jones, B. Anthracycline analogs: the past, present, and future. *Cancer Chemother. Pharmacol.* 1986, *18*, 185.
180. Zunino, F., Gambetta, R., Di Marco, A., Zaccara, A., Luoni, G. A comparison of the effects of daunomycin and adriamycin on various DNA polymerases. *Cancer Res.* 1975, *35*, 754–760.
181. Weiss, R. B. The anthracyclines: will we ever find a better doxorubicin?. *Semin. Oncol.* 1992, *19*(6), 670–686.
182. Rosenberg, B., Van Camp, L., Krigas, T. Inhibition of cell division in *Escherichia coli* by electrolysis products from a platinum electrode. *Nature* 1965, *205*, 698–699.
183. Bellon, S. F., Coleman, J. H., Lippard, S. J. DNA unwinding produced by site-specific intrastrand cross-links of the antitumor drug cis-diammine-dichloroplatinum (II). *Biochemistry* 1991, *30*, 8026–8035.
184. Israel, A., Bosl, G. J., Golbey, R. B., Whitmore, W., Jr., Martini, N. The results of chemotherapy for extragonadal germ-cell tumors in the cisplatin era: the Memorial Sloan-Kettering Cancer Center experience (1975 to 1982). *J. Clin. Oncol.* 1985, *3*, 1073–1078.
185. Bruckner, H. W., Ratner, L. H., Cohen, C. J., Wallach, R., Kabakow, B., Greenspan, E. M., Holland, J. F. Combination chemotherapy for ovarian carcinoma with cyclophosphamide, adriamycin and cis-dichlorodiammineplatinum (II) after failure of initial chemotherapy. *Cancer Treat. Rep.* 1978, *62*, 1021–1023.
186. Harrap, K. R. Preclinical studies identifying carboplatin as a viable cisplatin alternative. *Cancer Treat. Rev.* 1985, *12*(Suppl. A), 21–33.
187. Knox, R., Friedlos, F., Lydall, D., Roberts, J. Mechanism of cytotoxicity of anticancer platinum drugs: evidence that cis-diamminedichloroplatinum (II) and cis-diammine-(1, 1-cyclobutane dicarboxylato)platinum(II) differ only in the kinetics of their interaction with DNA. *Cancer Res.* 1986, *46*, 1972–1979.
188. Farrell, N. Non classical platinum antitumor agents: perspectives for design and development of new drugs complementary to cisplatin. *Cancer Invest.* 1993, *11*, 578–589.
189. Broomhead, J. A., Rendina, L. M., Webster, L. K. Dinuclear complexes of platinum having anticancer properties. DNA-binding studies and biological activity of bis(4,4′-dipyrazolylmethane-*N*,*N*′)-bis[dichloroplatinum(II)] and related complexes. *J. Inorg. Biochem.* 1993, *49*, 221–234.
190. Schiffer, C. A. BCR-ABL tyrosine kinase inhibitors for chronic myelogenous leukaemia. *New Engl. J. Med.* 2007, *357*, 258–265.
191. Buchdunger, E., Zimmermann, J., Mett, H., Meyer, T., Müller, M., Druker, B. J., Lydon, N. B. Inhibition of the Abl protein-tyrosine kinase *in vitro* and *in vivo* by a 2-phenylaminopyrimidine derivative. *Cancer Res.* 1996, *56*, 100–104.
192. Krause, D. S., Van Etten, R. A. Tyrosine kinases as targets for cancer therapy. *New Engl. J. Med.* 2005, *353*, 172–187.
193. Moreira, I. S., Fernandes, P. A., Ramos, M. J. Vascular endothelial growth factor (VEGF) inhibition – a critical review. *Anticancer Agents Med. Chem.* 2007, *7*, 223–245.
194. Folkman, J. Tumor angiogenesis: therapeutic implications. *New Engl. J. Med.* 1971, *285*, 1182–1186.
195. Ferrara, N., Hillan, K. J., Gerber, H. P., Novotny, W. Discovery and development of Bevacizumab, an anti-VEGF antibody for treating cancer. *Nat Rev. Drug Discov.* 2004, *3*, 391–400.
196. Sebolt-Leopold, J. S. J.M. English Mechanisms of drug inhibition of signalling molecules. *Nature* 2006, *441*, 457–462.
197. Chabner, B. A., Roberts, T. G. Chemotherapy and the war on cancer. *Nat. Rev. Cancer* 2005, *5*, 65–72.
198. Banting, F. G., Best, C. H., Collip, J. B., Campbell, W. R., Fletcher, A. A. Pancreatic extracts in the treatment of diabetes mellitus. *J. Can. Med. Assoc. J.* 1922, *12*, 141–146.
199. Gerich, J. E. Novel insulins: expanding options in diabetes management. *Am. J. Med.* 2002, *113*, 308–316.
200. Dunn, C. J., Plosker, G. L., Keating, G. M., McKeage, K., Scott, L. J. Insulin glargine: an updated review of its use in the management of diabetes mellitus. *Drugs* 2003, *63*, 1743–1778.
201. McMahon, G. T., Arky, R. A. Inhaled insulin for diabetes mellitus. *New Engl. J. Med.* 2007, *356*, 497–502.
202. Kopelman, P. G., Hitman, G. A. Diabetes. Exploding type II. *Lancet* 1998, *352*(Suppl. 4), SIV5.
203. Mehnert, H., Seitz, W. Wirkungsweise und Indikationsbereich blutzuckersenkender Biguanidderivate *Munch. Med. Wochenschr* 1958, *100*, 1849–1851.
204. Potier, P., Sasaki, A., Bakala, J., Garcia-Alvarez, M. C., Franck, G., Nhiri, N., Wang, Q., Ermolenko, L., Nguefeu, Y., Calvo, F., Gautier, J. F. Nouveaux aspects du diabète. *Ann. Pharm. Fr.* 2005, *63*, 371–384.
205. Janbon, M., Chaptal, J., Vedel, A., Schaap, J. Accidents hypoglycémiques graves par un sulfamidothiodiazol (le VK 57 ou 2254 RP). *Montpellier Med.* 1942, *441*, 21–22.
206. Duhault, J., Lavielle, R. History and evolution of the concept of oral therapy in diabetes. *Diabetes Res. Clin. Pract.* 1991, *14*(Suppl. 2), S9–s13.
207. Sohda, T., Mizuno, K., Tawada, H., Sugiyama, Y., Fujita, T. Y. Kawamatsu Studies on antidiabetic agents. I. Synthesis of

5-[4-(2-methyl-2-phenylpropoxy)-benzyl]thiazolidine-2,4-dione (AL-321) and related compounds,. *Chem. Pharm. Bull. (Tokyo)* 1982, *10*, 3563–3573.
208. Imoto, H., Imamiya, E., Momose, Y., Sugiyama, Y., Kimura, H., Sohda, T. Studies on non-thiazolidinedione antidiabetic agents. 1. Discovery of novel oxyiminoacetic acid derivatives. *Chem. Pharm. Bull. (Tokyo)* 2002, *50*, 1349–1357.
209. Dornhorst, A. Insulinotropic meglitinide analogues. *Lancet* 2001, *358*, 1709–1716.
210. Chang, F., Jaber, L. A., Berlie, H. D., O'Connell, M. B. Evolution of peroxisome proliferator-activated receptor agonists. *Ann. Pharmacother.* 2007, *41*, 973–983.
211. Levetan, C. Oral antidiabetic agents in type 2 diabetes. *Curr. Med. Res. Opin.* 2007, *23*, 945–952.
212. Goke, R., Fehmann, H. C., Linn, T., Schmidt, H., Krause, M., Eng, J., Goke, B. Exendin-4 is a high potency agonist and truncated exendin-(9–39)-amide an antagonist at the glucagon-like peptide 1-(7–36)-amide receptor of insulin-secreting beta-cells. *J. Biol. Chem.* 1993, *268*, 19650–19655.
213. Addison, T. *On the constitutional and local effects of disease of the suprarenal capsules*. S. Highley: London, 1855.
214. Mason, H. L., Myers, C. S., Kendall, E. C. The chemistry of crystalline substances isolated from the suprarenal gland. *J. Biol. Chem.* 1936, *114*, 613–631.
215. Reichstein, T. Über Bestandteile der Nebennieren-Rinde. VI.Trennungsmethoden sowie Isolierung der Substanzen Fa, H, und J.. *Helv. Chim. Acta.* 1936, *19*, 1107–1126.
216. Hench, P. S., Kendall, E. C., Slocumb, C. H., Polley, H. F. The effect of a hormone of the adrenal cortex (17-hydroxy-11-dehydrocorticosterone: compound E) and of pituitary adrenocorticotropic hormone on rheumatoid arthritis. *Proc. Mayo Clin.* 1949, *24*, 181–197.
217. O'Byrne, P. M., Parameswaran, K. Pharmacological management of mild or moderate persistent asthma. *Lancet* 2006, *368*, 794–803.
218. Medvei, V. C. *The History of Clinical Endocrinology*. The Parthenon Publishing Group: Lancashire and New York, 1993.
219. Allen, E., Doisy, E. A. The menstrual cycle of the monkey, *Macacus rhesus*: Observations on normal animals, the effects of removal of the ovaries and the effects of injection of ovarian and placental extracts into the spayed animals. *Contr. Embryol. Carneg. Instn.* 1927, *19*, 1–44.
220. Butenandt, A. F. J. Ueber die chemische Untersuchung der Sexualhormone. *Z. Angew. Chem.* 1931, *44*, 905–908.
221. Browne, J. S. L. Chemical and physiological properties of crystalline oestrogenic hormones. *Canad. J. Res.* 1933, *8*, 180–197.
222. Ruzicka, L., Wettstein, A. Ueber die kunstliche Herstellung des Testikelhormons Testosteron. (Androsten-3-on-17-ol). *Helv. Chim. Acta* 1935, *18*, 1264–1268.
223. Chast, F. Histoire de la contraception. *Pour la Science* 1997, *233*, 51–59.
224. Pincus, G. G., Chang, M. C. The effects of progesterone and related compounds on ovulation early development in the rabbit. *Acta Physiol. Latino-Am.* 1953, *3*, 177–183.
225. Djerassi, C. Steroid research at Syntex: "the pill" and cortisone. *Steroids* 1992, *57*, 631–641.
226. Djerassi, C. The mother of the pill. *Recent Prog. Horm. Res.* 1995, *50*, 1–17.
227. Colton, F. B. Steroids and "the pill": early steroid research at Searle. *Steroids* 1992, *57*, 624–630.
228. Black, J. W., Duncan, W. A., Durant, C. J., Ganellin, C. R., Parsons, E. M. Definition and antagonism of histamine H2-receptors. *Nature* 1972, *236*, 385–390.
229. Molinder, H. K. The development of cimetidine: 1964–1976. A human story. *J Clin. Gastroenterol.* 1994, *19*, 248–254.
230. Sachs, G., Collier, R. H., Schoemaker, R. L., Hirschowitz, B. I. The energy source of gastric acid secretion. *Biochim. Biophys. Acta* 1968, *162*, 210–219.
231. Pascaud, X. B., Errard, D. J., Blouin, M. M. The gastric antisecretory activity of pyridyl-2-thioacetamide (CMN 131). *Am. J. Dig. Dis.* 1974, *19*, 503–515.
232. Brandstrom, A., Lindberg, P., Junggren, U. Structure activity relationships of substituted benzimidazoles. *Scand. J. Gastroenterol. Suppl.* 1985, *108*, 15–22.
233. Fellenius, E., Berglindh, T., Sachs, G., Olbe, L., Elander, B., Sjostrand, S. E., Wallmark, B. Substituted benzimidazoles inhibit gastric acid secretion by blocking H^+, K^+-ATPase. *Nature* 1981, *290*, 159–161.
234. Lind, T., Cedeberg, C., Ekenved, G., Haglund, U., Olbe, L. Effect of omeprazole – a gastric proton pump inhibitor – on pentagastrin stimulated acid secretion in man. *Gut* 1983, *24*, 270–276.
235. Marshall, B. J., Warren, J. R. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet* 1984, *8390*, 1311–1315.
236. Konturek, S. J., Konturek, P. C., Brzozowski, T., Konturek, J. W., Pawlik, W. W. From nerves and hormones to bacteria in the stomach; Nobel Prize for achievements in gastrology during last century. *J. Physiol. Pharmacol.* 2005, *56*, 507–530.
237. Axon, A. T., O'Morain, C. A., Bardhan, K. D., Crowe, J. P., Beattie, A. D., Thompson, R. P., Smith, P. M., Hollanders, F. D., Baron, J. H., Lynch, D. A., Dixon, M. F., Tompkins, D. S., Birrell, H., Gillon, K. R. Randomised double blind controlled study of recurrence of gastric ulcer after treatment for eradication of *Helicobacter pylori* infection. *Br. Med. J.* 1997, *314*, 565–568.
238. Virchow, R. Cellular pathology. As based upon physiological and pathological histology. Lecture XVI – Atheromatous affection of arteries. *Nutr Rev (1989)* 1858, *47*, 23–25.
239. Olson, R. E. Discovery of the lipoproteins, their role in fat transport and their significance as risk factors. *J. Nutr.* 1998, *128*, 439S–443S.
240. Eghdamian, B., Ghose, K. Mode of action and adverse effects of lipid lowering drugs. *Drugs Today* 1998, *34*, 943–956.
241. Knappe, J., Ringelmann, E., Lynen, F. Über die β-Hydroxy-β-Methylglutaryl Reduktase der Hefe zur Biosynthese der Terpen. XI. *Biochem. Z.* 1959, *332*, 195–213.
242. Siperstein, M. D., Fagan, V. M. Feedback control of mevalonate synthesis by dietary cholesterol. *J. Biol. Chem.* 1966, *241*, 602–609.
243. Endo, A. The discovery and development of HMG-CoA reductase inhibitors. *J. Lipid Res.* 1992, 33, 1569–1582.
244. Brown, M. S., Goldstein, J. L. Lowering plasma cholesterol by raising LDL receptors. *New Engl. J. Med.* 1981, *305*, 515–517.
245. Williams, O., Jacks, A. M., Davis, J., Martinez, S. Case 10: Merck(A): Mevacor*. In *Innovation Management – Strategies, Implementation, and Profits* (Allan, Afuah, Ed.). Oxford University Press: New York, 1998.
246. Roth, B. D. The discovery and development of atorvastatin, a potent novel hypolipidemic agent. *Prog. Med. Chem.* 2002, *40*, 1–22.
247. Langley, J. N. On the reaction of cells and of nerve-endings to certain poisons, chiefly as regards the reaction of striated muscle to nicotine and to curari. *J. Physiol.* 1905, *33*, 374–413.
248. Clark, A. J. *The mode of action of drugs on cells*. Edward Arnold & Co: London, 1933.
249. Ariëns, E. J. Affinity and intrinsic activity in the theory of competitive inhibition. *Arch. Int. Pharma. Thera* 1954, *99*, 32–49.
250. Stephenson, R. P. A modification of receptor theory. *Br. J. Pharmacol. Chemother.* 1956, *11*, 379–393.
251. Parascandola, J. Origins of receptor theory. *Trends Pharmacol. Sci.* 1980, *1*, 189–192.
252. Dale, H. H. On some physiological actions of ergot. *J. Physiol.* 1906, *34*, 163–206.
253. Quirke, V. Putting theory into practice: James Black, Receptor theory and the development of the beta-blockers at ICI, 1958–1978. *Med. Hist.* 2006, *50*, 69–92.
254. Kebabian, J. W., Calne, D. B. Multiple receptors for dopamine. *Nature* 1979, *277*, 93–96.

255. Seeman, P. Dopamine receptor sequences. Therapeutic levels of neuroleptics occupy D2 receptors, clozapine occupies D4. *Neuropsychopharmacology* 1992, *7*, 261–284.
256. Rapport, M. M., Green, A. A., Page, I. H. Serum vasoconstrictor (serotonin). IV. Isolation and characterization. *J. Biol. Chem.* 1948, *176*, 1243–1251.
257. Hoyer, D., Hannon, J. P., Martin, G. R. Molecular, pharmacological and functional diversity of 5-HT receptors. *Pharmacol. Biochem. Behav.* 2002, *71*, 533–554.
258. Hesketh, P. J., Murphy, W. K., Lester, E. P., Gandara, D. R., Khojasteh, A., Tapazoglou, E., Sartiano, G. P., White, D. R., Werner, K., Chubb, J. M. GR 38032F (GR-C507/75): a novel compound effective in the prevention of acute cisplatin-induced emesis. *J. Clin. Oncol.* 1989, *7*, 700–705.
259. Sberna, G., Saez-Valero, J., Beyreuther, K., Masters, C. L., Small, D. H. The amyloid beta-protein of Alzheimer's disease increases acetylcholinesterase expression by increasing intracellular calcium in embryonal carcinoma P19 cells. *J. Neurochem.* 1997, *69*, 1177–1184.
260. Summers, W. K., Majovski, L. V., Marsh, G. M., Tachiki, K., Kling, A. Oral tetrahydroaminoacridine in long-term treatment of senile dementia, Alzheimer type. *New Engl. J. Med.* 1986, *315*, 1241–1245.
261. Summers, W. K. Tacrine, and Alzheimer's treatments. *J. Alzheimer's Dis.* 2006, *9*(Suppl. 3), 439–445.
262. Farlow, M. R. Update on rivastigmine. *Neurologist* 2003, *9*, 230–234.
263. Seltzer, B. Donepezil: an update. *Expert Opin. Pharmacother.* 2007, *8*, 1011–1023.
264. Brodaty, H., Woodward, M., Boundy, K., Barnes, N., Allen, G. Nature investigators a naturalistic study of galantamine for Alzheimer's disease. *CNS Drugs* 2006, *20*, 935–943.
265. Farlow, M. R., Cummings, J. L. Effective pharmacologic management of Alzheimer's disease. *Am. J. Med.* 2007, *120*, 388–397.
266. Birks, J. Cholinesterase inhibitors for Alzheimer's disease. *Cochrane Database Syst. Rev.* 2006, *25*, CD005593.
267. Sadowski, M., Wisniewski, T. Disease modifying approaches for Alzheimer's pathology. *Curr. Pharm. Des.* 2007, *13*, 1943–1954.
268. Mechoulam, R., Gaoni, Y. Hashish. IV. The isolation and structure of cannabinolic cannabidiolic and cannabigerolic acids. *Tetrahedron* 1965, *21*, 1223–1229.
269. Devane, W. A., Hanus, L., Breuer, A., Pertwee, R. G., Stevenson, L. A., Griffin, G., Gibson, D., Mandelbaum, A., Etinger, A., Mechoulam, R. Isolation and structure of a brain constituent that binds to the cannabinoid receptor. *Science* 1992, *258*, 1946–1949.
270. Axelrod, J., Felder, C. C. Cannabinoid receptors and their endogenous agonist, anandamide. *Neurochem. Res.* 1998, *23*, 575–581.
271. Högestätt, E. D., Jönsson, B. A., Ermund, A., Andersson, D. A., Björk, H., Alexander, J. P., Cravatt, B. F., Basbaum, A. I., Zygmunt, P. M. Conversion of acetaminophen to the bioactive *N*-acylphenolamine AM 404 via fatty acid amide hydrolase-dependent arachidonic acid conjugation in the nervous system. *J. Biol. Chem.* 2005, *280*, 31405–31412.
272. Giuffrida, A., Beltramo, M., Piomelli, D. Mechanisms of endocannabinoid inactivation: biochemistry and pharmacology. *J. Pharmacol. Exp. Ther.* 2001, *298*, 7–14.
273. Gelfand, E. V., Cannon, C. P. Rimonabant: a cannabinoid receptor type 1 blocker for management of multiple cardiometabolic risk factors. *J. Am. Coll. Cardiol.* 2006, *47*, 1919–1926.
274. Scheen, A. J., Finer, N., Hollander, P., Jensen, M. D., Van Gaal, L. F. Efficacy and tolerability of rimonabant in overweight or obese patients with type 2 diabetes: a randomised controlled study. *Lancet* 2006, *368*, 166–172. the Rio-Diabetes Study Group.
275. Despres, J. P., Golay, A., Sjöström, L. Effects of rimonabant on metabolic risk factors in overweight patients with dyslipidemia. *New Engl. J. Med.* 2005, *353*, 2121–2134. and the Rio-lipids Study Group
276. Gary-Bobo, M., Elachouri, G., Gallas, J. F., Janiak, P., Marini, P., Ravinet-Trillou, C., Chabbert, M., Cruccioli, N., Pfersdorff, C., Roque, C., Arnone, M., Croci, T., Soubrie, P., Oury-Donat, F., Maffrand, J. P., Scatton, B., Lacheretz, F., Le Fur, G., Herbert, J. M., Bensaid, M. Rimonabant reduces obesity-associated hepatic steatosis and features of metabolic syndrome in obese Zucker fa/fa rats. *Hepatology* 2007, *46*, 122–129.
277. Di Marzo, V., Bifulco, M., De Petrocellis, L. The endocannabinoid system and its therapeutic exploitation. *Nat. Rev. Drug Discov.* 2004, *3*, 771–784.
278. Persidis, A., Copen, R. M. Mental disorder drug discovery. *Nat. Biotechnol.* 1999, *17*, 307–309.
279. Loewi, O. Ueber humorale Uebertragbarkeit der Herznervenwirkung. *Pflüg. Arch. Ges. Physio*l. 1921, 1922, 1924, *189*, 239–242; *193*, 201–213; *203*, 408–412.
280. Triggle, D. J. Pharmacological receptors: a century of discovery – and more. *Pharm. Acta Helv.* 2000, *74*, 79–84.
281. Grof, P. Mogens Schou (1918–2005). *Neuropsychopharmacology* 2006, *31*, 891–892.
282. Burrows, G. D., Tiller, J. W. Cade's observation of the antimanic effect of lithium and early Australian research. *Aust. NZ. J. Psychiatry* 1999, *33*(Suppl. 1), S27–31.
283. Kozikowski, A. P., Gaisina, I. N., Yuan, H., Petukhov, P. A., Blond, S. Y., Fedolak, A., Caldarone, B., McGonigle, P. Structure based design leads to the Identification of lithium mimetics that block mania-like effects in rodents. Possible new GSK-3β therapies for bipolar disorders. *J. Am. Chem. Soc.* 2007, *129*, 8328–8332.
284. Staub, A. M. Recherches sur quelques bases synthétiques antagonistes de l'histamine. *Ann. Inst. Pasteur* 1939, *63*, 400–436.
285. Hamon, J., Paraire, J., Velluz, J. Remarques sur l'action du 4560 RP sur l'agitation maniaque. *Ann. Med. Psychol.* 1952, *10*, 332–335.
286. Domino, E. F. History of modern psychopharmacology: a personal view with an emphasis on antidepressants. *Psychosom. Med.* 1999, *61*, 591–598.
287. Chast, F. Hommage à Paul Janssen: deux découvertes par an (1926–2003). *Ann. Pharm. Fr.* 2004, *62*, 274–283.
288. Berger, F. M. Pharmacological properties of 2-methyl-2-propyl-l,3-propanediol dicarbamate (Miltown), a new interneural blocking agent. *J. Pharmacol. Exp. Ther.* 1954, *112*, 413–423.
289. Sternbach, L. H. The discovery of librium. *Agents Actions* 1972, *2*(4), 193–196.
290. Whitwam, J. G. Flumazenil and midazolam in anaesthesia. *Acta Anaesthesiol. Scand. Suppl.* 1995, *108*, 15–22.
291. Kuhn, R. The treatment of depressive states with G22355 (imipramine hydrochloride). *Am. J. Psychiatry* 1958, *115*, 459–464.
292. Iversen, L. L. The discovery of monoamine transporters and their role in CNS drug discovery. *Brain Res. Bull.* 1999, *50*, 379.
293. Carlsson, A., Fuxe, K., Ungerstedt, U. The effect of imipramine on central 5-hydroxytryptamine neurons. *J. Pharm. Pharmacol.* 1968, *20*, 150–151.
294. Wong, D. T., Perry, K. W., Bymaster, F. P. The discovery of fluoxetine hydrochloride (Prozac). *Nat. Rev. Drug Discov.* 2005, *4*, 764–774.
295. Shorter, E. *A History of Psychiatry: From the Era of the Asylum to the Age of Prozac*. John Wiley & Sons: New York, p.448, 1998.
296. Wells, J. G. *A History of the Discovery of the Applications of Nitrous Oxide Gas, Ether and Other Vapours to Surgical Operations*. Gaylord Perry: Hartford, CT, 1847.
297. Morton, W. T. G. *Remarks on the Proper Mode of Administering Sulphuric Ether by Inhalation*. Dutton & Wentworth: Boston, MA, 1847.
298. Bigelow, H. J. Insensibility during surgical operations produced by inhalation. *Boston Med. Surg. J.* 1846, *35*, 309–317. 379–382.
299. Snow, J. On the inhalation of chloroform and ether. With description of an apparatus. *Lancet* 1848, *1*, 177–180.
300. Baumann, E. Ueber Disulfone. *Berl. Dtsch. Chem. Ges.* 1886, *19*, 2806–2814.
301. Bernard, C. Cours de Médecine du Collège de France. Première leçon (29 Février 1856). In *Leçons sur les effets des substances toxiques et médicamenteuses*. J.-B. Baillière: Paris, 1857.

302. Grifith, H. R., Johnson, E. The use of curare in general anaesthesia. *Anesthesiology* 1942, *3*, 418–420.
303. Gray, T. C. J. Halton Technique for the use of d-tubocurarine chloride with balanced anaesthesia. *Br. Med. J.* 1946, *2*, 293–295.
304. Koller, C. Vorlaüfige Mittheilung über locale Anästhesirung am Auge. *Klin. Mbl. Augenheilk* 1884, *22*, 60–63. Beilageheft
305. Niemann, A. *Ueber eine neue organische Base in den Cocablättern*. E.A., Huth: Göttingen, 1860.
306. Bier, A. K. G. Versuche über Cocainisirung des Rückenmarkes. *Dtsch. Z. Chir.* 1899, *51*, 361–369.
307. Musto, D. F. Cocaine's history, especially the American experience. *Ciba Found. Symp.* 1992, *166*, 7–14.
308. Einhorn, A. Ueber die Chemie der localen Anesthaestica. *Munch. Med. Wschr.* 1899, *46*, 1218–1220. 1254–1256
309. Filehne, W. Die local Anästhesirende Wirkung von Benzoylderivaten. *Berliner Klinische Wochenschrift* 1887, *24*, 107–108.
310. Fourneau, E. Stovaïne, anesthésique local. *Bull. Soc. Pharmacol.* 1904, *10*, 141–148.
311. Dreifuss, F. E. Other antiepileptic drugs. In *Antiepileptic Drugs* (Levy, R. H., Mattson, R. H., Meldrum, B. S., Eds), 4th Edition. Raven Press: NewYork, 1995.
312. Fischer, E., Von Mering, J. Üeber eine neue Klasse von Schlafmitteln. *Ther. Gegenw.* 1903, *44*, 97–101.
313. Hauptmann, A. Luminal bei Epilepsie. *Münchner medizinische Wochenschrift* 1912, *59*, 1907–1909.
314. Jarman, R. History of intravenous anesthesia with six years experience in the use of Penthotal. *Postgrad. Med. J.* 1941, *17*, 70–80.
315. Glazko, A. J. Discovery of phenytoin. *Ther. Drug Monit.* 1986, *8*, 490–497.
316. Merritt, H. H., Putnam, T. J. Sodium diphenyl hydantoinate in the treatment of convulsive disorders. *J. Am. Med. Ass.* 1938, *111*, 1068–2016.
317. Burton, B. S. On the propyl derivatives and decomposition products of ethyl acetoacetate. *Am. Chem. J.* 1882, *3*, 385–395.
318. Meunier, H., Carraz, G., Meunier, Y., Eymard, P., Aimard, M. Propriétés pharmacodynamiques de l'acide *n*-dipropylacetique. *Therapie* 1963, *18*, 435–438.
319. McElroy, S. L., Keck, P. E., Jr. Antiepileptic drugs. In *Textbook of Psychopharmacology* (Schatzberg, A., Nemeroff, C. B., Eds), Vol. 17. American Psychiatric Press: Washington, DC, 1995, pp. 351–375.
320. Schindler, W. 5H-Dibenz[b,f]azepines. US pat. 2,948, 718 (1960 to Geigy). *Chemical Abstracts* 1960, *55*, 1671c.
321. Lambert, P. A., Carraz, G., Carbel, S. Action neuro-psychotrope d'un nouvel anti-épileptique: le dépamide. *Ann. Méd-Psycholog.* 1966, *124*, 707–710.
322. Cohen, A. F., Ashby, L., Crowley, D., Land, G., Peck, A. W., Miller, A. A. Lamotrigine (BW430C), a potential anticonvulsant. Effects on the central nervous system in comparison with phenytoin and diazepam. *Br. J. Clin. Pharmacol.* 1985, *20*, 619–629.
323. Saletu, B., Grunberger, J., Linzmayer, L. Evaluation of encephalotropic and psychotropic properties of gabapentin in man by pharmaco-EEG and psychometry. *Int. J. Clin. Pharmacol. Ther. Toxicol.* 1986, *24*, 362–373.
324. Pierce, M. W., Suzdak, P. D., Gustavson, L. E., Mengel, H. B., McKelvy, J. F., Mant, T. Tiagabine. *Epilepsy Res. Suppl.* 1991, *3*, 157–160.
325. Taylor, C. P., Angelotti, T., Fauman, E. Pharmacology and mechanism of action of pregabalin: the calcium channel alpha2-delta (alpha2-delta) subunit as a target for antiepileptic drug discovery. *Epilepsy Res.* 2007, *73c*, 137–150.
326. Raju, T. N. The Nobel chronicles. 1958: George Wells Beadle (1903–1989), Edward Lawrie Tatum (1909–1975) and Joshua Lederberg (b 1925). *Lancet* 1999, *353*, 2082.
327. Watson, J. D., Crick, F. Molecular structure of nucleic acids a structure for deoxyribose nucleic acid. *Nature* 1953, *171*, 737–738.
328. Bhushan, R., Reddy, G. P. Thin layer chromatography of dansyl and dinitrophenyl derivatives of amino acids. *Biomed. Chromatogr.* 1989, *3*, 233–240.
329. Kresge, N., Simoni, R. D., Hill, R. L. The solid phase synthesis of ribonuclease A by Robert Bruce Merrifield. *J. Biol. Chem.* 2006, *26*, e21–e23.
330. Anderson, W. F., Blaese, R. M., Culver, K. The ADA human gene therapy clinical protocol: Points to consider response with clinical protocol. *Hum. Gene Ther.* 1990, *1*, 331–362.
331. Zhang, X., Godbey, W. T. Viral vectors for gene delivery in tissue engineering. *Adv. Drug Deliv. Rev.* 2006, *7*, 515–534.
332. Porteus, M. H., Connelly, J. P., Pruett, S. M. A look to future directions in gene therapy research for monogenic diseases. *PLoS Genet.* 2006, *2*, e133.
333. Biswas, A., Hutchins, R. Embryonic stem cells. *Stem Cells Dev.* 2007, *16*, 213–222.
334. Kubinyi, H. Chance favors the prepared mind – from serendipity to rational drug design. *J. Recept. Signal Transduct. Res.* 1999, *19*, 15–39.
335. Wermuth, C. G. Selective optimization of side activities: the SOSA approach. *Drug Discov. Today* 2006, *11*, 160–164.
336. Drews, J., Ryser, S. The role of innovation in drug development. *Nat. Biotechnol.* 1997, *15*, 1318–1319.
337. Hopkins, A. L., Groom, C. R. The Druggable Genome. *Nat. Rev. Drug Discov.* 2002, *1*, 727–730.
338. Golden, J. B. Prioritizing the human genome: knowledge management for drug discovery. *Curr. Opin. Drug Discov. Dev.* 2003, *6*, 310–316.
339. Overington, J. P., Al-Lazikani, B., Hopkins, A. L. How many drug targets are there? *Nat. Rev. Drug Discov.* 2006, *5*, 993–996.
340. Clark, D. E. What has computer-aided molecular design ever done for drug discovery? *Expert Opin. Drug Discov.* 2006, *1*, 103–110.
341. Apic, G., Ignjatovic, T., Boyer, S., Russell, R. B. Illuminating drug discovery with biological pathways. *FEBS Lett.* 2005, *8*, 1872–1877.
342. Owens, J. 2006 drug approvals: finding the niche. *Nat. Rev. Drug Discov.* 2007, *6*, 99–101.

Chapter 2

Medicinal Chemistry: Definitions and Objectives, Drug Activity Phases, Drug Classification Systems

Peter Imming

I. DEFINITIONS AND OBJECTIVES
A. Medicinal chemistry and related disciplines and terms
B. Drugs and drug substances
C. Stages of drug development
II. DRUG ACTIVITY PHASES
A. The pharmaceutical phase
B. The pharmacokinetic phase
C. The pharmacodynamic phase
D. The road to successful drug development?
III. DRUG CLASSIFICATION SYSTEMS
A. Classification by target and mechanism of action
B. Other classification systems
REFERENCES

Medicinal chemistry remains a challenging science which provides profound satisfaction to its practitioners. It intrigues those of us who like to solve problems posed by nature. It verges increasingly on biochemistry and on all the physical, genetic and chemical riddles in animal physiology which bear on medicine. Medicinal chemists have a chance to participate in the fundamentals of prevention, therapy and understanding of diseases and thereby to contribute to a healthier and happier life.

A. Burger[1]

I. DEFINITIONS AND OBJECTIVES

A. Medicinal chemistry and related disciplines and terms

A definition of medicinal chemistry was given by a IUPAC specialized commission: "*Medicinal chemistry* concerns the discovery, the development, the identification and the interpretation of the mode of action of biologically active compounds at the molecular level. Emphasis is put on drugs, but the interests of the medicinal chemist are not restricted to drugs but include bioactive compounds in general. Medicinal chemistry is also concerned with the study, identification, and synthesis of the metabolic products of these drugs and related compounds."[2]

Drugs – natural and synthetic alike – are chemicals used for medicinal purposes. They interact with complex chemical systems of humans or animals. Medicinal chemistry is concerned with this interaction, focusing on the organic and biochemical reactions of drug substances with their targets. This is one aspect of drug chemistry. Other important aspects are the synthesis and the analysis of drug substances. The two latter aspects together are sometimes called *pharmaceutical chemistry*, but the synthesis of drugs is considered by some people – mainly chemists – to be part of medicinal chemistry, denoting analytical aspects as pharmaceutical chemistry. In German faculties of pharmacy, the literal translations of pharmaceutical and medicinal chemistry – Pharmazeutische and Medizinische Chemie – are used synonymically.

The general study of drugs is called *pharmacy* or *pharmacology*. A common narrower definition of pharmacology concentrates on the fate and effects of a drug in the body. *Clinical chemistry*, a different subject, is concerned with the determination of physiological and pathophysiological parameters in body fluids, for example, enzyme activities and metabolites in blood and urine. The term *biopharmacy* has

Copyright © 2008, Elsevier Ltd
All rights reserved.

been reserved for the investigation and control of absorption, distribution, metabolism, excretion and toxicology (ADMET) of drug substances.

Some further terms are more or less synonymous with medicinal chemistry: *(molecular) pharmacochemistry, drug design, selective toxicity*. The French equivalent to medicinal chemistry is "Chimie Thérapeutique" and the German ones are "Medizinische/Pharmazeutische Chemie" and "Arzneimittelforschung."

In the academia, medicinal chemistry is a major subject in most pharmacy faculties, both for undergraduates and in research, and in a growing number of chemistry faculties. In the pharmaceutical industry, medicinal chemistry is at the heart of finding new medicines.

The main activities of medicinal chemists appear clearly from the analysis of their most important scientific journals (*Journal of Medicinal Chemistry, European Jorunal of Medicinal Chemistry, Bioorganic and Medicinal Chemistry, Il Farmaco, Archiv der Pharmazie, Arzneimittelforschung, Chemical and Pharmaceutical Bulletin,* etc.).

The *objectives of medicinal chemistry* are as easily formulated as they are difficult to achieve: Find, develop and improve drug substances that cure or alleviate diseases (see below Section I.C.) and understand the causative and accompanying chemical processes (see below Section III.A.).

Medicinal chemistry is an interdisciplinary science covering a particularly wide domain situated at the interface of organic chemistry with life sciences, such as biochemistry, pharmacology, molecular biology, genetics, immunology, pharmacokinetics and toxicology on one side, and chemistry-based disciplines such as physical chemistry, crystallography, spectroscopy and computer-based techniques of simulation, data analysis and data visualization on the other side.

B. Drugs and drug substances

Drugs are composed of *drug substances* (syn. *active pharmaceutical ingredients, APIs*) and *excipients* (syn. *ancillary substances*). The combination of both is the work of pharmaceutical technology (syn. galenics) and denoted a *formulation*.

In 2007, the World Drug Index contained over 80,000 marketed and development drug substances.[3] In the United States, approximately 21,000 drug products were marketed in 2006; however, when duplicate active ingredients, salt forms, supplements, vitamins, imaging agents, etc. are removed, this number is reduced to only 1,357 unique drugs, of which 1,204 are small molecule drugs and 166 are biologicals.[4] In 2006 in Germany, approximately 8,800 drugs in 11,200 formulations contained approximately 2,400 APIs and 750 plant extracts.[5] The WHO Essential Medicines List held approximately 350 drug substances in 2007.[6]

What makes a chemical "druggable"? Because of the versatility of their molecular targets (see below), there can be no universal characteristic of drug substances. However, since the general structure of the target organisms is identical, generalizations as to drug substance structure are possible for biopharmacy.[7,8] For a chemical to be readily absorbed by the gut and distributed in the body, its size, hydrophilicity/lipophilicity ratio, stability toward acid medium and hydrolytical enzymes, etc. have to meet defined physicochemical criteria. A careful analysis of reasons for drug attrition revealed that only 5% were caused by pharmacokinetic difficulties whereas 46% were due to insufficient efficacy and 33% to adverse reactions in animals or humans.[9] Since both wanted and unwanted effects are due to the biological activity, 79% of drug candidates had unpredicted or wrongly predicted sum activities.

Predictions of toxicity from molecular features are very precarious.[10,11,12] Only rather general rules are for sure; such as avoidance of very reactive functional groups, for example, aldehyde because of oxidative instability and haptene nature; α,β-unsaturated carbonyl compounds and 2-halopyridines because of their unspecific reactivity as electrophiles. Torcetrapib was an antiatherosclerotic drug candidate promising to become a blockbuster when in latter phase III of clinical trials in 2006, an increased risk of mortality led the company to discontinue its development. It is not clear whether the effects were caused by the mechanism of action – inhibition of cholesteryl ester transfer protein – some other effect or an interaction with another drug. This is just one instance that "it isn't that simple [and] nothing's obvious and nothing's for certain" in rational drug development.[13]

C. Stages of drug development

Most drugs were discovered rather than developed.[14] That is why a large number of drug substances are natural products or derivatives thereof. It is a matter of debate if ethnic medicines or nature still hold gems as yet undiscovered by pharmacy.[15,16] Synthetic substance collections ("libraries") have been created through (automated) organic chemistry. The very high number and diversity of natural and synthetic chemical entities is faced with an equally growing number of potential reaction partners (targets) from biochemical and pathophysiological research.

In virtual, biochemical and cell-based testing, compound selections are run against an isolated or physiologically embedded target that may be involved in the disease process.[17] Compounds that exceed a certain threshold value in binding to the target or modulation of some functional signal behind it, are called *hits*. If the identity and purity of the compound and the assay result are confirmed in a multipoint activity determination, the compound raises to the status of *validated hit*. From this one hopes to develop leads. A *lead* is a compound or series of compounds with proven activity and selectivity in a screen and fulfills some drug development criteria such as originality, patentability and accessibility (by extraction or synthesis). Molecular variation

FIGURE 2.1 Example of an optimization algorithm. Source: Adapted from a presentation by Dr. U. Heiser, Probiodrug AG, Halle, Germany, reproduced with permission.

BOX 2.1 The Ideal New Drug Substance

- New chemical entity for patentability and registration.
- Maximum four-step synthesis with, for example, no heavy metal catalysts and no environmentally problematic waste; no chromatographic purification steps; purity >99%.
- Stable up to 70°C even in humid air and toward light.
- Solid-state properties (crystalline, not polymorphous, not hygroscopic) that make it a perfect partner for (tablet) compaction.
- Solubility in water sufficient for the production of stable blood-isotonic solutions.
- Oral bioavailability >90% with no interindividual variation.
- Very high activity and pharmacokinetic profile enable once-a-day-dosage at 5–10 mg.

hopefully tunes the physicochemical parameters so that it becomes suitable for ADME. An example of a small optimization algorithm is shown in Figure 2.1.

The resulting *optimized lead* (*preclinical candidate*), if it displays no toxicity in cell and animal models, becomes a *clinical candidate*. If this stands the tests of efficacy and safety in humans and overcomes marketing hurdles, a *new drug entity* will enter the treasure trove of pharmacy. The Box 2.1 will help to appreciate that activity is a necessary but not sufficient quality of medicines. There is, of course, no ideal drug in real world, but one has to find a relative optimum.

The role of medicinal chemistry in drug development is most prominent in the following three steps:

1. The discovery step, consisting of the choice of the therapeutic target (biochemical, cellular or *in vivo* model; see below) and the identification or discovery and production of new active substances interacting with the selected target.
2. The optimization step that deals with the improvement of an active compound. The optimization process takes primarily into account the increase in potency, selectivity and decrease in toxicity. Its characteristics are the establishment of structure–activity relationships, ideally based on an understanding of the molecular mode of action.
3. The development step, whose purpose is the continuation of the improvement of the pharmacokinetic properties and the fine-tuning of the pharmaceutic properties of active substances to render them suitable for clinical use. This can consist, to name a few instances, in the preparation of better-absorbed compounds, of sustained release formulations and of water-soluble derivatives or in the elimination of properties related to the patient's compliance (irritation, painful injection, undesirable organoleptic properties). For an example, see Figure 2.2.

The main tasks of medicinal chemistry in the optimization and development steps consist in the optimization of the following characteristics: (a) Higher affinity for better activity so the dosage and nonspecific side effects will be as low as possible. There are no examples of drugs that are dosed <10 mg/day that cause idiosyncratic adverse drug reactions.[18] For drug substances that have to be given in higher doses – the majority – medicinal chemistry tries to find active derivatives that will be metabolized in a safe way.[18] This includes assaying for inhibition of or reaction with key enzymes of biotransformation, for example, oxidases of the cytochrome type some of which are highly demanded by food constituents and xenobiotics including drug substances.[19] Medicinal chemistry tries to prepare drugs that are not metabolized by bottleneck enzymic pathways.[20] (b) Better selectivity may lead to a reduction of unwanted side effects. This entails that a sometimes very high number of other targets have to be assayed; for example, an antidepressive serotonin reuptake inhibitor has to be tested against all subtypes of serotonin, adrenaline and dopamine receptors at least.

FIGURE 2.2 An example of fine-tuning of pharmacologically active chemicals: Erythromycin 2′-ethylsuccinate and clarithromycin are semisynthetic derivatives of the macrolide antiinfective erythromycin. The small molecular change in the former leads to the elimination of bitterness which is important as this class of drugs is often used in pediatrics and administered as a syrup. In the latter, because hemiketal formation is no longer possible (arrow), clarithromycin is stable in the acidic milieu of the stomach (pH 2).

In spite of the high number of compounds, targets and assays, the development pipeline of new chemical entities as drug substances has not got fuller in the past 20 years; for possible explanations, see below the discussion of drug targets and Ref. [9].

II. DRUG ACTIVITY PHASES

The way of a drug into the body, to its target(s), and out again can be broken down into three mechanistically distinct phases, the second and third being partly simultaneous. During drug development, all three phases are investigated interdependently because structural changes required for one phase must not abolish suitability in another phase.

A. The pharmaceutical phase

Drug substances are applied orally (preferred mode) or parenterally (e.g. by subcutaneous or intravenous injection, rectally, by inhalation). A combination of the skills of medicinal chemists and pharmaceutical technologists has to provide the drug candidate in suitable formulations. For tablets, the drug substance needs to be crystalline and not have a low melting point; for injections, it should be water soluble, for example, as a salt. The required structural features must be compatible with the pharmacological activity, of course.

B. The pharmacokinetic phase

For this, medicinal chemists and biopharmacists work together to design a compound that will have suitable ADME parameters. Sufficient solubility in aqueous medium for absorption and blood transport has to be combined with sufficient lipophilicity for passage through cell membranes. If an active compound is too hydrophilic and at the same time contains a carboxylic acid group, for instance, conversion to a simple ester will facilitate absorption. Once in the blood, unspecific esterases will catalyse hydrolysis to the active carboxylic acid form. Such an ester is an instance of a *prodrug*.

Drug substances should remain active and in the body neither too short nor too long. For many drugs, a metabolic and/or excretion rate that enables "once a day" dosage is aimed at. Sometimes this needs the identification of sites in the molecule that will be metabolized quickly with concomitant loss of activity. The vasodilator iloprost, for instance, was developed from the endogeneous mediator prostacyclin that has very short half-life both *in vivo* and on the shelf. Modification of several chemically and metabolically vulnerable positions yielded a stable and active derivative – a highly sophisticated product of synthetic medicinal chemistry (Figure 2.3).[21] Vice versa, sometimes functionality is introduced for the acceleration of biotransformation and excretion. Articaine is a local anesthetic of the anilide type. Systemically, it interferes with heart rate – an unwelcome side effect in dentistry. That is why articaine contains an additional ester group. Once in the blood stream, this will be hydrolyzed quickly to an – in this case – inactive carboxylic acid (Figure 2.4).[22] Medicinal chemistry here has come full circle as anilide local anesthetics were developed from ester anesthetics like procain in order to prolong activity.

FIGURE 2.3 Prostacyclin and its synthetic analog, iloprost, that combines activity with sufficient *ex vivo* and *in vivo* stability.

FIGURE 2.4 Articaine, a common local anesthetic dentists use, and its inactive metabolite that is formed off the scene of painful action. The value for $t_{1/2}$ is from the reference Oertel, R.; Ebert, U., Rahn, R., Kirch, W. The effect of age on pharmacokinetics of the local anesthetic drug articaine. *Reg. Anesth. Pain Med.* 1999, *24*, 524–528.

C. The pharmacodynamic phase

While pharmacokinetics investigates what the body does to the drug, pharmacodynamics is concerned with what the drug does to the body. Most scientists who consider themselves to be medicinal chemists will be most comfortable with and interested in this phase. They will cooperate with biochemists and pharmacologists to elucidate mechanistic details of the interaction of the drug with its target(s), a topic we will treat in the Section III.

D. The road to successful drug development?

In the past years, many analyses have appeared that try to explain the dearth of new drug substances in the face of billions of dollars and assay data points, ten thousands of virtual and thousands of real hits that have been spent, accumulated and generated. In comparison, the Belgian medicinal chemist Paul Janssen and his relatively small group had tremendous success in the development of new drug entities and activities.[23] It was postulated that the individualization rather than integration of research guidelines into successive hypes (e.g. "as target subtype selective as possible;" "ADME rules have to be strictly adhered to;" "modeling programmes automatically give a correct representation of molecules;" "the more combinatorial ligands, the more hits") is responsible for the disappointing state of drug discovery. The "hypes" were likened to "games," that is sophisticated theories that work well in themselves but have lost contact with reality.[24] What is needed is to keep what we already know about how successful drugs were actually discovered or invented,[25] and provide an atmosphere of creativity in a team of scientists from various disciplines. Summarizing their long lasting experiences in antibacterial research, an industrial team concluded that for this therapeutic area at least, synthesizing novel chemical structures that interact with and block *established* targets in new ways is a robust strategy.[26]

So what does the increasing knowledge of targets mean for medicinal chemistry? This will be presented in the following paragraphs and discussed in detail in later chapters.

III. DRUG CLASSIFICATION SYSTEMS

Classification systems help with understanding what a drug actually does at the molecular level (classification by target),

and they are indispensable for categorizing the large number of drug substances (classification by clinical effect).

A. Classification by target and mechanism of action

1. *Targets*

Targets are molecular structures, chemically definable by at least a molecular mass, that will undergo a specific interaction with chemicals that we call drugs because they are administered to treat or diagnose a disease.[27] To be meaningful, the interaction has to have a connection with the clinical effect(s). It is very challenging to prove this.

A clinically relevant target might consist not of a single biochemical entity, but the simultaneous interference of a number of receptors. Only this will give a net clinical effect that might be considered beneficial. It is only by chance that the current *in vitro* screening techniques will identify drugs that work through such targets.

The number of targets presently "used" is still open to discussion in medicinal chemistry, but various approaches converge at a few hundred.[4,28] The number of potential targets, however, was estimated to be several hundred thousand in view of the manifold protein complexes, splicing variants and possible interventions with signaling pathways.[25] The problem with counting is mainly 2-fold: firstly, the identification of the reaction partners of drug substances in the body, and secondly, exactly what to define and count as the target. A target definition derived from the net effect rather than the direct chemical interaction will require input from systems biology, a nascent research field that promises to affect the drug discovery process significantly.[28] At the other end of the scale of precision, we can define some targets very precisely on the molecular level. For example, we can say that dihydropyridines block the CaV1.2a splicing variant in heart muscle cells of L-type high-voltage activated calcium channels.

The actual depth of detail used to define the target is primarily dependent on the amount of knowledge available about the target and its interactions with a drug. If the target structure has already been determined, it could still be that the molecular effect of the drug cannot be fully described by the interactions with one target protein alone. For example, antibacterial oxazolidinones interact with 23S-rRNA, tRNA and two polypeptides, ultimately leading to inhibition of protein synthesis.[29] In this case, a description of the mechanism of action that only includes interactions with the 23S-rRNA target would be too narrowly defined. In particular, in situations in which the dynamic actions of the drug substance stimulate, or inhibit, a biological process, it is necessary to move away from the descriptions of single proteins, receptors and so on and to view the entire signal chain as the target.

Lists that classify all marketed drug substances according to target, with references, were published, an excerpt is given in Table 2.1.[27]

2. *Mechanisms of action*

An effective drug target comprises a biochemical system rather than a single molecule. Present target definitions are static. We know this to be insufficient, but techniques to observe the dynamics of drug–target interactions are just being created. Most importantly, we are not able to gauge the interaction of the biochemical "ripples" that follow the drug's initial molecular effect. It has been pointed out that "two components are important to the mechanism of action … The first component is the initial mass-action-dependent interaction … The second component requires a coupled biochemical event to create a transition away from mass-action equilibrium" and "drug mechanisms that create transitions to a nonequilibrium state will be more efficient."[30]

Although the term "mechanism of action" itself implies a classification according to the dynamics of drug substance effects at the molecular level, the dynamics of these interactions are only speculative models at present, and so mechanism of action can currently only be used to describe static targets, as discussed above.

All drugs somehow interfere with signal transduction, receptor signaling and biochemical equilibria. For many drugs we know, and for most we suspect, that they interact with more than one target. So there will be simultaneous changes in several biochemical signals, and there will be feedback reactions of the pathways disturbed. In most cases, the net result will not be linearly deducible from single effects. For drug combinations, this is even more complicated. Awareness is also increasing of the nonlinear correlation of molecular interactions and clinical effects. For example, the importance of receptor–receptor interactions (receptor mosaics) was summarized for G-protein-coupled receptors (GPCRs), resulting in the hypothesis that cooperativity is important for the decoding of signals, including drug signals.[31] Table 2.2 lists examples of dynamic molecular mechanisms of drugs. Table 2.1 is the excerpt of an attempt at a complete list of drug targets. Notably, inhibitors and antagonists by far outnumber effectors, agonists and substitutes. It appears that reconstitution of biochemical and pharmacological balances is more easily achieved by blocking excessive or complementary pathways rather than by substitution or repair of deficient or defective biochemical input.

Greater knowledge of how drugs interact with the body (mechanisms of action, drug–target interactions) has led to a reduction of established drug doses and inspired the development of newer, highly specific drug substances with a known mechanism of action. However, a preoccupation with the molecular details has resulted in a tendency to focus only on this one aspect of the drug effects. For example, cumulative evidence now suggests that the proven influence of certain psychopharmaceuticals on neurotransmitter metabolism has little to do with the treatment of schizophrenia or the effectiveness of the drug for this indication.[32] With all our efforts to understand the molecular basis of drug action,

TABLE 2.1 The Main Drug Target Classes with Examples of Targets and Ligands. A Full List Can Be Found in Ref. [27]

Target class	Target subclass	Target example	Drug substance example (activity)
Enzymes	Oxidoreductases	Aldehyde dehydrogenase	Disulfiram (inhibitor)
	Transferases	Protein kinase C	Miltefosine (inhibitor)
	Hydrolases	Bacterial serine protease	β-Lactams (inhibitors)
	Lyases	DOPA decarboxylase	Carbidopa (inhibitor)
	Isomerases	Alanine racemase	D-cycloserine (inhibitor)
	Ligases (syn. synthases)	Dihydropteroate synthase	Sulphonamides (inhibitors)
Proteins	Growth factors	Vascular endothelial growth factor	Bevacizumab (antibody)
	Immunoglobulins	CD3	Muromonab-CD3 (antibody)
	Integrins	α4-Integrin subunit	Natalizumab (antibody)
	Tubulin	Human spindle	Vinca alkaloids (development inhibitors)
Substrates, metabolites	Substrate	Asparagine	Asparaginase (enhanced degradation)
	Metabolite	Urate	Rasburicase (enhanced degradation)
Receptors	Direct ligand-gated ion channel receptors	γ-Aminobutyric-acid (GABA)-A receptors	Barbiturates (allosteric agonists)
	G-protein-coupled receptors	Acetylcholine receptors	Pilocarpine (muscarinic receptor agonist)
		Opioid receptors	Buprenorphine (κ-opioid antagonist)
		Prostanoid receptors	Misoprostol (agonist)
	Cytokine receptors	TNFα receptors	Etanercept (receptor mimic)
	Integrin receptors	Glycoprotein IIb/IIIa receptor	Tirofiban (antagonist)
	Receptors associated with a tyrosine kinase	Insulin receptor	Insulin (agonist)
	Nuclear receptors, steroid hormone receptors	Mineralocorticoid receptor	Aldosterone (agonist)
	Nuclear receptors, other	Retinoic acid receptors	Isotretinoin (RARα agonist)
Ion channels	Voltage-gated Ca^{2+} channels	L-type channels	Dihydropyridines (inhibitors)
	K^+ channels	Epithelial K^+ channels	Diazoxide (opener)
	Na^+ channels	Voltage-gated Na^+ channels	Carbamazepine (inhibitor)
	Ryanodine-inositol 1,4,5-triphosphate receptor Ca^{2+} channel	Ryanodine receptors	Dantrolene (inhibitor)
	Transient receptor potential Ca^{2+} channel	TRPV1 receptors	Acetaminophen metabolite (inhibitor)
	Chloride channels	Mast cell chloride channels	Cromolyn sodium (inhibitor)
Transport proteins	Cation-chloride cotransporter family	Thiazide-sensitive NaCl symporter	Thiazide diuretics (inhibitors)
	Na^+/H^+ antiporters		Amiloride (inhibitor)
	Proton pumps	H^+/K^+ ATPase	Omeprazole (inhibitor)
	Eukaryotic (putative) sterol transporter (EST) family	Niemann-Pick C1 like 1 protein	Ezetimibe (inhibitor)
	Neurotransmitter/Na^+ symporter family	Serotonin/Na^+ symporter	Paroxetine (inhibitor)
DNA, RNA	Nucleic acids	Bacterial 16S-RNA	Aminoglycosides (protein synthesis inhibition)
	Ribosome	Bacterial 30S subunit	Tetracyclines (protein synthesis inhibition)
Physicochemical mechanism	Ion exchange	Hydroxide	Fluoride (enhanced acid stability of adamantine)
	Acid binding	In stomach	Hydrotalcite
	Adsorptive	In gut	Charcoal
	Surface-active	On oral mucosa	Chlorhexidine (disinfectant)
	Oxidative	On skin	Permanganate (disinfectant)
	Reductive	Disulphide bonds	N-acetylcysteine (mucolytic)
	Osmotically active	In gut	Lactulose (laxative)

TABLE 2.2 Examples of Dynamic (Process) Mechanisms of Drug Action

Dynamic mechanism	Example
Covalent modifications of the active center	Acylation of bacterial transpeptidases by β-lactam antibiotics
Drugs that require the receptor to adopt a certain conformation for binding and inhibition	Trapping of K^+ channels by methanesulphoanilide antiarrhythmic agents
Drugs that exert their effect indirectly and require a functional background	The catechol O-methyltransferase inhibitor entacapone, the effect of which is due to the accumulation of nonmetabolized dopamine
Antiinfectives that require the target organism to be in an active, growing state	β-Lactam antibacterials
Molecules requiring activation (prodrugs)	Enalaprilate, paracetamol
Modifications of a substrate or cofactor	Asparaginase, which depletes tumour cells of asparagine;[i] isoniazide, which is activated by mycobacteria leading to an inactive covalently modified NADH[ii]
Simultaneous modulation of several signaling systems	GPCR receptor mosaics for the decoding of drug signals
Fluctuations of physiological signalling molecules	Dopamine fluctuations after administration of cocaine, followed by a gradual increase in steady state dopamine concentration[iii]

[i] Graham, M. L. *Pegaspargase: a review of clinical studies*. Adv. Drug Deliv. Rev. **2003**, 55, 1,293–1,302.
[ii] Larsen, M. H., Vilchèze, C., Kremer, L., Besra, G. S., Parsons, L., Salfinger, M., Heifets, L., Hazbon, M. H., Alland, D., Sacchettini, J. C., Jacobs, W. R. *Overexpression of inhA, but not kasA, confers resistance to isoniazid and ethionamide in Mycobacterium smegmatis, M. bovis* BCG *and M. tuberculosis*. Mol. Microbiol. **2002**, 46, 453–466.
[iii] Heien, M. L., Khan, A. S., Ariansen, J. L., Cheer, J. F., Phillips, P. E., Wassum, K. M., Wightman, R. M. *Real-time measurement of dopamine fluctuations after cocaine in the brain of behaving rats*. Proc. Natl. Acad. Sci. USA **2005**, 102, 10,023–10,028.

we must not fall into the trap of reductionism. Indeed, the "one-drug-one-target" hypothesis (perhaps even adding "one disease," ignoring the complexity of medical diagnoses) may partly be responsible for the relative dearth of new drug substances.[25] For antibacterial research, multitargeting is now considered to be essential.[33] More generally, in recent years the limits of the reductionist approach in drug discovery have become painfully clear. Nobel laureate Roald Hoffmann put it this way: "Chemistry reduced to its simplest terms, is not physics. Medicine is not chemistry … knowledge of the specific physiological and eventually molecular sequence of events does not help us understand what [a] poet has to say to us."[34] The cartoon (Figure 2.5) illustrates this point. Although it is too early for systems biology to be provide clear-cut protocols for medicinal chemistry, "translational medicine"[35] and other integrative research efforts stress the functional as opposed to reductionist character of living systems, hopefully improving the success rate of drug research.[36]

B. Other classification systems

From a pharmaceutical standpoint there are many different criteria which can be used to classify medications: type of formulation, the frequency with which it is prescribed or recommended, price, refundibility, prescription or nonprescription medication, etc. If a classification of the APIs is undertaken, numerous possibilities are revealed, as well. At the end of the 19th century, drug substances were classified the same as other chemical entities; by nature of their primary elements, functional moieties or organic substance class. Recently, the idea of classifying drug substances strictly according to their chemical constitution or structure has been revived. Recent databases attempt to gather and organize information on existing or potential drug substances according to their chemical structure and diversity. The objective is to create substance "libraries," which contain pertinent information about possible ligands for new targets (e.g. an enzyme or receptor) of clinical interest,[37,38] and more importantly, to understand the systematics of molecular recognition (ligand–receptor).[39,40]

The most commonly used classification system for drug substances is the *ATC system*.[41] It was introduced in 1976 by the Nordic Council on Medicines as a method to carry out drug utilization studies throughout Scandinavia. In 1981, the World Health Organization recommended the use of the ATC classification for all global drug utilization studies and in 1982 founded the WHO Collaborating Centre for Drugs Statistics Methodology in Oslo to establish and develop the method. The ATC system categorizes drug substances at five different levels according to (1) the organ or system on which they act (Anatomy) (2) therapeutic and pharmacological properties and (3) chemical properties. The first level is comprised of the main anatomical groups, while the

FIGURE 2.5 Searching for molecular mechanisms … "The meaning of the message will not be found in the chemistry of the ink." Sperry, R. *Brain Circuits and Functions of the Mind*; Cambridge University Press: Cambridge, 1990. *Source*: Roger Sperry, neurophysiologist, Nobel Prize in Medicine 1981.

second level contains the pharmacologically relevant therapeutic subgroup. The third level consists of the pharmacological subgroup and the fourth the chemical subgroup. The fifth level represents the chemical substance (the actual drug entity). Drugs with multiple effects and different target organs can be found more than once within the system. The antiinflammatory agent diclofenac, for instance, has three ATC numbers, one of them being M01AB05. This key breaks down to: M01 (musculo-skeletal system; antiinflammatory and antirheumatic agents, nonsteroids), M01AB (acetic acid derivatives and related substances), 05 (diclofenac in M01AB). The two other keys classify diclofenac as a topical agent and its use for inflammation of sensory organs.

While ATC is better suited if the emphasis is on therapeutic use, the recently proposed[27,42] TCAT system puts the target chemistry first, suiting the medicinal chemical approach.

REFERENCES

1. Burger, A. Preface. In *Comprehensive Medicinal Chemistry* (Hansch, C., Sammes, P. G., Taylor, J. B., Eds). Pergamon Press: Oxford, 1990, p. 1.
2. Wermuth, C. G., Ganellin, C. R., Lindberg, P., Mitscher, L. A. Glossary of terms used in medicinal chemistry (IUPAC Recommendations, 1997). In *Annual Reports In Medicinal Chemistry* (Adam, J., Ed.). Academic Press: San Diego, CA, 1998, pp. 385–395.
3. World Drug Index, 2007. Database is available from Thomson Scientific at: http://scientific.thomson.com/products/wdi/.
4. Overington, J. P., Al-Lazikani, B., Hopkins, A. L. How many drug targets are there?. *Nat. Rev. Drug Discov.* 2006, *5*, 993–996.
5. Rote Liste 2006. Rote Liste Service GmbH: Frankfurt/Main, 2007.
6. World Health Organisation, WHO Model List of Essential Medicines, 2007 (http://www.who.int/medicines/publications/EML15.pdf).
7. Madden, J. C., Cronin, M. T. Structure-based methods for the prediction of drug metabolism. *Expert Opin. Drug Metab. Toxicol.* 2006, *2*, 545–557.
8. O'Brien, S. E., de Groot, M. J. Greater than the sum of its parts: combining models for useful ADMET prediction. *J. Med. Chem.* 2005, *48*, 1287–1291.
9. Schuster, D., Laggner, C., Langer, T. Why drugs fail – a study on side effects in new chemical entities. *Curr. Pharmaceut. Des.* 2005, *11*, 3545–3559.
10. Guengerich, F. P., MacDonald, J. S. Applying mechanisms of chemical toxicity to predict drug safety. *Chem. Res. Toxicol.* 2007, *20*, 344–369.
11. Cronin, M. T. Prediction of drug toxicity. *Farmaco* 2001, *56*, 149–151.
12. Wilson, A. G., White, A. C., Mueller, R. A. Role of predictive metabolism and toxicity modeling in drug discovery – a summary of some recent advancements. *Curr. Opin. Drug Discov. Dev.* 2003, *6*, 123–128.
13. Lowe, D. In the pipeline: trocetrapib. *Chemistry World.* 2007, *4*, 16.
14. Sneader, W. *Drug Discovery*. John Wiley & Sons: Chichester, 2005.
15. Jones, W. P., Chin, Y. W., Kinghorn, A. D. The role of pharmacognosy in modern medicine and pharmacy. *Curr. Drug Targets* 2006, *7*, 247–264.
16. Koehn, F. E., Carter, G. T. The evolving role of natural products in drug discovery. *Nat. Rev. Drug Discov.* 2005, *4*, 206–220.
17. Bleicher, K. H., Böhm, H. J., Müller, K., Alanine, A. I. Hit and lead generation: beyond high-throughput screening. *Nat. Rev. Drug Discov.* 2003, *2*, 369–378.
18. Kalgutkar, A. S., Soglia, J. R. Minimizing the potential for metabolic activation in drug discovery. *Expert Opin. Drug Metab. Toxicol.* 2005, *1*, 91–141.
19. Zlokarnik, G., Grootenhuis, P. D. J., Watson, J. B. High throughput P450 inhibition screen in early drug discovery. *Drug Discov. Today* 2005, *10*, 1443–1450.
20. Pelkonen, O., Raunio, H. *In vitro* screening of drug metabolism during drug development: can we trust the predictions?. *Exp. Opin. Drug Metab. Toxicol.* 2005, *1*, 49–59.
21. Skuballa, W., Schillinger, E., Stuerzebecher, C. S., Vorbrueggen, H. Prostaglandin analogs. Part 9. Synthesis of a new chemically and metabolically stable prostacyclin analog with high and long-lasting oral activity. *J. Med. Chem.* 1986, *29*, 313–315.
22. Oertel, R., Rahn, R., Kirch, W. Clinical pharmacokinetics of articaine. *Clin. Pharmacokin.* 1997, *33*, 417–425.

23. Black, J. A personal perspective on Dr. Paul Janssen. *J. Med. Chem.* 2005, *48*, 1687–1688.
24. Kubinyi, H. Drug research: myths, hype and reality. *Nat. Rev. Drug Discov.* 2003, *2*, 665–668.
25. Sneader, W. *Drug Prototypes and Their Exploitation*. John Wiley & Sons: Chichester, 1996.
26. Payne, D. J., Gwynn, M. N., Holmes, D. J., Pompliano, D. L. Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat. Rev. Drug Discov.* 2007, *6*, 29–40.
27. Imming, P., Sinning, C., Meyer, A. Drugs, their targets and the nature and number of drug targets. *Nat. Rev. Drug Discov.* 2006, *5*, 821–834.
28. Apic, G., Ignjatovic, T., Boyer, S., Russell, R. B. Illuminating drug discovery with biological pathways. *FEBS Lett.* 2005, *579*, 1872–1877.
29. Colca, J. R., McDonald, W. G., Waldon, D. J., Thomasco, L. M., Gadwood, R. C., Lund, E. T., Cavey, G. S., Mathews, W. R., Adams, L. D., Cecil, E. T., Pearson, J. D., Bock, J. H., Mott, J. E., Shinabarger, D. L., Xiong, L. Mankin. *A.S.J. Biol. Chem.* 2003, *278*, 21972–21979.
30. Swinney, D. C. Biochemical mechanisms of drug action: what does it take for success?. *Nat. Rev. Drug Discov.* 2004, *3*, 801–808.
31. Agnati, L. F., Fuxe, K., Ferré, S. How receptor mosaics decode transmitter signals. Possible relevance of cooperativity. *Trends Biochem. Sci.* 2005, *30*, 188–193.
32. Hyman, S. E., Fenton, W. S. What are the right targets for psychopharmacology? *Science* 2003, *299*, 350–351.
33. Silver, L. L. Multi-targeting by monotherapeutic antibacterials. *Nat. Rev. Drug Discov.* 2007, *6*, 41–55.
34. Roald Hoffmann, speech at the Nobel banquet, 1981 (http://nobelprize.org/nobel_prizes/chemistry/laureates/1981/hoffmann-speech.html).
35. FitzGerald, G. A. Anticipating change in drug development: the emerging era of translational medicine and therapeutics. *Nat. Rev. Drug Discov.* 2005, *4*, 815–818.
36. Walker, M. J. A., Soh, M. L. M. Challenges facing pharmacology – the *in vivo* situation. *Trends Pharmacol. Sci.* 2006, *27*, 125–126.
37. Schneider, G. Trends in virtual combinatorial library design. *Curr. Med. Chem.* 2002, *9*, 2095–2101.
38. Goodnow, R. A., Jr., Guba, W., Haap, W. Library design practices for success in lead generation with small molecule libraries. *Comb. Chem. High Throughput Screen* 2003, *6*, 649–660.
39. Hendlich, M., Bergner, A., Gunther, J., Klebe, G. Relibase: design and development of a database for comprehensive analysis of protein–ligand interactions. *J. Mol. Biol.* 2003, *326*, 607–620.
40. Gohlke, H., Klebe, G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angew. Chem. Int. Ed. Engl.* 2002, *41*, 2644–2676.
41. http://www.whocc.no/atcddd/
42. Imming, P., Buß, T., Dailey, L. A., Meyer, A., Morck, H., Ramadan, M., Rogosch, T. A classification of drug substances according to their mechanism of action. *Pharmazie* 2004, *59*, 579–589.

Chapter 3

Measurement and Expression of Drug Effects

Jean-Pierre Nowicki and Bernard Scatton

I. INTRODUCTION
II. *IN VITRO* EXPERIMENTS
A. Binding studies
B. Ligand–receptor interaction-induced functional effects
C. Allosteric interaction
D. Expression of functional effects for targets other than GPCRS
E. Cellular and tissular functional responses
III. *EX VIVO* EXPERIMENTS
IV. *IN VIVO* EXPERIMENTS
REFERENCES

Science is made of facts, just as houses are made of stones:
But a mere collection of facts is no more science
Than a pile of stones a house
Henri Poincare
(French mathematician 1854–1912)

I. INTRODUCTION

The biological activity of a drug is linked to its ability to bind to specific recognition sites located on intracellular proteins or on proteins that are integrated in the lipid bilayer of cellular membranes. Binding sites are specifically recognized by one or more endogenous molecules, generically named ligands or substrates (when enzymes are concerned), that can be small in size (e.g. glutamate) or very large (proteins, DNA). Interaction of the ligand with its recognition site stabilises a certain conformation of the protein with a resulting change (either an increase or a decrease) in protein function. These proteins possess, or are linked to other proteins that possess, the ability to generate an intracellular biochemical signal. Ultimately, cellular and tissular activity is modulated by the change in this signal induced by ligand-recognition site interaction. The diversity of signaling systems may be better appreciated when considering the following examples:

- *G-protein-coupled receptors (GPCRs)*: It has been estimated that approximately 25% of marketed drugs act by directly stimulating or blocking GPCRs[1] (see also Figure 3.2). Each GPCR is a single protein, with seven transmembrane domains, coupled intracellularly to a G-protein (sensitive to GTP) and to an effector protein, for example, the enzyme adenylyl cyclase, in which case cAMP is the intracellular signal. Examples of other effector proteins are phospholipase C (diacylglycerol (DAG) and inositol 1,4,5-triphosphate (IP3) as signals), Ca^{2+} and K^{+} channels (corresponding cations as signals).
- *Tyrosine kinase receptors*: These receptors exist in a monomer–dimer equilibrium. The dimer, which is stabilized upon ligand binding, is the signaling structure. Dimer formation stimulates catalytic activity and results in intermolecular autophosphorylation within the dimer and triggers signaling cascades that lead to the phosphorylation of cytoplasmic substrates (insulin receptor as example, Figure 3.1). The increase in phosphorylation of tyrosine residues of intracellular proteins either increases or decreases their activity, particularly that of protein kinases or protein phosphatases that often play a crucial role in the regulation of cellular function.

Copyright © 2008, Elsevier Inc.
All rights reserved.

- *Ligand-gated ion channels*: Examples include the *N*-methyl-d-aspartate (NMDA) subtype of glutamatergic receptors, the nicotinic subtypes of acetylcholine receptors or γ-aminobutyric acid (GABA)-A receptors. The signal is the particular ion that is allowed to flow across the membrane following ligand binding-induced opening of the ion channel.
- *Voltage-operated cation channels*: They open and close in response to changes in membrane potential. As in ligand-gated ion channels, intracellular signaling is generated by the flow of the corresponding ion (Ca^{2+}, K^+ or Na^+) across the membrane. Although voltage-activated channels have no ligand binding sites that control activity, they generally nevertheless possess other binding sites that allow pharmacological activation or blockade of the channel.
- *Enzymes*: They possess a recognition site, the catalytic site, where one or more substrates bind, and possibly also a few other modulatory sites. However, while ligand–receptor interaction generally leaves the ligand unaltered, the chemical structure of a substrate is changed following its interaction with the enzyme. Inhibition or activation of enzymatic activity changes substrate and product levels that constitute an intracellular signal, particularly when the enzyme substrates are proteins.
- *Transporters*: They allow cellular entry of various molecules against a concentration gradient with concomitant energy consumption. In the case of glucose transport, for example, tissues such as muscle and adipocytes can sense glucose and communicate changes in glucose flux to other tissues and thus glucose and its metabolites act as signaling molecules. One of the glucose transporters, GLUT4, is sequestered in intracellular vesicles in the absence of insulin. Upon insulin stimulation, GLUT4 vesicles translocate to and fuse with the plasma membrane thus increasing glucose flux.

FIGURE 3.1 Early biochemical steps following insulin receptor activation. Following insulin binding and insulin receptor autophosphorylation and activation, insulin receptor substrates-1 and -2 (IRS-1 and -2) are phosphorylated allowing binding of p85 and p110 phosphatidylinositol kinases and the generation of inositol triphosphate (IP3). *Source*: Redrawn from Thirone *et al.*[2]

Any protein belonging to the various classes described above and many others (chaperone proteins, nuclear receptors…) represents a mechanistic target for therapeutic research (Figure 3.2). How is a particular biological target

FIGURE 3.2 Number of launched drugs in each of the major class of biological targets. *Source*: Data from the Prous Integrity Drugs and Biologics database (June 21, 2007).

selected? Most often, increasing knowledge of the molecular basis of a given pathology (pathophysiology) unravels the potential implication of one or more proteins in the disease process and it is expected that pharmacological modification of the function of these proteins may afford treatment of the disease.

Once a new drug molecule has been synthesized, one has to verify that it possesses the functional activity and therapeutic efficacy expected. This is done by evaluating the properties of the drug in a logically ordered sequence of tests, the screening architecture, that starts with simple *in vitro* tests but later also includes sophisticated *in vivo* experiments; each result provides a piece of evidence confirming or infirming the potential interest of the compound. An idea of the various types of biological tests that may be included in a screening architecture, how the results are expressed and what they imply will be provided in the following, focusing primarily on GPCR receptors (Box 3.1).

II. *IN VITRO* EXPERIMENTS

A. Binding studies

The aim of binding experiments is to determine the affinity (the strength with which a compound binds to a site) of the compound for its biological target and to check its selectivity versus other binding sites or biological off-targets. Binding studies usually represent an initial step in compound characterization. Schematically, membranes are prepared from the tissue of interest (heart, bladder, brain …) or from mammalian cells that express the receptor of interest. The receptors can be native, that is, they are constitutively expressed by the cells or the tissue, or transduced, that is, a cDNA coding for the receptor isolated from any appropriate species has been inserted into the cell. Chinese hamster ovary (CHO) or human epithelial kidney 293 cells (HEK293) are generally used. In the latter, transfection can be stable and cells can proliferate while continuing to express the receptor, or it can be transient and cells rapidly loose their ability to express the receptor. Stable transfection in cell lines is often used to perform binding studies with human receptors since compound affinity may differ markedly between receptors isolated from animals and man.

An absolute requirement for binding experiments is a radioactive labeled ligand that specifically binds to the biological target under study. Most often, the ligand used is a synthetic molecule and not the endogenous ligand. For example, there are a large number of serotonin receptor subtypes but serotonin is used as a ligand to study only very few of them. The reasons for such a choice can be ease of use (greater affinity, better stability or solubility, smaller size than the endogenous ligand) or specificity (the fact that a ligand preferentially recognizes a given receptor subtype when compared to the others).

Typically, *displacement* experiments give rise to sigmoid curves similar to the one shown in Figure 3.3a. The drug concentration that displaces half of the maximum bound radioactive ligand represents the IC_{50}.

Alternatively, when membranes are incubated with various concentrations of the radiolabeled ligand, a plot of bound/free against bound ligand (Scatchard plot,[3] Figure 3.3b) generally gives rise to a straight line. In these *saturation* experiments, B_{max} (the maximal number of binding sites per unit of tissue or protein weight) is determined from the intercept of the line with the abscissa and K_d (the dissociation constant) from the negative reciprocal of the slope of the line. When such experiments are performed in the presence of various concentrations of the compound under study, they give rise to a family of lines. If B_{max} remains unchanged and the slope of the lines decreases with increasing concentrations of the compound, the displacement is competitive (i.e. the radiolabeled ligand and the compound occupy the same binding site). Unchanged slope and decreased B_{max} indicate that the displacement is non-competitive (binding of radiolabeled ligand and compound occurs at distinct proximal sites).

The lower the IC_{50} or K_d, the higher the affinity. Results can also be expressed as a K_i, the inhibitory (or affinity) constant of the displacer compound for the receptor. K_i and IC_{50} are not independent and are very simply related when the displacement is non-competitive ($K_i = IC_{50}$), but the relationship becomes more complicated (Cheng-Prusoff[4] equation) for a competitive displacement [$K_i = IC_{50}/(1 + [L]/K_d)$ where [L] is the concentration of the radioactive ligand].

When designing new drugs, high affinity is often sought and may represent a crucial parameter, particularly in cases where the affinity of the endogenous ligand for its binding site is very high (to be efficient, the compound has to displace the endogenous ligand). Generally, it is assumed

BOX 3.1 Expression of Drug Effects

In vitro

B_{max}	Maximal number of binding sites
K_d	Dissociation constant
K_A	Association constant
K_i	Inhibition constant
IC_{50}	Median inhibitory concentration
EC_{50}	Median effective concentration
pD_2	$-\log[EC_{50}]$
E_{max}	Maximal response of an agonist
pA_2	–log molar concentration of an antagonist producing a 2-fold shift of the concentration–response curve
MAC	Minimal active concentration

Ex vivo/in vivo

MAD	Minimal active dose
ID_{50}	Median inhibitory dose
ED_{50}	Median efficacious dose

FIGURE 3.3 (a) Displacement curve. A constant fraction of the membrane preparation is incubated with a fixed concentration of the radiolabeled ligand and various concentrations of the compound under study for a fixed period of time. Thereafter, free and membrane-bound ligands are separated by filtration and the radioactivity remaining on membranes is measured in a scintillation counter. Non-specific binding is obtained following incubation of the membrane preparation in the presence of a large excess of non-radioactive ligand. Specific binding = total – non-specific binding. (b) Scatchard plot.

that, when affinity is high, the compound is less likely to interfere with other, possibly unwanted off-target sites. However, this is not always true and selectivity has to be checked by evaluating the affinity of the compound for a large panel of receptors, enzymes and ion channels. It is obvious that selectivity has limits that depend on the size of the panel that has been investigated, but also on the scientific knowledge available at the time when the studies are performed. For these reasons, it is thus always possible that a compound considered to possess a high degree of specificity may nevertheless induce unexpected biological effects.

TABLE 3.1 Methods Used to Quantify Ligand–Receptor Interaction-Induced Changes in Intracellular 2nd Messengers

Intracellular 2nd messenger	Assay
cAMP	Immunoassay or methods based on fluorescence measurements (fluorescence polarization, fluorescence intensity)
IP3	Chromatography
Ca^{2+}	Indicator dyes and fluorescence measurements (Fluo-4/FLIPR) or bioluminescence readouts and proteins (aequorin)

B. Ligand–receptor interaction-induced functional effects

Binding experiments are performed in order to characterize the affinity of a compound for a receptor but they do not establish whether a compound behaves as an agonist, an antagonist or an inverse agonist. Such determinations necessarily involve functional measurements of ligand–receptor interaction-induced changes in an intracellular signal (Table 3.1).

Such experiments also represent the initial step in compound characterization as a transporter or enzyme inhibitor, or as a voltage-activated cation channel modulator. In this latter case, the compound potency in functional experiments is often much higher than that expected from the affinity determined in binding experiments but the reasons for this discrepancy are largely unknown to date.

Since antagonists block an existing ligand-activated functional effect, the receptor has to be incubated with a given concentration of an agonist and the effects of various concentrations of the putative antagonist are then studied. The curves are very similar to that depicted in Figure 3.3a and the results are expressed as an IC_{50}, the drug concentration that produces half of the maximal response (E_{max} in %) measured in the absence of the antagonist. Alternatively, the effects of the agonist in the presence of different concentrations of the antagonist can be studied (one concentration for each curve) thus giving rise to a family of curves (Figure 3.4) and allowing the calculation of a pA_2 (–log molar concentration of antagonist producing a 2-fold shift of the concentration–response curve, that is, a 2-fold increase in agonist concentration in order to obtain a similar effect). Competitive antagonists induce a parallel rightward displacement of the curves with increasing concentrations of the antagonist (Figure 3.4a), but, with non-competitive antagonists a rightward displacement of the curves with a decrease in E_{max} is observed (Figure 3.4b).

FIGURE 3.4 Effects of increasing concentrations of a competitive (a) and a non-competitive antagonist (b) on an agonist-induced functional response.

FIGURE 3.5 Effects of a full or a partial agonist.

FIGURE 3.6 Receptor theory assumes that the receptor can exist in at least two separate forms: one inactive form denoted by R and one active form denoted by R* that are in equilibrium. A full agonist has a much higher affinity for the active form of the receptor and will displace the equilibrium toward the active form and a full inverse agonist has a much higher affinity for the inactive form of the receptor. Neutral antagonists have a similar affinity for both receptor forms. *Source*: Redrawn from Brink *et al.*[7]

Agonists are characterized by incubating the receptor with the compound under study and the functional response is compared to that obtained in the presence of a ligand already identified as a full agonist. In order to fully characterize the effect of a drug it is necessary to take into account both the efficacy, the E_{max}, and the potency, the EC_{50}, that is, the effective concentration needed to reach 50% of the maximal effect. The results can also be expressed as a pD_2 ($pD_2 = -\log[EC_{50}]$).

Indeed, two agonists may possess a similar efficacy but one of them may be less potent than the other (rightward displacement of the curve). Conversely, two agonists may be equally potent but the efficacy of one of them can be lower (smaller maximal response). Ligands with an efficacy that is a fraction of the effects induced by a full agonist are named partial agonists (Figure 3.5).

A number of GPCRs display a measurable basal activity in the absence of any endogenous or exogenous agonist, either constitutively in the native state or following transfection with a mutated protein. The effects of full or partial agonists described above are unaffected by basal receptor activity. However, some ligands are able to decrease constitutive receptor activity, a property known as inverse agonism. In the absence of constitutive activity, inverse agonists behave as competitive antagonists but the mechanisms by which inverse agonists and neutral antagonists achieve their effects are different (Figure 3.6) (Box 3.2).

BOX 3.2 Antagonists or inverse agonists?

Although the existence of inverse agonists has been substantiated in experimental systems, the therapeutic relevance of this class of drugs is as yet unknown. However, a survey on the activity of 380 antagonists on 73 GPCRs indicates that 322 are inverse agonists, some of them being used clinically, and 58 (15%) are neutral antagonists[5]. Inverse agonists could be very useful in pathological situations where constitutively active receptor levels are increased and/or when the level of the endogenous ligand is low. In this latter case, antagonists are of no value due to the near absence of the ligand. There is, however, a risk of tolerance due to receptor up-regulation following chronic blockade.

Finally, a new functional property has recently been characterized: protean agonism.[6] Ligands belonging to this class of drugs act as partial agonists in quiescent silent systems and as inverse agonists in systems that show a high level of constitutive activity. The name protean comes from the Greek god Proteus who had the ability to change his shape at will. The reversal from agonism to inverse agonism would only occur when an agonist produces an active conformation of lower efficacy than a totally active conformation (in Figure 3.6, an other $R^{@}$ species distinct from R and R*). Therefore, the higher the constitutive activity, the greater chance to see this other conformation.

C. Allosteric interaction

The ligand-induced functional effects described above can occur when a drug binds to the site recognized by the endogenous ligand, the orthosteric site, leading to competitive interactions or to a site located extremely close to the orthosteric site inducing non-competitive interactions. However, the entire receptor surface (other than the orthosteric binding domain) can be considered as bearing potential binding sites for a drug. Such sites, distinct from the orthosteric binding domain, are allosteric sites and drugs that recognize these sites are allosteric modulators. When a drug binds to an allosteric site, protein conformation is altered, resulting in changes in the affinity between ligands and the orthosteric site. Although allosteric modulators were initially defined as ligands possessing no intrinsic agonist or inverse agonist properties, this assumption has been challenged and some allosteric modulators may give rise to agonist or inverse agonist effects in the absence of the orthosteric ligand.

Modulators are able to shift radioligand binding curves, but the allosteric nature of the interaction is revealed as progressively higher concentrations of antagonist fail to cause significant displacements of the radioligand saturation curve (Figure 3.7b), in contrast to what would theoretically be expected with an antagonist (Figure 3.7a).

FIGURE 3.7 Plots of fractional orthosteric ligand–receptor occupancy as a function of log [orthosteric ligand concentration]. Curves shifts induced by a competitive antagonist (a) or a negative allosteric modulator (b). Note the limits in curve shifts with an allosteric modulator (ceiling effect).

A common graphical method for assessing the relationship between radioligand saturation binding and antagonist concentration involves the determination of the affinity shift, that is, the ratio of radioligand affinity in the presence (K_{App}) to that obtained in the absence (K_A) of each concentration of antagonist. A plot of log (affinity shift–1) versus log [antagonist] should yield a straight line with a slope of 1 for a competitive interaction, but a curvilinear plot for an allosteric interaction. But an allosteric modulator can also alter the link between the orthosteric site and the functional response and therefore modify the efficacy of the orthosteric ligand. This parameter can sometimes be appreciated

FIGURE 3.8 The shift (EC_{50wo}/EC_{50w}) of functional concentration–response curves obtained in the absence or in the presence of the allosteric potentiator is a measure of the efficacy of the modulator.

by the shift between the EC_{50s} for functional concentration–response curves obtained in the absence or in the presence of the allosteric potentiator (Figure 3.8).

In general, the overall effect of an allosteric ligand results from the balance between the modulation of affinity and efficacy and it is usually necessary to also measure cooperativity factors and dissociation rates. A description of this rather complicated field is beyond the scope of this chapter and the interested reader is referred to some recent reviews.[8–11]

The use of allosteric ligands offers certain distinct advantages over orthosteric ligands. The first is a saturability of effect that is retained irrespective of the dose that is administered therapeutically. A second advantage of positive allosteric modulators relates to the fact that they do not replace the endogenous ligand to produce full receptor activation, but selectively "tune" tissue responses in those organs where the endogenous agonist exerts its physiological effects. Finally, a modulator may display the same affinity for each subtype of a receptor but still exert a selective effect by having different degrees of cooperativity at each subtype. Absolute subtype selectivity may therefore be obtained when a modulator remains neutrally cooperative at all receptor subtypes except the one targeted for therapeutic purposes. However, since the structure of allosteric sites is in most cases unknown, selectivity versus other receptors has to be carefully checked and this might not be such an easy task due to the probe dependence of allosteric phenomena (various radiolabeled ligands may induce different effects) and the difficulty in validating allosteric effects. Compilation of useful structure–activity relationship data for allosteric ligands is thus not simple, for example, in the case of muscarinic receptors (a basic subtype of cholinergic receptors), where there are multiple allosteric sites and complex interactions between them.

D. Expression of functional effects for targets other than GPCRS

Frequently, the effects of a compound that decreases the intracellular signal (an inhibitor for enzymes, a blocker for transporters and voltage-operated ion channels) will be characterized by an IC_{50}, as this value is obtained from a single experimental curve (similar to that depicted in Figure 3.3a) and it allows a relative ranking of the potency of a series of compound. However, since the IC_{50} depends on substrate (or ligand) and enzyme (or receptor) concentration, this comparison is only valid if IC_{50s} are determined under identical experimental conditions. A K_i (inhibition constant) can be calculated, particularly for enzymes, but the relation between K_i and IC_{50} vary with the type of inhibition and many types have been described (competitive, non-competitive, uncompetitive …).

E. Cellular and tissular functional responses

Ligand–receptor or drug–enzyme interaction is expected to alter cellular function but in intact cells, a number of functional events may interfere with the initial intracellular signaling and modify the final response. For example, receptor function may be under control of other, possibly ill-defined, regulatory mechanisms. The compound under evaluation may also interfere with other receptor subtypes that are unknown at the time the study is performed. Finally, since the targeted change in tissue function is generally the consequence of a cascade of intracellular events, many of the biochemical steps involved in this sequence may be subjected to tight regulatory mechanisms (Figure 3.9).

It is thus necessary to confirm the existence of modified cellular function following ligand–receptor or drug–enzyme interaction. Such *in vitro* experiments, performed on isolated cells, either native or transfected with the protein of interest, can be undertaken for mechanistic and/or therapeutic purposes. The data depicted in Figure 3.10 illustrate the fact that, in HEK293 cells in culture, inhibition of the enzyme glycogen synthetase kinase 3β (GSK-3β) decreases tau protein hyperphosphorylation, one of the anatomopathological hallmarks of Alzheimer's disease, and may thus represent a potential *therapeutic* approach for this disease. Alternatively, this experimental set-up can also be used for *mechanistic* purposes to characterize the efficacy of compounds as GSK-3β inhibitors. When compared to simple *in vitro* experiments in which the activity of the purified enzyme is measured in the presence of a

FIGURE 3.9 Regulation of vasodilatation by established and emerging GPCRs.[12] *Source*: Reprinted from Maguire J. J. and Davenport A. P. (2005), Copyright (2005), with permission from Elsevier.

FIGURE 3.10 Sodium nitroprusside increases protein tau phosphorylation in HEK293 cells transfected with mutated tau441 (B) when compared to untreated cells (A). GSK-3β is one of the protein kinases involved in tau phosphorylation. Inhibition of GSK-3β by LiCl markedly reduces tau phosphorylation (C). The results are expressed as the means ± S.D. ($N = 3$). *Source*: Redrawn from Zhang *et al.*[13]

drug, results similar to those shown in Figure 3.10 provide at least three important pieces of information:

1. The drug penetrates into the cell, a property that is rather difficult to assess directly.
2. Inhibition effectively takes place with the enzyme *in situ*.
3. The inhibitor does not possess any overt toxicity (cells remain viable).

Studies performed with isolated tissues (brain slices, vascular rings, isolated organs), that have been taken from a living animal, represent a more complicated situation in terms of the number and diversity of biochemical steps that link receptor stimulation and the final functional tissular response since it will integrate ligand–receptor interaction-induced changes in single cells, and the resulting interactions between many cells of the same type and cells of different types (e.g. endothelial and muscular cells in isolated vessels, see Figure 3.9 for indirect dilators).

In most cases, drug effects are expressed, as in binding experiments, as EC_{50}, pD_2, IC_{50} or pA_2. But in some rather sophisticated experiments the effects of only a very small number of drug concentrations will be evaluated and the final result will be expressed as a minimal active concentration (MAC). Although this value represents the first concentration that induces a statistically significant change when compared to control cells or tissue, no general definition can be given since it may also include other requirements specific to the experimental set-up, such as the fact that drug effect should be greater than 50%. Expressing results this way may appear unsatisfactory but this paucity of experimental data is generally dictated by practical reasons such as difficulties in obtaining the tissue (e.g. samples of human pathological

tissue) or technical difficulties such as in Figure 3.10 (a single concentration of LiCl was evaluated in three different experiments) where quantitative determination of tau phosphorylation is delicate, time-consuming and expensive.

III. EX VIVO EXPERIMENTS

Ex vivo experiments generally represent the next step in the characterization of drug effects although they cannot be undertaken with all biological targets. *Ex vivo* means that the drug has been administered by different routes (see below, "*In vivo*" part) to a living animal or to humans and that the evaluation of drug effects are performed *in vitro* with tissue samples or fluid aliquots (blood, cerebrospinal fluid) of the organism under study. An example of an *ex vivo* study is the inhibition of platelet aggregation. Putative inhibitors of platelet aggregation are administered systemically to the animal, drug effects take place inside the body of the animal, blood is sampled after a pre-determined period of time, platelets are isolated and aggregation is induced *in vitro* following addition of ADP.

In *ex vivo* studies, the drug concentration in the test tube is unknown (but can eventually be determined) and drug effects will basically be expressed as a function of the dose administered or time, depending on the aim of the study.

Since the only drug-related parameter known is the initial dose that has been administered, results can be expressed as a minimal active dose (MAD), that is, the lower administered dose that induces a statistically significant effect when compared to animals treated with the vehicle. If it has been possible to study a relatively large number of experimental groups treated with different drug doses, depending on the experimental set-up drug potency will be expressed as an ID_{50} (inhibitory dose 50%), that is, the dose that reduces by 50% the effect measured in control animals or an ED_{50} (efficacy dose 50%) the dose that induces an effect which is half the maximal effect that can be obtained. ED_{50} and ID_{50} are expressed in mg/kg, that is, the amount of drug (generally of the free base if the compound is a salt) per unit of body weight, and the route of administration is also specified (see below "*In vivo*" part). In the *ex vivo* binding experiment shown in Figure 3.11, data have been reported as % of receptor occupancy for each administered dose, which represents the fraction of H3 receptors occupied by the antagonist versus the total number of H3 receptors in the absence of the drug. In fact, what is actually measured is the number of receptors that remain free in each experimental condition and are thus able to bind the radioactive ligand. Receptor occupancy depends on the pharmacodynamic (affinity of the drug for the receptor) and pharmacokinetic (drug tissue concentrations) characteristics of the drug. This latter parameter is a crucial determinant of drug potency *ex vivo* and *in vivo*: it has been, for example, suggested that dopaminergic D2-receptor occupancy by antipsychotics

FIGURE 3.11 Inhibition of [^{3}H]R-α-methylhistamine *ex vivo* binding in rat cortex following oral administration of GSK189254 (a H3 histamine receptor antagonist) measured 2 h after dosing. *Source*: Redrawn from Medhurst *et al.*[15]

FIGURE 3.12 Time-related inhibition of nitric oxide synthase following a single administration of 10 mg/kg, i.p. of *N*ω-nitro-L-arginine (a slowly reversible enzyme inhibitor). If a 60% inhibition of enzymatic activity is considered biologically and statistically meaningful, then the duration of action is approximately 20 h at the administered dose. *Source*: Redrawn from Carreau *et al.*[16]

should lie in an optimal therapeutic window between ~65% and ~80% in order to gain a clinical response.[14]

Alternatively, the drug may be administered at a predetermined efficacious dose and drug effects are then studied as a function of time (Figure 3.12). If a functionally meaningful parameter is chosen, then the duration of action of the drug can be determined, that is, the time beyond which the drug will no longer be efficacious.

Ex vivo experiments are an important step in compound characterization as they investigate compound activity following systemic drug administration to a living animal. They provide a lot of important information concerning the fate of the drug following its administration. If the drug has been administered orally, drug activity implies that:

- *The drug has been absorbed*: Insufficient, or lack of, absorption (the fact that the drug passes from the

BOX 3.3 Drug Metabolism Can Also Occur Specifically into the Targeted Tissue

During the development of a compound aimed at treating a cerebral disease, it appeared that the amide was hydrolyzed into the corresponding acid after compound penetration into the brain. Both compounds had the same pharmacological properties but the acid was more toxic. Since the rate of metabolization could not be assessed in human brain (unknown risk for human volunteers), compound development was stopped.

gastro-intestinal tract into the blood) is often a problem when designing new drugs.

- *The drug has not been subjected to extensive metabolism*: Following metabolism the drug may loose its pharmacological properties or may no longer be able to penetrate into the tissue. But even if the drug is extensively metabolized, the expected functional change can sometimes take place due to the formation of an active metabolite.
- *The drug has reached, and penetrated into, the targeted tissue or cell and it has recognized the biological target (e.g. receptor or enzyme) of interest*: Achieving good tissue penetration may also be a problem, particularly when the brain is concerned since this organ is very efficiently protected from drug entry by the blood-brain barrier (Box 3.3).

IV. *IN VIVO* EXPERIMENTS

The aim of *in vivo* experiments is to confirm that the compound has the therapeutic efficacy expected, that is, that it will interfere with a pathological mechanism involved in an illness and induce beneficial effects. In preclinical *in vivo* studies, the compound under study is administered to an animal and drug effects are quantified by measuring either the behavior of the intact animal placed in a pathological situation (e.g. duration of kainic acid-induced seizures), a physiological parameter (blood pressure, heart rate) or drug-induced changes in an insult-related tissue alteration by biochemical or histological methods on tissue samples taken from the animal (e.g. size of tumors). Clearly, it is quite impossible to give an idea of a standard protocol, due to the very large number of experimental models that can be set up. Each global research field (oncology, cardiovascular research …) deals with field-related pathologies (e.g. anxiety or depression in psychiatric research) that require specialized experimental models that are aimed to mimic the pathology.

However, before being tested in highly sophisticated models, compound evaluation is generally done in relatively simple ones in a first instance. These models are most often performed on small laboratory rodents (mice, rats) and for those that are acute and technically simple, size and number of experimental groups should be large enough to express drug effects as an ED_{50} or ID_{50}. Sometimes results are expressed as a MAD and the potency of the compound is then compared to that of a given reference drug, if available, that may, or may not, have been included in the study, and drug effects may be expressed as simply as better, equal to or less interesting than the reference. Some experiments, such as those where the compound is administered for a very long time (months) to normal or transgenic mice with a chronic disease (i.e. diabetes) or a progressive neurodegenerative disorder (models of Alzheimer's disease) may appear simple but they are extremely time-consuming and usually very expensive, leading to a sharp reduction in the number of doses evaluated. The final result may well be that the drug is efficacious, or is not, at the dose investigated and if it is not efficacious, no further experiment will be undertaken. Sophisticated experiments can be performed in rats but also in dogs or primates. For a number of reasons, particularly ethical ones, the size and number of experimental groups are drastically reduced but in a number of cases, the results can nevertheless be expressed as described above. However, studies that are undertaken under strictly controlled physiological conditions, such as those performed in cardiovascular research, generate a large amount of data (blood pressure, heart rate, parameters of heart function, blood gases …) that need to be interpreted by a specialist in the field.

In a number of experimental set-ups, the calculation of an ED_{50} or ID_{50} is meaningless. In the example shown in Figure 3.13, the maximal drug-induced decrease in brain infarct volume would theoretically be 100% (infarct volume = $0\,mm^3$) but this is quite impossible. In this severe experimental model, every animal, either vehicle- or drug-treated, with no infarction should be discarded since it is very likely that the artery has not been properly occluded. Furthermore, again due to the severity of the model, maximal drug effects are not expected to exceed ~50% and the ED_{50} would represent the dose that induces an effect of ≤25% that is likely not to be statistically significant. The results of the study shown in Figure 3.13 will be presented as the effects at the maximally effective dose (−48% at 15 mg/kg, p.o.) together with experimental details that will influence drug efficacy (number of administrations, delay between artery occlusion and first drug administration, duration of artery occlusion…).

Important additional information arises from the shape of the dose–effect curves. Some drugs display inverted U-shaped curves, that is, drug efficacy increases with increasing doses up to a dose beyond which it decreases (Figure 3.13 for doses above 15 mg/kg). This progressive loss of efficacy is often indicative of a drug-induced deleterious mechanism (toxicity), generally unrelated to the main effect of the drug. The dose that induces the greater pharmacological effects is very important for clinical development

FIGURE 3.13 Decrease in the volume of the cerebral infarct induced by middle cerebral artery occlusion in the rat following oral administration of a mitogen-activated protein kinase (MAPK) inhibitor. *Source*: Redrawn from Barone *et al.*[17](*: $p < 0.05$; **: $p < 0.01$)

BOX 3.4 Drug effects in cancer research

In cancer research, experimental results are expressed in a rather unusual way (at least for those people not working in this research field). In xenograft models in immunodeficient mice, in which human tumoral cells are implanted subcutaneously, tumor growth is repeatedly evaluated by automated devices measuring tumor size on living mice and converted into tumor weight. In early stage tumors, results are expressed as T/C (%), the median tumor weight in treated group (T) versus median tumor weight in control (C) group when the latter is approximately 1,000 mg (as guidelines, a T/C < 10% implies a high antitumor activity; the drug is inactive with a T/C > 42%). In advanced stage tumors, tumor growth delay (T – C in days) is first evaluated, that is median tumor time for treated (T) minus control (C) tumor groups to reach a pre-determined size, but the results are generally expressed as log cell kill gross (lckg : (T – C)/[3.32 × tumor doubling time]), or for treatments over 10 days, as log cell kill net (lckn : (T – C)–duration of treatment/[3.32 × tumor doubling time]). A lckn < 0 means that the tumor grows under treatment; the drug is cytostatic (blocks the proliferation of tumor cells) with a lckn ≈ 0 and cytotoxic (kills tumor cells) with a lckn > 0 (a highly active antitumor drug will display a lckn > 2 for a treatment duration of 5–20 days). Expressing results as log cell kill allows quantitative comparison of antitumor efficacy between different treatments.

since, if a biological marker is available (for purposes of comparison between the animal and man), it may help the clinician to determine the dose that can be administered to humans that will display maximal efficacy and minimal drug-related risks (Box 3.4).

The route of administration is an important aspect of an *in vivo* experiment. Drugs may be administered in many ways but the most widely used are orally postoperative (p.o.), intraperitoneally (i.p., in the abdomen), intravenously (i.v.), subcutaneously (s.c.) and intracerebroventricularly (i.c.v., directly in the cerebrospinal fluid into the brain) although other routes (intra-thecally, trans-dermally …) may also be used. In early *in vivo* experiments the drug is generally administered i.p. since this route is easy to use in rodents, bypasses possible gastric absorption problems and is successful even for compounds with poor solubility. In more complex models, the choice of a route of administration depends on the targeted pathology, the physicochemical properties of the drug and aim of the study. For treating acute, life-threatening insults (heart or brain infarcts) the drug has to reach its site of action as quickly as possible and drugs will be injected i.v. This can be performed in the awake mouse but generally requires anesthesia or arterial catheterization in other species and the major issue is drug solubility. In most other pathologies, particularly those that require long-term treatment (e.g. depression or hypertension) the oral route will be selected, the drug being administered by oral gavage, gastric tubing or inclusion in the food. There are of course exceptions for drugs with a proven therapeutic utility and that are poorly absorbed and/or quickly metabolized (insulin for diabetes) and/or that display high systemic toxicity (anticancer drugs) in which case the s.c. or i.v. route will be selected. The i.c.v. route is devoted to proof-of-concept experiments for drugs acting on a cerebral target, that is, to ascertain that the drug has the mechanistic or therapeutic effects expected in the absence of any other interfering parameter (crossing of blood-brain barrier, absorption, metabolism …).

REFERENCES

1. Overington, J. P., Al-Lakizani, B., Hopkins, A. L. How many drug targets are there? *Nat. Rev. Drug Discov.* 2006, *5*, 993–996.
2. Thirone, A. C. P., Huang, C., Klip, A. Tissue-specific roles of IRS proteins in insulin signalling and glucose transport. *Trends Endocrinol. Metab.* 2006, *17*(2), 72–78.
3. Scatchard, G. The attraction of proteins for small molecules and ions. *Ann. NY Acad. Sci.* 1949, *51*, 660–672.
4. Cheng, Y. C., Prusoff, W. H. R. Relationship between the inhibition constant (K_i) and the concentration of inhibitor which causes 50 percent inhibition (I_{50}) of an enzymatic reaction. *Biochem. pharmacol.* 1973, *22*, 3099–3108.
5. Kenakin, T. Efficacy as a vector: the relative prevalence and paucity of inverse agonism. *Mol. Pharmacol.* 2004, *65*, 2–11.
6. Kenakin, T. Inverse, protean, and ligand-selective agonism: matters of receptor conformation. *FASEB J.* 2001, *15*, 598–611.
7. Brink, C. B., Harvey, B. H., Bodenstein, J., Venter, D. P., Oliver, D. W. Recent advances in drug action and therapeutics: relevance of novel concepts in G-protein-coupled receptor and signal transduction pharmacology. *Br. J. Clin. Pharmacol.* 2004, *57*, 373–387.
8. Christopoulos, A., Kenakin, T. G-protein-coupled receptor allosterism and complexing. *Pharmacol. Rev.* 2002, *54*, 323–374.

9. Kenakin, T. Allosteric modulators: the new generation of receptor antagonist. *Mol. Interv.* 2004, *4*, 222–229.
10. Raddatz, R., Schaffhauser, H., Marino, M. J. Allosteric approaches to the targeting of G-protein-coupled receptors for novel drug discovery: a critical assessment. *Biochem. Pharmacol.* 2007, *74*, 383–391.
11. May, L. T., Leach, K., Sexton, P. M., Christopoulos, A. Allosteric modulation of G protein-coupled receptors. *Annu. Rev. Pharmacol. Toxicol.* 2007, *47*, 1–51.
12. Maguire, J. J., Davenport, A. P. Regulation of vascular reactivity by established and emerging GPCRs. *Trends Pharmacol. Sci.* 2005, *26*, 448–454.
13. Zhang, Y. J., Xu, Y. F., Liu, Y. H., Yin, J., Wang, J. Z. Nitric oxide induces tau hyperphosphorylation via glycogen synthase kinase-3beta activation. *FEBS Lett.* 2005, *579*, 6230–6236.
14. Pani, L., Pira, L., Marchese, G. Antipsychotic efficacy: relationship to optimal D(2)-receptor occupancy. *Eur. Psychiatry*, (in press) 2007.
15. Medhurst, A. D., Atkins, A. R., Beresford, I. J., Brackenborough, K., Briggs, M. A., Calver, A. R., Cilia, J., Cluderay, J. E., Crook, B., Davis, J. B., Davis, R. K., Davis, R. P., Dawson, L. A., Foley, A. G., Gartlon, J., Gonzalez, M. I., Heslop, T., Hirst, W. D., Jennings, C., Jones, D. N., Lacroix, L. P., Martyn, A., Ociepka, S., Ray, A., Regan, C. M., Roberts, J. C., Schogger, J., Southam, E., Stean, T. O., Trail, B. K., Upton, N., Wadsworth, G., Wald, J., White, T., Witherington, J., Woolley, M. L., Worby, A., Wilson, D. M. GSK189254, a novel H3 receptor antagonist that binds to histamine H3 receptors in Alzheimer's disease brain and improves cognitive performance in preclinical models. *J. Pharmacol. Exp. Ther.* 2007, *321*, 1032–1045.
16. Carreau, A., Duval, D., Poignet, H., Scatton, B., Vigé, X., Nowicki, J. P. Neuroprotective efficacy of *N* omega-nitro-L-arginine after focal cerebral ischemia in the mouse and inhibition of cortical nitric oxide synthase. *Eur. J. Pharmacol.* 1994, *256*, 241–249.
17. Barone, F. C., Irving, E. A., Ray, A. M., Lee, J. C., Kassis, S., Kumar, S., Badger, A. M., White, R. F., McVey, M. J., Legos, J. J., Erhardt, J. A., Nelson, A. H., Ohlstein, E. H., Hunter, A. J., Ward, K., Smith, B. R., Adams, J. L., Parsons, A. A. SB 239063, a second-generation p38 mitogen-activated protein kinase inhibitor, reduces brain injury and neurological deficits in cerebral focal ischemia. *J. Pharmacol. Exp. Ther.* 2001, *296*, 312–321.

Chapter 4

Molecular Drug Targets

Jean-Pierre Gies and Yves Landry

I. INTRODUCTION
- A. How many drug targets for how many drugs?
- B. From the drug target to the response of the organism
- C. Drug binding, affinity and selectivity
- D. Various ligands for a single target

II. ENZYMES AS DRUG TARGETS
- A. Targeting human enzymes
- B. Targeting enzymes selective of invading organisms

III. MEMBRANE TRANSPORTERS AS DRUG TARGETS
- A. Established drug targets among membrane transporters
- B. Progress in the pharmacological control of membrane transporters

IV. VOLTAGE-GATED ION CHANNELS AS DRUG TARGETS
- A. Voltage-gated sodium channels (Na_V channels)
- B. Voltage-gated calcium channels (Ca_V channels)
- C. Potassium channels

V. NON-SELECTIVE CATION CHANNELS AS DRUG TARGETS

VI. DIRECT LIGAND-GATED ION CHANNELS (RECEPTORS WITH INTRINSIC ION CHANNEL)
- A. P2X-ATP receptors
- B. Glutamate-activated receptors
- C. The "Cys-loop receptor superfamily"

VII. RECEPTORS WITH INTRINSIC ENZYME ACTIVITY
- A. Receptors with guanylate cyclase activity
- B. Receptors with serine/threonine kinase activity
- C. Receptors with tyrosine kinase activity

VIII. RECEPTORS COUPLED TO VARIOUS CYTOSOLIC PROTEINS
- A. Receptors coupled to the cytosolic tyrosine kinase JAK and Btk tyrosine kinases (immunoreceptors)
- B. Receptors coupled to the cytosolic Src, Zap70/Syk
- C. Receptors coupled to the cytosolic serine/threonine kinase IRAK
- D. Receptors coupled to caspases and to NFκB
- E. Receptors of the cellular adhesion

IX. G-PROTEIN-COUPLED RECEPTORS
- A. How many druggable GPCRs?
- B. Diversity of G-proteins
- C. Diversity of GPCR-elicited signaling and related drug targets

X. NUCLEAR RECEPTORS AS DRUG TARGETS

REFERENCES

Only such substances can be anchored at any particular part of the organism, as fit into the molecules of the recipient complex like a piece of mosaic finds its place in a pattern.

Paul Ehrlich (1854–1915)

Copyright © 2008, Elsevier Inc.
All rights reserved.

I. INTRODUCTION

Most drug targets are cellular proteins undergoing a selective interaction with chemicals called drugs because they are administered to treat or diagnose a disease. These targets are human-genome-derived proteins, or belong to bacterial, viral, fungal or other pathogenic organisms. A limited set of drugs act through physicochemical mechanisms, or have unknown mechanism of action.

A. How many drug targets for how many drugs?

Analysis of the human genome in 2002 has led to the estimation of ±8,000 targets of pharmacological interest.[1] In 2006, Wishart *et al.* reported 14,000 targets for all approved and experimental drugs, although they revise this number to 6,000 targets on the DrugBank Database web site.[2] Only a small part of these targets relates to approved drugs. In 1996–1997, Drews[3,4] estimated the molecular drugs targets corresponding to all marketed drugs to be 483. This number was overestimated. In 2003, Golden proposed that all then-approved drugs acted through 273 proteins.[5] In 2006, Zheng *et al.* disclose 268 "successful" targets in the current version of the therapeutic Targets Database,[6,7] and Imming *et al.* cataloged 218 molecular targets for approved drug.[8] A consensus number of 324 drug targets for all classes of approved therapeutic drugs was proposed by Overington *et al.*[9] Of these, 266 are human-genome-derived proteins, and 58 are bacterial, viral, fungal or other pathogenic organism targets.

The discrepencies between these estimations arise from the criteria choose by each authors, such as including or not drugs under clinical trials but not yet approved, or considering or not the multiple relevant targets for a unique drug, including isoenzymes or different members of a receptor family. However, some interesting features can be drawn from such studies.

The analysis by Overington *et al.*[9] identifies in excess of 21,000 drug products marketed in US corresponding to 1,357 unique drugs, of which 1,204 are "small molecule drugs"(including 192 prodrugs) and 166 are "biological drugs." Twenty seven percent of these drugs bind to G-protein-coupled receptors (GPCR), 13% to nuclear receptors; 7.9% to ligands-gated ion channels, and 5.5% to voltage-gated ion channels. A selected target may have a unique approved drug, or a large number of me-too molecules.

The analysis by Imming *et al.*[8] gives an accurate view of the different biological classes of the 218 listed targets: 66 human enzymes and 20 bacterial, viral, fungal or parasital enzymes; 20 families of GPCR, each family including up to 5 members; 12 nuclear receptors for steroids and others; 7 cytokine receptors; and about 10 ions channels and 10 transport proteins of the plasma membrane. Altogether, these studies confirm that a very large number of putative drug targets remains to be explored.

B. From the drug target to the response of the organism

To understand drug actions, it is necessary to consider the effects induced by the drug on the biological system at various levels of complexity of organization. The main steps, can be designated as follows: binding to the cellular molecular target, signaling events leading to a cellular response (i.e. secretion, contraction and metabolism), integration at the level of tissues and organs corresponding to a modification of a physiological function (i.e. digestion, motricity, cardiovascular processes). Thus, drugs act by increasing or decreasing a normal function, but do not endow the organism with new functions. Although, gene therapy may soon challenge this principle, it remains valid for the immediate future.

The vast majority of drugs produce their effects by interacting with proteins, either with those on the surface of the cell comprising the plasma membrane, such as receptors of mediators, ionic channels and transporters (about 60% of drugs), or with components of the interior of the cell, such as enzymes and nuclear receptors.

Some others act extracellularly at non-cellular constituents of the body without involving a drug–receptor interaction. The simplest example is that of the neutralization of gastric acid by antacid drugs. In this reaction the excess of gastric acid is neutralized by a base such as sodium bicarbonate. This reaction is not considered as a drug–receptor interaction, since no macromolecular component is involved. Other types of extracellular mechanisms can be illustrated, for example, by the action of heparin which prevents blood coagulation.

Other mechanisms of drug action may occur at cellular sites and may involve macromolecular components, but the biological effects produced are non-specific consequences of the chemical properties of the drugs. Detergents, alcohol, oxidizing agents, phenol derivatives act by destroying the integrity of the cell through disrupting the cellular constituents.

A number of other molecular interactions between drugs and the components of the biological system may occur, such as the binding of drugs to plasma albumin. Serum albumin can transport drugs in the circulation to organs, and it can hold drugs up, preventing them from binding to their site of action. Those interactions affect the duration of the drug action or its rate of actions. Albumin might then be considered as an acceptor site for the drug rather than a target or receptor.

C. Drug binding, affinity and selectivity

Corpora non agunt nisi fixata
Compounds do not act unless bound
Paul Ehrlich (1854–1915)

The receptor concept was formulated by Langley and the term "receptor" was proposed by Ehrlich.[10] The concept of target binding or "receptor binding," *Corpora non agunt nisi fixata* (compounds do not act unless bound), has been subject to refinement but is still valid. The term "receptor" should be now restricted to the target of endogenous mediators but is often extented to the targets of exogenous compounds endowing various biochemical functions.

The various physicochemical interactions between a ligand and the target co-operate to establish the target–drug interaction:

- *Hydrophobic interactions* plays an important role in stabilizing the conformation of proteins and in the association of hydrophobic structure between the drug and its target.
- *Hydrogen bonding* is strongly directional and has considerable importance both in the maintaining the secondary and tertiary structure of the target itself and in the target–drug interaction.
- *Charge transfer complexes* formed between electron-rich donor molecules and electron-deficient acceptors are also often involved in drug–target interaction.
- *Ionic bonds* are of importance in the actions of ionizable drugs since they act across long distances; ionic bonds result from the electrostatic attraction that occurs between oppositely charged ions; most targets have a number of ionizable groups (COO^-, —O—, NH_3^+) at physiological pH that are available for the binding with charged drugs.
- *Covalent bonds* resulting in the formation of a long-lasting complex are less important in drug–target interaction. Although most drug–target interactions are readily reversible, some drugs, such as anticancer nitrogen mustards and alkylating compounds form reactive cationic intermediates (i.e. aziridinium ion) that can react with electron donor groups on the target.

These chemical interactions are related to the affinity of the drug for its target. The medium affinity of current small molecule drugs is about 20 nM, ranging from 200 mM to 10 pM[9]. A high affinity for a therapeutically relevant target is usually considered as a criterium for selectivity of the drug, with less risks to bind to targets inducing undesirable or toxic effects. But the strict application of this concept would eliminate any low affinity drugs whereas some have proved therapeutical interest. No drug can be considered *specific* of a single target, but only *selective*, according to the dose, *in vivo*, or the concentration, *in vitro*, used.[11]

D. Various ligands for a single target

Most of target types can be stimulated or inhibited depending of the ligand choosen. This leads to opposite regulations of related cellular functions. Terms used to characterize these different ligand types differ according to the biochemical nature of the targets.

Enzyme ligands more often lead to the inhibition of the enzyme activity, binding the active site with competition with the substrate (*competitive inhibitors*) or to allosteric sites (*non-competitive inhibitors*). Activation of an enzyme is more difficult to proceed unless giving or generating an excess of substrate or co-substrate. However some drugs are known to activate enzymes by direct binding, that is, forskolin for adenylyl cyclase.

Membrane transporters and ion channels permeability can be increased or decreased by direct binding of selected drugs termed *openers* and *inhibitors* (or *blockers*), respectively. However, such ligands are too often improperly referred to as *agonists* and *antagonists.*

Receptors of mediators are able to interact with a large diversity of ligand types (Figure 4.1):

- Agonists mimic the effects of endogenous mediators (neurotransmitters, hormones, cytokines …). Thus, mediators are considered the endogenous, or physiological, agonists of their receptors. Some exceptions to this concept are now known, some couples of mediators acting through the binding to a single receptor with agonist or antagonist properties respectively, i.e. interleukin 1 and IRAP, RANK-L and OPG, MSH and AGRP. Full agonists elicit a maximal response of the organism, usually similar to that of the mediator. Partial agonists elicit a partial response of the organism, and prevent the binding of the mediator. Thus the related function of the organism is decreased.
- Neutral antagonists prevent the binding of the mediator and thus abolish downstream signaling biochemical events and physiological responses. Most neutral antagonists bind to the agonist binding site.
- Inverse agonists also termed "negative antagonists" have been found among antagonists. Like neutral antagonists, they prevent the binding of agonists, including mediators, but elicit a response inverse to that of agonists. This has been first shown for ligands of the benzodiazepine binding site of GABA-A (γ-aminobutyric-acid) receptors,[12] whose "agonists" (classical benzodiazepines)

FIGURE 4.1 Mode of action of agonists, neutral antagonists and inverse agonists.

potentiate the opening of the intrinsic chloride channel elicited by GABA, whereas "inverse agonists" decrease it with opposite reponses, i.e. anxiolytic effect for agonists and anxiogenic effect for inverse agonists. This concept of inverse agonism has been first extended to opioid receptors[13] and then to others GPCR([14,15] for reviews), showing that such ligands decreased the constitutive activity of the receptor, e.g. its "activity" noticable in the absence of mediator.

Receptors of mediators including an intrinsic ion channels (ligand-gated ion channels such as nicotinic receptors), or an enzyme activity (i.e. a tyrosine kinase activity such as insulin receptors, or ganylyl cyclase activity such as ANF receptors) have ligands for their receptor part (agonists and antagonists) as well as for their ion channel (openers and inhibitors or blockers) or enzyme part (inhibitors).

The activity of these various target types can also be modulated indirectly through intracellular signaling, for instance by phosphorylation elicited by protein kinases or dephosphorylation involving protein phosphatases, or by protein–protein interactions such as regulations induced by interaction with the calciprotein calmoduline. This offers large alternatives to modify the status of a putative target through indirect ways when direct targeting has been unsuccessful.

II. ENZYMES AS DRUG TARGETS

At least 66 human enzymes and 20 bacterial, viral, fungal or parasital enzymes, are targets for approved drugs,[8] for example up to 40% of current targets. Note that several thousands enzymes are coded in the human genome, opening large opportunities to develop new drugs.

The basis of using enzyme inhibitors as drugs is that inhibition of a suitable selected enzyme leads to a build-up in concentration of substrate and a corresponding decrease in concentration of the metabolite, leading to a useful clinical response. Enzyme inhibiting processes may be divided into two main classes, reversible and irreversible, depending on the manner in which the drug is attached to the enzyme. Reversible inhibition occurs when the inhibitor is bound to the enzyme through a suitable combination of Van der Vaals', electrostatic, hydrogen bonding and hydrophobic attractive forces. However, there are also covalent but reversible inhibitors, for example, some alhehydes and activated ketones as serine protease inhibitors. Reversible inhibitors may be competitive, non-competitive, uncompetitive, or of mixed type. During irreversible inhibition, after initial binding of the inhibitor to the enzyme, covalent bonds are formed between a functional group on the enzyme and the drug. This is the case for the active-site-directed inhibitors (affinity labeling). However, many active-site-directed inhibitors (such as non-covalent inhibitors) are completely reversible.

A. Targeting human enzymes

The inhibitors used in therapy should possess a high selectivity toward the target enzyme, since inhibition of closely related enzymes may lead to a range of side effects.

This concept has led to market isoenzyme selective inhibitors, that is, monoamine oxidase (MAO) inhibitors (*moclobemide* for MAO-A as an antidepressive drug, *selegiline* for MAO-B in Parkinson disease), selective inhibitors for various cyclic nucleotide phosphodiesterases (*sildenafil* for PDE5), and selective cyclooxygenase inhibitors (*celecoxib* for cox2).

However, the claimed selectivity often remains quite low. For instance, *imatinib*, originally developed as a highly selective inhibitor of the tyrosine kinase activity of c-ABL (approved for chronic myeloid leukemia), has subsequently been discovered to also inhibit tyrosine kinase activity of c-KIT and PDGFR. In that case, the lack of selectivity offers the opportunity to extend the clinical utility of the drug.[16]

The inhibition of multifunctional enzymes can also have therapeutic interest. The 26S proteasome is a multicatalytic intracellular protease complex expressed in eukaryotic cells. This complex is responsible for selective degradation of intracellular proteins that are responsible for cell proliferation, growth, regulation of apoptosis and transcription of genes.[17] Thus, proteasome inhibition is a potential treatment option for cancer and diseases due to aberrant inflammation conditions. Bortezomib and PS-519 are the first proteasome inhibitors that have entered clinical trials. In multiple myeloma, both the FDA (United States

Food and Drug Administration) and EMEA (European Medicine Evaluation Agency) granted an approval for the use of bortezomib for the treatment of relapsed multiple myeloma. At present, several phase II and phase III trials in hematological malignancies and solid tumors are ongoing. PS-519 that focuses on inflammation, reperfusion injury and ischemia is currently under evaluation for the indication of acute stroke.[17]

B. Targeting enzymes selective of invading organisms

Interestingly, targeted enzymes of invading organisms may have no functional equivalent in human cells. For example, the unique properties of HIV-1 integrase makes it an ideal target for drug design. HIV-1 integrase is essential for retroviral replication, being involved in the integration of HIV DNA into host chromosomal DNA. HIV-1 integrase has been recently validated as a legitimate target and the results from the molecules like S-1360, JKT-303 which are under phase II/III clinical trials suggest synergistic effect with reverse transcriptase and protease inhibitors.[18]

Another recent example arises from the scrutiny of the 2-C-methyl-D-erythritol-4-phosphate pathway for isoprenoid biosynthesis where key enzyme is 1-deoxy-D-xylose 5-phosphate reductoisomerase (DXR). DXR have no functional equivalent in humans making it an attractive target for novel antimalarial, antibacterial and herbicidal agents.[19]

III. MEMBRANE TRANSPORTERS AS DRUG TARGETS

Membrane transporters constitute a rather small family of drug targets. Some of them have yet to reveal their relevant therapeutic interest, that is, as targets of diuretics or antidepressive drugs, but some others still resist to pharmacological control, that is, CFTR (Cystic Fibrosis Transmembrane conductance Regulator) for cystic fibrosis therapy and multidrug resistance transporter to improve cancer therapy.

Transporters genes encode proteins, generally constituted by 12 transmembrane spanning regions. These mediated Na^+ or H^+ dependent accumulation of small molecules such as neurotransmitters, antibiotics, ions and cationic amino-acid transporters into the cells or organelles. The transport is performed by different mechanisms: uniport, substrate-ion symport, substrate-ion antiport, substrate-substrate or ion-ion antiport, and ATP-dependent translocation.

A. Established drug targets among membrane transporters

Most success stories in this field concern old drugs whose targets have been often discovered after their efficient clinical use. They include: *cardiac glycosides* (sodium pump e.g. Na^+/K^+-ATPase); *omeprazole* and analogs (proton pump e.g. H^+/K^+ ATPase); *artemisinin* and derivatives (plasmodial sarcoplasmic and endoplasmic calcium ATPase, SERCA); diuretics (*thiazides* for the Na^+/Cl^- co-transporter, NCC; *furosemide* for the $Na^+/K^+/Cl^-$ co-transporter, NKCC); *reserpine, ephedrine* and *amphetamines* (vesicular monoamine transporter, VMAT); the antidepressant *paroxetine* for serotonin/Na^+ symporter, SERT); *cocaine* and the antidepressant *imipramine* for the norepinephrine, dopamine and serotonin/Na^+ symporters, NET, DAT and SERT. Let notice again the absence of selectivity of latter drugs for targets of a molecular and functional family, for example, monoamine transporters of neurons.

B. Progress in the pharmacological control of membrane transporters

ATP-binding cassette (ABC) transporters, including multidrug resistance transporters and CFTR are putative drug targets, but progress in finding drugs of clinical interest remains very slow.

Multidrug resistance is a serious impediment to improved healthcare. Multidrug resistance is most frequently due to active transporters, such as the P-glycoprotein (ABCB1) identified 30 years ago, that pump a broad spectrum of chemically distinct molecules out of the cells, including antibiotics, antimalarials and cancer chemotherapeutics in humans.[20]

Around 40% of human tumors develop resistance to chemotherapeutic drugs due to the overexpression of ABC proteins. Nonetheless, success in overcoming or circumventing multidrug resistance in a clinical setting has failed. A first approach has been to modify the structure of drugs so that they are no longer substrates for ABC transporters. But any modification to a drug that substantially reduces its affinity for a transporter also tends to reduce its ability to cross the cell membrane and bind to its target. The second approach to overcome multidrug resistance, the development of inhibitors of ABC transporters, has also proved unsatisfactory.[21]

CFTR discovered 20 years ago, is a cAMP-activated chloride channel, acting as an ATP-dependent pump with ATPase activity, expressed in epithelia in the lung, intestine, pancreas and other tissues, where it facilitates transepithelial fluid transport. In the intestine, CFTR provides the major route for chloride secretion in certain diarrheas. Mutations in CFTR cause the hereditary disease cystic fibrosis, where chronic lung infection and deterioration in lung function cause early death.

Small molecule modulators of CFTR function may be useful in the treatment of cystic fibrosis, secretory diarrhea and polycystic kidney disease. The most common mutation in the CFTR gene, F508 deletion (ΔF508), causes retention of F508-CFTR in the endoplasmic reticulum and

leads to the absence of CFTR Cl^- channels in the plasma membrane. Recently, *curcumin*[22] was shown to rescue ΔF508-CFTR localization and function. *Benzothiophene, phenylglycine and sulfonamide* potentiators were also identified[23] that correct the defective gating of ΔF508-CFTR chloride channels, and other small molecules that correct its defective cellular processing. Others mutations of CFTR, like G551D and G1349D (glycine to aspartic acid change at position 551 or 1349), cause only a gating defect. Antihypertensive 1,4-dihydropyridines, a class of drugs which block voltage-dependent calcium channels, have been identified as effective potentiators of CFTR gating, able to correct the defective activity of CFTR mutants.[24] Optimization of potency for CFTR versus calcium channels is in progress.[25]

IV. VOLTAGE-GATED ION CHANNELS AS DRUG TARGETS

Ion channels are essential for a wide range of functions such as neurotransmitters secretion and muscle contraction. Ion channels mediate Na^+, Ca^{2+}, K^+ and Cl^- conductance induced by membrane potential changes. These channels propagate action potentials in excitable cells and are also involved in the regulation of membrane potential and intracellular Ca^{2+} transients in most eukaryotic cells. About 300 genes code for subunits of voltage-gated ion channels. Na^+, Ca^{2+} and K^+ channels are drug targets. The pharmacology of Cl^- channels is not yet developed.

A. Voltage-gated sodium channels (Na_V channels)

Those play a critical role in initiating the action potential. The activation of the channels allows for the inward movement of Na^+ from the extracellular space of the cell. The Na_V channels from brain and striated muscles are heterooligomeric and composed of alpha and beta subunits (α- and β-subunit) (Figure 4.2). The α-subunit, with its 24 transmembrane helices, determines the major functional characteristics of Na_V channels. The human genome contains nine genes encoding the main α-subunit of Na_V channels and at least four genes encoding auxiliary β-subunits (1 transmembrane helix), the expression of which is tissue-specific.[26]

Natural toxins, like tetrodotoxin and saxitoxine, have not found any therapeutic application. Plant toxins, like pyrethrins and pyrethroids are currently used as insecticides. Numerous synthetic drugs, now proposed to be inhibitors of sodium channels, have been used before the determination of channels structure and diversity. This includes local anesthetics (*lidocaine* and analogs), class 1 antiarythmics (*disopyramide, flecaine and quinidine*) and some antiepileptics of first (*phenytoine and carbamazepine*) or second generation (*lamotrigine, topiramate and felbamate*). The selectivity

FIGURE 4.2 Structure of the voltage-gated sodium channels (Na_V channels).

of these drugs for the different sodium channels is not yet fully determined. The chemistry of selective ligands of Na_V channels deserves to be developed.

B. Voltage-gated calcium channels (Ca_V channels)

Ten different genes encode different α-subunits (24 transmembrane helices) composing the voltage-gated Ca^{2+} channels. Ca_V1 (α_{1S}, α_{1C}, α_{1D} and α_{1F}) mediate L-type Ca^{2+} currents; Ca_V2 (α_{1A}, α_{1B} and α_{1E}) mediate P/Q-type, N-type, and R-type Ca^{2+} currents, respectively; and Ca_V3 (α_{1G}, α_{1H}, α_{1I}) mediate T-type currents. The α1-subunits co-assemble with α2-, β-, δ- and γ-subunits to form functional channels in different tissues.

Ca_V1 channels (L-type) are targets of "calcium-channel blockers" or "calcium antagonists" which decrease the influx of Ca^{2+} in cardiac and smooth muscle vascular cells: dihydropyridines (*nifedipine* and analogs), phenylalkylamines (*verapamil*) and benzothiazepines (*diltiazem*), widely used as antihypertensive, antianginal and antiarrhythmic drugs. Ca_V1 openers, like Bay K 8644, have been synthesized but have not found any therapeutic interest.

Ca_V2 channels (N-type) control the release of neurotransmitters at the presynaptic level. It selective blocker, *ziconotide*, a synthetic peptide analog of an ω-conotoxin, has recently been approved for the intrathecal treatment of severe chronic pain.[27]

Ca_V3 channels (T-type) have only recently become targets of interest. The recent availability of cloned T-channels, facilitates identification of novel Ca_V3 blockers. Selective inhibition of T-channels may have clinical importance in cardiovascular diseases, some forms of epilepsy, sleep disorders, pain and possibly cancer.[28]

C. Potassium channels

Potassium channels are highly heterogeneous and are thus interesting drug targets. They are classified on the basis of the structure of the pore-containing unit (α-subunit) and/or of their regulatory processes (Figure 4.3).

Voltage-gated K^+ channels: K_V1 to K_V9, which α-subunit contains 6 transmembrane helices with a single pore, including also the slow delayed rectifiers $K_{V(s)}$ (or K_V LTQ), and the rapid delayed rectifiers $K_{V(r)}$ (or K_VEAG-like). The old experimental blocker of K_V channels, *4-aminopyridine*, is in phase 3 in the treament of multiple sclerosis.[29]

Voltage and G-protein-gated K^+ channels: K_{Ach} and K_M (6 helices) quite similar to $K_{V(s)}$ and $K_{V(r)}$, but which interact with Gi proteins coupled to M2 muscarinic acetylcholine receptors, or with Gq proteins coupled to M3 receptors, respectively.[30] Stimulation of M2 or M3 muscarinic receptors by acetylcholine in pacemaker cardiac

FIGURE 4.3 Structure of the α-subunit of the potassium channels.

cells activates K^+ currents, heart rate is thus slowed by hyperpolarization of the pacemaker depolarization potential as well as by block of tonic β-adrenergic stimulation of depolarizing pacemaking channels (see below, under KCN channels).

Voltage and calcium-activated K^+ channels: BK_{Ca} (7 helices), voltage sensitive and activated by direct binding of intracellular calcium; IK_{Ca} and SK_{Ca} (6 helices), poorly sensitive to voltage and activated by calcium through calmoduline. BK and IK channels openers are studied in view of the prevention and treatment of cardiovascular disorders.[31]

ATP-dependent K^+ channels (K_{ATP}) are composed of four inwardly rectifying K^+ channel subunits (Kir subunits, 2 helices) and four regulatory sulfonylurea receptors (SUR, 17 helices) (Figure 4.4). The various subunits are tissue-selective. These channels are inhibited by intracellular ATP and by sulfonylurea agents. Sulfonylureas (*tolbutamide, glibenclamide*) and glinides (*repaglinide*) stimulate insulin secretion via blockade of the pancreatic β-cell K_{ATP} channel, the induced depolarization of the cell membrane stimulating the opening of Ca_V channels. Pharmacological vasodilators such as *cromakalim*, *pinacidil* and *diazoxide* are openers, directly activating K_{ATP}. The associated membrane hyperpolarization closes Ca_V channels, leading to a reduction in intracellular Ca^{2+} and vasodilation.

Two-pore tandem K^+ channels are responsive for the background K^+ conductance in the cell at rest. Their α-subunit are arranged in 4 transmembrane helices determining $2K^+$ pores. Fifteen mammalian genes encode these channels including TASK1-3 (involved in chemoreception in respiratory motor neurons), TREK1-2 (expressed in neurons involved in thermogegulation), TWIK1-2 and other. They are controlled by several stimuli like oxygen tension, pH and mechanical stretch.

V. NON-SELECTIVE CATION CHANNELS AS DRUG TARGETS

These channels have 6 transmembrane helices and are considered non-elective for Na^+, K^+ and Ca^{++}, although their opening often correspond to membrane depolarization with Na^+

FIGURE 4.4 Structure of the ATP-dependent K^+ channels (K_{ATP}).

and Ca^{++} influx. *HCN channels* (hyperpolarization-activated cyclic nucleotide-gated cation channels) are present in the heart (HCN1, 2 and 4) and brain (HCN1 to 4). They are activated by hyperpolarization with potentiation induced by direct binding of cyclic AMP (cAMP) to their intracellular C-terminus. The opening of HCN channels of pacemaker cells increases cardiac rate (I_f current) due to cAMP generated through the activation of β-adrenergic receptors (see above the inverse effect of K_{Ach} and K_M channels). The If blocker *ivabradine*[32] has been recently approved to slow heart rate in angina.

CNG channels (cyclic nucleotide-gated ion channels) are activated by the binding of cyclic GMP. They mediate sensory signal transduction in photoreceptors and olfactory cells. Six mammalian CNG channel genes are known and some human visual disorders are caused by mutations in retinal rod or cone CNG genes.[33]

TRP channels (transient receptor potential ion channels) are poorly voltage sensitive. The 28 mammalian TRP channels include six main subfamilies: TRPC (canonical), TRPV (vanilloid), TRPM (melastatin), TRPP (polycystin), TRPML (mucolipin) and TRPA (ankyrin). TRP channels are expressed in almost every tissue and cell type and play an important role in the regulation of vascular tone, thermosensation, irritant stimuli sensing and flow sensing in the kidney.[34] Moreover, recent data concerning TRP vanilloid (TRPV) type 6, TRP melastatin (TRPM) type 1 and 8 channels indicate their relevance for common human cancer types.[35] Numerous ligands of TRP channels have been recently proposed,[36] such as analogs of capsaicin for TRPV in view to develop new peripheral analgesics.

VI. DIRECT LIGAND-GATED ION CHANNELS (RECEPTORS WITH INTRINSIC ION CHANNEL)

Direct ligand-gated ion channels are homomultimeric or heteromeric proteins that span the cell membrane and include both a binding site for neurotransmitters and an ion-conducting pore. These receptors transfer their signals by altering the cell's membrane potential and the cytoplasmic ionic composition. They control the fastest synaptic events in the nervous system by increasing transient permeabilities (Figure 4.5).

Excitatory neurotransmitters, like acetylcholine and glutamate, induce an opening of cation channels; these channels are relatively unselective for cations, but this results in a net Na^+ inwards current, which depolarizes the cell and increases the generation of action potentials occuring in a fraction of a millisecond; in this way the receptor converts a chemical signal (neurotransmitter) into an electrical signal (depolarization).

Inhibitory neurotransmitters, like GABA and glycine decrease the firing of the action potential by opening of anionic channels which results in an inwards flux of Cl^- with slight hyperpolarization.

This relatively small group of drug targets is especially important because, besides agonists and competitive antagonists, both positive and negative *allosteric effectors* have been developed with therapeutic relevancies.

Also, each type of these ion channels with receptor property exists in multiple molecular forms, offering large putative opportunities for selective ligands. They are divided into three main families according to the number

FIGURE 4.5 Gated ion channel receptors, receptors with intrinsic enzymatic activity and receptors associated with cytosolic tyrosine kinases.

FIGURE 4.6 Structure of different ligand-gated ion channels. (a) Schematic representation of a single subunit from three families of ligand-gated ion channels. (b) Proposed topography of the GABA-A receptor. Left: a cross section in the plane of the membrane. The receptor complex contained 20 membrane-spanning segments (5 × 4) surrounding the central ion channel. Right: schematic representation of a ligand-gated ion channel. The different subunits with the putative-binding sites for allosteric ligands are represented.

of transmembrane helices (2, 3 or 4) present in the subunits that form the channels (Figure 4.6a).

A. P2X-ATP receptors

P2X receptors for ATP (adenosine triphosphate) are formed from the homotrimeric or heterotrimeric assembly of seven different receptor subunits (P2X1–7) to give a range of phenotypes. Each subunit contains 2 transmembrane helices separated by a large extracellular loop. P2X are present in most cells like neurons and smooth muscle cells, and are putative drug targets.[37] Other purinergic receptors (ADP-P2 Y receptors and adenosine-A receptors) belong to the GPCR superfamily.

B. Glutamate-activated receptors

These cationic channels (also referred to as "ionotropic receptors" in contrast to "metabotropic receptors" which are GPCR activated by glutamate), include *N*-methyl-d-aspartate receptors (NMDAR), α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic-acid receptors (AMPAR) and kainate receptors. They are widely expressed in the central nervous system where they play key roles in excitatory

synaptic transmission, neuronal plasticity and long-term potentiation (LTP) involved in memory processes.

NMDAR are tetrameric complexes incorporating different subunits within a repertoire of three subtypes: NR1, NR2 and NR3. There are eight different NR1 subunits generated by alternative splicing from a single gene, four different NR2 subunits (A, B, C and D) and two NR3 subunits (A and B); the NR2 and NR3 subunits are encoded by six separate genes. Several approved drugs are non-competitive antagonists (channel blockers) of NMDAR: *phencyclidine* and ketamin (general anesthetics), amantadine (Parkinson) and memantine (Alzheimer). Given the growing body of evidence that diverse brain disorders implicate different NMDAR subtypes, such as NR2B in pain or NR3A in white matter injury, there is a growing interest in exploiting the pharmacological heterogeneity of NMDARs for the development of novel NMDAR subtype-selective compounds.[38]

AMPAR and kainate receptors are also multimeric. In recent years several classes of AMPA receptor potentiators have been reported including *pyrrolidones* (piracetam, aniracetam) and *benzothiazides* (cyclothiazide). Clinical and preclinical data have suggested that positive modulation of AMPAR may be therapeutically effective in the treatment of cognitive deficits.[39]

C. The "Cys-loop receptor superfamily"

They are so called because of a conserved cysteine loop in their extracellular domain. Structural and functionnal evidence support the view that these allosteric proteins are heteropentameric oligomers each subunit made up of an extracellular amino-terminal domain and four transmembrane segments.

This superfamilly includes:

- Nicotinic acetylcholine receptors, nAchR, (cationic), which transduce acetylcholine effects beside acetylcholine muscarinic receptors (M1 to M5) belonging to GPCRs. Nicotinic receptors of skeletal muscles are pentamers composed of four distinct subunits (α, β, γ or ε, and δ) in the stoichiometric ratio of 2:1:1:1. The amino acid residues in the acetylcholine site involve more than one subunit and produce a ligand pocket at the interface of subunits α–γ and α–δ. The neuronal type is a pentameric combination of several α and β subunits. Nine distinct α-subunits and four β-subunits have been cloned making the existence of a great number of nicotinic receptors possible. But their physiological significance still needs to be completely understood. Only ligands showing selectivity between $\alpha4\beta2$ and $\alpha7$ receptors have been obtained.[40] The $\alpha4\beta2$ partial agonists *cytisine* and *varenicline* are approved for the treatment of smoking addiction. Also, *altenicline* and *ispronicline* are $\alpha4\beta2$ agonists entered in phase II for Parkinson and age-associated memory impairment treatment, respectively.
- Serotoninergic 5-HT3 receptors (cationic) transduce the effects of serotonin (5-hydroxy tryptamine, 5-HT), altogether with others 5-HT receptors which are GPCR. Selective 5-HT3 antagonists, like *ondansetron* and analogs, are antiemetics used during cancer therapy.
- GABA-A receptors, (anionic) mediate most synaptic inhibition in CNS. (GABA-B receptors are GPCRs). The pentameric structure of GABA-A is homologous with the nAChR (Figure 4.6b). The GABA site ligands (α–β interface) include agonists, like muscimol, and antagonists like bicucullin. The plant convulsant picrotoxin is a blocker of the Cl^- channel. The third site (α–γ interface) is the "benzodiazepine site." Binding of different ligands to this site can either potentiate the opening of the Cl^- channel elicited by GABA ("agonist" benzodiazepines like diazepam with anxiolytic, anticonvulsivant and sedative effects.), or decrease this opening (experimental inverse agonists with anxiogenic and convulsant properties), or block the benzodiazepine effect ("antagonist" benzodiazepines like *flumazenil*). Other allosteric sites bind *barbiturates*, etomidate, n-octanol, ethanol, propofol, halothane, and neuroactive steroid, with increase of GABAergic neurotransmissions, and depending of the subunit composition of each GABA-A receptor.[41]
- *Strychnine-sensitive glycine receptors (GlyRs)* (anionic) mediate synaptic inhibition, besides GABA, in spinal cord, brainstem and other regions of the CNS. (Glycine is also a strychnine-insensitive co-agonist at NMDAR, with a binding site distinct to that of glutamate). GlyRs regulate not only the excitability of motor and sensory neurones, but are also essential for the processing of photoreceptor signals, neuronal development and inflammatory pain sensitization. GlyRs, subtype-selective compounds are expected to emerge that will allow dissection of specific GlyR isoform functions.[42]

VII. RECEPTORS WITH INTRINSIC ENZYME ACTIVITY

These membrane receptors are glycoproteins spanning the membrane only once with an intrinsic enzymatic activity (guanylate cyclase, serine/threonine kinase or tyrosine kinase), located intracellularly, activated following the extracellular agonist–receptor interaction (Figure 4.5). Their dimerization is usually considered in their active state. Drug targeting concern the agonist binding site and the enzyme entity.

A. Receptors with guanylate cyclase activity

The cyclase catalytic domain converts GTP to cyclic GMP (cGMP). These membrane receptors mediate the action

of the atrial natriuretic peptide (ANP) and its structural analogs, that is, the others natriopeptides BNP and CNP, guanylin peptides, and the heat-stable enterotoxin of *Escherichia coli.*[43] [A second family of guanylate cyclases is found in the cytosol. These enzymes are activated by NO, now considered as an endogenous mediator]. This small putative drug target family has not yet received large attention whereas progress in the therapy of cardiovascular, renal and intestinal diseases might be concerned. Currently, the only natriuretic peptide available commercially is BNP that is *nesiritide* to treat congestive heart failure.[44]

B. Receptors with serine/threonine kinase activity

Over 35 distinct transforming growth factor (TGF)-β members have been identified in the human genome, including TGF-βs, growth differentiation factors, bone morphogenetic proteins (BMP), activins, inhibins, and glial cell line-derived neurotrophic factor. Individual family members are structurally related to the prototypical-founding member TGF-β1. All family members have profound effects on developmental processes ranging from the development of soft tissues, including angiogenesis, to the development of the skeleton.

These mediators exert their effect by binding to specific serine/threonine kinase type I and type II receptor complexes. Seven type I receptors, also termed *activin receptor-like kinase* (ALK) 1 to 7 and five type II receptors have been identified. TGF-β has high affinity for the TGF-β type II receptor (TβRII), and on binding a specific TGF-β type I receptor (TβRI) is recruited. On heteromeric complex formation between type I and type II receptors, the type I receptor is transphosphorylated by the type II receptor. This results in activation of the type I receptor, which can subsequently propagate the signal inside the cell by the phosphorylation of Smad proteins. The receptor complexes (R-Smads) are presented to the type I receptor and phosphorylated. The common Smad (Smad4) subsequently forms heteromeric complexes with others Smads. These complexes then translocate to the nucleus and modulate gene expression. [A third type of TGF-β receptors (endoglin and betaglycan) is a transmembrane protein with short intracellular domains that lack an enzymatic motif. Betaglycan can present TGF-β to serine/threonine kinase receptors and thereby facilitates signaling. The role of endoglin is less well understood].

Very active researches concern this family of drug targets in the fields of cancer, angiogenesis and bone therapy. *Dibotermin alfa* (rhBMP2), a recombinant form of BMP is accepted for the treatment of acute tibia fractures in adults, as an adjunct to standard care using open fracture reduction and intramedullary nail fixation in patients in whom there is a substantial risk of non-union. Three platforms of TGF-β inhibitors have recently evolved: antisense oligonucleotides, monoclonal antibodies and small molecules, some of them are in phase I or II.[45,46]

C. Receptors with tyrosine kinase activity

These receptors mediate the actions of many growth factors such as insulin, insulin-like growth factor (IGF), vascular endothelial growth factor (VEGF) epidermal growth factor (EGF), nerve growth factor (NGF), platelet-derived growth factor (PDGF), and the stem cell factor (cKIT ligand). Their cytoplasmic domain contains the tyrosine kinase activity as well as tyrosine sites of autophosphorylation. While most of these receptors possess a single polypeptide chain, the IGF and insulin receptors have two chains, α and β arising from a single gene, linked as by disulfide bounds. In this case the α chains possess the ligand binding site and the β chains, the tyrosine kinase activity.

The tyrosine kinase domain seems to be similar among these receptors while the ligand binding domain shows very little sequence homology between the members of this family. The autophosphorylation on tyrosine residues allows for the association of various proteins characterized by SH2 domains : phospholipase Cγ (PLC-γ), and adaptor proteins such as Grb2. These first events initiate cascades of reactions including small G-proteins, phosphoinositide 3-kinase (PI3K), cytosolic tyrosine kinases and mitogen-activated protein kinases (MAP kinases). MAP kinases phosphorylate one or more transcription factors that initiate gene expression, resulting in a variety of cellular responses, including cell division and proliferation.

This large family of drug targets has received much attention in the last decade. The mediators themselves are of therapeutical interest, like insulin, and more recently, *becaplermin*, a recombinant form of PDGF approved to treat ulcers of the foot, ankle, or leg in patients with diabetes.

Several inhibitors of their tyrosine kinase activity (non-selective from one receptor to the other) are approved for cancer therapy (*imatinib, erlotinib, sorafenib and sunitinib*) and many others are clinically studied (*dasatinib, nilotinib, pazopanib, vatalanib, vandetanib*…).

Another strategy is the therapeutical use of either monoclonal antibodies directed against the receptors with blocking, antagonist-like effect, approved as anticancer drugs (*trastuzumab* and *cetuximab*, anti-EGFR (HER2), or monoclonal antibodies directed against the mediator itself (*bevacizumab*, anti-VEGF). Another antagonist-like effect is that of *pegaptanib*, a pegylated modified oligonucleotide which directly interacts with VEGF, preventing its binding to VEGFR (approved for the treatment of neovascular (wet) age-related macular degeneration). These families of drugs related to tyrosine kinase receptors are still growing.

VIII. RECEPTORS COUPLED TO VARIOUS CYTOSOLIC PROTEINS

These membrane receptors are glycoproteins spanning the membrane only once, but often occuring as dimers, and coupled to cytosolic proteins (enzymes or transcription factors), either directly, or via various adaptor proteins (Figure 4.5). Some of these receptors are heteromultimers.

A. Receptors coupled to the cytosolic tyrosine kinase JAK

These JAK/STAT-coupled receptors concern the effects of cytokines: the growth hormone somatotropin (GH), erythropoietin (EPO), prolactin (PRL), granulocyte-colony-stimulating-factor (G-CSF), granulocyte and macrophage-colony-stimulating-factor (GM-CSF), leptin, thrombopoietin, interferons α, β and γ, and interleukins 2, 3, 4, 5, 6, 7, 9, 10 and 15.

These receptors are quite similar to the above receptors with intrinsic tyrosine kinase activity, but the receptor and the enzyme entities are two separate proteins. They do not have intrinsic kinase activity but associate, when activated by ligand binding, with JAK tyrosine kinases, which are the first step in the kinase cascade. JAKs (JAK1 to JAK5, TYK2) constitute a family of the very large superfamily of cytosolic tyrosine kinase proteins. Activated JAKs phosphorylate the cytoplasmic domain of the receptor, thereby creating recruitment sites for of latent cytoplasmic transcription factors STATs (signal transducers and activators of transcription) (Figure 4.7). STATs phosphorylated by JAKs, dimerize and subsequently migrate to the nucleus where they regulate gene transcription.

The recombinant forms of some of these cytokines are marketed: GH, EPO, interferons, *aldesleukine* (interleukin 2), *filgrastime* (G-CSF), *molgramostime* (GM-CSF). *Pegvisomant* is an analog of human growth hormone (GH) with antagonist properties. Overproduction of growth hormone leads to abnormally high levels of IGF-I, which then causes acromegaly like symptoms. Thus, pegvisomant is approved in acromegaly. Also, the JAK/STAT pathway represents an excellent opportunity for targeted cancer therapy and active research concern its pharmacological control at the intracellular level.[47,48]

B. Receptors coupled to the cytosolic Src, Zap70/Syk and Btk tyrosine kinases (immunoreceptors)

These receptors are heteromultimeric, including antigen receptors present on T lymphocytes, TCR, B lymphocytes BCR, and on NK cells NCR, and receptors for the Fc

FIGURE 4.7 Receptors coupled to the cytosolic tyrosine kinase JAK.

portion of immunoglobulins located on various hemopoietic cell types, FcεRI (IgE receptor), FcαR (IgA receptor), FcγR (IgG rceptor), FcδR (IgD receptor) and FcμR (IgM receptors). CD20, present on lymphocytes B, is also coupled to Src, but its ligand remains unknown.

Immunoreceptors have no intrinsic tyrosine kinase activity, but their subunits bear immunoreceptor tyrosine-based activation motifs (ITAMs) in their intracellular domain. ITAMs initiate cellular activation by modulating three classes of tyrosine kinases: the Src, ZapP70/Syk and Btk families. This signal predicates all subsequent outcomes of cell activation, including PI3K and MAP kinases activation, leading to the activation of various transcription factors controlling differentiation and proliferation, and to secretion of various mediators (histamine, cytokines and arachidonic acid derivatives) involved in inflammatory allergic and immunological processes.

This knowledge has not yet generated significant therapeutic advance in the respective fields. *Omalizumab* is a recombinant DNA-derived humanized IgG1 monoclonal antibody that selectively binds to human IgE, used mainly in allergy-related asthma therapy, with the purpose of reducing allergic hypersensitivity. *Rituximab* and *ibritumomab* are anti-CD20 monoclonal antibodies. Rituximab is considered as single first-line therapy for patients with follicular lymphoma, and ibritumomab in various non-Hodkins lymphoma. Other membrane receptors, SLAM, SAP and CD31, are also involved in cytokine secretion during the development of innate and adaptive immune responses. The fast increasing informations on these new receptors might lead to consider them as potential focal targets for novel therapeutic approaches.[49]

C. Receptors coupled to the cytosolic serine/threonine kinase IRAK

This receptor family mediates the effects of interleukin 1, a major proinflammatory cytokine, and interleukin 18, and includes TLR receptors of macrophages like TLR2 which recognize peptigoglycans of gram-positive bacteria. These receptors are coupled to the cytosolic serine/threonine kinase IRAK (interleukin receptor-associated kinase). Following signals include the activation of transcription factors like NFκB, or others via MAP kinases.

Interestingly, an endogenous structural analog of interleulin 1 (*IRAP*: interleukin 1 receptor antagonist protein) play the role of antagonist of interleukin 1. *Anakinra* is a recombinant non-glycosylated form of IRAP indicated for the reduction in signs and symptoms of moderately to severely active rheumatoid arthritis. Another approach is the inhibition of the IL-1 converting enzyme (ICE) which converts proIL-1 into its mature, proinflammatory form, and the inhibition of the p38-MAP kinase which controls interleukin 1 and tumor necrosis factor α (TNFα) production.[50,51]

D. Receptors coupled to caspases and to NFκB

These receptors mediate the effects of TNFα (TNF receptors), and of the RANK ligand, RANK-L or receptor activator of nuclear factor-kB, and its endogenous antagonist, osteoprotegerin (OPG), RANK receptors.

TNFα is the founding member of 19 different proteins identified within this cytokine family, including the Fas-ligand and the tumor necrosis factor apoptosis-ligand (TRAIL). TNF family members exert their biological effects through the TNF receptors (TNFR1, TNFR2 and Fas) that share a stretch of 80 amino acids within their cytoplasmic region, the death domain (DD), critical for recruiting the death machinery. This includes the p38-MAP kinase and the transcription factor NFκB controlling inflammation processes, and the caspases cascade which convey activation and an apoptotic signal in a proteolytic pathway that degrades cellular proteins leading to cell death. *Tasonermin* is a recombinant form of human TNFα1a approved for cancer therapy. *Infliximab* and *adalizumab* are anti-TNFα monoclonal antibodies with immunosuppressive effects indicated in rheumatoid arthritis, psoriatic arthritis, ankylosing spondylitis and Crohn's disease. *Etanercept* is a recombinant fusion protein acting as an antagonist of TNF receptors, also approved for its immunosuppressive effects.

Another approach is the inhibition of the TNF-converting enzyme (TACE), which converts proTNF into its mature, proinflammatory form.[50,51] Also, numerous inhibitors of P38 and Erk-MAP kinases have been synthesized and some have reached the clinical trial stage. MAP-p38 kinase occupies a central role in the signaling network responsible for the up-regulation of proinflammatory cytokines like interleukin 1 and TNFα.[52] The pathway of Erk-MAP kinase is often up-regulated in human tumors and as such represents an attractive target for the development of anticancer drugs.[53]

Diverse drug targets in this field are at an early stage of development, such as endogenous "inhibitors of apoptosis proteins" (IAP), a family of caspase inhibitors that selectively bind and inhibit caspases-3, -7, and -9. The inhibition of these IAP might stimulate apoptosis of cancer cells with potential as a treatment of malignancy.[54]

RANK-L, the endogenous agonist of RANK receptors, and OPG, their endogenous antagonist, directly regulate osteoclast differentiation and osteolysis. RANK-L is a powerful inducer of bone resorption and OPG acts as a strong inhibitor of osteoclastic differentiation. RANK-L also induces BMP-2 expression in chondrocytes. Furthermore, recent data demonstrate that the OPG–RANK–RANK-L system modulates cancer cell migration. RANK-L promotes the activation of several intracellular signaling pathways, including stimulation of NF-κB/MAP kinase pathways , and the Akt/protein kinase B (PKB) pathway. *Denosumab*, studied in phase 3, is a humanized monoclonal antibody directed

against RANK-L with antiosteoclast activity, resulting in inhibition of osteoclast activity, a decrease in bone resorption, and an increase in bone mineral density. Denosumab also decreases bone turnover in advanced cancer.[55]

E. Receptors of the cellular adhesion

Cellular adhesion is performed by a large set of transmembrane proteins: integrins, cadherins, selectins and Ig superfamily members (Ig-N-Cam, ICAM …).

Integrins are heterodimer transmembrane receptors (24 known integrin heterodimers) for the extracellular matrix composed of an α- and a β- subunit. Natural integrin ligands include laminin, fibronectin, and vitronectin, but they also include fibrinogen and fibrin, thrombospondin, MMP-2, and fibroblast growth factor. *Natalizumab* is an anti-α_4 integrin (VLA4) monoclonal antibody targeting immune cells and approved for the treatment of multiple sclerosis. *Abciximab* is an anti-$\alpha_{IIb}\beta_3$ integrin (also termed anti-IIb/IIIa glycoprotein) monoclonal antibody. Its binding to the receptor prevents fibrinogen, von Willebrand factor, vitronectin, and other adhesive molecules from binding to the receptor, thereby inhibiting platelet aggregation. It is indicated as an adjunct to aspirin and heparin for the prevention of acute cardiac ischemic complications. *Eptifibatide*, a synthetic heptapeptide, and *tirofiban*, a piperidinyl derivative of tyrosine, also bind to the $\alpha_{IIb}\beta_3$ integrin with antagonist effect inhibiting platelet aggregation.

Integrin $\alpha_V\beta_3$, the vitronectin receptor, has been identified as a promising potential target for the treatment of osteoporosis, diabetic retinopathy and cancer. Three classes of integrin antagonists are currently in preclinical and clinical development: monoclonal antibodies targeting the extracellular domain of the heterodimer, *vitaxin*, synthetic peptides such as *cilengitide* and several peptidomimetics.[56]

IX. G-PROTEIN-COUPLED RECEPTORS

GPCRs are characterized by a common topology and by varying degrees of primary sequences similarities. All of them are formed by a single polypeptide chain of 350–1,200 residues, hydropathy plots revealing seven hydrophobic regions which are likely to correspond to transmembrane α-helices. Thus GPCRs are also termed 7-TM receptors or heptahelical receptors (Figure 4.8). The amino-terminal extracellular domain contains potential N-linked glycosylation sites in most receptors.[57] The carboxy-terminal cytoplasmic end is involved in the coupling to G-proteins and contain a palmitoylation site (Cys residue) and phosphorylation sites (Ser and Thr residues), both involved in the receptor desensitization. The three cytoplasmic loops are implicated in the coupling with heterotrimeric G-proteins (distinct from "small G-proteins" which are monomeric).

Almost 30% of all marketed drugs act on GPCRs. The most familiar GPCRs as historical drug targets are the muscarinic acetylcholine receptors, the α- and β- adrenergic, dopaminergic, histaminergic and opioids receptors. Some

FIGURE 4.8 Nuclear receptors and G-protein-coupled receptors.

others GPCR ligands have been developed as drugs during the last three decades, that is, serotonergic ligands, prostaglandins, leucotrienes, ADP and calcium receptors ligands. The actual top-selling GPCRs ligands are *clopidogrel* (ADP-P2Y12 antagonist, platelet antiaggregant), *olanzapin* (mixed serotonin-5HT2/dopamine-D2 antagonist, neuroleptic), *valsartan* (angiotensin-AT1 antagonist, antihypertensive), *fexofenadine* (histamine-H1 antagonist, antiallergic), *sumatriptan* (serotonin-5HT1D antagonist, antimigrainous), *leuprorelin* (GnRH/LH-RH agonist, anti hormone-dependent cancer).

Thus, GPCRs have been and remain very attractive targets. Only a small proportion of known GPCRs are currently targeted by therapeutics. This provides a great number of promising targets for the development of new medicines.

A. How many druggable GPCRs?

GPCRs are the largest class of receptors mediating the effects of small neurotransmitters, all known neuropeptides, many peptide hormones and inflammatory mediators, some lipids and even calcium for the control of its blood concentration.

According to recent analysis of the human genome about 780 to more than 860 genes encode GPCRs.[58,59] More than 50% of GPCRs are activated by sensory stimuli (8 by light, 22 by taste compounds and 388 to 460 by odorant stimuli). The full repertoire of receptors for endogenous ligands is likely to include 367 members.[59] Among the latter, about 180 GPCRs are activated by well characterized endogenous ligands (Table 4.1).

Note that one endogenous mediator may activate several GPCRs: 13 for serotonin; 9 for adrenaline and noradrenaline; 8 for glutamate; 5 for dopamine; 5 for acetylcholine; 4 for histamine; 2 for GABA. Also, some mediators activate GPCRs but also receptors with intrinsic ion channel.

Since 1995, 60 neuropeptides activating GPCRs have been discovered (nociceptin/orphanin, orexins/hypocretins, PRL-releasing peptide apelin, ghrelin, melanin concentrating hormone, urotensin II, neuromedin U, metastatin, prokineticin1/2, relaxin 3, neuropeptide B/W, neuropeptide S, relaxin 3, obestatin …).[60] Their receptors immediately became new putative drug targets. However, there are still more than 140 orphan GPCRs, and deciphering their function remains a priority for fundamental and clinical research. Research on orphan GPCRs has concentrated mainly on the identification of their natural ligands, whereas recent data suggest additional ligand-independent functions for these receptors.[61] This emerging concept is connected with the observation that orphan GPCRs can heterodimerize with GPCRs that have identified ligands, and by so doing regulate the function of the latter.

Some non-heptahelical receptors might also activate heterotrimeric G-proteins. This property has been assumed for some transmembrane proteins, with or without kinase activity on their cytosolic ending, and for some receptors

TABLE 4.1 Diversity of the G-Protein-Coupled Receptors Family

Endogenous ligand	Subtypes
Biogenic amines	
Acetylcholine	M_1, M_2, M_3, M_4, M_5
Adrenaline, noradrenaline	β_1, β_2, β_3, α_{1A}, α_{1B}, α_{1D}, α_{2A}, α_{2B}, α_{2C}
Dopamine	D_1, D_2, D_3, D_4, D_5
Histamine	H_1, H_2, H_3, H_4
Serotonin	5-$HT_{1A/B/D/E/F}$; 5-$HT_{2A/B/C}$; 5-$HT_{4/6/7}$; 5-$HT_{5A/B}$
Peptides/Proteins	
Adrenocorticotrophin (ACTH)	MC_2
Adrenomedullin	AM_1, AM_2
Amylin	$AMY_{1/2/3}$
Angiotensin II	AT_1, AT_2
Bradykinin	B_1, B_2
CC chemokines	CCR1-10
CXC chemokines	CXCR1-6
CX_3C chemokines	XCL1/2, CX3L1
Corticotropin-releasing factor	CRF_1, CRF_2
Endothelin-1,-2,-3	ET_A, ET_B
Follicle-stimulating hormone	FSH
Formyl-Met-Leu-Phe	fMLP
Gastric inhibitory peptide	GIP
Gastrin	CCK_2
Bombesin (gastrin-releasing peptide)	BB2
Motilin	GPR38
Neuropeptide FF and AF	NPFF1, NPFF2
Neuropeptide W-23, W-30	GPR7, GPR8
Neuropeptide Y	$Y_{1/2/4/5/6}$
Opioids	δ, κ, μ, ORL1
Orexin A/B	OX1, OX2
Relaxin	LGR7, LGR8
Substance P, neurokinin A/B	NK_1, NK_2, NK_3
Amino acids	
Glutamate	$mglu_{1/2/3/4/5/6/7/8}$
γ-aminobutyric acid (GABA)	$GABA_{B1/2}$
Lipids	
Leukotriene B_4	BLT
Leukotriene C_4, D_4	$CysLT_1$, $CysLT_2$
LXA_4	FPRL1
Lysophosphatidic acid	edg2/4/7
Lysophosphatidylcholine	LPC_1
Platelet-activating factor	PAF
Prostacyclin	IP
Prostaglandin D_2	DP
Prostaglandin E_2	$EP_{1/2/3/4}$
Prostaglandin $F_2\alpha$	FP
Thromboxane A_2	TP
Nucleotides/Nucleosides	
Adenosine	A_1; $A_{2A/B}$; A_3
ADP	$P2Y_1$; $P2Y_{12}$; $P2Y_{13}$
ATP	$P2Y_{1/2/4/11}$
UDP	$P2Y_6$
UTP	$P2Y_2$, $P2Y_4$

(Continued)

TABLE 4.1 (Continued)

Endogenous ligand	Subtypes
Proteases	
Thrombin and others	PAR1/3/4
Trypsin and others	PAR2
Ions	
Calcium	CaSR

belonging to the class of glycosylphosphatidylinositol (GPI)-anchored proteins. The relationship between these receptors and trimeric G-proteins remains controversial. The pentaspanin integrin-associated protein (IAP or CD47), a receptor for thrombospondins associated to integrins, mimicks heptahelical receptors. Its coupling to G-proteins is well demonstrated.[62,63] CD47 modulates a range of cell activities including platelet activation, leukocytes motility, adhesion and migration, monocytes/macrophages phagocytosis and secretion of inflammatory mediators. CD47 is considered as a valuable drug target.[64,65]

B. Diversity of G-proteins

G-proteins are located on the inner side of the plasma membrane. They are heterotrimeric with α, β, and γ subunits. Activated cell surface receptors initiate G-protein signaling, promoting the exchange of GDP, inducing the dissociation of the α-subunit from a high stable βγ dimer. In this dissociated state both α- and βγ subunits modulate the activity of an effector molecules.[63,66] Slow hydrolysis of bound GTP by the GTPase intrinsic to the α-subunit leads to reassociation of the oligomer and cessation of the signal (Figure 4.9).

The diversity of heterotrimeric G-proteins has been demonstrated, around 1,980, with the purification of Gs (s = stimulatory for adenylate cyclase), Gi (i = inhibitory for adenylate cyclase), and Gt (t = transducine which activate a cGMP-phosphodiesterase in retinal cells). G-protein subunits are highly homologuous in both primary sequence and tertiary structure. With the sequence of the human genome nearly complete, the number of subunits variants identified includes 27 Gα (39 to 52 kDa in size), 5 Gβ (36 kDa) and 14 Gγ subunits (7–8 kDa). This leads to a theoretical diversity of 27 × 5×14; 1,890 combinations of heterotrimers, questioning the selectivity of receptor–G-proteins interactions. The usual classification of G-proteins remains based on their α- subunits with four families, each α- subunit modulating selective effectors:

- Gs family including α_s, which activates all the isoforms of adenylate cyclase, but also Src tyrosine kinases, and α_{olf} (olf = olfactive; activation of adenylate cyclases);
- Gi family including α_{i1}, α_{i2} and α_{i3} (inhibits adenylate cyclase isoforms 1, 2, 3, 5 and 6, but activates Src tyrosine kinases), α_{t1} and α_{t2} (t = transducin, activate cGMP-phosphodiesterase), $\alpha_{o1/A}$, $\alpha_{o2/B}$,α_z and α_{gust} (gust = gustducin);
- Gq family including α_q, α_{11}, α_{14}, α_{15} and α_{16}, which activate phospholipase Cβ and Bruton's tyrosine kinase (Btk);
- G12 family including α12 (activates Btk), and Gap1, a Ras-GTPase-activating protein), and α13 (activates p115 RhoGEF).

The main effectors regulated by βγ dimers are phospholipases Cβ (activation); adenylate cyclase 1, 5 and 6 (inhibition) and adenylate cyclases 2, 4 and 7 (activation); GIRK/Kir3.1 and 3.4 potassium channels (activation); Ca_V2 calcium channels (inhibition) and PI3Kγ (activation).

C. Diversity of GPCR-elicited signaling and related drug targets

GPCRs are involved in all physiological processes, corresponding to a large diversity of their signaling pathways. Signals arising from GPCRs are never unique. Several parallel pathways may be activated in response to agonist stimulation of a receptor, from Gα and Gβγ subunits, or from the activation of two different G-proteins. Note that most pathways lead to protein phosphorylation and/or to calcium-sensitive protein activation that tightly control final cell responses such as secretion, contraction, general metabolism and protein synthesis through gene transcription. Effector enzymes have multiple subtypes that differ in tissue distribution. Thus, targeting such molecules may lead to organ-specific pharmacological regulation. However, most GPCR-elicited pathways (PLCs, PI3K, small G-proteins and MAP kinases) are also actors in signaling of other receptor families, decreasing their druggability.

Adenylate cyclases are transmembrane proteins of the plasma membrane transforming ATP to cAMP. Adenylate (or adenylyl) cyclases are activated through stimulation of Gs-coupled receptors, and some of them are inhibited through stimulation of Gi-coupled receptors. The plant terpenoid forskolin stimulates cAMP formation by acting directly on the adenylate cyclases. Water-soluble forskolin derivatives with high selectivity for type 5 (cardiac) adenylate cyclases have been proposed in the treatment of acute heart failure. Adenine analogs or P-site inhibitors are now utilized to develop isoform-specific inhibitors as well.[67] Targeting adenylate cyclase isoforms, either of isoform-specific stimulation or inhibition, may be a novel strategy to develop new drugs.

The main target of cAMP is protein kinase A (PKA) which phosphorylates various proteins on Ser and Thr residues, for instance myosin light chain kinase in smooth muscles, or Ca_V1 calcium channels in cardiac contractile cells. Beside PKA, cAMP can also bind to some other direct targets such as HCN cationic non-selective channels in cardiac pacemaker cells (see above).

FIGURE 4.9 Activation of G-proteins. G-proteins include three subunits (alpha (α), beta (β) and gamma (γ)). Interaction of the α-subunit to an agonist stimulated receptor *(I)* causes the exchange of the bound GDP with GTP *(II)*. The α-GTP complex and the dimer β–γ dissociated *(III)*. The α-GTP complex interacted with an effector (E1) and the dimer β–γ with an other effector (E2) *(IV–V)*. The α-subunit catalyses hydrolysis of the bound GTP to GDP *(VI)* and reassociated with the dimer β–γ *(VIII)*. This deactivation of signalling can be accelerated by proteins termed regulators of G-protein signaling (RGS) which have been shown to directly bind to the α-subunit of G-proteins. Heterotrimeric G-proteins (αβγ); A: agonist; R: receptor; E: effector.

cAMP is broken down by phosphodiesterases (PDEs) which hydrolyze the 3′-phosphate ester to give the common inactive metabolite, 5′-AMP. The PDEs superfamily currently includes 20 different genes subgrouped into different PDE families.[67] PDEs 5, 6, 9 and 11 are selective for the hydrolysis of cGMP, others families are selective for cAMP. Subtype-specific phosphodiesterase inhibitors, such as *sildenafil*, a PDE5 inhibitor, and *milrinone*, a PDE3 inhibitor, are now widely used in the treatment of erectile dysfunction and heart failure, respectively. The search for selective PDE inhibitors remains active.[68]

Phospholipases C (PLCs) are cytosolic enzymes transforming the membranous lipid phosphatidylinositol 4,5-bisphosphate (PIP_2) to diacylglycerol (DAG), which remains a membrane component, and to cytosolic inositol (1,4,5)-triphosphate (IP_3). DAG is the principal endogenous regulator of membrane bound protein kinases C (PKCs). IP_3 binds to endoplasmic membrane receptors and liberates calcium from sequestered stores (endoplasmic reticulum) inducing an increase of cytoplasmic calcium. PLC subtypes, β, γ and δ, have been characterized. Four β-, two γ-, four δ-isoforms, and multiple spliced variants have been described in mammals. The PLCβ family appears to be regulated by Gαq and by the Gβγ dimer usually arising from Gi dissociation. Moreover, PLCγ has a SH2 domain allowing its interaction with phosphorylated tyrosine residues. Thus, PLCγ belong to pathways elicited by receptor and cytosolic tyrosine kinases. The interest of PLCs as drug targets is poorly considered.

PI3Ks are a large family of intracellular signal transducers that have attracted much attention over the past 10 years. PI3Ks phosphorylate inositol lipids at the 3′ position of the inositol ring to generate the 3-phosphoinositide PI(3,4,5)P3, (PIP3). PI3Ks of class 1B are directly activated by Gβγ, but others PI3K also belong to pathways of receptors with or coupled to tyrosine kinase activity. PIP3 may recruit to the membrane various protein kinases, phosphoinositide-dependent kinase-1 (PDK-1), protein kinase B (PKB, or Akt), protein kinase Cξ (PKCξ), PLCγ and cytosolic tyrosine kinases such as the Bruton's tyrosine kinase (Btk).

Discovered in leukocytes, PI3K pathways has been recently studied in many cell types. Their main involvements may be in the control of cell development and differentiation. The central role of PI3K signaling in allowing cancer cells to bypass normal growth-limiting controls has led to the development of PI3K inhibitors.[69] Recent findings suggest an involvement of PI3K in the pathogenesis of others diseases including heart failure and autoimmune/inflammatory disorders. The tissue selectivity of PI3K isoforms has to be closely considered in the development of PI3K inhibitors. For instance, inhibitors of PI3K pathways, including tyrosine kinase inhibitors, used to treat cancer, may induce cardiopathies.[70] Interestingly, an inhibitor of the mammalian target of rapamycin (m-TOR), a downstream effector of PI3K, did not have adverse effects on the heart,[71] showing possible alternative targeting downstream effectors. Conversely, isoform-selective PI3K inhibitors are now proposed as novel therapeutics for the treatment of acute myocardial infarction.[72]

Interestingly, the polyphenol *resveratrol*, which has chemopreventive and chemotherapeutic properties thought to be related to histone deacetylase HDAC3 (sirtuin) inhibition, also inhibits PI3K.[73] This points out the difficulties to correlate

therapeutic properties of drugs to their known targets and remember the complementarity of new drug research strategies associating molecular and integrative processes.

X. NUCLEAR RECEPTORS AS DRUG TARGETS

The nuclear receptors are ligand-gated transcription factors that modulate gene expression, acting as homodimers or heterodimers. Their ligands are lipophilic molecules such as steroidal hormones, thyroid hormones, vitamin D, retinoids like vitamin A, lipid mediators including free fatty acid, and possibly a large range of unknown molecules (Figure 4.8). These ligands control major aspects of eukaryotic development, differentiation, reproduction, and metabolic homeostasis. Nuclear receptors are defined as a superfamily subdivided into three classes. The steroid receptor family, and the thyroid/retinoid family include targets of largely developed drugs. The orphan receptors family includes half of the members of the nuclear receptors superfamily. They are so-called "orphan" receptors because the identity of their ligand, if any, is unknown. Their druggability has recently attracted much attention.

The steroid receptor family includes androgen receptors (AR), estrogen receptors (ER, α and β), progesterone receptors (PR, A and B), glucocorticoid receptors (GR), and mineralocorticoid receptors (MR). Traditional models propose that, upon binding their hormonal ligand, the receptors release heat shock proteins like hsp90, translocate into the nucleus, and bind as homodimers to imperfect palindromic response elements at upstream promoter sites.

Their ligands are well known and widely used. The more recent advances in this field concern the so-called selective estrogen receptor modulators (SERMs) like *tamoxifen* and *raloxifene*. SERMs are used for the prevention and treatment of diseases such as osteoporosis and breast cancer. Ideally, it is presumed that SERMs should selectively act as an agonist in the bone and brain while simultaneously acting as an antagonist in the breast and uterus.[74]

The thyroid/retinoid receptor family includes the thyroid receptors (TR α and β), vitamin D receptors (VDR), retinoic acid receptors (RAR α, β and γ), retinoic X receptors (RXR, α, β and γ) and peroxisome proliferator-activated receptors (PPAR, α, β/δ and γ). These receptors typically function as heterodimers, often including RXR, which tend to stay bound to their response elements regardless of whether agonist ligands are present. In the absence of ligand, gene activation is prevented by corepressor interactions with the DNA-bound heterodimer. Upon binding ligand, corepressor proteins are released and coactivators are recruited, leading to transcriptional activation.

TR, VDR and RAR are targets for current medicines. PPARs are lipid-activated transcription factors that regulate the expression of genes involved in the control of lipid and lipoprotein metabolism, glucose homeostasis and inflammatory processes. Their wide range of potential therapeutic actions make them attractive targets for the development of oral agents targeting risk factors associated with the metabolic syndrome, type 2 diabetes and cardiovascular diseases. PPARα agonists belong to the fibrate class (*clofibrate, fenofibrate, bezafibrate, ciprofibrate and gemfibrozil*), widely prescribed to reduce triglycerides while increasing plasma HDL-cholesterol. PPARγ agonists *(pioglitazone and rosiglitazone)* have beneficial effects on glucose homeostasis by increasing insulin sensitivity and glucose disposal. However, they are under critical discussion due to the significantly increased risk of heart attack and cardiovascular death.[75] To date, no PPARβ/δ agonist has been fully developed and the clinical potential of targeting this isotype remains to be clearly determined.

Huge investments have been made in the last decade to develop new PPAR agonists with improved efficacy relative to the existing drugs, but a large set of preclinical and clinical adverse events have to be considered. PPAR agonists remain interesting drugs but they display side effects which limit their therapeutic use. Current strategies aim at reducing side effects by identifying selective PPAR modulators (SPPARMs) and the optimization of the selectivity ratio between the different PPAR isoforms.[76]

The orphan receptor family gathers nuclear receptor which cognate ligand, if any, is unknown, but belong to the steroid or the thyroid/retinoid family. This concerns half of the members of the nuclear receptors superfamily, a large reserve of drug targets. Some examples of these nuclear orphan, or deorphanized, receptors are RXR, ROR, LXR, FXR and ERR.

Retinoic X receptors (RXR, α, β and γ), also termed rexinoid receptors, are usually classified among orphan receptors, but belong to the thyroid/retinoid family. RXR is an obligatory partner dimerizing with other thyroid/retinoid receptors. RXR selectively bind the 9-*cis* isomer of retinoic acid, whereas RAR bind all-*trans* retinoic acid as well as its 9-*cis* isomer.[77] RXR selective agonists are termed rexinoids, like *bexarotene* used as an antineoplastic in the treatment of cutaneous T-cell lymphoma and the cutaneous lesions of T-cell lymphomas and Kaposi's sarcoma.

Retinoic acid receptor-related orphan receptor α (RORα) has been recently deorphanized, cholesterol been identified as its ligand. RORα is expressed in many tissues and is therefore a regulator of multiple biological processes. A beneficial modulatory role of RORα is proposed in the pathogenesis of dyslipidemia, inflammation and atherosclerosis.[78]

Liver X receptors (LXR α and β) are oxysterol receptors that regulate multiple target genes involved in cholesterol homeostasis. Recent studies also suggest that they may also be involved in glucose metabolism, inflammation and Alzheimer's disease. Although the prototypic LXR agonists induce liver triglyceride accumulation by regulating the hepatic lipogenesis pathway, it is hoped that a subtype-specific agonist or selective modulators would provide

the desired cardioprotection without the undesirable induction of lipogenesis.[79]

The farnesoid X receptor (FXR) is activated by the bile acids, chenodeoxycholic acid, lithocholic acid and deoxycholic acid. Upon activation FXR heterodimerises with RXR and regulates a cohort of genes involved in cholesterol catabolism and bile acids biosynthesis. Thus, development of potent FXR agonists might represent a new approach for the treatment of cholestastic disorders.[80] Interestingly, bile acids also activate TGR5, a G-protein-coupled receptor. Selective ligands are now available to differentiate genomic and non-genomic effects mediated by bile acids.[81]

The estrogen-related receptors (ERR, α, β and γ) is structurally most related to the canonical ER and has been shown to modulate estrogen signaling. These observations have heightened interest in ERR as a therapeutic target in both breast and ovarian cancer and in other estrogenopathies.[82]

REFERENCES

1. Burgess, J., Golden, J. Cracking the druggable genome. *Bio-IT World* [online], (2002) <http://www.bio.itworld.com/archive/100902/first-base.html> .
2. Wishart, D. S., Knox, C., Guo, A. C., Shrivastava, S., Hassanali, M., Stothard, P., Chang, Z., Woolsey, J. DrugBank: a comprehensive resource for in *silico drug* discovery and exploration. *Nucleic Acids Res.* 2006, *43*, D668–D672.
3. Drews, J. Genomic sciences and the medicine of tomorrow. *Nat. Biotechnol.* 1996, *14*, 1516–1518.
4. Drews, J., Ryser, S. Classic drug targets. *Nat. Biotechnol.* 1997, *15*, 1318–1319.
5. Golden, J. B. Prioritizing the human genome: knowledge management for drug discovery. *Curr. Opin. Drug Discov. Dev.* 2003, *6*, 310–316.
6. Zheng, C., Han, L., Yap, C. W., Xie, B., Chen, Y. Progress and problems in the exploration of therapeutic targets. *Drug Discov. Today* 2006, *11*, 412–420.
7. Zheng, C. J., Han, L. Y., Yap, C. W., Ji, Z. L., Cao, Z. W., Chen, Y. Z. Therapeutic targets: progress of their explorations and investigation of their characteristics. *Pharma. Rev.* 2006, *58*, 259–279.
8. Imming, P., Sinning, C., Meyer, A. Drugs, their targets and the nature and number of drug targets. *Nat. Rev. Drug Discov.* 2006, *5*, 821–834.
9. Overington, J. P., Al-Lazikani, B., Hopkins, A. L. How many drug targets are there?. *Nat. Rev. Drug Discov.* 2006, *5*, 993–996.
10. Ehrlich, P. Chemotherapeutics: scientific principles, methods and results. *Lancet* 1913, *2*, 445–451.
11. Landry, Y., Gies, J.-P. *Pharmacologie: des cibles vers l'indication thérapeutique*. Dunod: Paris, 2003. 596 pages
12. Braestrup, C., Schmiechen, R., Neef, G., Nielsen, M., Petersen, E. N. Interaction of convulsive ligands with benzodiazepine receptors. *Science* 1982, *216*, 1241–1243.
13. Costa, T., Herz, A. Antagonists with negative intrinsic activity at delta opioid receptors coupled to GTP-binding proteins. *Proc. Natl Acad. Sci. USA* 1989, *86*, 7321–7325.
14. Daeffler, L., Landry, Y. Inverse agonism at heptahelical receptors: concept, experimental approach and therapeutic potential. *Fundam. Clin. Pharmacol.* 2000, *14*, 73–87.
15. Bond, R. A., Ijzerman, A. P. Recent developments in constitutive receptor activity and inverse agonism, and their potential for GPCR drug discovery. *Trends Pharmacol. Sci.* 2006, *27*, 92–96.
16. Roussidis, A. E., Theocharis, A. D., Tzanakakis, G. N., Karamanos, N. K. The importance of c-Kit and PDGF receptors as potential targets for molecular therapy in breast cancer. *Curr. Med. Chem.* 2007, *14*, 735–743.
17. Zavrski, I., Kleeberg, L., Kaiser, M., Fleissner, C., Heider, U., Sterz, J., Jakob, C., Sezer, O. Proteasome as an emerging therapeutic target in cancer. *Curr. Pharm. Des.* 2007, *13*, 471–485.
18. Dubey, S., Satyanarayana, Y. D., Lavania, H. Development of integrase inhibitors for treatment of AIDS: an overview. *Eur. J. Med. Chem.* 2007, *XX*, 1–10. doi:10.1016/j.ejmech.2007.01.024.
19. Singh, N., Cheve, G., Avery, M. A., McCurdy, C. R. Targeting the methyl erythritol phosphate (MEP) pathway for novel antimalarial, antibacterial and herbicidal drug discovery: inhibition of 1-deoxy-D-xylulose-5-phosphate reductoisomerase (DXR) enzyme. *Curr. Pharm. Des.* 2007, *13*, 1161–1177.
20. Higgins, C. F. Multiple molecular mechanisms for multidrug resistance transporters. *Nature* 2007, *446*, 749–757.
21. Gottesman, M. M., Fojo, M., Bates, S. E. Multidrug resistance in cancer: role of ATP-dependent transporters. *Nat. Rev. Cancer.* 2002, *2*, 48–58.
22. Lipecka, J., Norez, C., Bensalem, N., Baudouin-Legros, M., Planelles, G., Becq, F., Edelman, A., Davezac, N. Rescue of DeltaF508-CFTR (cystic fibrosis transmembrane conductance regulator) by curcumin: involvement of the keratin 18 network. *J. Pharmacol. Exp. Ther.* 2006, *317*, 500–505.
23. Verkman, A. S., Lukacs, G. L., Galietta, L. J. CFTR chloride channel drug discovery – inhibitors as antidiarrheals and activators for therapy of cystic fibrosis. *Curr. Pharm. Des.* 2006, *12*, 2235–2247.
24. Pedemonte. T. Diena, N., Caci, E., Nieddu, E., Mazzei, M., Ravazzolo, R., Zegarra-Moran, O., Galietta, L. J. Antihypertensive 1,4-dihydropyridines as correctors of the cystic fibrosis transmembrane conductance regulator channel gating defect caused by cystic fibrosis mutations. *Mol. Pharmacol.* 2005, *68*, 1736–1746.
25. Pedemonte, N., Boido, D., Moran, O., Giampieri, M., Mazzei, M., Ravazzolo, R., Galietta, L. J. Structure-activity relationship of 1, 4-dihydropyridines as potentiators of the CFTR chloride channel. *Mol. Pharmacol.* 2007, *23*. [Epub ahead of print].
26. Camerino, D. C., Tricarico, D., Desaphy, J.-F. Ion channel pharmacology. *Neurotherapeutics* 2007, *4*, 184–198.
27. Klotz, U. Ziconotide – a novel neuron-specific calcium channel blocker for the intrathecal treatment of severe chronic pain – a short review. *Int. J. Clin. Pharmacol. Ther.* 2006, *44*, 478–483.
28. Lory, P., Chemin, J. Towards the discovery of novel T-type calcium channel blockers. *Expert Opin. Ther. Targets* 2007, *11*, 717–722.
29. Judge, S. I., Bever, C. T., Jr. Potassium channel blockers in multiple sclerosis: neuronal Kv channels and effects of symptomatic treatment. *Pharmacol. Ther.* 2006, *111*, 224–259.
30. Shi, H., Wang, H., Yang, B., Xu, D., Wang, Z. The M3 receptor-mediated K(+) current (IKM3), a G(q) protein-coupled K(+) channel. *J. Biol. Chem.* 2004, *279*, 21774–21778.
31. Ledoux, J., Werner, M. E., Brayden, J. E., Nelson, M. T. Calcium-activated potassium channels and the regulation of vascular tone. *Physiology (Bethesda)* 2006, *2111*, 69–78.
32. Ruzyllo, W., Tendera, M., Ford, I., Fox, K. M. Antianginal efficacy and safety of ivabradine compared with amlodipine in patients with stable effort angina pectoris: a 3-month randomised, double-blind, multicentre, noninferiority trial. *Drugs* 2007, *67*, 393–405.
33. Pifferi, S., Boccaccio, A., Menini, A. Cyclic nucleotide-gated ion channels in sensory transduction. *FEBS Lett.* 2006, *580*, 2853–2859.
34. Nilius, B., Owsianik, G., Voets, T., Peters, J. A. Transient receptor potential cation channels in disease. *Physiol. Rev.* 2007, *87*, 165–217.
35. Okuhara, D. Y., Hsia, A. Y., Xie, M. Transient receptor potential channels as drug targets. *Expert Opin. Ther. Tar.* 2007, *11*, 391–401.
36. Bodding, M. TRP proteins and cancer. *Cell Signal* 2007, *19*, 617–624.
37. Burnstock, G. Physiology and pathophysiology of purinergic neurotransmission. *Physiol. Rev.* 2007, *87*, 659–797.
38. Paoletti, P., Neyton, J. NMDA receptor subunits: function and pharmacology. *Curr. Opin. Pharmacol.* 2007, *7*, 39–47.

39. O'Neill, M. J., Witkin, J. M. AMPA receptor potentiators: application for depression and Parkinson's disease. *Curr. Drug Targets* 2007, *8*, 603–620.
40. Romanelli, M. N., Gratteri, P., Guandalini, L., Martini, E., Bonaccini, C., Gualtieri, F. Central nicotinic receptors: structure, function, ligands, and therapeutic potential. *Chem. Med. Chem.* 2007, *2*, 746–767.
41. Olsen, R. W., Chang, C. S., Li, G., Hanchar, H. J., Wallner, M. Fishing for allosteric sites on GABA(A) receptors. *Biochem. Pharmacol.* 2004, *68*, 1675–1684.
42. Betz, H., Laube, B. Glycine receptors: recent insights into their structural organization and functional diversity. *J. Neurochem.* 2006, *97*, 1600–1610.
43. Garbers, D. L., Chrisman, T. D., Wiegn, P., Katafuchi, T., Albanesi, J. P., Bielinski, V., Barylko, B., Reglield, M. M., Burnett, J. C., Jr. Membrane guanylyl cyclase receptors: an update. *Trends Endocrinol. Metab.* 2006, *17*, 251–258.
44. Vesely, D. L. Which of the cardiac natriuretic peptides is most effective for the treatment of congestive heart failure, renal failure and cancer?. *Clin. Exp. Pharmacol. Physiol.* 2006, *33*, 169–176.
45. Lahn, M., Kloeker, S., Berry, B. S. TGF-beta inhibitors for the treatment of cancer. *Expert Opin. Investig. Drugs.* 2005, *14*, 629–643.
46. Akhurst, R. J. Large- and small-molecule inhibitors of transforming growth factor-beta signaling. *Curr. Opin. Investig. Drugs.* 2006, *7*, 513–521.
47. Klampfer, L. Signal transducers and activators of transcription (STATs): novel targets of chemopreventive and chemotherapeutic drugs. *Curr. Cancer Drug Targets* 2006, *6*, 107–121.
48. Valentino, L., Pierre, J. JAK/STAT signal transduction: regulators and implication in hematological malignancies. *Biochem. Pharmacol.* 2006, *71*, 713–721.
49. Garcia, V. E., Chuluyan, H. E. SLAM and CD31: signaling molecules involved in cytokine secretion during the development of innate and adaptive immune responses. *Cytokine Growth Factor Rev.* 2007, *18*, 85–96.
50. Wagner, G., Laufer, S. Small molecular anti-cytokine agents. *Med. Res. Rev.* 2006, *26*, 1–62.
51. Le, C. T., Abbenante, G. Inhibitors of TACE and Caspase-1 as anti-inflammatory drugs. *Curr. Med. Chem.* 2005, *12*, 2963–2977.
52. Lee, M. R. Dominguez C.MAP kinase p38 inhibitors: clinical results and an intimate look at their interactions with p38alpha protein. *Curr. Med. Chem.* 2005, *12*, 2979–2994.
53. Kohno, M., Pouyssegur, J. Targeting the ERK signaling pathway in cancer therapy. *Ann. Med.* 2006, *38*, 200–211.
54. Dean, E. J., Ranson, M., Blackhall, F., Holt, S. V., Dive, C. Novel therapeutic targets in lung cancer: inhibitor of apoptosis proteins from laboratory to clinic. *Cancer Treat. Rev.* 2007, *33*, 203–212.
55. Hamdy, N. A. Targeting the RANK/RANKL/OPG signaling pathway: a novel approach in the management of osteoporosis. *Curr. Opin. Investig. Drugs* 2007, *8*, 299–303.
56. Stupp, R., Ruegg, C. Integrin inhibitors reaching the clinic. *J. Clin. Oncol.* 2007, *25*, 1637–1638.
57. Haddad, E. B., Landry, Y., Gies, J. P. Sialic acid residues as catalysts for M2-muscarinic agonist receptor interactions. *Mol. Pharmacol.* 1990, *37*, 682–688.
58. Fredriksson, R., Lagerström, M. C., Lundin, L.-G., Schiöth, H. B. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Molec. Pharmacol.* 2003, *63*, 1256–1272.
59. Vassilatis, D. K., Hohmann, J. G., Zeng, H., Li, F., Ranchalis, J. E., Mortrud, M. T., Brown, A., Rodriguez, S. S., Weller, J. R., Wright, A. C., Bergmann, J. E., Gaitanaris, G. A. The G protein-coupled receptor repertoires of human and mouse. *Proc. Natl. Acad. Sci. USA* 2003, *100*, 4903–4908.
60. Chung, S., Civelli, O. Orphan neuropeptides. Novel neuropeptides modulating sleep or feeding. *Neuropeptides* 2006, *40*, 233–243.
61. Levoye, A., Dam, J., Ayoub, M. A., Guillaume, J. L., Jockers, R. Do orphan G-protein-coupled receptors have ligand-independent functions? New insights from receptor heterodimers. *EMBO Rep.* 2006, *7*, 1094–1098.
62. Brown, E. J., Frazier, W. A. Integrin-associated protein (CD47) and its ligands. *Trends Cell Biol.* 2001, *11*, 130–135.
63. Landry, Y., Niederhoffer, N., Sick, E., Gies, J.-P. Heptahelical and other G-protein-coupled receptors (GPCRs) signalling. *Curr. Med. Chem.* 2006, *13*, 51–63.
64. Isenberg, J. S., Romeo, M. J., Abu-Asab, M., Tsokos, M., Oldenborg, A., Pappan, L., Wink, D. A., Frazier, W. A., Roberts, D. D. Increasing survival of ischemic tissue by targeting CD47. *Circ. Res.* 2007, *100*, 712–720.
65. Kaczorowski, D. J., Billiar, T. R. Targeting CD47: NO limit on therapeutic potential. *Circ. Res.* 2007, *100*, 602–603.
66. Landry, Y., Gies, J. P. Heterotrimeric G-proteins control diverse pathways of transmembrane signaling, a base for drug discovery. *Mini. Rev. Med. Chem.* 2002, *2*, 361–372.
67. Iwatsubo, K., Okumura, S., Ishikawa, Y. Drug therapy aimed at adenylyl cyclase to regulate cyclic nucleotide signalling. *Endocr. Metab. Immune Disord. Drug Targets* 2006, *6*, 239–247.
68. Ke, H., Wang, H. Crystal structures of phosphodiesterases and implications on substrate specificity and inhibitor selectivity. *Curr. Top. Med. Chem.* 2007, *7*, 391–403.
69. Wetzker, R., Rommel, C. Phosphoinositide 3-kinases as targets for therapeutic intervention. *Curr. Pharm. Des.* 2004, *10*, 1915–1922.
70. McMullen, J. R., Jay, P. Y. PI3K (p110alpha) inhibitors as anti-cancer agents: minding the heart. *Cell Cycle* 2007, *6*, 910–913.
71. Fingar, D. C., Blenis, J. Target of rapamycin (TOR): an integrator of nutrient and growth factor signals and coordinator of cell growth and cell cycle progression. *Oncogene* 2004, *23*, 3151–3171.
72. Doukas, J., Wrasidlo, W., Noronha, G., Dneprovskaia, E., Hood, J., Soll, R. Isoform-selective PI3K inhibitors as novel therapeutics for the treatment of acute myocardial infarction. *Biochem. Soc. Trans.* 2007, *35*, 204–206.
73. Frojdo, S., Cozzone, D., Vidal, H., Pirola, L. Resveratrol is a class IA phosphoinositide 3-kinase inhibitor. *Biochem. J.* 2007, 406, 511–518.
74. Musa, M. A., Khan, M. O., Cooperwood, J. S. Medicinal chemistry and emerging strategies applied to the development of selective estrogen receptor modulators (SERMs). *Curr. Med. Chem.* 2007, *14*, 1249–1261.
75. Rosen, C. J. The rosiglitazone story. Lessons from an FDA advisory committee meeting. *N. Engl. J. Med.* 2007, 357, 844–846.
76. Rubenstrunk, A., Hanf, R., Hum, D. W., Fruchart, J. C., Staels, B. Safety issues and prospects for future generations of PPAR modulators. *Biochim. Biophys. Acta* 2007, 1171, 1065–1081.
77. Desvergne, B. RXR: from partnership to leadership in metabolic regulations. *Vitam. Horm.* 2007, *75*, 1–32.
78. Jakel, H., Fruchart-Najib, J., Fruchart, J. C. Retinoic acid receptor-related orphan receptor alpha as a therapeutic target in the treatment of dyslipidemia and atherosclerosis. *Drug News Perspect.* 2006, *19*, 91–97.
79. Cao, G., Bales, K. P., DeMattos, R. B., Paul, S. M. Liver X receptor-mediated gene regulation and cholesterol homeostasis in brain: relevance to Alzheimer's disease therapeutics. *Curr. Alzheimer Res.* 2007, *4*, 179–184.
80. Rizzo, G., Renga, B., Mencarelli, A., Pellicciari, R., Fiorucci, S. Role of FXR in regulating bile acid homeostasis and relevance for human diseases. *Curr. Drug Targets Immune Endocr. Metabol. Disord.* 2005, *5*, 289–303.
81. Pellicciari, R., Sato, H., Gioiello, A., Costantino, G., Macchiarulo, A., Sadeghpour, B. M., Giorgi, G., Schoonjans, K., Auwerx, J. Nongenomic actions of bile acids. Synthesis and preliminary characterization of 23- and 6,23-alkyl-substituted bile acid derivatives as selective modulators for the G-protein coupled receptor TGR5. *J. Med. Chem.* 2007, 50, 4265–4268.
82. Stein, R. A., McDonnell, D. P. Estrogen-related receptor alpha as a therapeutic target in cancer. *Endocr. Relat. Cancer.* 2006, *13*(Suppl 1), S25–32.

Chapter 5

Drug Targets, Target identification, Validation and Screening

Kenton H. Zavitz, Paul L. Bartel and Adrian N. Hobden

I. INTRODUCTION
II. IMPROVING THE RESOLUTION OF DISEASE ETIOLOGY
III. BIOPHARMACEUTICAL THERAPIES
A. Passive immunotherapy
IV. DRUG TARGET IDENTIFICATION
A. Rare mutations leading to generalized therapies
B. Mining the proteome
C. Yeast two-hybrid systems
D. RNA interference
V. HIT-TO-LEAD
A. Cell-based screening
B. Intracellular receptors
C. Intracellular enzymes
D. G-protein-coupled receptors
E. Transgenic animals
F. Drug metabolism
G. Toxicology
VI. CLINICAL BIOMARKERS
VII. CONCLUSIONS
REFERENCES

Science may set limits to knowledge, but should not set limits to imagination.
Bertrand Russell, *British author, mathematician, & philosopher (1872–1970)*
[Source: http://www.quotationspage.com/quotes/Bertrand_Russell]

I. INTRODUCTION

Over the past 50 years, the pharmaceutical industry has been extremely successful in its search for new and improved medicines. However, a quick survey of the world's best-selling drugs reveals that the majority are small molecules which were discovered by using natural product screening, medicinal chemistry and animal testing but without the aid of modern molecular biology technology. If the traditional drug discovery paradigm was so successful, why do we need molecular biology? Of course, we should not forget that molecular biology is a relatively new science dating only from 1975[1] and the process of drug discovery, refinement and testing can take a long time. It is, therefore, not surprising that the current drugs are just beginning to reflect the revolution that has occurred in the pharmaceutical and biotechnology industries. It is very unlikely that any of tomorrow's drugs will not have benefited from molecular biology at some stage in their discovery. Indeed, for most new drugs, molecular biology technology will have been used, directly or indirectly, at all stages in the drug discovery process.

In its infancy, molecular biology or "genetic engineering" was considered to be useful only for the production of therapeutic proteins. Many companies, for example, Genentech and Biogen, were founded solely with that objective in mind. However, proteins do not make ideal drugs, being costly to produce, difficult to administer, rapidly cleared and potentially immunogenic. Despite these disadvantages, a rapidly increasing number of "biopharmaceuticals," including recombinant proteins, therapeutic monoclonal antibodies, polyclonal antibodies and even antisense oligonucleotides (e.g. Vitravene for cytomegalovirus (CMV) retenitis), have been approved by the US Food and Drug Administration (FDA) for indications ranging from metastatic breast cancer (Herceptin) to rheumatoid arthritis (Remicade, Enbrel).[2] These biopharmaceutical therapies have been made possible by advances in molecular biology that allow the routine cloning of genes, expression of the corresponding proteins and the purification of the resulting product in commercially viable quantities, as well as a favorable regulatory environment. Nonetheless, the pharmaceutical industry has begun to exploit an ever

Copyright © 2008, Elsevier Ltd
All rights reserved.

expanding array of new molecular technologies in the drug discovery process beyond simply the development and production of therapeutic proteins. In particular, the initial steps of drug target identification and validation as well as drug screening technologies have undergone the greatest paradigm shift over the last 10 years.

The intention of this chapter is not to describe, in great detail, the techniques of molecular biology. There are numerous specialized textbooks available to those who wish to learn them. Nor do we want to detail the process of drug discovery. That is covered elsewhere within this book. Rather, this chapter will illustrate the various uses of modern molecular biological technologies in the various stages of the drug discovery process with an emphasis on drug target identification, validation and screening. Some of these applications are well established, almost mature for such a young science, others are just now being applied and still more applications will be conceived of and brought to fruition in the future. The essence of pharmaceutical research is innovative thought and competition. The winners will be those who have the best ideas and can most rapidly exploit them by bringing a drug to market.

II. IMPROVING THE RESOLUTION OF DISEASE ETIOLOGY

Throughout the history of medicine, drugs have been designed to target the cause of a particular disease. Early in its history, the "target" may have been based on epidemiological evidence or superstition. As microbes were discovered and cultivated, the resolution of the target improved and lead to our first antibiotics and vaccines. The creation of electrophysiological techniques allowed researchers to identify and characterize the activity of certain proteins in relative isolation, using a patch clamp to measure ion channel currents. Again, the resolution of the target improved and drug discovery efforts created anti-psychotic, anti-hypertensive and cardiac drugs. The Genomic Era we are currently living in offers unprecedented resolution of our target. This resolution has become so great, that we can actually visualize, via X-ray crystallography, the substructures of proteins that we would like to modulate. This era was made possible largely through technical achievements in molecular biology. We now face a new problem in drug discovery as increased resolution creates greater complexity. Again, it is molecular biology techniques that will advance current and future drug discovery efforts: aiding in target identification, hit-to-lead selection, preclinical development and clinical efficacy.

Clearly, as the resolution of the target for a particular disease increased, investigators have been able intervene and treat its symptoms. Treatment of symptoms is indeed the foundation of medicine and even most of the cutting-edge therapies are symptomatic in nature. Anti-tumor strategies, in general, are symptomatic: an attempt to alleviate the symptom of unregulated cell proliferation. However, this does not diminish their importance to the patient. Molecular biology has played a critical role in the discovery of more effective and less toxic weapons against cancer. Determining the molecular motors behind a particular "subspecies" of cancer offers a targeted approach at intervention, not offered by non-specific cytotoxic agents which indiscriminately kill all rapidly dividing cells. Herceptin (trastuzumab) is a prime example of a molecular biology approach to cancer therapy. This human moncloncal antibody targeting the HER2/neu tyrosine kinase receptor was introduced by Genentech in 1998. After cloning the HER2 gene, genetic analysis of almost 200 primary breast tumors showed a 2- to greater than 20-fold gene amplification in 30% of the tumors. There was also a strong correlation between HER2 amplification of decreased survival and time to relapse in breast cancer patients.[3] While small molecule inhibitors to this target have proven unsuccessful in the clinic either due to lack of potency or specificity, Herceptin is now recommended for use in the approximately 25% of breast cancer patients that overexpress the HER2 receptor. This monoclonal antibody inhibits HER2 cell signaling by binding to the extracellular receptor, rather than targeting the intracellular domains like small molecules. Creation of this drug required great skill in a number of molecular biology techniques: generating a highly efficacious antibody specific to its target, humanization of this antibody so as to not illicit an immune response, and production of this antibody on a commercial scale.

Another example of a high impact symptomatic therapy is the use of acetylcholinesterase (AChE) inhibitors in treatment of Alzheimer's disease (AD). These drugs have offered relief to many AD patients by boosting the levels of a critical neurotransmitter, acetylcholine, in regions of the brain damaged by the advancing pathological hallmarks of AD, amyloid plaques and neurofibrillary tangles. To use our earlier analogy, the target resolution of this therapy is at the degenerating synapse. Symptomatic relief, however, is temporary because these AChE inhibitors do not attack the cause of the disease that creates this AD-specific neurodegeneration. Next generation AD therapies aim to modulate targets that prevent disease progression, rather than treat its symptoms. The development of disease modifying drugs may in fact become the greatest utility of a molecular biology approach to drug discovery. To date, few marketed drugs strictly satisfy the criteria necessary to be called disease modifiers on their package inserts, for example Enbrel (TNF-α antibody therapy) currently used to treat rheumatoid arthritis and interferon therapy for multiple sclerosis (MS). This low number of disease modifying drugs highlights the challenges we face in drug discovery to alter the course of human disease. Currently, a multitude of potentially disease modifying therapies, targeting a multitude of mechanisms, are in clinical trials for AD. As an example throughout this chapter, we will the highlight molecular biology approaches used to identify potential

causes of AD, bring AD drugs into clinical trials and determine efficacy of disease modification with biomarkers.

III. BIOPHARMACEUTICAL THERAPIES

As mentioned in the introduction, the ability to move DNA from man to bacteria, or indeed from bacteria to man, made it possible, suddenly, to do what had previously been impossible. Human proteins could be produced in sufficient quantities to make it possible to use them as drugs. The first commercial example was human insulin which has now taken over from porcine insulin as the drug of choice for type 1 diabetics.

The techniques of molecular biology started to reveal a whole range of proteins that could be used as drugs. But their structure did not always allow successful production in bacteria. As a general rule, *Escherichia coli* neither readily secrete proteins nor will they glycosylate them. As a consequence, if the protein required a large number of specific disulfide bonds or glycosylation for activity, *E. coli* were unsuitable hosts for their production. Although it was possible to produce the protein, it was unfolded and usually precipitated within the cell. No amount of protein re-folding *in vitro* could produce reasonable quantities of active product. It became necessary to use other protein expression systems. Today we have a vast array of systems from which to choose, each with its own advantages and disadvantages. For example, the yeast *Saccharomyces cerevisiae* is easy to grow and to manipulate genetically and will secrete proteins. However, quantities of secreted protein tend to be low and the glycosylation profile of proteins secreted from yeast is distinct from that of mammalian cells. Most of the therapeutic proteins currently on the market, for example, erythropoietin, G-CSF and tPA, are produced in mammalian cell expression systems. Obviously, these cells will secrete and glycosylate the protein in a manner similar to the natural protein. However, the cells are harder to manipulate and much more expensive to grow than their microbial counterparts. Furthermore, the expression levels have until recently been relatively low. Newer expression systems based on viruses have started to make expression in complex eukaryotic cells much more straightforward due to the ease of getting foreign DNA into the cells and the high level of expression of recombinant protein following infection of the cells. Particular systems of great merit are baculovirus[4] which will infect certain insect cells and the semliki forest virus system[5]which has a very broad host range allowing a large number of different cell lines to be used.

Whilst therapeutic proteins were an obvious use for the technology, it is evident that any protein can be produced provided the right system is chosen. Drug discovery requires that, if small molecules are the objective, they should work against the correct target, that is, the human protein or specific viral enzyme. Genetic engineering often provides the only mechanism to acquire sufficient protein for high-throughput screening or X-ray crystallography to facilitate rational drug design. The technology is in routine use to supply proteins for these purposes. However, expression of recombinant proteins is not always a neutral event for the host cell. As an example, it was attempted some years ago to express HIV protease in *E. coli* in order to acquire sufficient material for high-throughput screening. It proved impossible, however, to express large quantities since the moment the cell was induced to make the HIV protease, it stopped growing (Figure 5.1). If the active site aspartic acid was mutated to asparagine (making the protease inactive), then large quantities of protein were produced by actively growing cells. It became apparent that production of active protein prevented cell growth, presumably because of the protease activity of the recombinant product and it occurred to several research groups[6] that it was unnecessary to purify large quantities of HIV protease to use in some biochemical screen for inhibitors of the enzyme. The recombinant *E. coli* could act as the screen. The bacteria would grow whilst expressing HIV protease provided an inhibitor of the enzyme was present.

FIGURE 5.1 Growth response of *E. coli* to the expression of HIV protease. *E. coli* containing the HIV protease gene were either induced (open squares) or uninduced (solid squares) to express HIV protease at 4 h (arrow) and their subsequent growth measured. The growth phenotype of such an *E. coli* strain could be used as a screen for inhibitors of the HIV protease.

A. Passive immunotherapy

Passive immunotherapy originated over a century ago when it was discovered that sera from one patient affected with diphtheria toxin could cure the diphtheria of an another individual.[7] Antibody treatments have long been viewed as promising potential therapies for a wide variety of diseases

due to their ability to bind with high affinity and selectivity. Although this treatment approach has long been desired, the first antibody treatment to hit the market was in 1986 with Muromonab-CD3 for the treatment of transplant rejection. Currently, there are over 50 antibody therapies on the market with many in late stage development. Molecular biology techniques were essential in making these exciting treatments possible for patients.

There were two obstacles that antibody therapies had to overcome to get into the clinic. First, a production method would be needed to produce and purify antibodies on a massive scale. This has largely been accomplished by creating hybridoma cell lines able to grow to high density, in serum-free conditions, and other necessities critical for commercial production. Second, non-human antibodies will illicit an immune response regardless of the epitope. Therefore, antibodies needed to be produced by a rodent, but appear human to our immune system. Molecular biology techniques were used to create chimeric and later fully "humanized" antibodies, the distinction is characterized by the ratio of mouse and human DNA.

Drug discovery efforts in AD are particularly focused on this approach. Many different antibodies, all targeting various regions and forms of beta-amyloid (Aβ), the peptide fragment that aggregates and deposits as senile plaques in the brain of an AD patient,[8] are being tested in the clinic. The popularity of this approach largely originates from the clinical trial of an active immunotherapy approach (the AN-1792 vaccine targeting Aβ). Although the trial was halted due to an inflammatory response in a small subset of patients, continued analysis of previously dosed patients led to optimism that antibodies directed against Aβ are potentially effective in clearing amyloid plaques from the brain and may improve cognitive function.[9] Bapineuzumab, the most advanced passive immunotherapy targeting Aβ is currently being studied in Phase 3 clinical trials in patients with AD.

IV. DRUG TARGET IDENTIFICATION

A. Rare mutations leading to generalized therapies

Traditional genetic analysis of human subpopulations (families and isolated cultures) having rare inherited diseases was instrumental in the identification of molecular targets for certain diseases. Barrter's syndrome, for example, a rare inherited disorder is characterized by hypokalemia, alkalosis and low blood pressure. Four of the genetic lesions were identified in the late 1990s, isolating the cause of the disease to loop of Henle region of the kidney.[10] More importantly, this research led to the discovery of the function of each of these genes, that each gene's function was interdependent on the other, and that the concerted effort of all four was necessary for normal kidney physiology.

In AD, this type of mendelian analysis could be credited with discovering the mechanism for the pathophysiology of the disease. Early onset AD is as devastating as it is rare – less than 10% of AD patients contract this familiar form with an average onset age of around 55. Identification of genetic lesions in three genes, the amyloid precursor protein (APP), and the PS1 and PS2 components of the gamma secretase protease,[11–14] provided clues to the players involved. Characterizing the phenotypes of these mutated genes when expressed in cells and transgenic mice added to our current understanding of the "amyloid cascade hypothesis,"[15] a model for the cause of all AD, not just the early onset familial form.

B. Mining the proteome

At the time of press, the complete DNA sequences of 37 different eukaryotic genomes are publically available, including five species of primates. Genetic analysis alone has unlocked the potential role for many genes through the identification of common genetic domains or chromosome localization. The human "Kinome," an effort to identify and classify all kinases was achieved in such a manner.[16] Amazingly, field leader Tony Hunter's pregenomic prediction of a 1,001 kinases was not far off the mark (there are 518 in the human genome). However, all past and present drugs target proteins, not DNA. Genetic therapies are the obvious exception, but due to the infancy of the field of gene therapy and lack of an approved drug, will not be covered in this chapter. Researchers have invented cleaver strategies to analyze the literal product of genomic sequencing efforts- the proteome.

An example of an effort to bridge the gap between genes and protein targets is the pursuit of novel substrates for the kinases mentioned above. This class of enzymes is responsible for a majority of the signal transduction in the cell and aberrant cell signaling underlies many pathophysiological conditions. The addition of so many uncharacterized kinases brought forth an intense effort to determine their substrates. Conversely, many interesting and well-studied drug targets are regulated by unknown kinases. Obtaining the answers to each of these unknowns is a perfect example of target identification in drug discovery. Molecular biology techniques are essential for this mining of the proteome.

Attacking this problem necessitated new techniques to analyze thousands of different proteins, most importantly, isolating potentially a single protein from this mix. To achieve this goal, researches put a new twist on the classical viral phage display technique of genetic cloning. A lambda phage library was created to incorporate cDNA from specific tissues or tumors. Bacteria were then transformed to conditionally express a kinase substrate of interest. The final tool needed was a phospho-specific antibody to specifically detect when this protein was activated (phosphorylated). The bacteria was infected and chemically induced to

express both the substrate and the unknown cDNA inserted in the lambda phage. As the viral life cycle transitions from lysogenic to lytic, bacteria are lysed and plaques form on the agar. The proteins expressed in these plaques are then lifted onto membranes and probed to see if the interaction of the substrate and expressed cDNA in the bacterial cell resulted in substrate activation (phosphorylation). The critical element to this technique is that the research can rapidly and easily determine the identity of positive interactions by going back to the original plaque and sequencing the DNA it contains.[17] This method has also been used to identify substrates of related enzymes (phosphatases).[18] This experiment highlights the importance of functional genomics in drug discovery. Other examples of harvesting functional interactions between specific members of the proteome are addressed below.

C. Yeast two-hybrid systems

As an example of the application of these new approaches, we will examine the yeast two-hybrid system for the identification of protein–protein interactions within a cell. This system consists of a yeast genetic assay in which the physical association of two proteins is measured by the reconstitution of a functional transcriptional activator to drive a reporter gene such as β-galactosidase or an auxotrophic marker for selection[19,20] (Figure 5.2). In general, a protein of known function, such as a disease causing protein from an inherited syndrome, is fused to the DNA-binding domain (DBD) of a transcription factor, for example, GAL4. This hybrid "bait" protein is introduced into a yeast strain along with a library of human "prey" proteins fused to a transcriptional activation domain. Activation of the reporter gene indicates that a direct interaction has occurred between the "bait" protein and the selected "prey" protein which is then easily recovered and identified by DNA sequence. Since its introduction in 1989, a huge number of interacting proteins from a vast variety of studies have been used by small and large laboratories alike to piece together previously undiscovered biological pathways. For example, this methodology was instrumental in tying the APC colon cancer gene, into the Wnt/β-catenin signal transduction pathway. The advantages of the yeast two-hybrid system include its rapidity, low cost, robustness and applicability to virtually any protein of interest, and its adaptability to automation and high-throughput methodologies. The difficulty comes in attempting to determine which of this vast array of interactions are of biological relevance and furthermore which interactions can be linked to the disease state of interest. The necessary target validation steps developed thus far are of greatly lower-throughput. Current efforts to improve this situation include the development of human cell-based assays that provide a suitable biological readout (e.g. apoptosis or cell adhesion for carcinogenesis) coupled to automated methodologies to express full-length or

FIGURE 5.2 The yeast two-hybrid system. (a) Two DNA vectors are engineered to separately express hybrid proteins in yeast. The "bait" consists of the target protein fused to the DBD of the GAL4 transcription factor. The "prey" protein is fused to the GAL4 transcriptional activation domain (AD). This prey can be a known protein, as in (b), or a random cDNA from a library. If a high affinity protein–protein interaction occurs between the bait and prey, a functional transcriptional activator is created which results in β-galactosidase expression and a phenotypic color change of the host yeast cells (with the addition of colorometric substrate). However, if no interaction occurs, the colonies remain white.

smaller domains of the candidate proteins of interest.[9] These approaches will allow for the functional analysis of many candidate drug targets in parallel.

An additional development of enormous potential is that variations of the yeast two-hybrid system are now themselves being used as drug screens.[21] Such a screen is designed to detect small molecules capable of specifically affecting the association of the two target proteins in yeast.

(b) Target specificity confirmation

(c) Compound screening

FIGURE 5.2 (*Continued*) (b) To characterize the binding site between two proteins, a proposed functional domain was mutated (DM-Bait). Prey 1 will only bind to the wild-type bait while Prey 2 binds both the wild-type and domain mutant. This suggests that the Prey 1/bait interaction occurs at the mutated functional domain while the Prey 2/bait interaction lies outside of this region. (c) Compound screening is used to identify small molecules that disrupt a specific protein–protein interaction. To reduce "false positive" hits from compound-mediated cytotoxicity, a dual reporter system was created. Constitutive expression of the gusA gene causes measurable fluorescence (*x*-axis). β-Galactosidase activity (*y*-axis) results from binding of the two targets the screen is design to disrupt. Therefore, a low level of β-gal activity with a correspondingly high-level of gusA activity would denote a non-toxic disruption of the protein–protein interaction (compound A). Compound B would be considered a false positive as it inhibits the activity of both reporters and is therefore likely to be a cytotoxic agent. Source: Data provided by Siavash Ghaffari, personal communication.

Recently, a high-throughput yeast two-hybrid screen successfully identified small molecules that disrupt the interaction between the α1B and β3 subunits of the N-type calcium channel. These compounds were subsequently found to selectively inhibit the activity of the N-type calcium channel in neurons in culture and may thus serve as the basis for a structure–activity synthetic chemistry program.[22] The success of this approach raises the exciting and daunting possibility that, in addition to the traditional drug targets of G-protein-coupled receptors (GPCRs) (60% of drugs on the market), ion channels, nuclear hormone receptors and enzymes (proteases, kinases), virtually any protein, even those of unknown biochemical function, may serve as a viable target for therapeutic intervention. These novel classes of drug targets will certainly present new challenges to the practice of medicinal chemistry in the future.

D. RNA interference

RNA interference (RNAi) has found remarkable utility in the drug development process through its ability to affect gene expression in a specific manner. In some respects, it may be regarded as a surrogate for the pharmacological knockdown of protein activity. Because of its ability to silence the expression of specific genes it has found widespread application in both cellular and animal model systems. It has also been employed in large-scale screens to systematically knock down sets of genes and identify those that affect cellular phenotypes related to disease etiology. Application of RNAi technology has been an invaluable aid in target identification, target validation, the establishment of mechanism-based cellular models, and proof-of-principle experiments.[23–27] In addition to its application in the laboratory, significant effort is being made to exploit RNAi for therapeutic purposes.[28]

RNAi relies on the action of 19–25 base pair-long small interfering RNA strands (siRNA) that most often lead to the degradation of specific messenger RNAs (mRNA). One of the strands of an siRNA acts as a guide strand which is incorporated into the RNA-induced silencing complex (RISC) and pairs with the complementary strand of a target mRNA. This induces the cleavage of the mRNA by the argonaute protein, which is the catalytic component of the RISC complex. The cleaved mRNA is thus unavailable for translation of its encoded protein sequence (Figure 5.3).

RNAi was first described in plants but has been demonstrated in a range of eukaryotic organisms.[29–32] In cells from many eukaryotic organisms, including *C. elegans*, *Drosophila* and plants, long double-stranded RNA (dsRNA) molecules can be introduced to initiate RNAi. The enzyme Dicer cleaves these dsRNA molecules into short siRNA fragments of 20–25 base pair. In mammalian cells, long siRNAs induce an interferon response, and therefore, short dsRNAs are used instead. These dsRNA molecules are introduced

into cells either as 21–23 base pair siRNAs or as viral- or plasmid-encoded short hairpin RNA (shRNA) constructs. shRNA constructs direct the expression of a pair of complementary sequences connected by a hairpin linker region.[33] The hairpin is cleaved from the transcribed shRNA by Dicer to generate a functional siRNA. Whether introduced exogenously or expressed endogenously these siRNAs engage RISC and lead to silencing of the target gene.

Chemically synthesized siRNAs and shRNA-expressing vectors are available from a number of commercial sources. Sequences may be available premade or can be designed using proprietary or publically available algorithms. An alternative to synthesized siRNAs is esiRNA in which short RNA fragments are generated from *in vitro* synthesized dsRNA by RNase III digestion.[34] This approach generates an overlapping collection of siRNA molecules and may offer reduced off-target effects.[35] The selection of any particular approach is dependent on the nature of the experiment that will be conducted.

Within the drug development process, RNAi technology has had the greatest impact on target identification. One of the first publications utilizing siRNA technology identified the Tsg101 protein as essential for viral budding and created a new target, with a new mechanism of action for HIV therapy.[36] Prior to RNAi, the identification of new drug targets was largely dependent on the association of gene/protein expression differences or genetic changes with disease. RNAi-based cellular screens allow the systematic evaluation of the effect of protein suppression in disease-relevant cellular models. Screens have been successfully conducted with libraries of shRNA-expressing vectors,[37–39,27] chemically synthesized siRNAs[40,24] and esiRNAs.[41,42] One example is a screen for host proteins that are required for HIV infection. Pools of four siRNAs for each of over 21,000 genes were transfected into TZM-bl cells.[24] After 72 h, HIV virus was added and 48 h later the cells were stained for HIV p24 protein expression. An aliquot of supernatant was also applied to cells carrying a tat-dependent β-galactosidase reporter gene to capture effects on late events such as viral assembly and budding. Pools that showed an effect on viral replication in the absence of effects on cell viability were confirmed by rescreening the four individual siRNAs from each pool in the assay. From this approach, over 250 HIV-dependency factors were identified. Brass *et al.* further characterized the roles of proteins involved in retrograde vesicular transport (Rab6A and Vps53), nuclear import (TNPO3) and the mediator complex (Med28) in HIV infection.[24]

An elegant integrated approach to target identification involved the combination of an shRNA screen with DNA copy number analysis, a kinase overexpression screen, and a number of pathway-directed assays to implicate IKKε as a breast cancer oncogene.[27] A screen of shRNAs targeting 1,200 genes for ones that are required for tumor cell proliferation identified IKKε, among others, as important for the proliferation and viability of MCF-7 breast cancer cells. DNA copy number analysis of tumor cell lines demonstrated amplification of 1q32, which harbors the IKKε locus, in 8 of 49 breast cancer cell lines. Finally, an overexpression screen in HEK cells engineered to overexpress a constitutively active MEK1 demonstrated that myristoylation sequence-tagged IKKε could cooperate to induce anchorage-independent growth of the cells. The authors then further characterized the effect of IKKε modulation, both by overexpression and siRNA, on NF-κB signaling in tumor cells.

The further validation of drug targets *in vivo* often makes use of pharmacological tools and may require the generation of genetically modified animal strains. These tools may be expensive, difficult or even impossible to obtain. Increasingly, RNAi approaches are being used in *in vivo* settings to confirm the role of particular genes in the disease process. In cancer programs, xenografts of shRNA-transfected cells are commonly used to examine the effect of target suppression on tumor development. For example, shRNA suppression of PKCε in MDA-MB231 breast cancer cells was shown to decrease cell proliferation, invasion and motility.[25] In an orthotopic mouse model of breast cancer, grafts of cells harboring PKCε shRNA had reduced tumor volumes and a lower incidence of lung metastasis as compared to controls. In a different application of target validation applied to AD, lentiviral vector-encoded shRNAs directed against BACE1 were delivered by intercranial injection into the hippocampus of APP transgenic mice.[26] Treatment with BACE1 shRNA resulted in lower levels of BACE1 protein, Aβ peptides, and fewer amyloid plaque deposits. In addition, animals injected with BACE1 shRNA showed amelioration of performance deficits in a water maze test.

One limitation to RNAi technology is the prospect of generating off-target effects. While an siRNA reagent can be designed to avoid complementarity to untargeted genes, short stretches of homology to other genes may allow an siRNA to act as a microRNA (miRNA) which inhibits translation from selected mRNA molecules. Several publications describe the off-target effects of siRNA by looking at mRNA levels using RNA chip technology. The degree to which an mRNA is knocked down varies with the specific siRNA sequence. While some siRNAs directed at a target gene may quite potently reduce its mRNA expression others may fail to reduce the message at all. Often researchers will experiment with a larger number (four or more) of siRNAs for a particular protein target to better understand off- versus on-target effects and to improve the odds of identifying potent RNAi reagents. The degree of knockdown is also influenced by how well the siRNA can be transfected or electroporated into the target cell line. Introduction of synthetic siRNA into certain cell lines may be inefficient thus limiting its effect on protein and mRNA levels. Thus, the level of protein knockdown may vary by siRNA, cell

type, and transfection/electroporation protocol and these differences may translate into distinct phenotypes.[43]

While siRNA may be utilized as a surrogate approach for pharmacological inhibition these methods differ in a number of respects.[44] Pharmacological targeting of a protein may specifically affect one activity of the protein such as enzymatic catalysis. Treatment with siRNA reduces the level of the entire target protein, reducing not only its enzymatic activity but also affecting its interactions with other proteins. This may impair the assembly of entire protein complexes in a way that is not recapitulated by a small molecule inhibitor. In cells, siRNA may take hours or days to affect gene silencing as opposed to the rapid inhibition provided by a small molecule. In addition, the RNAi effect may not be as durable as pharmacological inhibition; by 4–5 days post-treatment the siRNA effect may be greatly diminished. This may be an important issue in longer-term assays. If longer-term gene suppression is required, use of an shRNA-expressing vector may be more appropriate.

When employed properly, and coupled with data generated using orthogonal approaches such as protein overexpression, dominant negatives, and pharmacological modulation, RNAi approaches have aided in the identification of numerous potential disease targets. Many of these targets are advancing through the drug development process. Because of its many applications, it is likely that RNAi will contribute significantly to the development of novel therapeutic approaches in the future.

V. HIT-TO-LEAD

Equally important to the identification of important validated targets for drug intervention, is the identification of the actual drug. The creation of a drug is a long and iterative process, which originates from a simple "hit" often from an *in vitro* biochemical assay. In this section, we will address the impact that molecular biology has had in hit identification, as well as important cellular and animal models used to define the mechanism of action. These types of studies, in the in current era of rational drug design, are critical as a drug development program matures, and have become an important hurdle to clear before progression into the clinic.

A. Cell-based screening

As target resolution increases, the mechanism of action in drug discovery has become much more complex. Assay development methods have used molecular biology techniques to keep pace. Cells expressing recombinant proteins can act as biosensors for real-time analysis of target inhibition in the cell instead of relying on traditional cytotoxicity. These target-, rather than phenotypic-, based assays permit a much more accurate design to the structural–activity relationship (SAR) studies performed by medicinal chemists. High-throughput imaging assays can now be combined with siRNA library transfections to identify genetic targets that display a desired phenotype, like altered spindle fiber formation (Figure 5.3).

B. Intracellular receptors

We have come to recognize that the intracellular receptor gene family is both large and diverse. Its best-characterized members are the sex hormone receptors, for estrogen, progesterone and testosterone, but also included are receptors for corticosteroids, vitamin D_3, thyroxine and retinoids. In addition, molecular biology has revealed a number of "orphan" receptors, that is, proteins known to be produced that carry a sequence motif suggestive of the ability to bind a small molecule, but for which the ligand is currently unknown.

This family of receptors expressed, as their name suggests, within the cell are already the targets for many drug discovery programs. For example, tamoxifen is widely used for the treatment of breast cancer and is an antagonist of the estrogen receptor. Many synthetic analogs of corticosteroids are used in asthma treatment. The estrogen receptor is present in the cytoplasm in association with HSP90, a heat shock protein. Upon binding estrogen, this complex dissociates and the receptor enters the nucleus where it binds to specific DNA sequences and activates transcription of certain genes. This chain of events has been reconstructed in the yeast, *S. cerevisiae.*[45] The estrogen responsive DNA sequence was inserted into a yeast promoter upstream of a reporter gene. The reporter, in this case β-galactosidase, is usually an enzyme whose presence can be detected simply by a colorimetric indicator. The effect of inserting the DNA sequence into the yeast promoter is to render it inactive until bound by an estrogen receptor/estradiol complex. Obviously, therefore, the yeast must also express the receptor. With this combination of receptor, responsive element and indicator (Figure 5.4), the yeast is ready to be used as a screen for estrogen agonists or antagonists. Similar systems have been reported for corticosteroids[46] and androgens.[47]

Of course, the above is a rather simplistic description of the screen. In reality, the screener is seeking to achieve stability and sensitivity in the screen. The recombinant yeast must, therefore, be "fine-tuned" to ensure that the "foreign" DNA is not lost upon frequent growth of the cells and the concentration of estrogen receptor is sufficient, but not too high, so as to detect small quantities of active material. Once a therapeutic opportunity has been defined for the growing number of newly discovered orphan receptors, it is likely that agonists and antagonists will be sought using this technology.

FIGURE 5.3 RNAi relies on the action of 19–25 base pair-long siRNA that lead to the degradation of specific mRNA. siRNA can either be chemically synthesized or generated enzymatically by the endonuclease Dicer. One of the strands of an siRNA acts as a guide strand that is incorporated into the RISC and pairs with the complementary strand of a target mRNA. This induces the cleavage of the mRNA by the argonaute protein, which is the catalytic component of the RISC complex. The cleaved mRNA is thus unavailable for translation of its encoded protein sequence. Target knockdown can be monitored by Western analysis or mRNA quantification. *Source*: Data provided by Aaron Rogers, personal communication. Protein knockdown may lead to a cellular phenotype; shown above is an image selected from an siRNA screen for abnormal mitotic spindle morphology (blue = nuclei, green = microtubules, red = phospho histone H3). *Source*: Joshua Jones, personal communication.

FIGURE 5.4 Yeast-based screen for agonists or antagonists of the human estrogen receptor. The hormone estrogen (estradiol) binds to the estrogen receptor which is expressed from a gene driven by the PGK promoter. The hormone-receptor complex binds to an estrogen responsive element (ERE) that controls the expression of the β-galactosidase reporter gene. The assay measures the activity of the enzyme using a substrate that forms a colored product on conversion.

C. Intracellular enzymes

It was mentioned earlier that the expression of HIV protease within *E. coli* gives rise to a phenotype. In a similar fashion, it has been observed that phosphodiesterases (PDEs) when expressed in yeast affect the cells. These enzymes function to modulate intracellular concentrations of the cyclic mononucleotides cAMP or cGMP. Yeast has two endogenous genes encoding PDEs which, when deleted, lead to elevated levels of cAMP within the cell. The consequence to the yeast of elevated cAMP is increased sensitivity to heat shock and inability to utilize acetate as sole carbon source. These yeast mutants may be complemented by the human PDE gene and the phenotype reversed (Figure 5.5). The use of such yeast in the search for inhibitors of PDEs with utility in, for example, asthma has been proposed[48] and certainly works with the known type IV PDEs inhibitor, rolipram.

In a similar fashion, it is evident that the estrogen screen described above, could be modified to include enzymes required for the synthesis or degradation of estradiol. An alternative therapeutic objective for estradiol inhibition might be to prevent its synthesis. Thus, a yeast strain already built to be sensitive to estradiol could be supplied instead with the precursor to estradiol, 19-nortestosterone, and the enzyme, aromatase, required for its conversion to estradiol. An inhibitor of the enzyme would, therefore, lead to the inability to synthesize estradiol and the loss of production of β-galactosidase.

A major potential objection to the above approaches is that the compound is required to cross the yeast cell wall and membrane. Failure of a compound to do so would lead it to not being identified in this type of screen. Obviously, an *in vitro* biochemical screen does not suffer from this constraint. There is no simple argument to counter this objection but a series of observations should allow the reader to make some judgment on the relative merits of the two approaches. Firstly, biochemical assays can be expensive and complicated preventing their use in high-throughput screening. Secondly, screening of random compounds rarely results in a complete failure to identify leads. Rather, it is often difficult to decide which, of a series of structurally diverse but relatively inactive leads, should progress into medicinal chemistry. The mechanism by which compounds enter cells is poorly defined but there is considerable overlap in their ability to cross microbial and mammalian cell membranes. Starting with a compound already able to cross the membrane may well be advantageous to the medicinal chemist.

FIGURE 5.5 Yeast-based screen for inhibitors of human PDE IV. A PDE-deficient yeast (PMY) will not utilize acetate as a sole carbon source and is sensitive to heat shock (55°C). Complementation with a human type IV PDE (PMY + PDE) expressed from a copper-dependent (CUP1) promoter reverses the mutant phenotype. Addition of type IV PDE inhibitor (rolipram) to the complemented yeast restores the mutant phenotype (PMY + PDE + rolipram).

There are, of course, a number of targets for drug discovery that are not located within the cell. Rather, they are located within the cytoplasmic membrane where they serve to tell the cell about its environment. They are the cell surface receptors and have been, over the years, the targets of many of the world's best-selling drugs.

D. G-protein-coupled receptors

The GPCRs are a super-family of structurally related proteins, located in the cell membrane and consisting of 7-transmembrane segments. Their primary amino acid sequence, however, can be quite diverse. Agonists or antagonists acting at these receptors constitute a large number of today's best-selling pharmaceuticals. Examples include the H2 antagonists for ulcer therapy, β-blockers for hypertension, β-agonists for asthma and serotonin agonists for migraine.

In addition to the extensive families of these receptors that have small molecules as their agonists (e.g. histamine, prostaglandins, acetylcholine), many have peptides or even proteins as their ligand, (e.g. angiotensin II, gastrin, luteinizing hormone). There is, in addition, an extensive collection of "orphan" 7-transmembrane receptors, identified by molecular biology techniques but for which a ligand has not yet been identified.

There is enormous activity worldwide seeking to identify non-peptide agonists or antagonists for both the peptide receptors and the orphans, since it is expected that this will be a fruitful area for drug discovery. Witness the success of the non-peptide angiotensin II receptor antagonists (collectively known as the "sartans") approved several years ago for the treatment of hypertension.

The standard approach to finding such molecules has been to express the cloned human receptor in mammalian cells and look for molecules able to inhibit ligand binding. This method can be successful, as with the angiotensin II receptor antagonists, however, it is most useful for identifying antagonists (rather than agonists) and requires both the ligand to be known and for a radio-labeled ligand derivative to be

available. Recently, two novel approaches have been reported which potentially should facilitate the whole process.

The first system again makes use of yeast. It has been known for some time that *S. cerevisiae* can exist as two sexual types, a cells and α cells, which communicate with each other via sex pheromones, a-factor and α-factor. The receptors for these two pheromones are members of the 7-transmembrane family, although their amino acid sequences are quite distinct from their mammalian counterparts. The consequence of the binding of the pheromone to its receptor is to set in motion a complex set of biochemical events that lead, ultimately, to mating of the two opposite cell types. However, there are two principal events that can readily be detected. The cells undergo rapid, but transient, cell cycle arrest and express on their cell surface a variety of proteins that aid in fusion of the mating types. Unlike their mammalian counterparts, the intracellular signal is transmitted via the β- and γ-subunits of the trimeric G-protein complex and not by the α-subunit. A detailed description of this pathway can be found elsewhere.[49] More importantly, from the point of view of this chapter, the system has been engineered so that the yeast express the human β_2-adrenergic receptor and its cognate Gsα subunit instead of the yeast homologs.[50] The yeast respond to the presence of a β_2-agonist by inducing the FUS1 promoter which, in turn, has been connected to a β-galactosidase reporter gene. The yeast, therefore, turn blue in the presence of the indicator 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) (Figure 5.6).

It is evident that this yeast strain could be used to look for agonists or antagonists of this receptor and, because yeast cells can be grown rapidly and inexpensively, such a system has the potential to be used for very high-throughput screens. Indeed, a series of novel and selective peptide agonists of an orphan GPCR were identified in a screen conducted in a similar yeast system designed to couple receptor activation to histidine prototrophy as a selectable marker.[51] Many companies are seeking to exploit similar technology for their favorite 7-transmembrane receptors,

FIGURE 5.6 Yeast-based screen for β_2-agonists. The β_2-adrenergic receptor (B Ar) expressed from a GAL1 promoter links to the mating type response via the human Gs α subunit expressed off the CUP1 promoter, by complementation of GPA-1. The detection of signaling is by induction of the FUS promoter linked to a β-galactosidase reporter gene.

however, it is worth pointing out that these are very complex yeast strains to construct and not all receptors will be amenable to this approach.

As an alternative approach, the second system uses frog melanophores. This is an immortalized cell line derived from frog melanophores that responds to melanocyte-stimulating hormone or melatonin by, respectively, a dispersal or aggregation of pigment.[52] The consequence of addition of these ligands, which act at 7-transmembrane receptors, is that the cells change color within 30 min. Indeed, a dose–response curve for agonists or antagonists can be constructed using these cells in a 96-well plate in combination with an ELISA plate reader. Human GPCRs, such as the β_2-adrenergic receptor, as well as chemokine receptors and tyrosine kinase receptors have been expressed in these cells to generate a screenable bioassay for the identification and characterization of novel agonists and antagonists of these receptors.[53]

The system is especially powerful when one wishes to determine SAR for compounds derived in a medicinal chemistry program. Dose–response curves can be constructed versus the human receptor in less than 1 h. However, the system is not without its problems. Firstly, the cells need special conditions to grow. They are amphibian and, therefore, need lower temperatures and a frog-derived growth factor supplement. Secondly, they have an endogenous background of a variety of frog receptors that may complicate analysis. Thirdly, they are difficult to transfect with exogenous DNA. Nonetheless, the system has great promise and may even have application for high-throughput screening of random compounds.

In the process of drug discovery, one can imagine running the initial lead discovery part of the program in yeast, then switching to frog melanophores for the lead refinement stage. There is still, of course, a requirement that the compound will work *in vivo*. This is a combination of drug absorption, excretion and metabolism activity of the compound at its target *in vivo* and a lack of other activities, that is, toxicity. Molecular biology techniques are starting to address all these issues.

E. Transgenic animals

Advances in molecular biology and genetic transfer protocols have caused an explosion of genetically modified animals to investigate the role of specific genes in development, physiology and pathophysiology. Surprising new phenotypes have resulted, uncovering new roles for some genes in complex biological systems. For example, a knock-out mouse model for an ion transport protein ($Na^+/K^+/Cl^-$ cotransporter (NKCC)) involved in salt reabsorption in the kidney and fluid secretion in the gastrointestinal (GI) track displayed a puzzling phenotype of profound deafness. This finding led not only to the implication of NKCC activity as critical for the generation of endolymph in the inner ear, but also explained ototoxicity seen in some patients after the administration of loop diurectics for hypertension.[54,55]

Animal models, largely developed in mice, are an asset in drug discovery for more than just target identification. "Knock-in" molecular biology approaches, can replace the rodent target with its human homolog. Although small genetic changes between homologous mouse and human genes will not usually result in functional changes, they can drastically alter the affinity (and efficacy) of these rationally designed molecules for their target. It is therefore imperative to analyze *in vivo* efficacy using a human target expressed in rodent models whenever possible.

Unfortunately, examples of one genetic disruption equaling one disease, for example, cystic fibrosis, are exceedingly rare. Both academia and industry are involved in creating more accurate animal models to recreate multi-factorial disease states, a large leap forward from the first generation "knock-in" technology. Conditional knock-outs/knock-ins (external regulation of the timing for genetic disruption), *trans*-gene technologies and tissue or cell-specific expression/disruption will all be used to generate the most accurate reproduction of a human disease.

A good example of this effort to recreate human pathophysiology in a rodent model would be AD: the most common cause of dementia and afflicting 24 million people worldwide. With a large amount of funding, large number of investigators, and tremendous breadth and depth of research into the causes of AD, it may seem surprising that failure rate of AD drugs in the clinic is much higher than other disease indications. Not a single drug has been approved that slows the rate of progression of the disease (disease modification). One reason for this may be the lack of a clear animal model that recreates all of the histopathological and neurodegenerative hallmarks of AD. Animal models for AD largely exhibit only one of the two histological hallmarks for AD: amyloid-laden plaques and tau-based, neurofibrillary tangles. Genetic manipulation to increase one of these characteristics does not concurrently increase levels of the other, suggesting a complicated relationship. Until recently, the failure to merge tau and amyloid-based pathologies into a single animal model has unfortunately bifurcated AD drug discovery research. Hopefully, newer animals models that incorporate both pathological characteristics[56] will help to join this rift and provide a model for drug testing that is more predictive of human clinical success.

While not yet ideal, AD animal models have been instrumental in the development of potentially disease modifying drug therapies currently in clinical trials. Researchers focused on amyloid-lowering therapies, for example, will use an animal model designed to produce high levels of Aβ) resulting in the development of amyloid plaque pathology and deficits in learning and memory. Treating these mice with the compound tarenflurbil, an experimental drug shown to modulate the activity of the protease γ-secretase

to produce shorter, less toxic forms of Aβ, reduced brain levels of amyloid[57] and prevented the learning and memory deficits.[58] These studies helped to provide the rationale for the clinical development of the compound and tarenflurbil is currently in Phase 3 trials in patients with AD. Importantly, for neurodegenerative diseases, animal models must also be able to measure macro-scale neurodegeneration. An active area of research in drug development is to identify behavioral and cognitive models that test the activity level of specific regions of the brain (e.g. the hippocampus) primarily affected in AD, but not other dementias. Progress in both the creation and behavioral/cognitive assessment of AD mouse models over the last decade has directly resulted in drugs being brought into clinical trails.[59,60]

Advances in this field are of vital importance for drug development as animal models are often the last hurdle to clear before costly clinical trails. However, this technology is now being used earlier in the drug development process as well. Lower organisms, such as Drosophila (fly), *C. elegans* (worm) and more recently the Danio rerio (zebrafish), are being used to screen compounds for efficacy and toxicity. These techniques allow researchers to harness much more information regarding the mechanism of action when compared to traditional cell-based screening. The fast reproduction time, ease of husbandry and strain preservation, and low cost make these models important tools for preclinical drug research. Hopefully the advances being achieved today, implemented using molecular biology tools, will dramatically increase the success rate of future therapies in humans.

F. Drug metabolism

A major problem in all drug discovery programs is to discover compounds with good pharmacokinetics. Although it is possible to examine the metabolism of the drug in animals, it has often been difficult to predict what would happen in man. The obvious implications of drug metabolism are an effect on half-life *in vivo* and the production of toxic metabolic products.

In seeking to establish an effective dose for a new drug, the clinician needs to know what ranges of abilities humans will have to metabolize the drug and what effect the drug will have on the metabolizing enzymes. Failure to metabolize the drug may lead to overdose, whereas rapid metabolism could lead to lack of clinical benefit. Equally, inhibition of the metabolism of another drug could cause problems in a patient receiving several medications.

A large proportion of the metabolizing enzymes are members of the P450 superfamily[61] and a large number of these genes have now been cloned and their metabolic potential determined. Increasingly, the enzymes are being expressed in microbial systems, for example yeast, where their ability to metabolize the drug can be evaluated. In a few years, it would not be surprising if all new drugs were "typed" for their complete P450 metabolism profile. Equally, their metabolic products can be identified and their biological activity and toxicity determined.

An additional application is likely to be the P450 genotyping of patients. As "poor" metabolizers become recognized in the population, the problem is often found to be mutations in one or more of their P450 genes. Once identified, such mutations are easily screened for and it is entirely likely that some degree of P450 "profiling" will take place for patients in the future. Armed with knowledge on the metabolic fate of new drugs, the physician will then be able to prescribe the best drug for an individual depending on their P450 profile. This individualization of drug therapy based on genetic information is known as pharmacogenomics. There is a massive effort currently underway to identify and characterize polymorphisms in a wide variety of genes, including drug receptors and effectors, in addition to drug metabolizing enzymes. It is hoped that the correlation of these polymorphisms with clinical outcomes and drug effects across a population will allow for the prediction of the safety, efficacy and toxicity of both established drugs and new drugs in development and, thereby, a reduction in the size and expense of clinical trials.[62,63]

G. Toxicology

Toxicity testing for new drugs is a legal requirement of drug discovery and, of course, reflects our ignorance of biological processes. Toxicity is the unwanted effects of the molecule. Whilst it is hard to imagine that long-term testing of compounds in animals will not always need to be performed, molecular biology is starting to impact on genetic toxicology, that is, the ability of compounds to induce mutations in DNA and, thus, to act as potential carcinogens. Systems have been constructed which permit identification of genetic mutation *in vitro* and *in vivo* extremely rapidly, therefore, a compound's potential as a carcinogen can be identified with concomitant savings in numbers of animals, human effort and the supply of compound needed for the larger scale animal studies.

Most systems reported so far depend on the detection of mutations either in an indicator gene, for example β-galactosidase, or a gene controlling the expression of the indicator gene.[62,63] The bacteriophage λ, which normally infects and lyses the bacteria *E. coli*, has been altered genetically such that the β-galactosidase gene is contained within its DNA. This gene will only be expressed when the phage infects *E. coli.* The phage λ DNA is then incorporated into the mouse genome such that it is inherited in subsequent generations of mice. Since, the phage λ DNA is not capable of expressing any proteins in the mouse, it is effectively neutral in the mouse's growth and development. However, the λ DNA may be rescued from mouse by extracting

total DNA and adding a "λ packaging extract" *in vitro*. This complex, which is commercially available, finds and extracts the λ DNA and packages it into infectious phage particles. The phages are then used to infect *E. coli* where they replicate and lyse the bacteria. The lysed "plaques" can be stained for β-galactosidase. Mutations within the λ DNA are scored by the proportion of plaques scoring negative for β-galactosidase.

In practice, the transgenic mice are given the new drug over a period of a few days before sacrifice. DNA is then extracted from a variety of organs and the mutation frequency scored by counting λ plaques. It therefore becomes possible to demonstrate the ability of the drug to induce mutations and determine whether it shows any tissue selectivity. Furthermore, since the λ packaging reaction can be repeated several times for each DNA sample, far fewer animals are required to obtain a statistically significant result.

It can be expected that several refinements to this system will be developed with time allowing, for example, the scoring of mutation frequency to be automated.

VI. CLINICAL BIOMARKERS

The increased target resolution and well-defined mechanisms of action in drug development today, not only offer the potential of more specific and more efficacious drugs, but also increase the chances of measuring efficacy in the clinic. For example, measurement of cholesterol in the blood is now used to demonstrate efficacy of newer statin drugs. This has important implications for both the patient and the drug development companies. For the patient, cholesterol lowering is necessary to justify the administration of a drug with the potential for liver toxicity. Immediate feedback of the drug's efficacy will also increase compliance. The resolve to remain on drugs designed to prevent a life-threatening event (e.g. heart attacks) can diminish over time, especially if the drug possess unwanted side effects. Biomarker data can also have a large impact on clinical trial design. For statins, it is very costly in time and resources to design a clinical trial that shows diminished incidence of heart attacks over a 5–10-year period. Biomarker data allows investigators to directly measure the efficacy of cholesterol lowering in months not years.

Biomarkers development for the diagnosis of AD and to follow the potential treatment effects of new AD therapies is a very active area of research for these same reasons. If a set of biomarkers existed that could accurately diagnose AD or could be closely correlated with the progression of the severity of the disease, AD drug development would be completely restructured. Current diagnostic tools such as cognitive measures have difficulty in accurately identifying patients with the earliest form of the disease. However, this is the precise patient population that is likely to see the greatest benefit from new disease modifying drugs that may slow the onset of symptoms. Ideally, a biomarker would identify which patients will develop AD while they are still in the presymptomatic stage of the disease. AD biomarkers would also significantly speed up drug development. As with statin drug trials, future AD trials might last only a few months and measure a validated biomarker surrogate, perhaps a decrease in Aβ levels. Currently, disease modifying clinical trials are designed with long durations of treatment (18-months or more of therapy) with difficult to measure outcomes quantifying rates of cognitive decline. Finally, as with the statin drugs, a marketed AD drug with a biomarker correlated with clinical benefit will increase patient compliance. It is likely that the validation of biomarker surrogates will be used to speed drug development for many different medical indications in the future.

VII. CONCLUSIONS

The development of modern molecular biology has already had an enormous impact on the process of drug discovery and its influence will certainly increase in the future. The power of these new technologies such as RNAi will facilitate the discovery and development of novel pharmaceuticals and shorten the time and cost from idea to market. The pharmaceutical industry is very competitive and companies will prosper through a combination of hard work, innovation and serendipity. The companies that fully adopt the diverse and evolving techniques of molecular biology at every stage of the drug discovery process and use them in imaginative and inventive ways will ultimately find that serendipity has a less important role to play.

REFERENCES

1. Cohen, S. N., Boyer, H. W. Process for producing biologically functional molecular chimeras. *US Patent 4,237,224*, 1980.
2. Reichert, J. M. New biopharmaceuticals in the USA: trends in development and marketing approvals 1995–1999. *Trends Biotechnol.* 2000, *18*(9), 364–369.
3. Slamon, D. J., Clark, G. M., Wong, S. G., Levin, W. J., Ullrich, A., McGuire, W. L. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* 1987, *235*(4785), 177–182.
4. Miller, L. K. Insect baculoviruses: powerful gene expression vectors. *Bioessays* 1989, *11*(4), 91–95.
5. Liljestrom, P., Garoff, H. A new generation of animal cell expression vectors based on the Semliki Forest virus replicon. *Biotechnology (NY)* 1991, *9*(12), 1356–1361.
6. Baum, E. Z., Bebernitz, G. A., Gluzman, Y. Isolation of mutants of human immunodeficiency virus protease based on the toxicity of the enzyme in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 1990, *87*(14), 5573–5577.
7. Wei, M. Q., Metharom, P., Ellem, K. A., Barth, S. Search for "weapons of mass destruction" for cancer – immuno/gene therapy comes of age. *Cell Mol. Immunol.* 2005, *2*(5), 351–357.
8. Hawkes, C. A., McLaurin, J. Immunotherapy as treatment for Alzheimer's disease. *Expert Rev. Neurother.* 2007, *7*(11), 1535–1548.

9. Senior, K. Dosing in phase II trial of Alzheimer's vaccine suspended. *Lancet Neurol.* 2002, *1*(1), 3.
10. Gamba, G. Molecular physiology and pathophysiology of electroneutral cation-chloride cotransporters. *Physiol. Rev.* 2005, *85*(2), 423–493.
11. Lemere, C. A., Lopera, F., Kosik, K. S., Lendon, C. L., Ossa, J., Saido, T. C., Yamaguchi, H., Ruiz, A., Martinez, A., Madrigal, L., Hincapie, L., Arango, J. C., Anthony, D. C., Koo, E. H., Goate, A. M., Selkoe, D. J. The E280A presenilin 1 Alzheimer mutation produces increased A beta 42 deposition and severe cerebellar pathology. *Nat. Med.* 1996, *2*(10), 1146–1150.
12. Mullan, M., Crawford, F., Axelman, K., Houlden, H., Lilius, L., Winblad, B., Lannfelt, L. A pathogenic mutation for probable Alzheimer's disease in the APP gene at the N-terminus of beta-amyloid. *Nat. Genet.* 1992, *1*(5), 345–347.
13. Murrell, J., Farlow, M., Ghetti, B., Benson, M. D. A mutation in the amyloid precursor protein associated with hereditary Alzheimer's disease. *Science* 1991, *254*(5028), 97–99.
14. Scheuner, D., Eckman, C., Jensen, M., Song, X., Citron, M., Suzuki, N., Bird, T. D., Hardy, J., Hutton, M., Kukull, W., Larson, E., Levy-Lahad, E., Viitanen, M., Peskind, E., Poorkaj, P., Schellenberg, G., Tanzi, R., Wasco, W., Lannfelt, L., Selkoe, D., Younkin, S. Secreted amyloid beta-protein similar to that in the senile plaques of Alzheimer's disease is increased *in vivo* by the presenilin 1 and 2 and APP mutations linked to familial Alzheimer's disease. *Nat. Med.* 1996, *2*(8), 864–870.
15. Hardy, J., Selkoe, D. J. The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* 2002, *297*(5580), 353–356.
16. Manning, G., Whyte, D. B., Martinez, R., Hunter, T., Sudarsanam, S. The protein kinase complement of the human genome. *Science* 2002, *298*(5600), 1912–1934.
17. Matsuo, R., Ochiai, W., Nakashima, K., Taga, T. A new expression cloning strategy for isolation of substrate-specific kinases by using phosphorylation site-specific antibody. *J. Immunol. Meth.* 2001, *247*(1–2), 141–151.
18. Lock, P., Abram, C. L., Gibson, T., Courtneidge, S. A. A new method for isolating tyrosine kinase substrates used to identify fish, an SH3 and PX domain-containing protein, and Src substrate. *EMBO J.* 1998, *17*(15), 4346–4357.
19. Fields, S., Song, O. A novel genetic system to detect protein–protein interactions. *Nature* 1989, *340*(6230), 245–246.
20. Bartel, P. L. Fields S., *The Yeast Two-Hybrid System*. Oxford University Press: New York, 1997.
21. Vidal, M., Endoh, H. Prospects for drug screening using the reverse two-hybrid system. *Trends Biotechnol.* 1999, *17*(9), 374–381.
22. Young, K., Lin, S., Sun, L., Lee, E., Modi, M., Hellings, S., Husbands, M., Ozenberger, B., Franco, R. Identification of a calcium channel modulator using a high throughput yeast two-hybrid screen. *Nat. Biotechnol.* 1998, *16*(10), 946–950.
23. Iorns, E., Lord, C. J., Turner, N., Ashworth, A. Utilizing RNA interference to enhance cancer drug discovery. *Nat. Rev. Drug Discov.* 2007, *6*(7), 556–568.
24. Brass, A. L., Dykxhoorn, D. M., Benita, Y., Yan, N., Engelman, A., Xavier, R. J., Lieberman, J., Elledge, S. J. Identification of host proteins required for HIV infection through a functional genomic screen. *Science* 2008.
25. Pan, Q., Bao, L. W., Kleer, C. G., Sabel, M. S., Griffith, K. A., Teknos, T. N., Merajver, S. D. Protein kinase C epsilon is a predictive biomarker of aggressive breast cancer and a validated target for RNA interference anticancer therapy. *Cancer Res.* 2005, *65*(18), 8366–8371.
26. Singer, O., Marr, R. A., Rockenstein, E., Crews, L., Coufal, N. G., Gage, F. H., Verma, I. M., Masliah, E. Targeting BACE1 with siRNAs ameliorates Alzheimer's disease neuropathology in a transgenic model. *Nat. Neurosci.* 2005, *8*(10), 1343–1349.
27. Boehm, J. S., Zhao, J. J., Yao, J., Kim, S. Y., Firestein, R., Dunn, I. F., Sjostrom, S. K., Garraway, L. A., Weremowicz, S., Richardson, A. L., Greulich, H., Stewart, C. J., Mulvey, L. A., Shen, R. R., Ambrogio, L., Hirozane-Kishikawa, T., Hill, D. E., Vidal, M., Meyerson, M., Grenier, J. K., Hinkle, G., Root, D. E., Roberts, T. M., Lander, E. S., Polyak, K., Hahn, W. C. Integrative genomic approaches identify IKBKE as a breast cancer oncogene. *Cell* 2007, *129*(6), 1065–1079.
28. de Fougerolles, A., Vornlocher, H. P., Maraganore, J., Lieberman, J. Interfering with disease: a progress report on siRNA-based therapeutics. *Nat. Rev. Drug Discov.* 2007, *6*(6), 443–453.
29. Napoli, C., Lemieux, C., Jorgensen, R. Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans. *Plant Cell* 1990, *2*(4), 279–289.
30. Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., Mello, C. C. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 1998, *391*(6669), 806–811.
31. Tuschl, T., Zamore, P. D., Lehmann, R., Bartel, D. P., Sharp, P. A. Targeted mRNA degradation by double-stranded RNA *in vitro*. *Genes Dev.* 1999, *13*(24), 3191–3197.
32. Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K., Tuschl, T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* 2001, *411*(6836), 494–498.
33. Paddison, P. J., Caudy, A. A., Bernstein, E., Hannon, G. J., Conklin, D. S. Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev.* 2002, *16*(8), 948–958.
34. Yang, D., Buchholz, F., Huang, Z., Goga, A., Chen, C. Y., Brodsky, F. M., Bishop, J. M. Short RNA duplexes produced by hydrolysis with *Escherichia coli* RNase III mediate effective RNA interference in mammalian cells. *Proc. Natl. Acad. Sci. USA* 2002, *99*(15), 9942–9947.
35. Kittler, R., Surendranath, V., Heninger, A. K., Slabicki, M., Theis, M., Putz, G., Franke, K., Caldarelli, A., Grabner, H., Kozak, K., Wagner, J., Rees, E., Korn, B., Frenzel, C., Sachse, C., Sonnichsen, B., Guo, J., Schelter, J., Burchard, J., Linsley, P. S., Jackson, A. L., Habermann, B., Buchholz, F. Genome-wide resources of endoribonuclease-prepared short interfering RNAs for specific loss-of-function studies. *Nat. Meth.* 2007, *4*(4), 337–344.
36. Garrus, J. E., von Schwedler, U. K., Pornillos, O. W., Morham, S. G., Zavitz, K. H., Wang, H. E., Wettstein, D. A., Stray, K. M., Cote, M., Rich, R. L., Myszka, D. G., Sundquist, W. I. Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* 2001, *107*(1), 55–65.
37. Paddison, P. J., Silva, J. M., Conklin, D. S., Schlabach, M., Li, M., Aruleba, S., Balija, V., O'Shaughnessy, A., Gnoj, L., Scobie, K., Chang, K., Westbrook, T., Cleary, M., Sachidanandam, R., McCombie, W. R., Elledge, S. J., Hannon, G. J. A resource for large-scale RNA-interference-based screens in mammals. *Nature* 2004, *428*(6981), 427–431.
38. Popov, N., Wanzel, M., Madiredjo, M., Zhang, D., Beijersbergen, R., Bernards, R., Moll, R., Elledge, S. J., Eilers, M. The ubiquitin-specific protease USP28 is required for MYC stability. *Nat. Cell Biol.* 2007, *9*(7), 765–774.
39. Ngo, V. N., Davis, R. E., Lamy, L., Yu, X., Zhao, H., Lenz, G., Lam, L. T., Dave, S., Yang, L., Powell, J., Staudt, L. M. A loss-of-function RNA interference screen for molecular targets in cancer. *Nature* 2006, *441*(7089), 106–110.
40. MacKeigan, J. P., Murphy, L. O., Blenis, J. Sensitized RNAi screen of human kinases and phosphatases identifies new regulators of apoptosis and chemoresistance. *Nat. Cell Biol.* 2005, *7*(6), 591–600.
41. Kittler, R., Buchholz, F. Functional genomic analysis of cell division by endoribonuclease-prepared siRNAs. *Cell Cycle* 2005, *4*(4), 564–567.
42. Zhu, C., Zhao, J., Bibikova, M., Leverson, J. D., Bossy-Wetzel, E., Fan, J. B., Abraham, R. T., Jiang, W. Functional analysis of human microtubule-based motor proteins, the kinesins and dyneins, in mitosis/

cytokinesis using RNA interference. *Mol. Biol. Cell* 2005, *16*(7), 3187–3199.
43. Hemann, M. T., Fridman, J. S., Zilfou, J. T., Hernando, E., Paddison, P. J., Cordon-Cardo, C., Hannon, G. J., Lowe, S. W. An epi-allelic series of p53 hypomorphs created by stable RNAi produces distinct tumor phenotypes *in vivo*. *Nat. Genet.* 2003, *33*(3), 396–400.
44. Knight, Z. A., Shokat, K. M. Chemical genetics: where genetics and pharmacology meet. *Cell* 2007, *128*(3), 425–430.
45. Metzger, D., White, J. H., Chambon, P. The human oestrogen receptor functions in yeast. *Nature* 1988, *334*(6177), 31–36.
46. Schena, M., Yamamoto, K. R. Mammalian glucocorticoid receptor derivatives enhance transcription in yeast. *Science* 1988, *241*(4868), 965–967.
47. Purvis, I. J., Chotai, D., Dykes, C. W., Lubahn, D. B., French, F. S., Wilson, E. M., Hobden, A. N. An androgen-inducible expression system for *Saccharomyces cerevisiae*. *Gene* 1991, *106*(1), 35–42.
48. McHale, M. M., Cieslinski, L. B., Eng, W. K., Johnson, R. K., Torphy, T. J., Livi, G. P. Expression of human recombinant cAMP phosphodiesterase isozyme IV reverses growth arrest phenotypes in phosphodiesterase-deficient yeast. *Mol. Pharmacol.* 1991, *39*(2), 109–113.
49. Konopka, J. B., Fields, S. The pheromone signal pathway in *Saccharomyces cerevisiae*. *Antonie Van Leeuwenhoek* 1992, *62*(1-2), 95–108.
50. King, K., Dohlman, H. G., Thorner, J., Caron, M. G., Lefkowitz, R. J. Control of yeast mating signal transduction by a mammalian beta 2-adrenergic receptor and Gs alpha subunit. *Science* 1990, *250*(4977), 121–123.
51. Klein, C., Paul, J. I., Sauve, K., Schmidt, M. M., Arcangeli, L., Ransom, J., Trueheart, J., Manfredi, J. P., Broach, J. R., Murphy, A. J. Identification of surrogate agonists for the human FPRL-1 receptor by autocrine selection in yeast. *Nat. Biotechnol.* 1998, *16*(13), 1334–1337.
52. Potenza, M. N., Lerner, M. R. A rapid quantitative bioassay for evaluating the effects of ligands upon receptors that modulate cAMP levels in a melanophore cell line. *Pigment Cell Res.* 1992, *5*(6), 372–378.
53. Carrithers, M. D., Marotti, L. A., Yoshimura, A., Lerner, M. R. A melanophore-based screening assay for erythropoietin receptors. *J. Biomol. Screen.* 1999, *4*(1), 9–14.
54. Humes, H. D. Insights into ototoxicity. Analogies to nephrotoxicity. *Ann NYAcad. Sci.* 1999, *884*, 15–18.
55. Delpire, E., Lu, J., England, R., Dull, C., Thorne, T. Deafness and imbalance associated with inactivation of the secretory Na-K-2Cl co-transporter. *Nat. Genet.* 1999, *22*(2), 192–195.
56. Oddo, S., Caccamo, A., Shepherd, J. D., Murphy, M. P., Golde, T. E., Kayed, R., Metherate, R., Mattson, M. P., Akbari, Y., LaFerla, F. M. Triple-transgenic model of Alzheimer's disease with plaques and tangles: intracellular Abeta and synaptic dysfunction. *Neuron* 2003, *39*(3), 409–421.
57. Eriksen, J. L., Sagi, S. A., Smith, T. E., Weggen, S., Das, P., McLendon, D. C., Ozols, V. V., Jessing, K. W., Zavitz, K. H., Koo, E. H., Golde, T. E. NSAIDs and enantiomers of flurbiprofen target gamma-secretase and lower Abeta 42 *in vivo*. *J. Clin. Invest.* 2003, *112*(3), 440–449.
58. Kukar, T., Prescott, S., Eriksen, J. L., Holloway, V., Murphy, M. P., Koo, E. H., Golde, T. E., Nicolle, M. M. Chronic administration of R-flurbiprofen attenuates learning impairments in transgenic amyloid precursor protein mice. *BMC Neurosci.* 2007, *8*, 54.
59. Morgan, D., Diamond, D. M., Gottschall, P. E., Ugen, K. E., Dickey, C., Hardy, J., Duff, K., Jantzen, P., DiCarlo, G., Wilcock, D., Connor, K., Hatcher, J., Hope, C., Gordon, M., Arendash, G. W. A beta peptide vaccination prevents memory loss in an animal model of Alzheimer's disease. *Nature* 2000, *408*(6815), 982–985.
60. Schenk, D., Barbour, R., Dunn, W., Gordon, G., Grajeda, H., Guido, T., Hu, K., Huang, J., Johnson-Wood, K., Khan, K., Kholodenko, D., Lee, M., Liao, Z., Lieberburg, I., Motter, R., Mutter, L., Soriano, F., Shopp, G., Vasquez, N., Vandevert, C., Walker, S., Wogulis, M., Yednock, T., Games, D., Seubert, P. Immunization with amyloid-beta attenuates Alzheimer-disease-like pathology in the PDAPP mouse. *Nature* 1999, *400*(6740), 173–177.
61. Nebert, D. W. Proposed role of drug-metabolizing enzymes: regulation of steady state levels of the ligands that effect growth, homeostasis, differentiation, and neuroendocrine functions. *Mol. Endocrinol.* 1991, *5*(9), 1203–1214.
62. Myhr, B. C. Validation studies with Muta Mouse: a transgenic mouse model for detecting mutations *in vivo*. *Environ. Mol. Mutagen.* 1991, *18*(4), 308–315.
63. Shephard, S. E., Sengstag, C., Lutz, W. K., Schlatter, C. Mutations in liver DNA of lacI transgenic mice (Big Blue) following subchronic exposure to 2-acetylaminofluorene. *Mutat. Res.* 1993, *302*(2), 91–96.

Part II

Lead Compound Discovery Strategies

John R. Proudfoot

Section Editor

Chapter 6

Strategies in the Search for New Lead Compounds or Original Working Hypotheses

Camille G. Wermuth

I. INTRODUCTION
A. Hits and leads
B. The main hit or lead finding strategies
II. FIRST STRATEGY: ANALOG DESIGN
A. Typical examples
B. The different categories of analogs
C. Pros and cons of analog design
III. SECOND STRATEGY: SYSTEMATIC SCREENING
A. Extensive screening
B. Random screening
C. High-throughput screening
D. Screening of synthesis intermediates
E. New leads from old drugs: the SOSA approach
IV. THIRD STRATEGY: EXPLOITATION OF BIOLOGICAL INFORMATION
A. Exploitation of observations made in humans
B. Exploitation of observations made in animals
C. Exploitation of observations made in the plant kingdom and in microbiology
V. FOURTH STRATEGY: PLANNED RESEARCH AND RATIONAL APPROACHES
A. L-DOPA and Parkinsonism
B. Inhibitors of the ACE
C. Discovery of the H_2-receptor antagonists
VI. CONCLUSION
REFERENCES

So ist denn in der Strategie alles sehr einfach, aber darum nicht auch alles sehr leicht.
(Thus in the strategy everything is very simple, but not necessarily very easy)

Carl von Clausewitz[1]

I. INTRODUCTION

This chapter deals with the various strategies leading to active compounds and active compounds collections. The objective is to identify original starting points for therapeutic discovery programs. Such programs typically begin with the search for "hits."

A. Hits and leads

A hit is an active substance having a preferential activity for the target and which satisfies all of the following criteria[2]: (1) reproducible activity in a relevant bioassay, (2) confirmed structure and high purity, (3) specificity for the target under study, (4) confirmed potential for novelty and (5) chemically tractable structure, that is, molecules presenting a certain affinity for a target.

Identifying hits for a new target usually involves screening of a wide range of structurally diverse small molecules in an *in vitro* bioassay. Alternatively, small molecules can be screened for their potential to modulate a biological process thought to be critical in disease or in which the target is thought to play a major role. Thanks to miniaturization and robotics, the number of compounds that can be

Copyright © 2008, Elsevier Ltd
All rights reserved.

screened has greatly increased and several thousand compounds can be screened in 1 day.

Once a hit is discovered, its activity must be confirmed and validated. Typical hit validation criteria are as follows: (1) the hit must be active *in vivo*, (2) the hit must not display human ether-a-go-go-related (hERG) toxicity, (3) the analogs of the hit must display clear structure–activity relationships (SAR), (4) the hit must not contain chemically reactive functions and (5) the hit must provide patent opportunities. Only then it becomes a lead substance, commonly named "lead."

If a lead molecule emerges from these additional studies on SAR, absorption, distribution, metabolism, and excretion (ADME) and toxicity, it acquires the "clinical drug candidate" status. After a short toxicological study it fulfills the criteria required for administration to humans for initial clinical studies.

B. The main hit or lead finding strategies

A retrospective analysis of the ways leading to discovery of new drugs suggests that four successful strategies can yield new hits and/or lead compounds:[3,4] The first strategy is based on the modification and improvement of already existing active molecules. The second one consists of the systematic screening of sets of arbitrarily chosen compounds on selected biological assays. The third approach resides in the retroactive exploitation of various pieces of biological information which result sometimes from new discoveries made in biology and medicine, and sometimes are just the fruits of more or less fortuitous observations. The fourth route to new active compounds is a rational design based on the knowledge of the molecular cause of the pathological dysfunction.

II. FIRST STRATEGY: ANALOG DESIGN

The most popular strategy in drug design is the synthesis of analogs of existing active molecules. The objective is to start with known active principles and, by various chemical transformations, prepare new molecules (sometimes referred to as "me-too compounds") for which an increase in potency, a better specific activity profile, improved safety or a formulation that is easier-to-handle by physicians and nurses or more acceptable to the patient are claimed.

A. Typical examples

A typical illustration of this approach is found in the series of losartan analogs (Figure 6.1) or in the conazole series

FIGURE 6.1 Angiotensin AT1 receptor antagonists derived from losartan. Despite their structural similarity of the structures, it can be assumed that the corresponding discoveries were made independently. The first year under parentheses is the basic patent year, the second one is the year of the first launch.

FIGURE 6.2 An example of me-too compounds (full analogs) is given by micoconazole-derived fungistatics which act by inhibition of the ergosterol biosynthesis. The first year under parentheses is the basic patent year, the second one is the year of the first launch.

(Figure 6.2). All the compounds show similar structures and similar affinity for the angiotensin II receptor. As such they can be considered as "full" analogs.

In the pharmaceutical industry, motivations for analog design are often driven by competitive and economic factors. Indeed, if the sales of a given medicine are high and the company is found in a monopolistic situation, protected by patents and trade marks, other companies will want to produce similar medicines, if possible with some therapeutic improvements. They will therefore use the already commercialized drug as a lead compound and search for ways to modify its structure and some of its physical and chemical properties while retaining or improving its therapeutic properties.

B. The different categories of analogs

The term analogy, derived from the latin and greek *analogia*, is used in natural sciences since 1791 to describe structural *and* functional similarity.[5] Extended to drugs, this definition implies that the analog of an existing drug molecule shares chemical and therapeutical similarities with the original compound. Formally, this definition allows anticipating three categories of drug analogs: (1) analogs presenting chemical and pharmacological similarity, (2) analogs presenting only chemical similarity and (3) analogs displaying similar pharmacological properties but presenting totally different chemical structures.

Analogs of the first category, presenting at the same time chemical and pharmacological similarities, can be considered as "full" or "true" analogs (Figures 6.1 and 6.2). These analogs correspond to the class of drugs often referred to as "me-too compounds." Usually, they are improved versions of a pioneer drug over which they present a pharmacological, pharmacodynamic or biopharmaceutical advantage. Other examples are the angiotensin-converting enzyme (ACE) inhibitors derived from captopril, the histamine H_2 antagonists derived from cimetidine, and the hydroxymethylglutaryl-CoA reductase (HMG-CoA reductase) inhibitors derived from mevinolin, etc. Such analogs are designed for industrial and marketing reasons with the same justifications

FIGURE 6.3 Some examples of structural analogs. Despite their structural analogy these compounds present different pharmacological activities.

FIGURE 6.4 Zopiclone and zolpidem are selective benzodiazepine receptor agonists not related chemically to benzodiazepines.[7,8]

as those which are valid for any other industrial products such as laptop computers or automobiles.

The second class of analogs made of chemically resembling molecules and for which we propose the term "structural analogs," contains compounds originally prepared as close and patentable analogs of a novel lead, but for which the biological assays revealed totally unexpected pharmacological properties. A historical, example of the emergence of a new activity is provided by the discovery of the antidepressant properties of imipramine which was originally designed as an analog of the potent neuroleptic drug chlorpromazine. Observation of an "emergent" activity can be purely fortuitous or can result from a voluntary and systematic investigation.

Another example, illustrating that chemical similarity does not necessarily mean biological similarity, is found for steroid hormones: testosterone and progesterone, although being chemically very close, have totally different biological functions (Figure 6.3). Similarly, minaprine is a dopaminergic drug whereas its cyano analog SR 95191 is a potent MAO-A inhibitor.[6]

For the third class of analogous compounds chemical similarity is not observed, however, they share common biological properties. We propose the term "functional analogs" for such compounds. Examples are the neuroleptics chlorpromazine and haloperidol or the tranquillizers diazepam and zopiclone (Figure 6.4). Despite totally different chemical structures, they show similar affinities for the dopamine and the benzodiazepine receptors, respectively. The design of such drugs is presently facilitated, thanks to virtual screening of large libraries of diverse structures.

C. Pros and cons of analog design

Analog design lacks originality and has often been a source of criticism of the pharmaceutical industry.[9] Each laboratory wants to have its own antiulcer drug, its own antihypertensive, etc. These drug copies are called "me-too products." Generally, the owner firm of the original drug continues to prepare new analogs, to insure both a maximum perimeter of protection of its patents and to remain the leader in a given area. For these reasons, the chemical transformation of known active molecules constitutes the most widespread practice in the pharmaceutical research.

A reassuring aspect obtained by making a therapeutic copy resides in the certainty to end with an active drug in the desired therapeutic area. It is indeed extremely rare, and practically improbable, that a given biological activity is unique to a single molecule. Molecular modifications allow the preparation of additional products for which one can expect, if the investigation has been sufficiently prolonged, a comparable activity to that of the copied model, perhaps even a better one. This factor is comforting for the copier as well as for the financiers that subsidize him. It is necessary, however, to keep in mind that the original inventor of a new medicine possesses a technological and scientific advantage over the copier and that, he too, has been able to design a certain number of copies of his own compound before he selected the molecule insuring the best compromises between activity, secondary effects, toxicity and invested money.

A second element favoring the copy derives from the information already gained which then facilitates subsequent pharmacological and clinical studies. As soon as the pharmacological models that served to identify the activity profile of a new prototype are known, it suffices to apply them to the therapeutic copies. In other terms, the pharmacologist will know in advance to what kind of activity he desires and which tests he will have to apply to select the wanted activity

FIGURE 6.5 The striking analogy between the vasodilator drug flosequinan and the quinolone antibiotic norfloxacin.

profile. In addition, during clinical studies, the original research, undertaken with the lead compound, will serve as a reference and can be transposed as unchanged to the evaluation of the copy. Criticism of this approach is a result of the obvious fact that, in selecting a new active molecule by means of the same pharmacological models as were used for the original compound, one will inevitably end with a compound presenting an identical activity profile and thus the innovative character of such a research is practically nil.

Finally, financial arguments may play in favor of the therapeutic copy. Thus, it may be important, and even vital, for a pharmaceutical company or for a national industry, to have its own drugs rather than to subcontract a license. Indeed, in paying dues of license, an industry deprives its own researches. Moreover, the financial profitability of a research based on me-too drugs can appear to be higher, because no investment in fundamental research is required. The counterpart is that the placement on the market of the copy will naturally occur later than that of the original drug and thus it will make it more difficult to achieve a high sales ranking, all the more so because the me-too drug will be in competition with other copies targeting a similar market.

In reality the situation is more subtle because very often the synthesis of me-too drugs is justified by a desire to improve the existing drug. Thus, for penicillins, the chemical structure that surrounds the β-lactamic cycle is still being modified. Current antibiotics that have been derived from this research (the cephalosporins for example) are more selective, more active on resistant strains and can be administered by the oral route. They are as different from the parent molecule as a recent car compared to a 40-year-old model. In other terms, innovation can result from the sum of a great number of stepwise improvements as well as from a major breakthrough.

It can also happen that during the pharmacological or clinical studies of a me-too compound a totally new property, not present in the original molecule, appears unexpectedly. Thanks to the emergence of such a new activity, the therapeutic copy becomes in turn a new lead structure. That was the case for imipramine, initially synthesized as an analog of chlorpromazine and presented to the clinical investigators for the study of its antipsychotic profile.[10] During its clinical evaluation this substance demonstrated much more activity against depressive states than against psychoses. Imipramine has truly opened, since 1954, a therapeutic avenue for the pharmacological treatment of depression. On its way to becoming Viagra, the compound UK-92,480, prepared in 1989 by the Pfizer scientists in Sandwich, England went first from a drug for hypertension to a drug for angina. Then it changed again when a 10-day toleration study in Wales turned up its unusual side effect: penile erections.[11] It seems probable that a similar emergence of a new activity occurred with flosequinan which is a sulfoxide bioisostere of the quinolone antibiotics (Figure 6.5). This compound turned out to be a vasodilator and cardiotonic drug having totally lost any antibiotic activity.[12]

III. SECOND STRATEGY: SYSTEMATIC SCREENING

This method consists in screening new molecules, whether they are synthetic or of natural origin, on an animal model or on any biological test without having in mind, the hypotheses on its pharmacological or therapeutic potential. It rests on the systematic use of selective batteries of experimental models destined to mimic closely the pathological events. The trend is to undertake *in vitro* rather than *in vivo* tests: binding assays, enzyme inhibition measurements, activity on isolated organs or cell cultures, etc. In practice, systematic screening can be achieved in two different manners. The first one is to apply to a small number of chemically sophisticated and original molecules, a very exhaustive pharmacological investigation: it is called "extensive screening." The second one, on the contrary, strives to find, among a great number of molecules (several hundreds or thousands), one that could be active in a given indication: this is "random screening."

A. Extensive screening

Extensive screening is generally applied to totally new chemical entities coming from an original effort of chemical research or from a laborious extraction from a natural source. For such molecules the high investment in synthetic or extractive chemistry justifies an extensive pharmacological study (central nervous, cardiovascular, pulmonary, and

FIGURE 6.6 Drugs discovered by random screening.

digestive systems, antiviral, antibacterial or chemotherapeutic properties, etc.) to detect if there exists an interesting potential linked to these new structures. In summary, a limited number of molecules is studied in a thorough manner (vertical screening). It is by such an approach that the antihistaminic, and later on the neuroleptic properties of the amines derived from phenothiazine were identified. Initially, these compounds had been submitted, with negative results, to a limited screening study only directed toward possible chemotherapeutic, antimalarial, trypanocidal and anthelmintic activities.

Original chemical research is also at the origin of the discovery of the benzodiazepines by Sternbach.[13] By the way, this author specifies that the class of compounds he was seeking had to fulfill the following criteria: (1) the chemical series had to be relatively unexplored, (2) it had to be easily accessible, (3) it had to allow a great number of variations and transformations, (4) it had to offer some challenging chemical problems and (5) it had to "look" as if it could lead to biologically active products.

The extensive screening approach has often led to original molecules, it is however highly dependent on the skill and the intuition of the medicinal chemist and even more, on the talents of the pharmacologist who has to be able to adapt and to orient his tests as soon as his findings evolve to reveal the real therapeutic potential of the molecule under study.

More recent examples are seen by the discovery, thanks to systematic screening programs, of the cyclopyrrolones, for example, zopiclone (Figure 6.6), as ligands for the central benzodiazepine receptor,[7,14] or of taxol as an original and potent anticancer drug (for a review see Suffness[15]).

B. Random screening

In this case, the therapeutic objective is fixed in advance and, contrary to the preceding case, a great number (several thousands) of molecules is tested, but on a limited number of experimental models only. By this method one practices the so-called random screening. This method has been used for the discovery of new antibiotics. By submitting samples of earth collected in countries from all over the world to a selective antibacterial and antifungal screening, the rich arsenal of anti-infectious drugs that are presently at the disposal of the clinicians, was developed. During the World War II, an avarian model in chickens infected with *Plasmodium gallinaceum*, was used for the massive screening of thousands of potential antimalarials. The objective was then to solve, by finding a synthetic antimalarial, the problem of the shortage of quinine. Unfortunately, no satisfying drug was found. Massive screening was implemented in Europe and the United States to discover new anticancer[16] and antiepileptic drugs. Here again the problem is to select some predictive, but cheap cellular or animal models. A common criticism of these methods is that they constitute, by the absence of a rational lead, a sort of fishing. Besides, the results are very variable: nil for the discovery of new antimalarials, rather weak for the anticancer drugs but excellent, in their time, for the discovery of antibiotics.

Among more recent successes of this approach one can mention the discovery of lovastatin, also called mevinolin (Figure 6.7),[17,18] which was the basis of a new generation of hypocholesterolemic agents, acting by inhibition of HMG-CoA reductase.

Sometimes unexpected findings result from systematic screening applied in an unprejudiced manner. An example is found in the tetracyclic compound BMS-192548 extracted from Aspergillus niger WB2346 (Figure 6.8).

For any medicinal chemist or pharmacologist the similarity of this compound with the antibiotic tetracycline is striking. However, none of them would *a priori* forecast that BMS-192548 exhibits central nervous system (CNS) activities. Actually, the compound turns out to be a ligand for the neuropeptide Y receptor preparations.[19]

C. High-throughput screening

Since the 1980s, with the arrival of robotics and with the miniaturization of the *in vitro* testing methods, it became possible to combine the two preceding approaches. In other words, screen thousands of compounds on a large number of biological targets. This high-throughput screening is usually applied to the displacement of radioligands and to the

FIGURE 6.7 The natural compounds compactin (mevastatin) and lovastatin block the cholesterol biosynthesis in inhibiting the enzyme HMG-CoA reductase. The later developed compounds simvastatin and pravastatin are semi-synthetic analogs. The open-ring derivative pravastatin is less lipophilic and therefore presents less central side effects. For all these compounds the ring-opened form is the actual active form *in vivo*.

FIGURE 6.8 Unexpected CNS activity of the tetracycline analog BMS-192548.

inhibition of enzymes. The present trend is to replace radioligand-based assays by fluorescence-based measurements. As it now is possible for a pharmaceutical company to screen several thousand molecules simultaneously on 30–50 different biochemical tests, the problem becomes to feed the robots with interesting molecules. Primary sources are chemicals coming from in-house libraries or from commercial collections, but the samples can also be crudely purified vegetal extracts or fermentation fluids. In this latter case, one proceeds to the isolation and to the identification of the responsible active principle[20,21] only when an interesting activity, coming up after the screening, is observed. High-throughput screening will be treated in Chapter 7.

D. Screening of synthesis intermediates

As synthesis intermediates are chemically connected to final products and as they often present some common groupings with them, it is not excluded that they share equally some pharmacological properties. For this reason, it is always prudent to submit also these compounds to a pharmacological evaluation. Among drugs discovered this way one finds the tuberculostatic semicarbazones: they were initially used in the synthesis of antibacterial sulfathiazoles. Subsequent testing of isonicotinic acid hydrazide, destined for the synthesis of a particular thiosemicarbazone,

FIGURE 6.9 Departing from methotrexate, simple intermediates led to new drugs. Mercaptopurine and azathioprine are immunosuppressants and allopurinol is used in the treatment of gout.

revealed the powerful tuberculostatic activity of the precursor that has become since then a major antitubercular drug (isoniazide).

Inhibitors of the enzyme *dihydrofolate-reductase* such as methotrexate (Figure 6.9) are used in the treatment of leukemia's. During the search for methotrexate analogs a

FIGURE 6.10 A successful SOSA approach allowed the identification of the antibacterial sulfonamide sulfathiazole as a ligand of the endothelin ET_A receptor and its optimization to the selective and potent compounds BMS-182874, BMS-193884 and BMS-207940.[29,30]

very simple intermediate, mercaptopurine, was also submitted to testing. It turned to be active but relatively toxic. Subsequent optimization led to azathioprine, a prodrug releasing mercaptopurine *in vivo*.

Azathioprine was found to be more potent as immunosuppressive agent than the previously used corticoids and was systematically used in all organ transplantations until the advent of cyclosporine. Another intermediate in this series, *allopurinol*, inhibits xanthine-oxidase and therefore is used in the treatment of gout.[22]

E. New leads from old drugs: the SOSA approach

The SOSA approach (SOSA = **S**elective **O**ptimization of **S**ide **A**ctivities) represents an original alternative to high-throughput screening (HTS).[3,23–27] It consists of two steps:

1. Screening on newly identified pharmacological targets of a limited set (approximately 1,000 compounds) of well-known drug molecules for which bioavailability and toxicity studies have already been performed and which have proven usefulness in human therapy. By definition, in using such a library, all hits that are found are drug-like!
2. Optimize hits (by means of traditional, parallel or combinatorial chemistry) in order to increase the affinity for the new target and decrease the affinity for the other targets. The objective is to prepare analogs of the hit molecule in order to transform the observed "side activity" into the main effect and to strongly reduce or abolish the initial pharmacological activity.

The rationale behind the SOSA approach lies in the fact that, in addition to their main activity, almost all drugs used in human therapy show one or several side effects. In other words, if they are able to exert a strong interaction with the main target, they exert also less strong interaction with some other biological targets. Most of these targets are unrelated to the primary therapeutic activity of the compound. The objective of the medicinal chemists is then to proceed to a reversal of the affinities, the identified side effect becoming the main effect and vice versa. Many cases of activity profile reversals by means of the SOSA approach have been published.

A typical illustration of the SOSA approach is given by the development of *selective ligands for the endotheline ET_A receptors* by scientists from Bristol-Myers-Squibb.[28,29] Starting from an in-house library, the antibacterial compound sulfathiazole (Figure 6.10) was an initial, but weak, hit (IC_{50} = 69 μM). Testing of related sulfonamides identified the more potent sulfisoxazole (IC_{50} = 0.78 μM). Systematic variations led finally to the potent and selective ligand BMS-182874. *In vivo*, this compound was orally active and produces a long-lasting hypotensive effect.

Further optimization guided by pharmacokinetic considerations led the BMS scientists to replace the naphtalene ring by a diphenyl system.[29] Among the

prepared compounds, BMS-193884 (ET_AK_i = 1.4 nM; ET_BK_i = 18,700 nM) showed promising hemodynamic effects in a phase II clinical trial for congestive heart failure. More recent studies led to the extremely potent antagonist BMS-207940 (edonentan; ET_AK_i = 10 pM) presenting an 80,000-fold selectivity for ET_A versus ET_B. The bioavailability of this compound is 100% in rats and it exhibits oral activity already at a 3 μM/kg dosing.[29]

Another example is the antidepressant minaprine (Figure 6.11). In addition to reinforcing serotoninergic and dopaminergic transmission, this amino-pyridazine possesses weak affinity for muscarinic M_1 receptors (K_i = 17 μM). Simple chemical variations allowed to abolish the dopaminergic and serotoninergic activities and to boost the cholinergic activity up to nanomolar concentrations.[31–33]

Similarly, chemical variations of the D_2/D_3 non-selective benzamide sulpiride (Figure 6.12) led to compound Do 897, a selective and potent D_3 receptor partial agonist.[34]

As mentioned above, a differentiating peculiarity of this type of library is that it is constituted by compounds that have already been safely given to humans. Thus, if a compound were to "hit" with sufficient potency on an orphan target, there is a high chance that it could rapidly be tested in patients for Proof of Principle. Alternatively, if one or more compounds hit, but with insufficient potency, optimized analogs can be synthesized and the chances that these analogs will be good candidate drugs for further development are much higher than if the initial lead is toxic or not bioavailable. One of these new-type of chemical libraries, the Prestwick Chemical Library, is available.[35] It contains 1,120 biologically active compounds with high chemical and pharmacological diversity as well as known bioavailability and safety in humans. Over 90% of the compounds are well-established drugs, and 10% are bioactive alkaloids. For scientists interested in drug-likeness such a library fulfills certainly in the most convincing way the quest for "drug-like" leads!

FIGURE 6.11 Progressive passage from minaprine to a potent and selective partial muscarinic M1 agonist.[31,32,33]

FIGURE 6.12 The progressive change from the D_2/D_3 receptor non-selective antagonists to the highly D_3-selective compound Do 897.[34] The numbers between parentheses indicate the D_2/D_3 affinity ratio.

IV. THIRD STRATEGY: EXPLOITATION OF BIOLOGICAL INFORMATION

A major contribution to the discovery of new active principles comes from the exploitation of biological information. By this meant information which relates to a given biological effect (fortuitous or voluntary) provoked by some substances in humans, in animals or even in plants or bacteria. When such information becomes accessible to the medicinal chemist, it can serve to initiate a specific line of therapeutic research. Originally, the observed biological effect can simply be noticed without any rational knowledge on how it works.

A. Exploitation of observations made in humans

The activity of exogenous chemical substances on the human organism can be observed under various circumstances: ethnopharmacology, popular medicines, clinical observation of secondary effects or adverse events, fortuitous observation of activities of industrial chemical products, etc. As in all cases, the harvested information is observed directly in man, this approach presents a notable advantage.

1. *Study of indigenous medicines (ethnopharmacology)*

Natural substances were for a long time the unique source of medicines. At present, they constitute 30% of the used active principles and probably more (approximately 50%) if one considers the number of prescriptions that utilize them, particularly since use of antibiotics plays a major role.[36] Behind most of these substances one finds indigenous medicines. As a consequence, *ethnopharmacology* represents a useful source of lead compounds. Historically, we are indebted to this approach for the identification of the cardiotonic digitalis glucosides of the digital, the opiates and the cinchona alkaloids. Curare was obtained from a South American plant used for a long time by natives to make arrow poisons. The cardiotonic glucosides of the *Strophantus* seeds, and the alkaloid eserine from the Calabar beans are other examples of active drugs originally used by natives as poisons. The *Rauwolfia serpentina* has been used since centuries in India before western medicine became interested in its tranquillizing properties and extracted reserpine from it. Atropine, pilocarpine, nicotine, ephedrine, cocaine, theophylline and innumerable other medicines have thus been extracted from plants to which the popular medicine attributed therapeutic virtues.

Despite its extremely useful contributions to the modern pharmacopoeia such as artemisin and huperzine, folk medicine is a rather unreliable guide in the search for new medicines. This is illustrated by the example of antifertility agents: according to natives of some islands of the Pacific, approximately 200 plants would be efficient in reducing male or female fertility. Extracts have been prepared from 80 of these plants and have been administered at high dosings to rats during periods of 4 weeks and more, without observing the slightest effect upon pregnancies or litter sizes.[37] When ethnopharmacology and the natural substance chemistry end in the discovery of a new active substance, this latter is first reproduced by total synthesis. It is then the object of systematic modifications and simplifications that aim to recognize by trial and error the minimal requirements that are responsible for the biological activity.

2. *Clinical observation of side effects of medicines*

The clinical observations of entirely unexpected side effects constitute a quasi-inexhaustible source of tracks for the search of lead compound. Indeed, beside the wanted therapeutic action, most drugs possess side effects. These are accepted either from the beginning as a necessary evil, or recognized only after some years of utilization. When side effects present a medical interest by themselves, a planned objective can be the dissociation of the primary from the side effect activities: enhance the activity originally considered as secondary and diminish or cancel the activity that initially was dominant. Promethazine for example, an antihistaminic derivative of phenothiazine, is burdened with important sedative effects. The merit of a clinician such as Laborit[38] has been to promote the utilization of this side effect and to direct research toward better profiled analogs. This impulse was the origin of the birth of chlorpromazine, the prototype of a new therapeutic series, the neuroleptics, whose existence was unsuspected until then and that has revolutionized the practice of psychiatry.[10,39] Innumerable other examples can be found in the literature, such as the hypoglycemic effect of some antibacterial sulfamides, the uricosuric effect of the coronaro-dilating drug benziodarone, the antidepressant effect of isoniazide, an antitubercular drug, and the hypotensive effect of β-blocking agents, etc.

This last example is beautifully illustrated by the discovery of the potassium channel activator cromakalim.[40] Cromakalim is the first antihypertensive agent to be shown to act exclusively through potassium channel activation.[41] This novel mechanism of action involves an increase in the outward movement of potassium ions through channels in the membranes of vascular smooth muscle cells, leading to relaxation of the smooth muscle. The discovery of this compound can be summarized as follows: β-adrenergic receptor blocking drugs were not thought to have antihypertensive effects when they were first investigated. However, pronethalol, a drug that was never marketed, was found to reduce arterial blood pressure in hypertensive patients with *angina pectoris*. This antihypertensive effect was subsequently demonstrated for propranolol and all other β-adrenergic antagonists.[42] Later on, there were some doubts that blockade of the β-adrenergic receptors was responsible for the

hypotensive activity and attempts were made to dissociate, in the classical β-blocking molecules, the β-blockade from the antihypertensive activity. Among the various conceivable molecular variations which are possible for the flexible β-blockers, it was found that conformational restriction obtained in cyclizing the carbon atom bearing the terminal amino group onto the aromatic ring yielded derivatives devoid of β-blocking activity, but retaining the antihypertensive activity (Figure 6.13).

One of the first compounds prepared (compound **1**, Figure 6.13) was indeed found to lower blood pressure in hypertensive rats by a direct peripheral vasodilator mechanism; no β-blocking activity was observed. Optimization of the activity led to the 6-cyano-4-pyrrolidinylbenzopyran (compound **2**), which was more than a 100-fold potent than the nitro derivative. The replacement of the pyrrolidine by a pyrrolidinone (which is the active metabolite) produced a 3-fold increase in activity and the optical resolution led to the (−)-3*R*, 4*S* enantiomer of cromakalim (BRL 38227) that concentrates almost exclusively the hypotensive activity.[40,43,44]

3. *New uses for old drugs*

In some cases, a new clinical activity observed for an old drug is sufficiently potent and interesting to justify the immediate use of the drug in the new indication, this is illustrated hereafter.

Amiodarone, for example (Figure 6.14), was introduced as a coronary dilator for angina. Concern about corneal deposits, discoloration of skin exposed to sunlight and thyroid disorders led to the withdrawal of the drug in 1967. However, in 1974 it was discovered that amiodarone was highly effective in the treatment of a rare type of arrhythmia known as the Wolff-Parkinson-White syndrome. Accordingly, amiodarone was reintroduced specifically for that purpose.[45]

Benziodarone, initially used in Europe as a coronary dilator, proved later on to be a useful uricosuric agent. Presently, it is withdrawn from the market due to several cases of jaundice associated with its use. The corresponding brominated analog, benzbromarone was specifically marketed for its uricosuric properties.

Thalidomide, was initially launched as a sedative/hypnotic drug (Figure 6.15), but withdrawn because of its extreme teratogenicity.

Under restricted conditions (no administration during pregnancy or to any woman of childbearing age), it found a new use as immunomodulator. Particularly it seems efficacious for the treatment of *erythema nodosum leprosum*, a possible complication of the chemotherapy of leprosy.[46]

FIGURE 6.13 The clinical observation of the hypotensive activity of the "open" (and therefore flexible) β-blocking agents was the initial lead to cyclized analogs devoid of β-blocking activity, but retaining the antihypertensive activity (Stemp and Evans[40]).

FIGURE 6.14 Structures of the arones.

In 1978, the synthesis of the *indenoisoquinoline* NSC 314622 (Figure 6.16) was reported as the result of an unexpected transformation during a synthesis of nitidine chloride. Given its weak antitumor activity, it was not investigated further.

Twenty years later, NSC 314622 resurfaced as a potential topoisomerase I (top I) inhibitor and served as lead structure for the design of cytotoxic non-camptothecin top I inhibitors such as the compound "**19a**."[47]

In 2001, the antimalarial drug *quinacrine* and the antipsychotic drug *chlorpromazine* (Figure 6.17) were shown to inhibit prion infection in cells. Prusiner *et al.*[48] identified the drugs independently and found that they inhibit conversion of normal prion protein into infectious prions and clear prions from infected cells.

Both drugs can cross over from the bloodstream to the brain, where the prion diseases are localized.

FIGURE 6.15 Structure of thalidomide. The marketed compound is the racemate.

A more recent example is provided by the discovery of the use of sildenafil (Viagra®, Figure 6.18), a phosphodiesterase type 5 (PDE5) inhibitor, as an efficacious, orally active agent for the treatment of male erectile dysfunction.[49,50] Initially, this compound was brought to the clinic as an hypotensive and cardiotonic substance and its usefulness in male erectile dysfunction resulted clearly from the clinical observations.

In many therapeutic families each generation of compounds induces the birth of the following one. This happened in the past for the sulfamides, penicillins, steroids, prostaglandins, and tricyclic psychotropics families, and one can draw real genealogical trees representing the progeny of the discoveries. More recent examples are found in the domain of ACE inhibitors and in the family of histaminergic H_2 antagonists.

Research programs based on the exploitation of side effects are of great interest in the discovery of new tracks as far as they depend on information about activities *observed directly in man* and not in animals. On the other hand, they allow to detect new therapeutic activities *even when no pharmacological models in animals do exist*.

4. *The fortuitous discovery of activities of industrial chemical products*

During the industrial manufacture of nitroglycerin toxic manifestations due to this compound, particularly strong

FIGURE 6.16 The indenoisoquinoline NSC 314622 resurfaced 20 years after its first testing as a top 1 inhibitor.[47]

FIGURE 6.17 Old drugs, new use. The antimalarial drug quinacrine and the antipsychotic drug chlorpromazine are able to inhibit prion infection.[48]

FIGURE 6.18 Structure of the PDE5 inhibitor sildenafil.[49,50]

vasodilating properties, were observed in workers. There from came the utilization of this substance, and later on of other nitric esters of aliphatic alcohols, in *angina pectoris* and as cerebral vasodilators. In an analogous manner it was observed during the manufacture of the sulfa drug sulfathiazole that 2-amino-thiazole, one of the starting materials, was endowed with antithyroidal properties. This observation fostered the use of this compound, and of amino-thiazoles in general, for the treatment of thyroid gland hyperactivity. Tetraethylthiurame disulfide was originally used as antioxidant in the rubber industry. After having manipulated it, workers felt an intolerance to alcohol. Therefore, this product was proposed for ethylic alcohol withdrawal cures (disulfiram). On the molecular level, the mode of action of disulfiram rests on the inhibition of the enzyme aldehyde-dehydrogenase that normally insures the oxidation of acetaldehyde into acetic acid. The intake of alcohol under disulfiram provokes an accumulation of acetaldehyde that achieves a real intoxication of the patient. Another example of a fortuitous discovery is given by the example of probucol. This antihyperlipoproteinemic compound was originally synthesized as an antioxidant for plastics and rubber.[51,52]

B. Exploitation of observations made in animals

We find here all the research done by physiologists that has been the basis of the discovery of vitamins, hormones and neurotransmitters and the fall-outs of various pharmacological studies, when they were performed *in vivo*. Other observations made on animals, often in a more or less fortuitous manner, have led to useful discoveries. An example is provided by the dicoumarol-derived anticoagulants.

The discovery of the anticancer properties of the alkaloids of *Vinca rosea* constitutes a particularly beautiful example of pharmacological feed-back. Preparations from this plant had the reputation in some popular medicines to possess antidiabetic virtues. During a controlled pharmacological test, these extracts were proven to be devoid of hypoglycemic activity. On the other hand, it was frequently observed that the treated rats died from acute septicemia. A study of this phenomenon showed that it was due to massive leukopenia. In taking the leukocytes count as the activity end-point criterion, it became possible to isolate the main alkaloid, vinblastine.[53] At the same time, in an other laboratory, routine anticancer screening had revealed the activity of the crude extract on the murine leukemia.[54] Subsequently, the antileukemic activity became a screening tool. Out of 30 alkaloids isolated from various periwinkles, four (vinblastine, vinleurosine, vincristine and vinrosidine) were found active in human leukemias.[55]

Analogs of L-arginine with modifications at the terminal guanidino nitrogen and/or the carboxyl terminus of the molecule have been widely used for their ability to inhibit the production of nitric oxide (NO) and are thought to be competitive antagonists of nitric synthase. In studies designed to elucidate the role of NO in the gastrointestinal tract, an inhibitory effect of N^G-nitro-L-arginine methyl ester (L-NAME) on cholinergic neural responses was sometimes observed. This inhibitory effect was shown to be consistent with a blockade of the muscarinic receptors.[56]

Remember too that it was the research of insecticides that led to the discovery of the organophosphorus acetylcholinesterase inhibitors by Schrader at the Bayer laboratory.[57] The study of their mechanism of action has shown that they act by acylation of a serine hydroxyl in the catalytic site of the enzyme. This was one of the first examples describing a molecular mechanism for an enzymatic inhibition.

Replacement by Janssen *et al.*[58] of the *N*-methyl group of pethidine by various propio- and butyrophenones led to potent analgesics such as R951 and R1187 (Figure 6.19). During their pharmacological study it was noted that mice which had been injected with these drugs became progressively calm and sedated. The resemblance of the sedation with that produced by chlorpromazine encouraged Janssen to synthesize analogs of R1187 in the hope that one might be devoid of analgesic activity whilst retaining tranquilizing activity.

From this effort, haloperidol emerged in 1958 as the most potent tranquillizer yet to have been discovered. It is 50–100 times as potent as chlorpromazine, with fewer side effects.[58,59]

C. Exploitation of observations made in the plant kingdom and in microbiology

Among the numerous discoveries that we owe to the botanists and the pharmacognosts, the precocious interest for tryptophan metabolites has to be evoked, especially the interest for indolylacetic acid.[60] This compound acts as growth hormone in plants. *Para*-chlorinated phenoxyacetic

FIGURE 6.19 The passage from pethidine-related opiate analgesics to the dopaminergic antagonist haloperidol.

acids (MCPA or methoxone; 2,4-D or chloroxone) are mimics of indolylacetic acid (bioisostery) and show similar phytohormonal properties: at high doses they serve as weeders. Ring-chlorinated phenoxyacetic acids have been later on introduced in molecules, as varied as meclofenoxate (cerebral metabolism), clofibrate (lipid metabolism) and ethacrynic acid (diuretic).

The 5-hydroxylated analog of indolacetic acid is the principal urinary metabolite of serotonin. On the basis of two biochemical observations, the possible role of serotonin in inflammatory processes and the increase of urinary metabolites of tryptophan in rheumatic patients, Shen, from the Merck Laboratories, designed anti-inflammatory compounds derived from indolacetic acid. Among them he found indomethacin in 1963, one of the most powerful non-steroidal anti-inflammatory drugs currently known.[61]

A particularly rich contribution of this approach in the therapeutic area has been the discovery and the development of penicillin (see Chapter 1). It initiated the discovery of many other major antibiotics such as chloramphenicol, streptomycin, tetracyclines, rifampicine, etc.

In conclusion, whatever its origin may have been, the use of biological information constitutes a preferential source for original molecule research. It presents the advantage to offer creative approaches, not resting on the exploitation of routine pharmacological models. Once the lead molecule is identified, it will immediately be the object of thorough studies to elucidate its molecular mechanism of action. Simultaneously, one will proceed to the synthesis of structural analogs, as well as to the establishment of structure–activity relationships, and to the optimization of all indispensable parameters for its development: potency, selectivity, metabolism, bioavailability, toxicity, cost price, etc. In other terms, even if the initial discovery was purely fortuitous, subsequent research must be marked by a very important effort of rationalization.

V. FOURTH STRATEGY: PLANNED RESEARCH AND RATIONAL APPROACHES

The approaches that we have described up to now allow a great place to chance (screening, fortuitous discoveries) or they lack originality (therapeutic copies). A more scientific approach is based on the knowledge of the incriminated molecular target: enzyme, receptor, ion channel, signaling protein, transport protein or DNA. The progresses in molecular and structural biology allowed the identification and characterization of several hundreds of new molecular targets and made it possible to envisage the design of drugs at a more scientific level.

A. L-DOPA and Parkinsonism

A historical example in which the key-information that rendered possible a rational approach to drug design is the discovery of the usefulness of L-3,4-dihydroxy-phenylalanine (L-DOPA) in the treatment of Parkinson's disease. Thus, since it was observed in patients suffering from parkinsonism that the dopamine levels in the basal ganglions were much lower than those found in the brains of healthy persons,[62] a symptomatic, but rational, therapy became possible. This therapy consists of administering to patients the L-DOPA; this amino acid is able to cross the blood-brain barrier, and is then decarboxylated into dopamine by brain DOPA-decarboxylase. Initial clinical studies were undertaken by Cotzias, Van Woert and Schiffer.[63] Several hundred thousand patients have benefited from this treatment. However, 95% of the DOPA administered by the oral route is decarboxylated in the periphery before having crossed the blood-brain barrier. To preserve the peripheral DOPA from this unwanted precocious degradation, a peripheral

FIGURE 6.20 Scheme of the renin-angiotensin and of the kallikrein–kinin systems. The converting enzyme (a carboxy-dipeptidyl-hydrolase) is common to the two systems.

inhibitor of DOPA-decarboxylase is usually added to the treatment. An additional improvement of the treatment is the simultaneous addition of an inhibitor of catechol *O*-methyltransferase such as tolcapone or entacapone (see the Section II.A. in Chapter 14).

Other examples of rational approach in pharmacology are the discovery of inhibitors of the ACE or that of antagonists of histaminergic H_2-receptors.

B. Inhibitors of the ACE

The ACE catalyses two reactions which are supposed to play an important role in the regulation of the arterial pressure: (1) conversion of angiotensin I, which is an inactive decapeptide, into angiotensin II, an octapeptide with a very potent vasoconstrictor activity and (2) inactivation of the nonapeptide, bradykinin, which is a potent vasodilator (Figure 6.20).

An inhibitor of the converting enzyme would therefore constitute a good candidate for the treatment of patients suffering from hypertension. The first substance developed in this sense has been teprotide, a nonapeptide presenting an identical sequence to that of some peptides isolated in 1965 by Ferreira from the venom of *Bothrops jararaca*, a Brazilian viper (Figure 6.21).

Teprotide inhibits in a competitive manner the degradation of angiotensin I by the converting enzyme. The presence of four prolines and a pyroglutamate renders this peptide relatively resistant to hydrolysis, but not to a sufficient degree to allow its oral administration. In the search for a molecule offering better bioavailability, the reasoning of the Squibb scientists rested on the analogy of the ACE with the bovine carboxypeptidase A.[64] In fact, both enzymes are carboxypeptidases; carboxypeptidase A detaches only one C-terminal amino acid while the converting enzyme detaches two. Furthermore, it was known that the active site of carboxypeptidase A comprises three important elements for the interaction with the substrate (Figure 6.22): an electrophilic center, establishing an ionic bond with a carboxylic function, a site capable to establish a hydrogen bond with a peptidic C-terminal function, and an atom of zinc, solidly fixed on the enzyme and serving to form a coordinating bond with the carbonyl group of the penultimate (the scissile) peptidic function.

pyro-Glu-Trp-Pro-Arg-Pro-Glu-Ile-Pro-Pro-OH

FIGURE 6.21 The structure of the nonapeptide teprotide.

By identifying that the conversion enzyme had a similar function, however altered by one amino acid unit (cleavage of the second peptidic bond instead of the first, departing from the terminal carboxyl group), scientists of the Squibb company have imagined the model drawn on Figure 6.23.

According to this model, *N*-succinyl amino acids such as the succinyl prolines shown in Figure 6.23 (right) should be able to interact with each of the above-mentioned sites based on, first their proline carboxyl (ionic bond), their amide function (hydrogen bond) and then on the carboxyl of the succinyl moiety (coordination with the zinc atom). These compounds should then be able to act as competitive and specific inhibitors of the converting enzyme. Therefore, a series of *N*-succinyl amino acids was prepared and the *N*-L-proline derivative **1** (Figure 6.24) showed some activity (IC_{50} = 330 μM). Amino acids other than L-proline lead to less active succinyl derivatives; this result is in agreement with the fact that several peptidic inhibitors (notably teprotide) possess also a proline in the C-terminal position. In the present example, *N*-succinyl-L-proline was selected as lead compound. The next task was to optimize its activity and this was done by researching the best interaction with the active site of the enzyme. Two steps were decisive in this quest: "the fishing for hydrophobic pockets" and the research for a better coordinant for the zinc atom (Figure 6.24). The exploration of hydrophobic pockets was achieved by substituting the succinyl moiety with methyl groups (four possibilities taking into account the regio- and the stereoisomers). Structure **2**, methylated at position β to the amide appeared clearly more active than **1** (IC_{50} = 22 instead of 330 μM). In this process, one observes an important stereoselectivity, since the IC_{50} value of epimer **3** of the compound **2** drops to 1,480 μM. The best coordination with the zinc was achieved in replacing the carboxyl function by a mercapto group. The gain resulting from this modification has been extremely important as shown by the comparison of compounds **1** and **4** or also **2** and **5**. Compound **5** (SQ 14225) with an IC_{50} of 0.023 μM, and a

FIGURE 6.22 Interactions between carboxypeptidase A and a substrate (left) or an inhibitor (right). *Source*: Adapted after Cushman *et al.*[64]

FIGURE 6.23 Interactions between ACE (a dipeptidyl carboxypeptidase) and a substrate (left) or inhibitors (right). *Source*: Adapted after Cushman *et al.*[64]

FIGURE 6.24 Structures of some key compounds in the development of captopril and enalapril.[64]

K_i of 0.0017 μM is active by the oral route and has been introduced in therapy under the designation of captopril.

It is interesting to observe that the loss in affinity caused by the replacement of the mercapto function by a carboxyl rest was compensated, thanks to an additional hydrophobic interaction. Thus, scientists from Merck developed enalaprilat **6a**, a compound of comparable effectiveness and for which the additional hydrophobic interaction is provided by a phenethyl rest. But enalaprilat is poorly absorbed orally; therefore the commercial compound is enalapril **6b**, the corresponding ethyl ester.

C. Discovery of the H_2-receptor antagonists

Research to develop specific antagonists of the H_2 histamine receptor in view of the treatment of gastric ulcers has also proceeded through a rationally thought process.[65,66] Starting from the observation that the antihistaminic compounds known at that time (antagonists of H_1-receptors) were not capable to antagonize the gastric secretion provoked by histamine, Black and his collaborators envisaged the existence of an unknown subclass of the histamine receptor (the future H_2-receptor). From 1964 on, they initiated a program of systematic research of specific antagonists for this receptor.

The starting point was the guanyl-histamine (Figure 6.25) that possesses weak antagonistic properties against the gastric secretion induced by histamine. The lengthening of the side chain of this compound increased clearly the H_2-antagonistic activity, but a residual agonist effect remained. In replacing the strongly basic guanidino function by a neutral thiourea, *burimamide* was obtained. Although very active, this compound was rejected for its low oral bioavailability.

The addition of a methyl group in position 4 of the imidazolic ring, followed by the introduction of an electron-withdrawing sulfur atom in the side chain, led finally to a compound that was both very active and less ionized, properties which improved its absorption by the oral route. The derivative thus obtained, *metiamide* was excellent and, moreover, 10 times more potent than burimamide. However, metiamide, because of its thiourea grouping, was tainted with side effects (agranulocytosis, nephrotoxicity), that would limit its clinical use. The replacement of the thiourea by an isosteric grouping having the same pK_a (*N*-cyanoguanidine) led finally to *cimetidine* that became a medicine of choice in the treatment of gastric ulcers. Later on, it appeared that the imidazolic ring, present in histamine and in all H_2-antagonists discussed hitherto, was not indispensable to the H_2-antagonistic activity. Thus, *ranitidine*, which possesses a furan ring, has appeared to be even more active than cimetidine. The same proved to be true for *famotidine* and *roxatidine*.

FIGURE 6.25 Structures of some key compounds in the development of H_2-receptor antagonists.

VI. CONCLUSION

The means leading to the discovery of new lead compounds, and possibly to new drugs, can be schematically classified into four approaches. These consist of the improvement of already existing drugs, of systematic screening, of retroactive exploitation of biological information, and of attempts toward rational design. Depending on which of these four strategies they apply, medicinal chemists can be seen as copiers, industrious, intuitive and deductive. It would be imprudent to compare hastily the merit of each of these characteristics. Indeed a "poor" research can end with a universally recognized medicine and, conversely, a brilliant rational demonstration can remain sterile. It is therefore of highest importance, given the random character of discovery and the virtual impossibility of planned invention of new active principles, that decision-makers in the pharmaceutical industry appeal to all the four strategies that have been described and that they realize that they are not mutually exclusive. On the other hand, it would be inappropriate, once a lead compound is discovered and characterized, not to study its molecular mechanism of action. Every possible effort should be made in this direction. In conclusion, all strategies resulting in identification of lead compounds are *a priori* equally good and advisable, provided that the research they induce afterwards is done in a rational manner.

REFERENCES

1. Clausewitz von, C. *Vom Kriege, Drittes Buch, 1. Kapitel Strategie*, 18-1973 ed. Ferdinand Dümmlers Verlag: Bonn, 1832.
2. Ramesha, C. S. How many leads from HTS? – Comment. *Drug Discov. Today* 2000, *5*, 43–44.
3. Wermuth, C. G. Selective optimization of side activities: another way for drug discovery. *J. Med. Chem.* 2004, *47*, 1303–1314.
4. Hopkins, A., Lanfear, J., Lipinski, C. A., Beeley, L. Chemical tools for indications discovery. In *Annual Reports in Medicinal Chemistry* (Doherty, A., Ed.), Vol. 40. Academic Press: San Diego, CA, 2005, pp. 339–348.
5. Rey, A. *Dictionnaire historique de la langue française*, Vol. 1. Dictionnaires Le Robert: Paris, 1992.
6. Worms, P., Kan, J.-P., Wermuth, C. G., Roncucci, R., Bizière, K. SR 95191 a selective inhibitor of type A monoamine oxidase with dopaminergic properties. I. Psychopharmacological profile in rodents. *J. Pharmacol. Expt. Ther.* 1987, *240*, 241–250.
7. Jeanmart, C., Cotrel, C. Synthèse de (chloro-5 pyridyl-2)-6 méthyl-4 pipérazinyl-1) carbonyloxy-5 oxo-7 dihydro-6,7 5*H*-pyrrolo [3,4-*b*] pyrazine. *Compt. Rend. Acad. Sci. Paris Ser. C* 1978, *287*, 377–378.
8. Doble, A., Canto, T., Piot, O., Zundel, J.-L., Stutzmann, J.-M., Cotrel, C., Blanchard, J.-C. The pharmacology of cyclopyrrolone derivatives acting at the GABAA/benzodiazepine receptor. *Adv. Biochem. Psychopharmacol.* 1992, *47*, 407–418.
9. Angell, M. *The truth about the drug companies – how they deceive us and what to do about it*. Random House: New York, 2004.
10. Thuilier, J. *Les dix ans qui ont changé la folie*. Robert Laffont: Paris, 1981, pp. 253–257.
11. Kling, J. From hypertension to angina to Viagra. *Mod. Drug Discov.* 1988, 31–38.
12. Mannhold, R. BTS-49465. *Drug. Fut.* 1986, *11*, 177–178.
13. Sternbach, L. H. The benzodiazepine story. *J. Med. Chem.* 1979, *22*, 1–7.
14. Blanchard, J. C., Boireau, A., Garret, C., Julou, L. *In vitro* and *in vivo* inhibition by zopiclone of benzodiazepine binding to rodent brain receptors. *Life Sci.* 1979, *24*, 2417–2420.
15. Suffness, M. Taxol: from discovery to therapeutic use. In *Annual Reports in Medical Chemistry* (Bristol, J. A., Ed.), Vol. 28. Academic Press: San Diego, 1993, pp. 305–314.
16. Boyd, *Drug Dev. Res.* 1995, *34*, 91–109.
17. Endo, A. Compactin (ML-236B) and related compounds as potential cholesterol-lowering agents that inhibit HMG-CoA reductase. *J. Med. Chem.* 1985, *28*, 401–405.
18. Lee, T. J. Synthesis, SARs and therapeutic potential of HMG-CoA reductase inhibitors. *Trends Pharmacol. Sci.* 1987, *8*, 206–208.
19. Shu, Y. Z., Cutrone, J. Q., Klohr, S. E., Huang, S. BMS-192548, a tetracyclic binding inhibitor of neuropeptide Y receptors, from Aspergillus niger WB2346. II. Physico-chemical properties and structural characterization. *J. Antibiot.* 1995, *48*, 1060–1065.
20. Nisbet, L. J., Westley, J. W. Developments in microbial products screening. In *Annu. Rep. Med. Chem.* (Bristol, J., Ed.), Vol. 21. Academic Press: San Diego, CA, 1986, pp. 149–157.
21. Hylands, P. J., Nisbet, L. J. The search for molecular diversity (I). In *Annu. Rep. Med. Chem.* (Bristol, J., Ed.), Vol. 26. Academic Press: San Diego, CA, 1991, pp. 259–269.
22. Böhm, H.-J., Klebe, G., Kubinyi, H. *Wirkstoffdesign, p.140*. Spektrum Akademischer Verlag: Heidelberg, 1996, p. 599.
23. Wermuth, C. G. Search for new lead compounds: the example of the chemical and pharmacological dissection of aminopyridazines. *J. Heterocycl. Chem.* 1998, *35*, 1091–1100.
24. Wermuth, C. G., Clarence-Smith, K. "Drug-like" leads: bigger is not always better. *Pharmaceutical News*, 2000, *7*, 53–57.
25. Poroikov, V., Akimov, D., Shabelnikova, E., Filimonov, D. Top 200 medicines: can new actions be discovered through computer-aided prediction? *SAR and QSAR in environmental research*. 2001, *12*, 327–344.
26. Wermuth, C. G. The "SOSA" approach: an alternative to high-throughput screening. *Med. Chem. Res.* 2001, *10*, 431–439.
27. Wermuth, C. G. Possible alternatives to high-throughput screening. In *Drug Discov. Dev.* (Chorgade, M. S., Ed.), Vol. 1. John Wiley & Sons: Hoboken, NJ, 2006, pp. 213–232.
28. Stein, P. D., Hunt, J. T., Floyd, D. M., Moreland, S., Dickinson, K. E. J., Mitchell, C., Liu, E. C.-K., Webb, M. L., Murugesan, N., Dickey, J., McMullen, D., Zhang, R., Lee, V. G., Serdino, R., Delaney, C., Schaeffer, T. R., Kozlowski, M. The discovery of sulfonamide endothelin antagonists and the development of the orally active ETA antagonist 5-(dimethylamino)-*N*-(3,4-dimethyl-5-isoxazolyl)-l-naphthalenesulfonamide. *J. Med. Chem.* 1994, *37*, 329–331.
29. Murugesan, N., Gu, Z., Spergel, S., Young, M., Chen, P., Mathur, A., Leith, L., Hermsmeier, M., Liu, E. C.-K., Zhang, R., Bird, E., Waldron, T., Marino, A., Koplowitz, B., Humphreys, W. G., Chong, S., Morrison, R. A., Webb, M. L., Moreland, S., Trippodo, N., Barrish, J. C. Biphenylsulfonamide endothelin receptor antagonists. 4. Discovery of *N*-[[(4,5-dimethyl-3-isoxazolyl)amino]sulfonyl]-4-(2-oxazolyl)[1,1′-biphenyl]-2-yl]methyl]-*N*,3,3-trimethylbutanamide (BMS-207940, a highly potent and orally active ETA selective antagonist. *J. Med. Chem.* 2003, *46*, 125–137.
30. Stein, P. D., Hunt, J. T., Floyd, D. M., Moreland, S., Dickinson, K. E. J., Mitchell, C., Liu, E. C.-K., Webb, M. L., Murugesan, N., Dickey, J., McMullen, D., Zhang, R., Lee, V. G., Serdino, R., Delaney, C., Schaeffer, T. R., Kozlowski, M. The discovery of sulfonamide endothelin antagonists and the development of the orally active ET_A antagonist 5-(dimethylamino)-*N*-(3,4-dimethyl-5-isoxazoly1)-l-naphthalenesulfonamide. *J. Med. Chem.* 1994, *37*, 329–331.
31. Wermuth, C. G., Schlewer, G., Bourguignon, J.-J., Maghioros, G., Bouchet, M.-J., Moire, C., Kan, J.-P., Worms, P., Bizière, K. 3-Aminopyridazine derivatives with atypical antidepressant, serotonergic and dopaminergic activities. *J. Med. Chem.* 1989, *32*, 528–537.

32. Wermuth, C. G. Aminopyridazines – an alternative route to potent muscarinic agonists with no cholinergic syndrome. *Il Farmaco* 1993, *48*, 253–274.
33. Wermuth, C. G., Bourguignon, J.-J., Hoffmann, R., Boigegrain, R., Brodin, R., Kan, J.-P., Soubrié, P. SR 46559A and related aminopyridazines are potent muscarinic agonist with no cholinergic syndrome. *Bioorg. Med. Chem. Lett.* 1992, *2*, 833–836.
34. Pilla, M., Perachon, S., Sautel, F., Garrido, F., Mann, A., Wermuth, C. G., Schwartz, J.-C., Everitt, B. J., Sokoloff, P. Selective inhibition of cocaine-seeking behaviour by a partial dopamine D_3 agonist. *Nature* 1999, *400*, 371–375.
35. Prestwick Chemical Inc., K. S. N. *Suite 1475*, Washington, DC, 20006-1202; www.prestwickchemical.com, In.
36. Kleemann, A., Engel, J. *Pharmazeutische Wirkstoffe – Synthese, Patente, Anwendungen (Preface)*. Georg Thieme Verlag: Stuttgart, 1982.
37. Price, J. R. Antifertility agents of plant origin. In *A Symposium on Agents Affecting Fertility* (Austin, C. R., Perry, J. S., Eds). Little, Brown and Co.: Boston, MA, 1965, pp. 3–17.
38. Laborit, H., Huguenard, P., Alluaume, R. Un nouveau stabilisateur végétatif, le 4560 R.P. *Presse Méd.* 1952, *60*, 206–208.
39. Maxwell, R. A., Eckhardt, S. B. *Drug Discovery – A Casebook and Analysis*. Humana Press: Clifton, NJ, 1990.
40. Stemp, G., Evans, J. M. Discovery and development of cromakalim and related potassium channel activators. In *Medicinal chemistry, the role of organic chemistry in drug research, 2nd Edition* (Ganellin, C. R., Roberts, S. M., Eds). Academic Press: London, 1993, pp. 141–162.
41. Hamilton, T. C., Weir, S. W., Weston, A. H. Comparison of the effects of BRL 34915 and verapamil on electrical and mechanical activity in rat portal vein. *Brit. J. Pharmacol.* 1986, *88*, 103–111.
42. Gerber, J. G., Nies, A. S. Antihypertensive agents and the drug therapy of hypertension. In *Goodman and Gilman's The Pharmacological Basis of Therapeutics, 8th Edition* (Goodman-Gilman, A., Rall, T. W., Nies, A. S., Taylor, P., Eds). Pergamon Press: New York, 1990, pp. 784–813.
43. Evans, J. M., Fake, C. S., Hamilton, T. C., Poyser, R. H., Watts, E. A. Synthesis and antihypertensive activity of substituted *trans*-4-amino-3, 4-dihydro-2,2-dimethyl-2*H*-1-benzopyran-3-ols. *J. Med. Chem.* 1983, *26*, 1582–1589.
44. Evans, J. M., Fake, C. S., Hamilton, T. C., Poyser, R. H., Showell, G. A. Synthesis and antihypertensive activity of 6,7-disubstituted *trans*-4-amino-3,4-dihydro-2,2-dimethyl-2*H*-1-benzopyran-3-ols. *J. Med. Chem.* 1984, *27*, 1127–1131.
45. Sneader, W. *Drug Prototypes and Their Exploitation*. John Wiley & Sons: Chichester, 1996, p. 242.
46. Iyer, C. G. S., Languillon, J., Ramanujam, K. WHO coordinated short-term double-blind trial with thalidomide in the treatment of acute lepra reactions in male lepromatus patients. *Bull. WHO* 1971, *45*, 719–732.
47. Cushman, M., Jayaraman, M., Vroman, J. A., Fukunaga, A. K., Fox, B. M., Kohlhagen, G., Strumberg, D., Pommier, Y. Synthesis of new indeno[1,2-*c*]isoquinolines: cytotoxic non-camptothecin topoisomerase I inhibitors. *J. Med. Chem.* 2000, *43*, 3688–3698.
48. Korth, C., May, B. C. H., Cohen, F. E., Prusiner, S. B. Acridine and phenothiazine derivatives as pharmacotherapeutics for prion disease. *Proc. Natl. Acad. Sci. USA* 2001, *98*, 9836–9841.
49. Terret, N. K., Bell, A. S., Brown, D., Ellis, P. Sildenafil (Viagra), a potent and selective inhibitor of type 5 cGMP phosphodiesterase with utility for the treatment of male erectile dysfunction. *Bioorg. Med. Chem. Lett.* 1996, *6*, 1819–1824.
50. Boolell, M., Allen, M. J., Ballard, S. A., Gepi-Attee, S., Muirhead, G. J., Naylor, A. M., Osterloh, I. H., Gingell, C. Sildenafil: an orally active-type 5 cyclic GMP-specific phosphodiesterase inhibitor for the treatment of penile erectile dysfunction. *Int. J. Urol. Res.* 1996, *8*, 47–52.
51. Foye, W. O., Lemke, T. L., Williams, D. A. *Principles of medicinal chemistry, p. 519*. Lea & Febiger: Baltimore, MD, 1995, p. 995.
52. Neuworth, M. B., Laufer, R. J., Barnhart, J. W., Sefranka, J. A., McIntosh, D. D. Synthesis and hypocholesterolemic activity of alkylidenedithio bisphenols. *J. Med. Chem.* 1970, *13*, 722–725.
53. Noble, R. L., Beer, C. T., Cutts, J. H. Role of chance-observations in chemotherapy: *Vinca rosea*. *Ann. NY Acad. Sci.* 1958, *76*, 882–894.
54. Johnson, I. S., Wright, H. F., Svoboda, G. H., Vlantis, J. Antitumor principles derived from *Vinca rosea* Linn. I. Vincaleukoblastine and leurosine. *Cancer Res.* 1960, *20*, 1016–1022.
55. Johnson, I. S., Armstrong, J. G., Gorman, M., Burnett, J. P., Jr. The vinca alkaloids: a new class of oncolytic agents. *Cancer Res.* 1963, *23*, 1390–1427.
56. Buxton, I. L. O., Cheek, D. J., Eckman, D., Westfall, D. P., Sanders, K. M., Keef, K. D. N^G-Nitro-L-arginine methyl ester and other alkyl esters of arginine are muscarinic antagonists. *Circ. Res.* 1993, *72*, 387–395.
57. Schrader, G. *Die Entwicklung neuer Insektizide auf Grundlage von Organischen Fluor und Phosphorverbindungen. Monographie No.62*. Verlag Chemie: Weinheim, 1952.
58. Janssen, P. A. J., Van de Westeringh, C., Jagneau, A. W. M., Demoen, P. J. A., Hermans, B. K. F., Van Daele, G. H. P., Shellek, K. H. L., Van der Eycken, C. A. M., Niemegers, C. J. E. Chemistry and pharmacology of CNS depressants related to 4-(4-hydroxy-4-phenylpiperidino) butyrophenone. Part I – synthesis and screening data in mice. *J. Med. Pharm. Chem.* 1959, *1*, 281–297.
59. Janssen, P. A. J. The evolution of the butyrophenones, haloperidol and trifluperidol, from meperidine-like 4-phenylpiperidines. *Int. Rev. Neurobiol.* 1965, *8*, 221–263.
60. Albert, A. *Selective Toxicity*, 6 Edition. Chapman and Hall: London, 1979, p. 221.
61. Shen, T. Y. Perspectives in non-steroidal antiinflammatory agents. *Angew. Chem. Int. Edit.* 1972, *11*, 460–472.
62. Ehringer, H., Hornykiewicz, O. Verteilung von Noradrenalin und Dopamin (3-hydroxytyramin) im Gehirn des Menschen und ihr Verhalten bei Erkrankungen des extrapyramidalen systems. *Klin. Wochenschr.* 1960, *38*, 1236–1239.
63. Cotzias, G. E., Van Woert, M. H., Schiffer, L. M. Aromatic amino acids and modification of Parkinsonism. *New Engl. J. Med.* 1967, *276*, 374–379.
64. Cushman, D. W., Cheung, H. S., Sabo, E. F., Ondetti, M. A. Design of potent competitive inhibitors of angiotensin-converting enzyme. Carboxyalkanoyl and mercaptoalkanoyl amino acids. *Biochemistry* 1977, *16*, 5485–5491.
65. Black, J. W., Duncan, W. A. M., Durant, J. C., Ganellin, C. R., Parsons, M. E. Definition and antagonism of histamine H_2-receptors. *Nature* 1972, *236*, 385–390.
66. Ganellin, C. R. Cimetidine. In *Chronicles of Drug Discovery* (Bindra, J. S., Lednicer, D., Eds), Vol. 1. John Wiley and Sons: New York, 1982, pp. 1–38.

Chapter 7

High-Throughput Screening and Drug Discovery

John R. Proudfoot

I. INTRODUCTION
II. HISTORICAL BACKGROUND
III. FROM SCREEN TO LEAD
A. Compound collections
B. Assays
C. Hit-to-lead process
IV. EXAMPLES OF DRUGS DERIVED FROM SCREENING LEADS
A. Reverse transcriptase inhibitors, nevirapine, efavirenz, and delavirdine
B. Endothelin antagonists, bosentan, sitaxentan, edonentan, and ambrisentan
C. Raf kinase inhibitor, sorafenib
V. PRACTICAL APPLICATION RECENT EXAMPLE
A. IKK inhibitors
VI. CONCLUSION
REFERENCES

Ever tried Ever failed. No matter. Try again. Fail again. Fail better.

Samuel Beckett, "Worstward Ho"

I. INTRODUCTION

Until about 1980, information on drug targets at the molecular level was scarce and drug discovery was mostly driven by data obtained from testing relatively small numbers of compounds in pharmacological models.[1–4] Marketed drugs generally originated from lead structures that had well-defined medicinal properties, such as natural products or other drugs. In 1988, for example, Kurt Freter noted that all of the new drugs that had been approved by the Food and Drug Administration (FDA) in 1985 appeared to be the result of analog-based approaches.[5] The number of exploited lead structures was relatively small: Walter Sneader categorized some 244 drug prototypes, fewer than 140 of which would be considered drug-like as currently defined and, among these, only 25 originated from screening processes.[6] Screening, either random or directed, was a low throughput, manually intensive process generally conducted in animal models and usually directed toward the identification of drug rather than lead candidates.

Even through the end of the 1980s a screening capacity of hundreds of samples per week was deemed high-throughput. Over the past 20 years drug discovery has moved to a target-based focus[7] that has been enabled by advances in molecular biology, automation, combinatorial chemistry, and informatics. Many thousands of compounds can now be screened rapidly against a biological molecular target or cellular process and, in most drug discovery organizations, high-throughput screening (HTS) or ultra-high-throughput screening (uHTS) is a central paradigm for the identification of novel lead structures. Although HTS approaches are now also applied during lead optimization (LO) to the assessment of properties such as solubility and cytochrome P450 inhibition, the focus here is on impact related to lead discovery rather than LO.

II. HISTORICAL BACKGROUND

Over the years, various screening strategies have been applied to the identification of drug and lead candidates. The screening of natural product extracts for bioactivity followed by the isolation of the active principle or principles,

Copyright © 2008, Elsevier Ltd
All rights reserved.

which may be classified as the screening of compound mixtures against multiple biological targets simultaneously, has been a highly productive source of drug and lead discovery. The identification of cyclosporine A as an effective immunesupressant is illustrative of this process.[8] Screening of a fungal extract for antimicrobial activity led to the isolation of a number of cyclosporins that were subsequently found to possess immunosuppressive properties. The molecular target mediating the pharmacological activity was only identified several years after marketing approval was obtained for the drug.[9]

Screening of discrete synthetic compounds *in vivo* or *in vitro* for a targeted pharmacological or phenotypic effect, essentially the testing of individual compounds against multiple targets simultaneously, constitutes a second approach. This approach was employed by Ehrlich in the early part of the last century to identify arsphenamine, an important early anti-infective drug.[10] From the 1950s onward, the National Cancer Institute has screened many thousands of samples per year *in vivo* and *in vitro*[11] and identified anticancer agents such as hydroxyurea[12] and the lead structure for carmustine[13] using this approach.

A third approach, screening large numbers of individual compounds or defined pools of compounds against discrete biological targets has been effectively enabled in recent years and constitutes a core concept in HTS. Although there are certainly historical examples of LO campaigns driven by test data derived against isolated targets,[14,15] the capacity for such testing in a high-throughput manner to identify novel lead structures was previously limited by the relatively small numbers of synthetic compounds available for screening and the lack of well-characterized biological targets. Advances in molecular biology provided access to many potential drug targets as pure or overexpressed proteins and made them available for molecular or cellular assays. Improved automation and informatics provided tools for the organization of screening libraries and the collection and interpretation of the large quantities of data generated by screening campaigns. Finally, combinatorial chemistry and high-throughput synthesis methods provided large collections of compounds to fuel the screening process. In all, the combined application of these new technologies has enabled a HTS approach to lead identification, and now large numbers of discrete small molecules can be assessed for activity against a well-defined biological target within a relatively short period of time.

The widespread adoption of HTS and associated technologies for lead identification led the expectation that an increased number of drug candidates would progress through the clinic; however, it is only relatively recently that the positive impact of HTS on drug discovery has become apparent. As shown in Table 7.1, from 2005 onward a significant number of approved small molecule drugs has emerged from screening leads, whereas in the decade before, one or two new drugs per year at most originated from any screening approach. One key reason for this time-lag or delay in impact resides in the long timelines of the modern drug discovery process, often 12 years or more from project initiation to drug approval. However, also contributing to this delayed impact is the time required to develop processes that effectively leverage the application of these new technologies to lead identification, in other words – a learning curve (Box 7.1). For instance, Chris Lipinski and colleagues noted in 1997 that HTS campaigns tended to produce relatively large, lipophilic lead molecules.[16,17] Since it is common during LO to enhance potency through the addition of lipophilic substituents, many discovery campaigns based on screening leads provided large, insoluble lipophilic drug candidates-molecules difficult to progress successfully through the clinic.[18] The resulting "Rule of 5" provided a correlation between molecular properties (MW, Log *P*, number of

TABLE 7.1 Drugs Derived from Screening Leads

Drug	Approved	Lead source
Nevirapine[73]	1996	Corporate historical
Delavirdine[74]	1997	Corporate historical
Efavirenz[75]	1998	Corporate historical
Tirofiban[76]	1998	Directed screening
Bosentan[58]	2001	Corporate historical
Gefitinib[77]	2002	Computational screening
Sivelestat[78]	2002	Corporate historical
Aprepitant[79]	2003	Different company
Cinacalcet[80]	2004	Drug
Sorafenib[81]	2005	Commercial acquisition
Tipranavir[82]	2005	Drug
Conivaptan[83,84]	2005	Different company
Mozavaptan[85]	2006	
Sunitinib[86]	2006	
Dasatinib[87]	2006	Target switch
Sitaxentan[49,50,51,52]	2006	Drug/directed screening
Sitagliptin[88]	2006	
Ambrisentan[60]	2007	Agrochemical
Maraviroc[89]	2007	

BOX 7.1 The Evolution of New Technologies

Source: James C. Bezdek: Fuzzy models – What are they and why? *IEEE Trans. Fuzzy Syst*. 1993, 1. © 1993 IEEE, Reprinted with permission.

H-bond donors and acceptors) and the potential to achieve physical properties consistent with acceptable oral absorption, and spurred many subsequent efforts to qualify drug-like molecular properties. Further refinements, based on analyses of lead and corresponding drug molecules, led to the proposal of distinct lead-like molecular properties (e.g. MW < 300, Log $P < 3$)[19–21] and assessments of drug and lead-like characteristics are now a routine part of the progression from screening-hit to lead series. Almost simultaneously, also in response to the challenges associated with identifying high quality lead structures from screening campaigns, a formalized "hit-to-lead" process distinct from LO emerged.[22] Today, in most large pharmaceutical organizations, dedicated teams are responsible for the progression of screening hits to the start of LO.

III. FROM SCREEN TO LEAD

Fundamentally, the quality of the lead structures obtained from screening will depend on the nature of the compounds in the screening collection, the quality of the assay system, and the processes that are in place to progress from the assessment of active samples to the delivery of a lead series.[23]

A. Compound collections

Corporate screening collections now often exceed one million compounds.[24] Ideas on the optimal size for a collection range from the suggestion that two to three million suitable compounds should deliver multiple starting points from any given screen[25] to an estimate that up to 24 million compounds would be needed to ensure potent hits for all targets.[26] Collections consist of varying ratios of compounds originating from a number of sources such as previous drug discovery campaigns, combinatorial or high-speed chemistry, and acquisition or purchase from academic or commercial vendors. Natural products of plant, animal or microbial origin, either pure materials or extracts or mixtures, are also frequently part of a screening collection.

Compounds derived from previous drug discovery campaigns, historical or heritage compounds, may lack structural diversity particularly if prior research was focused in specific, limited target areas. Recent analyses of screening collections, prompted by mergers and consolidation, also indicate that a significant fraction of historical compounds might not be suitable for screening because of chemical degradation during long-term storage under nonideal conditions.[27,28] Indeed, confirmation of the chemical structure and composition of active samples is one of the standard early tasks involved in assessing hits from a screening campaign.

Compounds produced by early combinatorial chemistry methods were designed based on synthetic accessibility and diversity of structure and tended to be high molecular weight, lipophilic structures of questionable purity. The current emphasis on the quality of compounds produced by high-throughput synthesis, rather than on the sheer numbers, has led to libraries where drug-like properties are considered before synthesis is initiated and where design is often focused on providing leads against particular targets or target classes.

Compounds acquired commercially to augment screening collections are selected for diversity of structure and drug- or lead-like properties. However, since they originate in the public domain, these same compounds, or close analogs, are often present in many different screening collections and, if identified as lead structures, do not carry a satisfactory intellectual property position without substantial structural modification.[29]

Natural products tend to be structurally diverse and complex and, from a synthesis perspective, are often difficult to modify at sites that are relevant for SAR studies. As is apparent from Table 7.1, none of the screening-derived drugs that were approved over the past decade came from leads provided by the natural product world. A possible rationale is provided by Hann *et al.* who postulate an inverse relationship between molecular complexity beyond a certain level and the likelihood of encountering a productive binding event with a biological target. In other words, very complex molecules are less likely to forge a sufficient number of productive interactions with a target to overcome the many likely negative or unproductive ones.[20]

B. Assays

In the early 1990s, HTS was largely a manual process[30] with a throughput on the order of hundreds of samples

per week. Screening capacity has grown to the extent that collections of a million compounds or more can now be assessed in a month or less. The increased capacity has been facilitated by automation, high density plates and the movement of detection technology to fluorescence-based techniques. Statistical methods have been developed to assess the quality of any particular screen; the parameter *Z* which gauges signal reproducibility across the dynamic range of the assay is frequently used to quantify the robustness of an assay.[31] With 384 and 1,536 well plates now in routine use, miniaturization to low volume (uL) assays means that only minute amounts of compound are required for an individual test, and low milligram sample quantities can last for hundreds of screening campaigns.

Structural elements that often result in false and misleading positive activity are now well recognized; such functional groups include those that are chemically reactive and can act as alkylating, acylating or reducing agents. A composite listing of such groups is shown in Table 7.2, along with some additional structural features that are generally considered to be undesirable in screening samples.[28,32–34] These features may confer chelating, detergent-like or aggregation properties on a test sample[35,36] and can result in false positive activity through nonselective mechanisms. In all, a substantial experience base has emerged that now enables robust assays and the early identification of spurious activity based on artifacts.[37,38]

TABLE 7.2 Functional Groups and Structural Features Undesirable for Molecules in a Screening Collection*

Reactivity	Structural
RO—OR, RS—SR, RN—NR,	Any element other than H, C, N, O, S, F, Cl, Br, I
RN—OR, RN—SR, RS—OR	More than six F, more than three Cl, Br, I
Anhydride, acyl halide,	Epoxide, aziridine or thiirane
Activated ester or thioester	More than one nitro group
Alkyl chloride, bromide, iodide	Any ring larger than eight-membered
Sulfonyl halide, sulfonate ester	Crown ethers
N=C=N, N=C=S, N=C=O, isonitrile	Linear polycyclic aromatic systems
Nitroso, diazo, thiocyanate	≥linear $(CH_2)_6$
Aldehyde, cyanohydrin, imine, chloramidine	>2 Ar—NH_2 groups (Ar = phenyl or naphthyl)
Michael acceptors	>2 Ar—OH groups (Ar = phenyl or naphthyl)
Alkyl—SH	Diacetylene, polyene
1,2- or 1,4-quinone	Trihydroxyphenyl
N-Halogen or S-Halogen	≥4 Acidic groups
Activated 2-haloheterocycles	≥4 Basic Nitrogen atoms
β-Lactam	>2 Quaternary amine
1,2-dicarbonyl	
S, O, Cl, or I atom carrying a positive charge	
Phosphoramide, phosphorane	

*Source: Compiled from Refs 28, 32–37

C. Hit-to-lead process

It would not be unusual for a screening campaign involving one million samples to generate several thousand samples that display activity above a meaningful threshold. Confirmation of both the activity and the identity of the active structures provides a set of confirmed screening hits. Whereas in the past identification of structures with acceptable molecular potency, selectivity, and patentability might have been sufficient to initiate LO,[22] it is now customary to provide a more detailed assessment of the liabilities and opportunities associated with any intended lead series via a process frequently referred to as "hit to lead." An important part of this process is the demonstration that an appropriate pharmacokinetic (PK) or pharmacological profile can be achieved in addition to satisfactory molecular potency.

Obstacles to a targeted PK profile, for example, poor permeability or rapid metabolism, can be identified through surrogates such as permeability in a Caco-2 assay or by *in vitro* metabolism as measured in microsomal or hepatocyte preparations. It is also increasingly common to provide *in vivo* PK data on representative structures.[39] Off-target liabilities such as hERG or CYP inhibition[40,41] should also be identified at this early stage – it is not necessary to fix all the issues that are identified, but data should indicate that there is a path forward during LO (Box 7.2).

IV. EXAMPLES OF DRUGS DERIVED FROM SCREENING LEADS

By its nature, a screening approach to lead identification has the potential to identify not only structurally novel lead compounds, but also unexpected modes of action or allosteric inhibition.[42] In cases for which there is no prior experience with drugging a particular biological target or where the natural ligand is not amenable to rational medicinal chemistry processes, screening may provide the most effective method for lead identification. However, because of the

BOX 7.2 Progression from Screen to Lead

		Number	
Hit identification	**Compounds screened**	10^6	
	Confirmed active samples	10^3–10^4	
	Counter screens Dose-responsive Druglike or leadlike properties, no toxicophores Confirmed structure and purity		
	Selective confirmed hits	10^1–10^3	
Hit to lead	Demonstrated, exploitable SAR Tractable synthesis Solubility, ADME, toxicity profiling Demonstrated efficacy in cellular and/or pharmacological models Expanded selectivity profiling Clear patent strategy		
Lead optimization	**Lead series**	3–5	~1 year

similarity of screening collections and the attractiveness of particular targets and target classes, the potential to discover structurally and mechanistically novel leads is coupled with the possible identification of similar lead structures by competing research groups. The examples selected below highlight some of the opportunities and challenges offered by leads obtained from screening processes.

A. Reverse transcriptase inhibitors, nevirapine, efavirenz, and delavirdine

Through the 1990s a handful of drugs derived from leads identified by screening or HTS reached the market. Three of these, efavirenz (**2**), delavirdine (**6**) and nevirapine (**9**) (Figure 7.1) are AIDS therapeutics and target the viral reverse transcriptase enzyme (HIV-1 RT) via a novel allosteric mechanism of inhibition. The lead structures, all derived from historical corporate collections, are also shown in Figure 7.1 along with a summary of the major issues, in addition to improving potency, that were addressed during the LO campaigns.

A number patents and publications describe sedative or antidepressant properties for quinazolinethiones such as **3** related to the efavirenz lead **1**.[43] Chemical instability of the lead structure **1**, due to the masked ketone at the 4-position, was addressed by replacing the ethoxy group with a carbon linked substituent.[44] A focus on replacing the thiourea functionality because of potential toxicity led to urea analogs, and subsequent efforts were directed toward solving the low metabolic stability of the *N*-methyl group. This was attained by a switch to the benzoxazinone system present in efavirenz, in which –O— is replaced –*N*(Me)–; however, this scaffold change was only enabled after considerable SAR studies had identified 4-position substituents that retained potent enzyme inhibition across the scaffold switch.

The lead **4** for delavirdine (**6**) was discovered in a screened set of 1,500 computationally diverse representatives of the Upjohn compound collection. There is only one literature ref.[45] to the delavirdine lead structural type, exemplified by compound **7**, prior to the disclosure of RT inhibitory activity for this class. Rapid SAR expansion of the lead was enabled by *N*-benzyl connectivity and many alkylated and acylated variations of the upper portion of the piperazine scaffold were explored. Ultimately the acylindole, initially bearing a 5-methoxy substituent as in the first clinical candidate atevirdine (**5**), emerged as preferred. This was found to be metabolically labile and was subsequently replaced with the methylsulfonamide group. Early work also identified the *N*-ethyl substituent of the lead as a potential metabolic liability and, although this pattern was retained in the first clinical candidate, it was replaced by the *N*-isopropyl substituent in the approved drug, delavirdine (**6**).

Structures related to the nevirapine lead **8** are well represented in the patent and scientific literature, since the core system is similar to that in the approved drug pirenzepine (**10**). Initial SAR efforts were driven largely by metabolic instability associated with each of the *N*-alkyl substituents. An acceptable profile was achieved with two changes. First, by modifying the attachment point of the methyl group from the 5- to the 4-position the extent of

FIGURE 7.1 HIV-1 reverse transcriptase inhibitors.

metabolism was significantly decreased and this change also led to an improvement in potency. Replacement of the *N*-11 ethyl substituent with a cyclopropyl group also provided an improved profile. Additionally, the introduction of a second nitrogen atom in the tricyclic ring system gave a further boost in potency and also improved solubility and provided the drug candidate nevirapine, **9**.[46]

It is important to emphasize that, for these three examples, the lead structures, which act by a unique allosteric mode of inhibition, would not have been discovered by any other method available at the time, that is, were it not for the application of a screening approach to lead identification this class of drugs might not have emerged to find use in the clinic (Box 7.3).

BOX 7.3 An Alternative Representation or the Evolution of the Nevirapine Series

This illustration of property change over time highlights the importance of corporate compound collection for the initial rapid progression of the SAR studies. There was a wealth of potential structure/activity information available even before synthesis was initiated. The attractive lead-like properties of the initial structure allowed substantial latitude for increases in MW and lipophilicity during LO. After the identification of nevirapine, subsequent medicinal chemistry efforts revolved around generating activity against multiple resistant mutant RT enzymes.[47,48] Improved potency profiles were achieved through additional substituents on the dipyridodiazepane ring system, although these usually came at the cost of metabolic or other liabilities.

B. Endothelin antagonists, bosentan, sitaxentan, edonentan, and ambrisentan

The examples above illustrate a situation where screening against a target gives multiple structurally distinct lead classes. Alternatively, screening may give identical or very similar lead structures to different organizations. This is not an unlikely scenario, given that a substantial portion of any screening collection may be composed of commercially acquired compounds. In the examples shown in Figure 7.2 from endothelin antagonist programs, screening based approaches identified the lead structures that resulted in sitaxentan, bosentan, ambrisentan (all approved drugs) and edonentan (a clinical candidate). Two of these discovery campaigns began from very similar lead structures.

A directed screening approach, in which computational searching based on the pharmacophoric elements apparent in the natural ligand was used to direct the purchase or selection of compounds for screening, identified sulfisoxazole (**11**) as a moderately effective inhibitor of binding of

FIGURE 7.2 Endothelin antagonists.

radiolabeled endothelin 1 to the ET_A receptor.[49] SAR studies showed that the aniline moiety was not required for activity and the phenyl ring was replaced with the isosteric thiophene.[50] Subsequent elaboration on this ring gave molecules, such as the amide **12**, which demonstrated inhibition at the low nM level.[51] These compounds, however, demonstrated poor bioavailability due to cleavage of the amide bond and a search for a stable metabolically stable replacement eventually identified the ketone linker present in sitaxentan (**13**).[52]

The same lead structure, sulfisoxazole (**11**), was also identified by Bristol Meyers Squibb, via the screening hit sulfathiazole (**14**).[53] In this instance modification of the aniline led, via a naphthalene ring, to the substituted biphenyl **15**;[54] however, preclinical PK studies indicated that the attached isobutyl group was subject to extensive hydroxylation.[55] Replacement of this group with an isoxazole ring imparted metabolic stability, and expansion of the SAR at the tolerant 2′-position provided the clinical candidate edonentan (**16**).[56,57]

Two additional approved endothelin antagonists which derived from HTS leads have progressed to the market – bosentan (**19**) and ambrisentan (**21**). The lead **18** for bosentan,[58] although not a sulfa drug, bears a structural resemblance to sulfisoxazole and sulfathiazole, although in more elaborated form. Substances closely related to **18**, differing only in the chloro and trifluoromethyl substituents, are exemplified in the patent literature from Hoffman La Roche as blood sugar lowering or antidiabetic agents devoid of antibacterial activity [59] and it is possible that the lead was originally synthesized based on a sulfa drug precursor structure. It is noteworthy that the lead remained part of the screening collection for more than 20 years before it emerged as the starting point for bosentan. Since the lead selection criteria included demonstrated bioavailability and *in vivo* activity, the optimization process focused on improving potency which was attained by the modifications shown.[58]

In the case of ambrisentan (**21**), the lead structure **20** was originally synthesized as an herbicide and is similar to many such compounds patented in the late 1980s. In addition to improving potency, one early goal was to simplify the lead structure by incorporating identical substituents at one of the two chiral centers. The strategy proved remarkably successful in that a straightforward diphenyl substitution at the left-hand side chiral center provided both simplification and improved potency. Ambrisentan is a rare example of an optimized drug that is smaller and less complex than the original screening lead.[60]

C. Raf kinase inhibitor, sorafenib

One notable example of HTS providing a lead structure with a novel chemotype and a novel mechanism of action is presented in the discovery of sorafenib (**28**), Figure 7.3, which was recently approved as an anticancer agent. An HTS conducted against Raf-1, a kinase implicated in cancer cell proliferation, provided the thienylurea **22**, a commercially available compound, as a lead structure (Figure 7.3).[61] Among many early structural changes, it was found that methyl substitution at the 4-position of the right-hand phenyl ring, as in **23** provided enhanced potency, although it proved difficult to make further progress with the thiophene left-hand ring in place. Since the structural class was amenable to combinatorial chemistry methods, many combinations of left- and right-hand side variations were made and tested. A significant discovery emerging from this approach was that the thiophene could be replaced with an isoxazole ring system as in **26**. The particular combination of left- and right-hand side fragments in **26** deviated from predicted activity in that the corresponding analogs with the single point changes, **24** and **25**, showed little activity against the target. This is an instructive example of the nonlinear nature of SAR progression and provides a clear example of the power of combinatorial chemistry applied during LO. In this instance, it provided the opportunity to discover an unexpected synergism between pharmacophoric elements. Introduction of a pyridine ring at the right-hand side improved solubility, and incorporation of a disubstituted phenyl ring on the left side provided additional potency as in **27**. Finally, introduction of the carboxamide substituent on the pyridine ring as shown for sorafenib (**28**) provided a further boost in potency thought to be due to a favorable hydrogen bonding interaction with the target enzyme.

At the time of discovery, the thienylurea lead structure **22** represented a novel structural class for kinase inhibition. Concurrently, similar structures were identified as inhibitors of P38 MAP kinase,[62,63] an enzyme involved in the regulation of inflammation. Subsequently, it was shown that sorafenib inhibits Raf-1 by binding to an inactive conformation of the enzyme, a mechanism of action that has also been observed for other kinase inhibitors.[64,65]

V. PRACTICAL APPLICATION, RECENT EXAMPLE

A. IKK inhibitors

The examples above represent successful exploitation of leads generated from various screening processes. Of necessity, since many of the drugs described have reached the market, they represent discovery processes that were in use a decade or more ago. An instructive example of more recently applied processes is illustrated in Figure 7.4. Screens against IKKβ, a kinase involved in the regulation of the gene transcription factor NF-κB, produced very similar, if not

FIGURE 7.3 Raf-1 kinase inhibitors.

identical lead structures for at least four different organizations, Boehringer Ingelheim,[66] AstraZeneca,[41,67] SmithKline Beecham[68] and Pharmacia[69] and two publications have appeared detailing the processes used to evolve the screening hit into lead series.

At AstraZeneca, the aminothiophenecarboxamide screening hit **28**, a commercially available compound, and the urea analog **29** were confirmed as a viable hit structures. Liabilities identified in progressing toward lead status were poor solubility and poor metabolic stability. Combining structural features of the two hits gave molecules with substantially improved molecular potency as in **30**. Incorporation of central ring heterocycles predicted to improve solubility, as in compounds **31** and **32**, instead gave substantially decreased potency. Although the solubility profile could not be significantly improved, representative final molecules showed acceptable *in vivo* exposure, and poor solubility was ultimately not an impediment to achieving LO status. Substitution on the amide nitrogen abolished potency, SAR information that was consistent with a binding model in which this group engages a key hinge-binding interaction with the target enzyme. Additional SAR studies focused on improving metabolic stability and identified electron withdrawing substitution on the phenyl ring that conferred an acceptable profile, exemplified by the *p*-fluorophenyl derivative **33**. Overall this molecule fulfilled all lead criteria aside from acceptable solubility, but since the solubility profile did not adversely affect oral bioavailability the series progressed to LO.

Screening at Boehringer Ingelheim identified the same hit **28** along with the fluorophenyl analog **34**. Criteria applied to the selection of these compounds as validated hits included selectivity, drug-likeness, tractability of syntheses and support for interaction with the target. Bicyclic analogs (**35**) and (**36**) clarified the important hydrogen bonding interactions with the hinge region of the kinase and, along with additional analogs, provided sufficient information to construct a useful pharmacophore model.[66] Although monocyclic scaffolds analogous to those examined by AstraZeneca were also made and tested, the hit-to-lead effort ultimately focused on bicyclic scaffolds exemplified by **37** as offering the best opportunities for exploring

FIGURE 7.4 IKK inhibitors.

the potential interactions identified by the pharmacophore model. Subsequent improvement in potency was achieved by substitution on this thienopyridine scaffold. Overall, this scaffold offered substantially improved metabolic stability and CYP-450 inhibition profiles over the original hit thiophene scaffold, along with a clearer IP position, and progressed to LO.

VI. CONCLUSION

Notwithstanding its relatively recent introduction, it was anticipated that HTS, in conjunction with combinatorial chemistry, would provide a large positive impact on drug discovery productivity and perhaps even identify potential drug candidates directly from large libraries of compounds.

TABLE 7.3 The Impact of High Throughput Screening

*Patent**	1985–1990	1990–1995	1995–2001
Drugs	Nevirapine	Tirofiban	Gefitinib
	Delavirdine	Bosentan	Cinacalcet
	Mozavaptan	Efavirenz	Sorafenib
	Sivelestat	Conivaptan	Dasatinib
		Tipranavir	Sitagliptin
		Sitaxentan	Sunitinib
		Aprepitant	Ambrisentan
			Maraviroc
Targets	HIV-RT, elastase	HIV-protease, ET_A, NK-1,	DPP-IV, Kinases,
	Vasopressin V2	Fibrinogen receptor	Ca^{2+} receptor, CCR5
First in class	1	3	6
Technologies	Membrane preparations	Fluorescence	Cell-based fluorescence
	Radioisotopic detection		
	Colorimetric detection		
Numbers	25,000 @ 200 week	100,000 as mixtures	200,000
Lead sources	Corporate historical	Drug	Compound acquisition

*Priority date.

However, through the 1990s only a small number of approved drugs originated in screening processes. For example, most drugs approved in 2000 derived from analog-based discovery approaches,[70] a situation that was not significantly different from that described earlier by Freter.[5]

Table 7.3 provides an alternative perspective on the list of HTS-derived drugs presented in Table 7.1. The timeline here is focused on drug discovery rather than patent approval date and additional target and technology information is included. It is clear that the many of the targets that have been successfully addressed by lead identification through screening approaches have provided first-in-class drug candidates. This compilation does not reflect the capability of current screening technologies and does not incorporate the impact of such initiatives as the recently implemented "New Pathways to Discovery" component of the National Institutes of Health (NIH) roadmap. This initiative which will expand access to HTS capacity beyond pharmaceutical companies to basic research institutions and is expected to impact not only the availability of tool compounds but also the delivery of therapies for rare diseases which are often not attractive to the private sector. Additionally, while further increases in screening capacity against individual targets can be expected, high-content screening which quantifies cellular and subcellular events through fluorescence microscopy and image analysis is also progressing into the high throughput world and promises to further expand the quantity and quality of information that will be available to assist lead identification.[71]

In the past, natural products, receptor agonizts, enzyme substrates, and literature compounds were the main sources of starting points for drug discovery campaigns. However for many biological targets the relevant enzyme substrates or receptor ligands are not attractive starting points for medicinal chemistry and, in these cases, HTS is proving to be a key resource for the generation of novel tractable lead series.[72]

REFERENCES

1. Lombardino, J. G., Lowe, J. A. The role of medicinal chemistry in drug discovery – then and now. *Nat. Rev. Drug Discov.* 2004, *3*, 853–862.
2. Drews, J. Drug discovery: a historical perspective. *Science* 2000, *287*, 1960–1963.
3. For example, during the first four years of the discovery campaign that led to cimetidine, some 200 compounds were synthesized and assessed. Ganellin, C. R. Discovery of the antiulcer drug tagamet. In *Drug Discovery and Development* (Chorghade, M. S., Ed.), John Wiley & Sons, Inc, Hoboken, New Jersey, Vol. 1, p. 300.
4. Smith, W. G. Pharmacological screening tests. *Progr. Med. Chem.* 1964, *1*, 1–33.
5. Freter, K. R. Drug discovery – today and tomorrow: the role of medicinal chemistry. *Pharmaceut. Res.* 1988, *5*, 397–400.
6. Sneader, W. S. *Drug Prototypes and Their Exploitation*. John Wiley & Sons: Chichester, 1996. p. 9
7. Sams-Dodd, F. Drug discovery: selecting the optimal approach. *Drug Discov. Today* 2006, *11*, 465–472.
8. Von Wartburg, A., Traber, R. Cyclosporins, fungal metabolites with immunosuppressive activities. *Progr. Med. Chem.* 1988, *25*, 1–33.
9. Takahashi, N., Hayano, T., Suzuki, M. Peptidyl-prolyl cis-trans isomerase is the cyclosporin A-binding protein cyclophilin. *Nature* 1989, *337*, 473–475.
10. Sneader, W. S. *Drug Prototypes and Their Exploitation*. John Wiley & Sons: Chichester, 1996. p. 39.
11. Shoemaker, R. H., Scudiero, D. A., Melillo, G., Currens, M. J., Monks, A. P., Rabow, A. A., Covell, D. G., Sausville, E. A. Application of high-throughput, molecular-targeted screening to anticancer drug discovery. *Curr. Top. Med. Chem.* 2002, *2*, 229–246.
12. Stearns, B., Losee, K. A., Bernstein, J. Hydroxyurea, a new type of potential antitumor agent. *J. Med. Chem.* 1963, *6*, 201.
13. Schabel, F. M., Jr, Johnston, T. P., McCaleb, G. S., Montgomery, J. A., Laster, W. R., Skipper, H. E. Experimental evaluation of potential anticancer agents. VIII. Effects of certain nitrosoureas on intracerebral L1210 leukemia. *Cancer Res.* 1963, *23*, 725–733.
14. Taylor, J. D., Wykes, A. A., Gladish, Y. C., Martin, W. B. New inhibitor of monoamine oxidase. *Nature* 1960, *187*, 941.

15. Hitchings, G. H. Chemotherapy and comparative biochemistry: G. H. A. Clowes memorial lecture. *Cancer Res.* 1969, *29*, 1895–1903.
16. Lipinski, C. A., Lombardo, F., Dominy, B. W., Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 1997, *23*, 3–25.
17. Lipinski, C. A. Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharmacol. Toxicol. Meth.* 2001, *44*, 235–249.
18. Wenlock, M. C., Austin, R. P., Barton, P., Davis, A. M., Leeson, P. D. A comparison of physiochemical property profiles of development and marketed oral drugs. *J. Med. Chem.* 2003, *46*, 1250–1256.
19. Teague, S. J., Davis, A. M., Leeson, P. D., Oprea, T. The design of leadlike combinatorial libraries. *Angew. Chem. Int. Ed.* 1999, *38*, 3742–3748.
20. Hann, M. M., Leach, A. R., Harper, G. Molecular complexity and the impact on finding leads for drug discovery. *J. Chem. Inform. Comput. Sci.* 2001, *41*, 856–864.
21. Oprea, T. Current trends in lead discovery: Are we looking for the appropriate properties? *J. Comput.-Aided Mol. Des.* 2002, *16*, 325–334.
22. Michne, W. F. Hit-to-lead chemistry: A key element in new lead generation. *Pharmaceut. News* 1996, *3*, 19–21.
23. Gribbon, P., Sewing, A. High-throughput drug discovery: What can we expect from HTS?. *Drug Discov. Today* 2005, *10*, 17–22.
24. Bleicher, K. H., Böhm, H. J., Müller, K., Alanine, A. I. A guide to drug discovery: Hit and lead generation: beyond high-throughput screening. *Nat. Rev. Drug Discov.* 2003, *2*, 369–378.
25. Harper, G., Pickett, S. D., Green, D. V. S. Design of a compound screening selection for use in high-throughput screening. *Combin. Chem. High Throughput Screening* 2004, *7*, 63–71.
26. Wintner, E. A., Moallemi, C. C. Quantized surface complementarity diversity (QSCD): a model based small molecule-target complementarity. *J. Med. Chem.* 2000, *43*, 1993–2006.
27. Lane, S. J., Eggleston, D. S., Brinded, K. A., Hollerton, J. C., Taylor, N. L., Readshaw, S. A. Defining and maintaining a high quality screening collection: the GSK experience. *Drug Discov. Today* 2006, *11*, 267–272.
28. Davis, A. M., Keeling, D. J., Steele, J., Tomkinson, N. P., Tinker, A. C. Components of successful lead generation. *Curr. Top. Med. Chem.* 2005, *5*, 421–439.
29. Richardson, R. Compound collections: acquisition, annotation and access. In *Exploiting Chemical Diversity for Drug Discovery* (Bartlett, P. A., Entzeroth, M., Eds). The Royal Society of Chemistry: Cambridge, 2006, pp. 113–136.
30. Anon, Redesigning drug discovery. *Nature* 1996, *384*, 1–5. Supplement.
31. Zhang, J-H., Chung, T. D. Y., Oldenburg, K. R. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J. Biomol. Screen* 1999, *4*, 67–73.
32. Rishton, G. M. Nonleadlikeness and leadlikeness in biochemical screening. *Drug Discov. Today* 2002, *8*, 86–96.
33. Pearce, B. C., Sofia, M. J., Good, A. C., Drexler, D. M., Stock, D. A. An empirical process for the design of high-throughput screening deck filters. *J. Chem. Inform. Model* 2006, *46*, 1060–1068.
34. Hann, M., Hudson, B., Lewell, X., Lifely, R., Miller, L., Ramsden, N. Strategic pooling of compounds for high-throughput screening. *J. Chem. Inform. Comp. Sci.* 1999, *39*, 897–902.
35. Soichet, B. K. Screening in a spirit haunted world. *Drug Discov. Today* 2006, *11*, 607–615.
36. McGovern, S. L., Caselli, E., Grigorieff, N., Shoichet, B. K. A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *J. Med. Chem.* 2002, *45*, 1712–1722.
37. Roche, O., Schneider, P., Zuegge, J., Guba, W., Kansy, M., Alanine, A., Bleicher, K., Danel, F., Gutknecht, E.-M., Rogers-Evans, M., Neidhart, W., Stalder, H., Dillon, M., Sjoegren, E., Fotouhi, N., Gillespie, P., Goodnow, R., Harris, W., Jones, P., Taniguchi, M., Tsujii, S., von Saal, W., Zimmermann, G., Schneider, G. Development of a virtual screening method for identification of "frequent hitters" in compound libraries. *J. Med. Chem.* 2002, *45*, 137–142.
38. Macarron, R. Critical review of the role of HTS in drug discovery. *Drug Discov. Today* 2006, *11*, 277–279.
39. Steinmeyer, A. The hit-to-lead process at Schering AG: strategic aspects. *ChemMedChem* 2006, *1*, 31–36.
40. Wunberg, T., Hendrix, M., Hillisch, A., Lobell, M., Meier, H., Schmeck, C., Wild, H., Hinzen, B. Improving the hit-to-lead process: data-driven assessment of drug-like and lead-like screening hits. *Drug Discov. Today* 2006, *11*, 175–180.
41. Baxter, A., Brough, S., Cooper, A., Floettmann, E., Foster, S., Harding, C., Kettle, J., McInally, T., Martin, C., Mobbs, M., Needham, M., Newham, P., Paine, S., St-Gallay, S., Salter, S., Unitt, J., Xue, Y. Hit-to-lead studies: the discovery of potent, orally active, thiophenecarboxamide IKK-2 inhibitors. *Bioorg. Med. Chem. Lett.* 2004, *14*, 2817–2822.
42. Oefner, C., Binggeli, A., Breu, V., Bur, D., Clozel, J-P., D'Arcy, A., Dorn, A., Fischli, W., Grüninger, F., Güller G. Hirth, R., Märki, H. P., Mathews, S., Müller, M., Ridley, R. G., Stadier, H., Vieira, E., Wilhelm, M., Winkler, F. K., Wostl, W. Renin inhibition by substituted piperidines: A novel paradigm for the inhibition of monomeric aspartic proteinases? *Chem. Biol.* 1999, *6*, 127–131.
43. Britton, T. C., Trepanier, D. L. 4-Hydroxy-3,4-dihydro-2(1H)-quinazolinethiones. US 4,093,805 (1978 to Dow Chemical Co.).
44. Tucker, T. J., Lumma, W. C., Culberson, J. C. Development of non-nucleoside HIV reverse transcriptase inhibitors. *Meth. Enzymol.* 1996, *275*, 440–473.
45. McCall, J. M., Ayer, D. E., Jacobsen, E. J., Van Doornik, F. J., Palmer, J. R. Preparation of pharmaceutically active heterocyclic amines and their use for treating head injury, spinal trauma, stroke, etc. WO8808424 (1988 Upjohn Co.).
46. Hargrave, K. D., Proudfoot, J. R., Grozinger, K. G., Cullen, E., Kapadia, S. R., Patel, U. R., Fuchs, V. U., Mauldin, S. C., Vitous, J., Behnke, M. L., Klunder, J. M., Pal, K., Skiles, J. W., McNeil, D. W., Rose, J. M., Chow, G. C., Skoog, M. T., Wu, J. C., Schmidt, G., Engel, W. W., Eberlein, W. G., Saboe, T. D., Campbell, S. J., Rosenthal, A. S., Adams, J. Novel non-nucleoside inhibitors of HIV-1 reverse transcriptase.1. Tricyclic pyridobenzo- and dipyridodiazepinones. *J. Med. Chem.* 1991, *34*, 2231–2241.
47. Proudfoot, J. R., Hargrave, K. D., Kapadia, S. R., Patel, U. R., Grozinger, K. G., McNeil, D. W., Cullen, E., Cardozo, M., Tong, L., Kelly, T. A., Rose, J., David, E., Mauldin, S. C., Fuchs, V. U., Vitous, J., Hoermann, M., Klunder, J. M., Raghavan, P., Skiles, J. W., Mui, P., Richman, D. D., Sullivan, J. L., Shih, C.-K., Grob, P., Adams, J. Novel non-nucleoside inhibitors of HIV-1 reverse transcriptase. 4. 2-Substituted dipyridodiazepinones are potent inhibitors of both wild type and cysteine-181 HIV-1 reverse transcriptase enzymes. *J. Med. Chem.* 1995, *38*, 4830–4838.
48. Cywin, C. L., Klunder, J. M., Hoermann, M., Brickwood, J. R., David, E., Grob, P. M., Schwartz, R., Pauletti, D., Barringer, K. J., Shih, C.-K., Sorge, C. L., Erickson, D. A., Joseph, D. P., Hattox, S. E. Novel nonnucleoside inhibitors of HIV-1 reverse transcriptase. 8. 8-Aryloxymethyl- and 8-arylthiomethyldipyridodiazepinones. *J. Med. Chem.* 1998, *41*, 2972–2984.
49. Chan, M. F., Okun, I., Stavros, F. L., Hwang, E. *Biochem. Biophys. Res. Comm.* 1994, *201*, 228–234.
50. Chan, M. F., Raju, B., Kois, A., Castillo, R. S., Verner, E. J., Wu, C., Hwang, E., Okun, I., Stavros, F., Balaji, V. N. Halogen substitution at the isoxazole ring enhances the activity of *N*-(isoxazolyl)sulfonamide endothelin antagonists. *Bioorg. Med. Chem. Lett.* 1996, *6*, 2393–2398.
51. Wu, C., Chan, M. F., Stavros, F., Raju, B., Okun, I., Castillo, R. S. Structure-activity relationships of N^2-Aryl-3-(isoxazolylsulfamoyl)-2-thiophenecarboxamides as selective endothelin receptor-A antagonists. *J. Med. Chem.* 1997, *40*, 1682–1689.
52. Wu, C., Chan, M. F., Stavros, F., Raju, B., Okun, I., Mong, S., Keller, K. M., Brock, T., Kogan, T. P., Dixon, R. A. F. Discovery of

TBC11251, a potent, long acting, orally active endothelin receptor-A selective antagonist. *J. Med. Chem.* 1997, *40*, 1690–1697.

53. Stein, P. D., Hunt, J. T., Floyd, D. M., Moreland, S., Dickinson, K. E. J., Mitchell, C., Liu, E. C. K., Webb, M. L., Murugesan, N., Dickey, J., McMullen, D., Zhang, R., Lee, V. G., Serdino, R., Delaney, C., Schaeffer, T. R., Kozlowskil, M. The discovery of sulfonamide endothelin antagonists and the development of the orally active ET_A antagonist 5-(dimethylamino)-*N*-(3,4-dimethyl-5- isoxazolyl)-1-naphthalenesulfonamide. *J. Med. Chem.* 1994, *37*, 329–331.
54. Murugesan, N., Gu, Z., Stein, P. D., Bisaha, S., Spergel, S., Girotra, R., Lee, V. G., Lloyd, J., Misra, R. N., Schmidt, J., Mathur, A., Stratton, L., Kelly, Y. F., Bird, E., Waldron, T., Liu, E. C.-K., Zhang, R., Lee, H., Serafino, R., Abboa-Offei, B., Mathers, P., Giancarli, M., Seymour, A. A., Webb, M. L., Moreland, S., Barrish, J. C., Hunt, J. T. Biphenylsulfonamide endothelin antagonists: structure-activity relationships of a series of mono- and disubstituted analogues and pharmacology of the orally active endothelin antagonist 2'-Amino-*N*-(3,4-dimethyl-5-isoxazolyl)-4'-(2-methylpropyl)[1,1'-biphenyl]-2-sulfonamide (BMS-187308). *J. Med. Chem.* 1998, *41*, 5198–5218.
55. Murugesan, N., Gu, Z., Stein, P. D., Spergel, S., Mathur, A., Leith, L., Liu, E. C.-K., Zhang, R., Bird, E., Waldron, T., Marino, A., Morrison, R. A., Webb, M. L., Moreland, S., Barrish, J. C. Biphenylsulfonamide endothelin receptor antagonists. 2. Discovery of 4'-oxazolyl biphenylsulfonamides as a new class of potent, highly selective ET_A antagonists. *J. Med. Chem.* 2000, *43*, 3111–3117.
56. Murugesan, N., Gu, Z., Stein, P. D., Spergel, S., Bisaha, S., Liu, E. C.-K., Zhang, R., Webb, M. L., Moreland, S., Barrish, J. C. Biphenylsulfonamide endothelin receptor antagonists. Part 3: Structure–activity relationship of 4'-heterocyclic biphenylsulfonamides. *Bioorg. Med. Chem. Lett.* 2002, *22*, 517–520.
57. Murugesan, N., Gu, Z., Spergel, S., Young, M., Chen, P., Mathur, A., Leith, L., Hermsmeier, M., Liu, E. C.-K., Zhang, R., Bird, E., Waldron, T., Marino, A., Koplowitz, B., Humphreys, W. G., Chong, S., Morrison, R. A., Webb, M. L., Moreland, S., Trippodo, N., Barrish, J. C. Biphenylsulfonamide endothelin receptor antagonists. 4. Discovery of *N*-[[2'-[[(4,5-dimethyl-3-isoxazolyl)amino]sulfonyl]-4-(2-oxazolyl)[1,1'-biphenyl]-2-yl]methyl]-*N*,3,3-trimethylbutanamide (BMS-207940), a highly potent and orally active ET_A selective antagonist. *J. Med. Chem.* 2003, *46*, 125–137.
58. Neidhart, W., Breu, V., Bur, D., Burri, K., Clozel, M., Hirth, G., Mueller, M., Wessel, H. P., Ramuz, H. The discovery of nonpeptide endothelin receptor antagonists. Progression towards bosentan. *Chimia* 1996, *50*, 519–524.
59. Grussner, A., Lorch, E., Schnider, O. 4-Benzenesulfamidopyrimidines. US 3,317,536 (May 2 1967 to Hoffmann La Roche).
60. Riechers, H., Albrecht, H.-P., Aberg, W., Baumann, E., Bernard, H., Bohm, H.-J., Klinge, D., Kling, A., Muller, S., Raschack, M., Unger, L., Walker, N., Wernet, W. Discovery and optimization of a novel class of orally active nonpeptidic Endothelin-A receptor antagonists. *J. Med. Chem.* 1996, *39*, 2123–2128.
61. Lowinger, T. B., Riedl, B., Dumas, J., Smith, R. A. Design and discovery of small molecules targeting Raf-1 kinase. *Curr. Pharmaceut. Des.* 2002, *8*, 2269–2278.
62. Dumas, J., Sibley, R., Riedl, B., Monahan, M. K., Lee, W., Lowinger, T. B., Redman, A. M., Johnson, J. S., Kingery-Wood, J., Scott, W. J., Smith, R. A., Bobko, M., Schoenleber, R., Ranges, G. E., Housley, T. J., Bhargava, A., Wilhelm, S. M., Shrikhande, A. Discovery of a new class of p38 kinase inhibitors. *Bioorg. Med. Chem. Lett.* 2000, *10*, 2047–2050.
63. Regan, J., Breitfelder, S., Cirillo, P., Gilmore, T., Graham, A. G., Hickey, E., Klaus, B., Madwed, J., Moriak, M., Moss, N., Pargellis, C., Pav, S., Proto, A., Swinamer, A., Tong, L., Torcellini, A. *J. Med. Chem.* 2002, *45*, 2994–3008.
64. Liu, Y., Gray, N. S. Rational design of inhibitors that bind to inactive kinase conformations. *Nat. Chem. Biol.* 2006, *2*, 358–364.
65. Wan, P. T. C., Garnett, M. J., Roe, S. M., Lee, S., Niculescu-Duvaz, D., Good, V. M., Jones, M. C., Marshall, C. J., Springer, C. J., Barford, D., Marais, R. Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF. *Cell*. 2004, *116*, 855–867.
66. Morwick, T., Berry, A., Brickwood, J., Cardozo, M., Catron, K., DeTuri, M., Emeigh, J., Homon, C., Hrapchak, M., Jacober, S., Jakes, S., Kaplita, P., Kelly, T. A., Ksiazek, J., Liuzzi, M., Magolda, R., Mao, C., Marshall, M., McNeil, D., Prokopowicz, A., Sarko, C., Scouten, E., Sledziona, C., Sun, S., Watrous, J., Wu, J. P., Cywin, C. L. Evolution of the thienopyridine class of inhibitors of IκB kinase beta: Part I: Hit-to-Lead strategies. *J. Med. Chem.* 2006, *49*, 2898–2908.
67. Baxter, A., Brough, S., Faull, A., Johnstone, C., Mcinally, T. Preparation of thiophenecarboxamides as inhibitors of the enzyme IKK-2 WO2001058890 (Astrazeneca).
68. Wan, Z., Burgess, J. L., Callahan, J. F. Preparation of aminothiophenes as NF-KB inhibitors. WO2003029242 (2003 Smithkline Beecham Corporation, USA). Parrish, C. A., Callahan, J. F., Wan, Z., Burgess, J. L., Stavenger, R. A., Holt, D. A. Preparation of 3-ureidothiophenes as angiogenesis and Chk1 kinase inhibitors for treating various forms of cancer and hyperproliferative disorders. WO2003028731 (Smithkline Beecham Corporation).
69. (a) Kishore, N., Sommers, C., Mathialagan, S., Guzova, J., Yao, M., Hauser, S., Huynh, K., Bonar, S., Mielke, C., Albee, L., Weier, R., Graneto, M., Hanau, C., Perry, T., Tripp, C. S. A selective IKK-2 inhibitor blocks NF-B-dependent gene expression in Interleukin-1-stimulated synovial fibroblasts. *J. Biol. Chem.* 2003, *278*, 32861–32871. (b) Graneto, M., Hanau, C. E., Perry, T. D., Heteroaromatic carboxamide derivatives, particularly 3-aminothiophene-2-carboxamides, useful as protein kinase inhibitors, for the treatment of cancer. inflammation, and inflammation-related disorders. (Pharmacia Corporation. USA). WO2003037886. (c) Hagen, T. J., Weier, R. M., Xu, X., Houdek, S. C., Clare, M. Preparation of substituted thiophene carboxamide compounds for the treatment of inflammation. (Pharmacia Corporation, USA). WO2004009582.
70. Proudfoot, J. R. Drugs, leads, and drug-likeness: an analysis of some recently launched drugs. *Bioorg. Med. Chem. Lett.* 2002, *12*, 1647–1650.
71. Haney, S. A., LaPan, P., Pan, J., Zhang, J. High content screening moves to the front of the line. *Drug Discov. Today.* 2006, *11*, 889–894.
72. Schnecke, V., Bostrom, J. Computational chemistry driven decision making in lead generation. *Drug Discov. Today.* 2006, *11*, 43.
73. Merluzzi, V. J., Hargrave, K. D., Labadia, M., Grozinger, K., Skoog, M., Wu, J. C., Shih, C. K., Eckner, K., Hattox, S., Adams, J., Rosenthal, A. S. Inhibition of HIV-1 replication by a nonnucleoside reverse transcriptase inhibitor. *Science.* 1990, *250*, 1411–1413.
74. Romero, D. L., Morge, R. A., Biles, C., Berrios-Pena, N., May, P. D., Palmer, J. R., Johnson, P. D., Smith, H. W., Busso, M., Tan, C.-K., Voorman, R. L., Reusser, F., Althaus, I. W., Downey, K. M., So, A. G., Resnick, L., Tarpley, W. G., Aristoff, P. A. Discovery, synthesis, and bioactivity of bis(heteroaryl)piperazines. 1. A novel class of non-nucleoside HIV-1 reverse transcriptase inhibitors. *J. Med. Chem.* 1994, *37*, 999–1014.
75. Tucker, T. J., Lyle, T. A., Wiscount, C. M., Britcher, S. F., Young, S. D., Sanders, W. M., Lumma, W. C., Goldman, M. E., O'Brien, J. A., Ball, R. G., Homnick, C. F., Schleif, W. A., Emini, E. A., Huff, J. R., Anderson, P. A. Synthesis of a series of 4-(arylethynyl)-6-chloro-4-cyclopropyl-3,4-dihydroquinazolin-2(1 H)-ones as novel non-nucleoside HIV-1 reverse transcriptase inhibitors. *J. Med. Chem.* 1994, *37*, 2437–2444.
76. Hartman, G. D., Egbertson, M. S., Halczenko, W., Laswell, W. L., Duggan, M. E., Smith, R. L., Naylor, A. M., Manno, P. D., Lynch, R. J., Zhang, G., Chang, C. T.-C., Gould, R. Non-peptide fibrinogen receptor antagonists. 1. Discovery and design of exosite inhibitors. *J. Med. Chem.* 1992, *35*, 4640–4642.
77. Ward, W. H. J., Cook, P. N., Slater, A. M., Davies, D. H., Holdgate, G. A., Green, L. R. Epidermal growth factor receptor tyrosine kinase: investigation of catalytic mechanism, structure-based

searching and discovery of a potent inhibitor. *Biochem. Pharmacol.* 1994, *48*, 659–666.

78. Imaki, K., Okada, T., Nakayama, Y., Nagao, Y., Kobayashi, K., Sakai, Y., Mohri, T., Amino, T., Nakai, H., Kawamura, M. Non-peptidic inhibitors of human neutrophil elastase: the design and synthesis of sulfonanilide-containing inhibitors. *Bioorg. Med. Chem.* 1996, *4*, 2115–2134.
79. Swain, C. J., Cascieri, M. A., Owens, A., Saari, W., Sadowski, S., Strader, C., Teall, M., Van Niel, M. B., Williams, B. J. Acyclic NK1 antagonists: replacements for the benzhydryl group. *Bioorg. Med. Chem. Lett.* 1994, *4*, 2161–2164.
80. Nemeth, E. F., Steffey, M. E., Hammerland, L. G., Hung, B. C. P., Van Wagenen, B. C., Delmar, E. G., Balandrin, M. F. Calcimimetics with potent and selective activity on the parathyroid calcium receptor. *Proc. Natl. Acad. Sci. USA.* 1998, *95*, 4040–4045.
81. Smith, R. A., Barbosa, J., Blum, C. L., Bobko, M. A., Caringal, Y. V., Dally, R., Johnson, J. S., Katz, M. E., Kennure, N., Kingery-Wood, J., Lee, W., Lowinger, T. B., Lyons, J., Marsh, V., Rogers, D. H., Swartz, S., Walling, T., Wild, H. Discovery of heterocyclic ureas as a new class of raf kinase inhibitors: identification of a second generation lead by a combinatorial chemistry approach. *Bioorg. Med. Chem. Lett.* 2001, *11*, 2775–2778.
82. Skulnick, H. I., Johnson, P. D., Howe, W. J., Tomich, P. K., Chong, K.-T., Watenpaugh, K. D., Janakiraman, M. N., Dolak, L. A., McGrath, J. P., Lynn, J. C., Horng, M.-M., Hinshaw, R. R., Zipp, G. L., Ruwart, M. J., Schwende, F. J., Zhong, W. Z., Padbury, G. E., Dalga, R. J., Shiou, L., Possert, P. L., Rush, B. D., Wilkinson, K. F., Howard, G. M., Toth, L. N., Williams, M. G., Kakuk, T. J., Cole, S. L., Zaya, R. M., Lovasz, K. D., Morris, J. K., Romines, K. R., Thaisrivongs, S., Aristoff, P. A. Structure-based design of sulfonamide-substituted non-peptidic HIV protease inhibitors. *J. Med. Chem.* 1995, *38*, 4968–4971.
83. Yamamura, Y., Ogawa, H., Chihara, T., Kondo, K., Onogawa, T., Nakamura, S., Mori, T., Tominaga, M., Yabuuchi, Y. OPC-21268, an orally effective, nonpeptide vasopressin V1 receptor antagonist. *Science* 1991, *252*, 572–574.
84. Matsuhisa, A., Taniguchi, N., Koshio, H., Yatsu, T., Tanaka, A. Nonpeptide arginine vasopressin antagonists for both V1A and V2 receptors: synthesis and pharmacological properties of 4′-(1,4,5,6-tetrahydroimidazo[4,5-d][1]benzazepine-6-carbonyl)benzanilide derivatives and 4’-(5,6-dihydro-4H-thiazolo[5,4-d][1]benzazepine-6-carbonyl)benzanilide derivative. *Chem. Pharmaceut. Bull.* 2000, *48*, 21–31.
85. Ogawa, H., Yamamura, Y., Miyamoto, H., Kondo, K., Yamashita, H., Nakaya, K., Chihara, T., Mori, T., Tominaga, M., Yabuuchi, Y. Orally active, nonpeptide vasopressin V1 antagonists. A novel series of 1-(1-substituted 4-piperidyl)-3,4-dihydro-2(1H)-quinolinones. *J. Med. Chem.* 1993, *36*, 2011–2017.
86. Sun, L., Tran, N., Tang, F., App, H., Hirth, P., McMahon, G., Tang, C. Synthesis and biological evaluations of 3-substituted indolin-2-ones: a novel class of tyrosine kinase inhibitors that exhibit selectivity toward particular receptor tyrosine kinases. *J. Med. Chem.* 1998, *41*, 2588–2603.
87. Lombardo, L. J., Lee, F. Y., Chen, P., Norris, D., Barrish, J. C., Behnia, K., Castaneda, S., Cornelius, L. A. M., Das, J., Doweyko, A. M., Fairchild, C., Hunt, J. T., Inigo, I., Johnston, K., Kamath, A., Kan, D., Klei, H., Marathe, P., Pang, S., Peterson, R., Pitt, S., Schieven, G. L., Schmidt, R. J., Tokarski, J., Wen, M.-L., Wityak, J., Borzilleri, R. M. Discovery of *N*-(2-chloro-6-methyl-phenyl)-2-(6-(4-(2-hydroxyethyl)-piperazin-1-yl)-2-methylpyrimidin-4-ylamino)thiazole-5-carboxamide (BMS-354825), a dual Src/Abl kinase inhibitor with potent antitumor activity in preclinical assays. *J. Med. Chem.* 2004, *47*, 6658–6661.
88. Xu, J., Ok, H. O., Gonzalez, E. J., Colwell, L. F., Jr, Habulihaz, B., He, H., Leiting, B., Lyons, K. A., Marsilio, F., Patel, R. A., Wu, J. K., Thornberry, N. A., Weber, A. E., Parmee, E. R. Discovery of potent and selective β-homophenylalanine based dipeptidyl peptidase IV inhibitors. *Bioorg. Med. Chem. Lett.* 2004, *14*, 4759–4762.
89. Wood, A., Armour, D. The discovery of the CCR5 receptor antagonist, UK-427,857, a new agent for the treatment of HIV infection and AIDS. *Progr. Med. Chem.* 2005, *43*, 239–271.

Chapter 8

Natural Products as Pharmaceuticals and Sources for Lead Structures

David J. Newman, Gordon M. Cragg and David G. I. Kingston[1]

I. INTRODUCTION
II. THE IMPORTANCE OF NATURAL PRODUCTS IN DRUG DISCOVERY AND DEVELOPMENT
A. The origin of natural products
B. The uniqueness of the natural products approach
C. The impact of new screening methods
III. THE DESIGN OF AN EFFECTIVE NATURAL-PRODUCTS-BASED APPROACH TO DRUG DISCOVERY
A. Acquisition of biomass
B. The unexplored potential of microbial diversity
C. Extraction
D. Screening methods
E. Isolation of active compounds
F. Structure elucidation
G. Further biological assessment
H. Procurement of large-scale supplies
I. Determination of structure–activity relationships
IV. EXAMPLES OF NATURAL PRODUCTS OR ANALOGS AS DRUGS
A. Antihypertensives
B. Anticholesterolemics
C. Immunosuppressives
D. Antibiotics
E. Microbial anticancer agents
F. Anticancer agents from plants
G. Anticancer agents from marine organisms
H. Antimalarial agents
I. Other natural products
V. FUTURE DIRECTIONS IN NATURAL PRODUCTS AS DRUGS AND DRUG DESIGN TEMPLATES
A. Introduction
B. Combinatorial chemistry
C. Natural products as design templates
D. Interactions of microbial sources, genomics, and synthetic chemistry
VI. SUMMARY
REFERENCES

Accuse not Nature, she hath done her part; do thou but thine

Milton, Paradise Lost

I. INTRODUCTION

Throughout the ages humans have relied on Nature to cater for their basic needs, not the least of which are medicines for the treatment of a myriad of diseases. Plants, in particular, have formed the basis of sophisticated traditional medicine systems, with the earliest records documenting the uses of approximately 1,000 plant-derived substances in Mesopotamia, and the "Ebers Papyrus" dating from 1500 BCE, documenting over 700 drugs, mostly of plant origin.[1] The first record of the Chinese Materia Medica documenting 52 prescriptions dates from about 1100 BCE, and was followed by works such as the Shennong Herbal (~100 BCE; 365 drugs) and the Tang Herbal (659 CE; 850 drugs).[2] Documentation of the Indian Ayurvedic system also dates from before 1000 BCE (Charaka; Sushruta and Samhitas with 341 and 516 drugs, respectively).[3,4]

The Greeks and Romans contributed substantially to the rational development of the use of herbal drugs in the ancient Western world. Dioscorides, a Greek physician

[1]Note: This chapter reflects the opinions of the authors, not necessarily those of the US Government.

Copyright © 2008, Elsevier Ltd
All rights reserved.

(100 CE), accurately recorded the collection, storage, and use of medicinal herbs during his travels with Roman armies throughout the then "known world," whilst Galen (130–200 CE), a practitioner and teacher of pharmacy and medicine in Rome, is well known for his complex prescriptions and formulae used in compounding drugs. However, it was the Arabs who preserved much of the Greco-Roman expertise during the Dark and Middle Ages (5th–12th centuries), and who expanded it to include the use of their own resources, together with Chinese and Indian herbs unknown to the Greco-Roman world. A comprehensive review of the history of medicine may be found on the web site of the National Library of Medicine (NLM), US National Institutes of Health (NIH), at www.nlm.nih.gov/hmd/medieval/arabic.html.

Plant-based systems continue to play an essential role in healthcare, and their use by different cultures has been extensively documented.[5,6] It has been estimated by the World Health Organization that approximately 80% of the world's inhabitants rely mainly on traditional medicines for their primary healthcare, while plant products also play an important role in the healthcare systems of the remaining 20% of the population, mainly residing in developed countries.[7]

II. THE IMPORTANCE OF NATURAL PRODUCTS IN DRUG DISCOVERY AND DEVELOPMENT

The continuing value of natural products as sources of potential chemotherapeutic agents has been reviewed by Newman and Cragg.[8] An analysis of the sources of new drugs over the period January 1981–June 2006 classified these compounds as N (an unmodified natural product), ND (a modified natural product), S (a synthetic compound with no natural product conception), S*, S*/NM (a synthetic compound with a natural product pharmacophore; /NM indicating competitive inhibition), and S/NM (a synthetic compound showing competitive inhibition of the natural product substrate). This analysis indicated that 66% of the 974 small molecule, new chemical entities (NCEs) are formally synthetic, but 17% correspond to synthetic molecules containing pharmacophores derived directly from natural products classified as S* and S*/NM. Furthermore, 12% are actually modeled on a natural product inhibitor of the molecular target of interest, or mimic (i.e. competitively inhibit) the endogenous substrate of the active site, such as adenosine triphosphate (ATP) (S/NM). Thus, only 37% of the 974 NCEs can be classified as truly synthetic (i.e. devoid of natural inspiration) in origin (S) (Figure 8.1). In the area of anti-infectives (anti-bacterial, -fungal, -parasitic, and -viral), close to 70% are naturally derived or inspired (N; ND; S*; S*/NM; S/NM), while in the cancer treatment area 77.8% are in this category, with the figure being 63% if the S/NM category is excluded.

FIGURE 8.1 Sources of drugs.

In recent years, a steady decline in the output of the R&D programs of the pharmaceutical industry has been reported, with the number of new active substances, also known as NCEs, hitting a 20-year low of 37 in 2001.[9] Furthermore, this drop in productivity was reflected by the report that only 16 New Drug Applications had been received by the US Food and Drug Administration (FDA) in 2001, down from 24 the previous year. While various factors have been held to blame for this downturn, it is significant that the past 10–15 years has seen a decline in interest in natural products on the part of major pharmaceutical companies in favor of reliance on new chemical techniques, such as combinatorial chemistry, for generating molecular libraries. The realization that the number of NCEs in drug development pipelines is declining may have led to the rekindling of interest in "rediscovering natural products,"[10] as well as the heightened appreciation of the value of natural product-like models in "improving efficiency" in so-called diversity-oriented synthesis.[11] The urgent need for the discovery and development of new pharmaceuticals for the treatment of cancer, AIDS and infectious diseases, as well as a host of other diseases, demands that all approaches to drug discovery be exploited aggressively, and it is clear that nature has played, and will continue to play, a vital role in the drug discovery process. As stated by Berkowitz in 2003 commenting on natural products,[10]"We would not have the top-selling drug class today, the statins; the whole field of angiotensin antagonists and angiotensin-converting enzyme inhibitors; the whole area of immunosuppressives; nor most of the anticancer and antibacterial drugs. Imagine all of those drugs not being available to physicians or patients today." Or, as was eloquently stated by Danishefsky in 2002, "a small collection of smart compounds may be more valuable than a much larger hodgepodge collection mindlessly assembled."[12] Recently, he and a coauthor restated this theme in their review on the applications of total synthesis to problems in neurodegeneration: "We close with the hope and expectation that enterprising and hearty organic chemists will not pass up the unique head start that natural products provide in the quest for new agents and new directions in medicinal discovery. We would chance to predict that even as the currently fashionable

"telephone directory" mode of research is subjected to much overdue scrutiny and performance-based assessment, organic chemists in concert with biologists and even clinicians will be enjoying as well as exploiting the rich troves provided by nature's small molecules."[13]

A. The origin of natural products

While the contributions of natural secondary metabolites (all non-proteinaceous natural products would fall under this term) to modern medicine are abundantly clear, the question of their origins has long intrigued chemists and biochemists. Six major hypotheses have been proposed, and these have been well summarized by Haslam.[14] (1) They are simply waste products with no particular physiological role. (2) They are compounds that at one time had a functional metabolic role, which has now been lost. (3) They are products of random mutations, and have no real function in the organism. (4) They are an example of "evolution in progress," and provide a pool of compounds out of which new biochemical processes can emerge. (5) Production provides a way of enabling the enzymes of primary metabolism to function when they are not needed for their primary purpose. "It is the *processes* of secondary metabolism, rather than the *products* (secondary metabolites) which are important." (6) They play a key role in the organism's survival, providing defensive substances or other physiologically important compounds.

Although each of the above has (or has had) its supporters, Williams *et al.*[15] and Harborne[16] amongst others, argue convincingly that the weight of the evidence is behind the sixth hypothesis. Indeed, it seems reasonable to assume that, in many instances, the production of these complex and often toxic chemicals has evolved over eons as a means of chemical defense by essentially stationary organisms, such as plants and many marine invertebrates, against predation and consumption (e.g. herbivory). For instance, pupae of the coccinellid beetle, *Epilachna borealis*, appear to exert a chemical defensive mechanism against predators through the secretion of droplets from their glandular hairs containing a library of hundreds of large-ring (up to 98 members) macrocyclic polyamines.[17] These libraries are built up from three simple (2-hydroxyethylamino)-alkanoic acid precursors, and are clear evidence that combinatorial chemistry has been pioneered and widely used in nature for the synthesis of biologically active compound libraries. A further example is that of the venom composed of combinatorial libraries of several hundred peptides and injected by species of the cone snail genus, *Conus*, to stun their prey prior to capture.[18] One component of this mixture has been developed as Ziconotide, a non-narcotic analgesic that is currently marketed as Prialt®.[19]

Microorganisms are reported to kill sensitive strains of the same or related microbial species through excretion of antimicrobial toxins,[20] which resembles the process of allelopathy whereby plants release toxic compounds in order to suppress the growth of neighboring plants.[21,22] Bacteria also use a cell to cell "chemical language" as a signaling mechanism known as quorum sensing, involving the excretion of quorum-sensing compounds, to control their density of population growth and so-called biofilm formation. The best studied of these compounds, the acylhomoserine lactones (AHLs) exemplified by compounds such as *N*-3-oxohexanoyl-l-homoserine lactone (Figure 8.2) from *Vibrio fisheri,* and a furanone boronate diester that appears to be a universal signal (Figure 8.2) promoting the activation of genes promoting virulence, spore formation, biofilm formation, and other phenomena.[23,24] A solid-phase synthetic route, adaptable to the synthesis of combinatorial libraries of AHL analogs, has been developed, and two such analogs have been identified which inhibit the formation of biofilms of *Pseudomonas aeruginosa*, the organism responsible for lung infections in cystic fibrosis patients which can often prove fatal.[25]

FIGURE 8.2 Quorum-sensing compounds.

B. The uniqueness of the natural products approach

Natural products are generally complex chemical structures, whether they are cyclic peptides like cyclosporin A, or complex diterpenes like paclitaxel. Inspection of the structures that are discussed in Section IV is usually enough to convince any skeptic that few of them would have been discovered without application of natural products chemistry. Recognition and appreciation of the value of natural product-like models in "improving efficiency" in so-called diversity-oriented synthesis has already been mentioned.[11]

Structural diversity is not the only reason why natural products are of interest to drug development, since they often provide highly selective and specific biological activities based on mechanisms of action. Two very good examples of this are the β-hydroxy-β-methylglutary-CoA reductase (HMG-CoA reductase) inhibition exhibited by lovastatin, and the tubulin-assembly promotion activity of paclitaxel. These activities would not have been discovered without the natural product leads and investigation of their mechanisms of action. A striking illustration of the influence of natural products on many of the enzymatic processes operative in

FIGURE 8.3 Natural products and the cell cycle. *Source*: Modified from Meijer[26] (original used by permission from Springer-Verlag).

cell cycle progression may be found at the web site of the Roscoff Biological Station (http://www.sb-roscoff.fr/CyCell/Frames80.htm) which covers diagrams originally published by Meijer[26] on natural products and the cell cycle, with a modified version shown in Figure 8.3. The bioactivity of natural products stems from the previously discussed hypothesis that essentially all natural products have some receptor-binding activity; the problem is to find which receptor a given natural product is binding to. Viewed another way, a given organism provides the investigator with a complex library of unique bioactive constituents, analogous to the library of crude synthetic products initially produced by combinatorial chemistry techniques. The natural products approach can thus be seen as complementary to the synthetic approach, each providing access to (initially) different lead structures. In addition, development of an active natural product structure by combinatorially directed synthesis is an extremely powerful tool. The task of the natural products researcher is thus to select those compounds of pharmacological interest from the "natural combinatorial libraries" produced by extraction of organisms. Fortunately, the means to do this efficiently are now at hand.

C. The impact of new screening methods

In the early days of natural products research, new compounds were simply isolated at random, or at best by the use of simple broad-based bioactivity screens based on antimicrobial or cytotoxic activities. Although these screens did result in the isolation of many bioactive compounds,[27] they were considered to be too non-specific for the next generation of drugs. Fortunately, a large number of robust and specific biochemical and genetics-based screens using transformed cells, a key regulatory intermediate in a biochemical or genetic pathway, or a receptor–ligand interaction (often derived from the explosion in genomic information since the middle 1990s), are now in routine use. These screens will permit the detection of bioactive compounds in the complex matrices that are natural product extracts with greater precision.

One interesting feature of such screens has increased the attractiveness of natural products to the pharmaceutical industry. The screens themselves are all highly automated and high throughput (upwards of 50,000 assay points per day in a number of cases). Because of this, the screening

capacity at many companies is significantly larger than the potential input from in-house chemical libraries. Since screening capacity is no longer the rate-limiting-step, many major pharmaceutical companies are becoming very interested in screening natural products (either as crude extracts or as prefractionated "peak libraries") as a low-cost means of discovering novel lead compounds. This is well illustrated by the discovery of a new antibiotic, platensimycin, by a team of scientists from Merck Research Laboratories. It has *in vitro* activity against several drug-resistant bacteria, is a selective FabF inhibitor, and was discovered through the testing of a library of 250,000 natural product extracts in a custom-designed assay involving an engineered strain of *Staphylococcus aureus* incorporating the fatty acid synthase pathway enzyme, FabF.[28] Such promise has also spawned small companies such as Merlion Pharmaceuticals in Singapore which has a library of many thousands of natural products derived from a variety of sources which it exposes to validated drug targets provided by pharmaceutical companies, with the goal of generating drug leads.[29]

III. THE DESIGN OF AN EFFECTIVE NATURAL-PRODUCTS-BASED APPROACH TO DRUG DISCOVERY

There are four major elements in the design of any successful natural-products-based drug discovery program: acquisition of biomass, effective screening, bioactivity-driven fractionation, and rapid and effective structure elucidation (which includes dereplication). Although some of these have been mentioned earlier, it is instructive to bring them together here.

A. Acquisition of biomass

The acquisition of biomass has undergone a very significant transition from the days when drug companies and others routinely collected organisms without any thought of ownership by, or reimbursement to, the country of origin. Today, thanks to the Convention on Biodiversity (or CBD) and similar documents and agreements such as the US National Cancer Institute's Letter of Collection (NCI's LOC: http://ttc.nci.nih.gov/forms/loc.doc), all ethical biomass acquisitions now include provisions for the country of origin to be recompensed in some way for the use of its biomass. It should be noted that the LOC predated the CBD by 3 years; its tenets, as a minimum, must be adhered to by any investigator who has his or her collections funded by the NCI/NIH. Such recompense to the country of origin is best provided through formal agreements with government organizations and collectors in the host country, with such agreements providing not only for reimbursement of collecting expenses, but also for further benefits (often in the form of milestone and/or royalty payments) in the event that a drug is developed from a collected sample. Agreements often include terms related to the training of host country scientists and transfer of technologies involved in the early drug discovery process. Recognition of the role played by indigenous peoples through the stewardship of resources in their region and/or the sharing of their ethnopharmacological information in guiding the selection of materials for collection is important in determining the distribution of such benefits. There have been sample legal agreements,[30] and discussions as to methods used by various groups published in the last few years.[31–33]

It is axiomatic that all samples collected, irrespective of type of source, must if at all possible be fully identified to genus and species. Such identification is usually possible for all plant species, but it is not always possible for microbes and marine organisms. Voucher specimens should be provided to an appropriate depository in the host country as well as to a similar operation in the home country of the collector.

The selection of plant samples often raises the question of the ethnobotanical/ethnopharmacological approach versus a random approach. The former method, which usually involves the selection of plants that have a documented (written or oral) use by native healers, is attractive in that it can tap into the empirical knowledge developed over centuries of use by large numbers of people. In addition, the bioactive constituents may be considered as having had a form of continuing clinical trial in man. The benefits of this approach have been extolled in several relatively recent articles,[34–36] and one author provides personal experience of the effectiveness of some jungle medicines.[37] The weakness of the ethnobotanical approach has always been that it is slow, requiring careful interviewing of native healers by skilled scientists, including ethnobotanists, anthropologists, trained physicians, and pharmacologists. In addition, the quoted folkloric activity in the collected plant(s) may not be detectable, given the particular screens in use by the screening laboratory. Where ethnobotanical approaches have the highest possibility of success is in studies related to overt diseases/conditions such as parasitic infections, fungal sores, and contraception/conception to name a few. In such cases, there are adequate controls, even on the same patient. Where there does not yet appear to be any successful relationship is in diseases such as cancer and AIDS-related conditions, where extensive testing of the patient is required for an accurate diagnosis.

1. *Classical natural sources: untapped potential*

Despite the intensive investigation of terrestrial flora, it is estimated that only 5–15% of the approximately 300,000 species of higher plants have been systematically investigated, chemically and pharmacologically,[38,39] while the potential of the marine environment as a source of novel

FIGURE 8.4 Natural products from novel sources.

drugs remains virtually unexplored.[40,41] Until recently, the investigation of the marine environment has largely been restricted to tropical and subtropical regions, but colder climes are now being explored, and the isolation of the cytotoxic macrolide palmerolide A (Figure 8.4) from an Antarctic tunicate has recently been reported.[42] Its structure has recently been revised and it has been synthesized.[43] The novel pyrido-pyrrolo-pyrimidine derivatives, the variolins (Figure 8.4) were isolated a few years earlier,[44,45] and this work was followed by total synthesis of these compounds and derivatives by chemists at PharmaMar a decade later.[46]

Exposure of the roots of hydroponically grown plants to chemical elicitors has been reported to selectively and reproducibly induce the production of bioactive compounds,[47] while feeding of seedlings with derivatives of selected biosynthetic precursors can lead to the production of non-natural analogs of the natural metabolites. This has been demonstrated in the production of non-natural terpene indole alkaloids related to the vinca alkaloids through the feeding of seedlings of *Catharanthus roseus* with various tryptamine analogs.[48]

B. The unexplored potential of microbial diversity

Until recently, the study of natural microbial ecosystems has been severely limited due to an inability to cultivate most naturally occurring microorganisms, and it has been estimated that much less than 1% of microorganisms seen microscopically have been cultivated. Given that "a handful of soil contain billions of microbial organisms,"[49] and the assertion that "the workings of the biosphere depend absolutely on the activities of the microbial world,"[50] it seems clear that the microbial universe presents a vast untapped resource for drug discovery. In addition, greatly enhanced understanding of the gene clusters that encode the multimodular enzymes, such as polyketide synthases (PKSs) and/or nonribosomal peptide synthetases (NRPSs), both of which are involved in the biosynthesis of a multitude of microbial secondary metabolites, has enabled the sequencing and detailed analysis of the genomes of long-studied microbes such as *Streptomyces avermitilis*. Through such studies, the presence of additional PKS and NRPS clusters has been revealed, leading to the discovery of novel secondary metabolites not detected in standard fermentation isolation processes.[51] Genome mining has been used in the discovery of a novel peptide, coelichelin, from the soil bacterium, *Streptomyces coelicolor*[52] and this concept is further expanded on in the discussion in Section V.D.

1. *Improved culturing procedures*

Recent developments of procedures for cultivating and identifying microorganisms are aiding microbiologists in their assessment of the earth's full range of microbial diversity. For example, in an application of a technique pioneered by a small, now defunct biotechnology company known as "One-Cell Systems" in the late 1980s, "nutrient-sparse" media simulating the original natural environment have been used for the massive parallel cultivation of gel-encapsulated single cells (gel micro-droplets (GMDs)) derived from microbes separated from environmental samples (sea water and soil).[53] This has permitted "the simultaneous and relatively non-competitive growth of both slow- and fast-growing microorganisms." This process prevents the overgrowth by fast-growing "microbial weeds," and has led to the identification of previously undetected species (using 16S rRNA gene sequencing), and the culturing and scale-up cultivation of previously uncultivated microbes. To add to this, recently Moore's group at the Scripps Institute of Oceanography has reported[54] the initial results of sequencing *Salinispora tropica* where they found that approximately 10% of the genome coded for potential secondary metabolites. If one couples this work to the recent paper on cultivation of Gram-positive marine microbes by Gontang *et al.*[55] then the potential for novel agents is immense.

2. *Extremophiles*

Extremophilic microbes (extremophiles) abound in extreme habitats, such as acidophiles (acidic sulfurous hot springs), alkalophiles (alkaline lakes), halophiles (salt lakes), piezo (baro)- and (hyper)thermophiles (deep-sea vents),[56–60] and

FIGURE 8.5 Natural products from extremophiles and endophytes.

psychrophiles (arctic and antarctic waters, alpine lakes).[61] While investigations thus far have focused on the isolation of thermophilic and hyperthermophilic enzymes (extremozymes),[62–66] these extreme environments will also indubitably yield novel bioactive chemotypes.

An unusual group of acidophiles which thrive in acidic, metal-rich waters has been found in abandoned mine waste disposal sites, polluted environments which are generally toxic to most prokaryotic and eukaryotic organisms.[67] In this work, the novel sesquiterpenoid and polyketide–terpenoid metabolites berkeleydione and berkeleytrione (Figure 8.5) showing activity against metalloproteinase-3 and caspase-1, activities relevant to cancer, Huntington's disease and other diseases, have been discovered from *Penicillium* species found in the surface waters of Berkeley Pit Lake in Montana.[67–69]

3. Endophytes

While plants have received extensive study as sources of bioactive metabolites, the endophytic microbes which reside in the tissues between living plant cells have received little attention. Endophytes and their host plants may have relationships varying from symbiotic to pathogenic, and limited studies have revealed an interesting realm of novel chemistry.[70–72] Amongst the new bioactive molecules discovered are novel wide-spectrum antibiotics, kakadumycins, isolated from an endophytic *Streptomycete* associated with the fern leafed grevillea (*Grevillea pteridifolia*) from the Northern Territory of Australia,[73] ambuic acid (Figure 8.5), an antifungal agent, which has been recently described from several isolates of *Pestalotiopsis microspora* found in many of the world's rainforests,[74] peptide antibiotics, the coronamycins, from a *Streptomyces* species associated with an epiphytic vine (*Monastera* species) found in the Peruvian Amazon[75] and aspochalasins I, J, and K (Figure 8.5),[76] from endophytes of plants from the southwestern desert regions of the United States.

In the case of endophytic fungi, recent reports (see below) of the isolation of important plant-derived anticancer drugs have served to focus attention on these sources. A recent genomic analysis of the fungus *Aspergillus nidulans* reported that "Sequence alignments suggest that *A. nidulans* has the potential to generate up to 27 polyketides, 14 nonribosomal peptides (NRPs), one terpene, and two indole alkaloids; similar predictions can be made from the *A. fumigatus* and *A. oryzae*" as a result of the analysis of the potential number of secondary metabolite clusters in these fungi.[77] This analysis demonstrated not only the presence of "clustered" secondary metabolite genes in this fungus, but also identified the potential "controller" of expression of these clusters and demonstrated it by expressing terrequinone A (Figure 8.5), a compound not previously reported from this species.[77] A recent review expands the discussion on control of secondary metabolites in fungi.[78]

As mentioned above, in the last few years fungi have been isolated from plants that have produced small quantities of various important anticancer agents. Examples are Taxol® from *Taxomyces*[79] and many *Pestalotiopsis* species,[80] camptothecin,[81,82] podophyllotoxin,[83,84] vinblastine,[85] and vincristine[86,87] from endophytic fungi isolated from the producing plants. It has been demonstrated that these compounds are not artifacts, and so the identification of the gene/gene product controlling metabolite production by

FIGURE 8.6 Examples of novel microbial natural products.

these microbes could provide an entry into greatly increased production of key bioactive natural products.

4. *Marine microbes*

Deep ocean sediments are proving to be a valuable source of new actinomycete bacteria that are unique to the marine environment,[88] and based on a combination of culture and phylogenetic approaches, the first truly marine actinomycete genus named *Salinospora* has been described.[55,89] Members of the genus are ubiquitous, and are found in sediments on tropical ocean bottoms and in more shallow waters, often reaching concentrations up to 10^4 per cc of sediment, as well as appearing on the surfaces of numerous marine plants and animals. Culturing using the appropriate selective isolation techniques has led to the observation of significant antibiotic and cytotoxic activity, leading to the isolation of a potent cytotoxin, salinosporamide A (Figure 8.6), a very potent proteasome inhibitor (IC_{50} = 1.3 nM),[90] currently in Phase I clinical trials. More recent studies have led to the isolation and cultivation of another new actinomycete genus, named *Marinispora*, which is also yielding rich new chemistry. Novel macrolides called marinomycins have been isolated, and marinomycins A–D (Figure 8.6) show potent activity against drug-resistant bacterial pathogens and some melanomas.[91] Publications by the Fenical group on the novel and diverse chemistry of these new microbial genera continue to appear regularly.

5. Microbial symbionts

There is mounting evidence that many bioactive compounds isolated from various macroorganisms are actually metabolites synthesized by symbiotic bacteria.[92] These include the anticancer maytansanoids (Figure 8.6), originally isolated from several plant genera of the Celastraceae family,[93] the pederin class of antitumor compounds (Figure 8.6) isolated from beetles of genera *Paederus* and *Paederidus* which have also been isolated from several marine sponges,[94–96] and a range of antitumor agents isolated from marine organisms which closely resemble bacterial metabolites.[40]

An interesting example of endo-symbiosis between a fungus and a bacterium has been discovered in the case of rice seedling blight where the toxic metabolite, rhizoxin (Figure 8.6), originally isolated from the contaminating *Rhizopus* fungus, has actually been found to be produced by a symbiotic *Burkholderia* bacterial species.[97] This unexpected finding reveals a complex symbiotic–pathogenic relationship, extending the fungal–plant interaction to a third, key bacterial player, thereby offering potentially new avenues for pest control. In addition, rhizoxin exhibits potent antitumor activity, but toxicity problems have precluded its further development as an anticancer drug. The cultivation of the bacterium independently of the fungal host has enabled the isolation of rhizoxin as well as rhizoxin analogs which may have significant implications in development of agents with improved pharmacological properties.

6. Combinatorial biosynthesis

Great advances have been made in the understanding of the role of multifunctional polyketide synthase enzymes (PKSs) in bacterial aromatic polyketide biosynthesis, and many such enzymes have been identified, together with their encoding genes.[98–101] The same applies to NRPSs responsible for the biosynthesis of NRPs.[100] Through the rapidly increasing analysis of microbial genomes, a multitude of gene clusters encoding for polyketides, NRPs, and hybrid polyketide-NRP metabolites have been identified, thereby providing the tools for engineering the biosynthesis of novel "non-natural" natural products through gene shuffling, domain deletions, and mutations.[100,102] Examples of novel analogs of anthracyclines, ansamitocins, epothilones, enediynes, and aminocoumarins produced by combinatorial biosynthesis of the relevant biosynthetic pathways have recently been reviewed by Shen *et al*.[103]

A recent example of the power of this technique when applied to natural products is the development of an efficient method for scale-up production of epothilone D (Figure 8.6), which entered clinical trials as a potential anticancer agent but has now been discontinued in favor of a congener, 9,10-didehydroepothilone D. Epothilone D was the most active of the epothilone series isolated from the myxobacterium, *Sorangium cellulosum*, and is the des-epoxy precursor of epothilone B. The isolation and sequencing of the polyketide gene cluster producing epothilone B from two *S. cellulosum* strains has been reported,[104,105] and the role of the last gene in the cluster, *epo*K, encoding a cytochrome P_{450}, in the epoxidation of epothilone D to epothilone B has been demonstrated. Heterologous expression of the gene cluster minus the *epo*K in *Myxococcus xanthus* resulted in large-scale production of crystalline epothilone D.[106] Further discussion on the integration of this technology into investigations of natural products is given in Section V.

C. Extraction

In the case of microbes and marine organisms, extraction is normally carried out on the whole organism (though now some groups are isolating the commensal/associated microbes from marine invertebrates before a formal extraction). With plants however, which may be large and have well-differentiated parts, it is common to take multiple samples from one organism and to extract them separately. The methods used by the NCI are summarized at http://npsg.ncifcrf.gov/. The procedures used for extraction vary with the nature of the sample, and in some cases are dependent on the nature of the ultimate assay. Thus, a number of screens are sensitive to the tannins and complex carbohydrates that are extractable from a variety of organisms, and systems have been developed that permit easy removal of such "nuisance" compounds before assay.[107,108] In the simplest cases, however, extraction with a lower alcohol (methanol to isopropanol) will bring out compounds of interest, though in most cases a sequential extraction system is utilized with compounds being extracted with solvents of ever-increasing polarity.

D. Screening methods

As mentioned earlier, the advent of new and robust high-throughput screens has had and continues to have a major impact on natural products research in the pharmaceutical industry. Most of the screens used are proprietary and published information is rare, although general summaries of this approach have been published.[109] One screen that has been described in detail is the NCI's 60-cell line cytotoxicity screen for anticancer agents.[110] Although this is not a true receptor-based screen, it has now been developed into a system whereby a large number of molecular targets within the cell lines may be identified by informatics techniques, and refinements are continuing. Information can be obtained from the following URL; http://dtp.nci.nih.gov as to the current status of the screens involved. An assay based on differential susceptibility to genetically modified yeast strains has been described,[111] and has led to many screens based on genetically modified yeasts, but at times, the low permeability of the unmodified yeast cell wall to chemical compounds

has been overlooked. Thus, data from such screens, particularly those designed with gene deletions, must be carefully scrutinized since a large number are based on hosts without a modified cell wall. In addition there are simple but robust assays that can be utilized by workers in academia that do not have access to or may not need high-throughput screens. Examples are the brine shrimp and potato disc assays[112,113] or the still useful disc-based antimicrobial assays.

High-throughput assays, where large numbers of samples can be screened in a short period of time, are becoming less expensive, and such assays are moving from the industrial or industrial/academic consortium-based groups to academia in general, with specific expression systems being employed as targets for natural product lead discovery.[114] The application of new techniques, including new fluorescent assays, NMR, affinity chromatography and DNA microarrays, has led to significant advances in the effectiveness of high-throughput screening.[115,116]

E. Isolation of active compounds

The isolation of the bioactive constituent(s) from a given biomass can be a challenging task, particularly if the active constituent of interest is present in very low amounts. The actual procedure will depend to a large extent on the nature of the sample extract: a marine sample,[117] for example, may well require a somewhat different extraction and purification process from that derived from a plant sample.[118] Nevertheless, the essential feature in all of these methods is the use of an appropriate and reproducible bioassay to guide the isolation of the active compound. It is also extremely important that compounds that are known to inhibit a particular assay, or those that are nuisance compounds be dereplicated (identified and eliminated) as early in the process as possible. Procedures for doing this have been discussed,[119–124] and various new approaches to isolation and structure elucidation have been reviewed.[125–127]

F. Structure elucidation

Structure elucidation of the bioactive constituent depends almost exclusively on the application of modern instrumental methods, particularly high-field NMR and MS. These powerful techniques, coupled in some cases with selective chemical manipulations, are usually adequate to solve the structures of most secondary metabolites up to 2 kD molecular weight. X-ray crystallography is also a valuable tool if crystallization of the material can be induced, and in some cases, it is the only method to unambiguously assign absolute configurations. Nowadays, the determination of the amino acid sequences of polypeptides or peptide-containing natural products up to 10–12 kD is a relatively straightforward task, requiring less than 5 mg of a polypeptide. In addition, MS techniques have developed to the stage where polypeptides containing unusual amino acids not recognized by conventional sequence techniques can be sequenced entirely by MS.

G. Further biological assessment

Once the bioactive component has been obtained in pure form and shown to be either novel in structure or to exhibit a previously unknown function (if it is a compound that is in the literature), then it must be assessed in a series of biological assays to determine its efficacy, potency, toxicity, and pharmacokinetics. These will help to position the new compound's spectrum of activity within the portfolio of compounds that a group may be judging for their utility as either drug candidates or leads thereto. If an idea can be gained as to its putative mechanism of action (MOA) (assuming that the screening techniques used to discover it were not MOA-driven) at this stage, then it too can help as a discriminator in the prioritizing process.

H. Procurement of large-scale supplies

Once a compound successfully completes evaluation in the initial biological assays then larger amounts of material will be required for the studies necessary if activity and utility are maintained as the compound proceeds along the path from "Hit" to a "Drug Lead" and then to a "Clinical Candidate." The supplies could be made available by cultivation of the plant or marine starting material, or by large-scale fermentation in the case of a microbial product. Chemical synthesis or partial synthesis may also be possible if the structure of the active compound is amenable to large-scale synthesis. The example of paclitaxel is instructive here: after initial large-scale production by direct extraction from *Taxus brevifolia* bark, it is now generally produced by a semi-synthetic procedure starting from the more readily available precursor 10-deacetylbaccatin III.[128]

Another method of obtaining adequate supplies of a natural plant product is by utilizing plant tissue culture methods. Although there are a few examples of the commercial production of secondary metabolites by plant cell culture (shikonin being perhaps the best known one[129]), the application of this technique to commercial production of pharmaceuticals had not found general acceptance, primarily for economic reasons.[130] However, the development of viable methods for the large-scale production of paclitaxel (Taxol®), has illustrated that this technology can now be successfully applied to the production of a major drug for commercial purposes.[131]

The discovery that several major anticancer drugs, originally isolated from plants, are produced by associated endophytic fungi (Section III.B.3.) opens up further avenues for exploring the large-scale production of plant-derived pharmaceuticals. Likewise, the probable role of microbial symbionts in the production of bioactive agents from marine macroorganisms (Section III.B.4.) offers

similar opportunities for scaling up the production of marine-derived pharmaceuticals. It is interesting to note that the anticancer agent, Yondelis® (ecteinascidin 743), originally isolated from the tunicate, *Ecteinascidia turbinata*, is now produced on a large scale by semi-synthesis from the antibiotic, cyanosafracin B, produced through fermentation of the bacterium, *Pseudomonas fluorescens* (Section IV.G.1.). As noted in Section III.B.6. an efficient method for scale-up production of epothilone D (Figure 8.6), isolated from the myxobacterium, *S. cellulosum* was developed through the manipulation of the polyketide gene cluster producing the epoxy analog, epothilone B, and heterologous expression of the modified gene cluster in *M. xanthus*.

In a relatively few cases, total synthesis has provided a viable route to large-scale production of important bioactive agents. A good example was the marine-derived anticancer agent, discodermolide[132] which entered Phase I clinical trials but currently is not progressing to later phases due to toxicity. However, in contrast, the modification of the halichondrin B skeleton to produce E7389 (eribulin; Section IV.G.2.) by total synthesis is an excellent example of modification of a very complex molecule to a slightly less complex agent that is now in Phase III trials for breast cancer.[133]

I. Determination of structure–activity relationships

The initial hit isolated from the biomass, irrespective of source, is not necessarily the lead required for further development into a drug. It may be too insoluble, not potent enough, or be broadly rather than specifically active. Once the structure has been determined, then synthetic chemistry, involving both conventional and combinatorial methods, may be used in order to generate derivatives/analogs that have the more desirable characteristics of a potential drug. Several examples of these types of processes are shown in Section IV.

The use of natural product-like compounds as scaffolds is leading to the generation of smaller, more meaningful combinatorial libraries. This is exemplified by the work of the Schreiber group who have combined the simultaneous reaction of maximal combinations of sets of natural-product-like core structures ("latent intermediates") with peripheral groups ("skeletal information elements") in the synthesis of libraries of over 1,000 compounds bearing significant structural and chiral diversity.[134,135] As stated in an article by Borman,[136] "an initial emphasis on creating mixtures of very large numbers of compounds has largely given way in industry to a more measured approach based on arrays of fewer, well-characterized compounds" with "a particularly strong move toward the synthesis of complex natural-product-like compounds – molecules that bear a close structural resemblance to approved natural-product-based drugs." Borman further emphasized this point in a second article,[137] in which he stated that "the natural product-like compounds produced in diversity oriented synthesis (DOS) have a much better shot at interacting with the desired molecular targets and exhibiting interesting biological activity."

Detailed analyses of active natural product skeletons have led to the identification of relatively simple key precursor molecules which form the building blocks for use in combinatorial synthetic schemes that have produced numbers of potent molecules, thereby enabling structure activity relationships to be probed. Thus, in the study of the structure–activity relationships of the epothilones, solid-phase synthesis of combinatorial libraries was used to probe regions of the molecule important to retention or improvement of activity.[138]

The use of an active natural product as the central scaffold in the combinatorial approach can also be applied to the generation of large numbers of analogs for structure–activity studies, the so-called parallel synthetic approach. This is embodied in the concept of "privileged structures," originally proposed by Evans *et al.*[139] and then advanced further by Nicolaou *et al.*[140–142] and the Waldmann *et al.*[143,144]

IV. EXAMPLES OF NATURAL PRODUCTS OR ANALOGS AS DRUGS

A. Antihypertensives

1. *Angiotensin-converting enzyme inhibitors (captopril and derivatives)*

The synthetic Angiotensin-Converting Enzyme (ACE) inhibitors were derived from studies of principles in the venom of the pit viper, *Bothrops jararaca*, that inhibited the degradation of the nonapeptide, bradykinin.[145] This active principle was shown to be the simple nonapeptide, teprotide with specific activity as an ACE inhibitor and with hypotensive activity in clinical trials.[146,147]. The prototypical ACE drug, Captopril®[146,147] (Figure 8.7) was then developed based on the C-terminal proline structure of all known peptidic ACE inhibitors.[146,147] In the last 20-plus years, 13 more compounds based on this original discovery have become approved drugs acting on this target, with the latest compounds being Benazepril® (Figure 8.7) and Cilazepril® (Figure 8.7).[148]

B. Anticholesterolemics

1. *Lovastatin*

An elevated serum cholesterol level is an important risk factor in cardiac disease (and in hypertension), and thus a drug which could lower this level would be an important prophylactic against cardiovascular diseases in general. Humans synthesize about 50% of their cholesterol requirement with

FIGURE 8.7 Natural product-based ACE inhibitors.

FIGURE 8.8 Natural and natural-product-based anticholesterolemics.

the rest coming from diet, thus if an inhibitor of cholesterol uptake/absorption is available, then a two-pronged approach might be feasible.

A potential site for inhibition of cholesterol biosynthesis is at the rate-limiting-step in the system, the reduction of hydroxymethylglutaryl coenzyme A by HMG-CoA reductase to produce mevalonic acid. Following the original identification in 1975 of compactin (Figure 8.8) from a fermentation beer of *Penicillium citrinum* as an inhibitor of HMG-CoA reductase by Sankyo,[149,150] it was also reported as an antifungal agent the next year by Brown *et al.*[151] from *Penicillium brevicompactum*. Using the HMG-CoA reductase inhibitor assay, Endo isolated the 7-methyl derivative of compactin, mevinolin (Figure 8.8), from *Monascus ruber* and submitted a patent for its biological activity to the Japanese Patent Office but without a structure under the name Monacolin K.[152,153] Concomitantly, Merck discovered the same material using a similar assay from an extract of *Aspergillus terreus*. It was reported in 1980[154] with an US Patent issued in the same year.[155]

Following a significant amount of development work, mevinolin (Lovastatin®) became the first commercialized HMG-CoA-reductase inhibitor in 1987.[156,157] Further work using either chemical modification of the basic structure or by use of biotransformation techniques led to two slightly modified compounds; one from converting the 2-methylbutanoate side-chain into 2,2-dimethylbutanoate (Simvastatin®) (Figure 8.8) and the second, by opening of the exocyclic lactone to give the free hydroxy acid (Pravastatin®) (Figure 8.8). Further development starting with compounds such as these has led to a number of totally synthetic "statins," with atorvastin (Lipitor®) (Figure 8.8) shown above. What is significant about these synthetic compounds, irrespective of which company has synthesized and/or developed them as commercial drugs, is that in every case, their "operative ends" are the dihydroxy-heptenoic

FIGURE 8.9 Immunosuppressives.

acid side-chain (or its reduced form) from the fungal products linked to a lipophilic ring structure. All of these compounds demonstrate the intrinsic value of natural products as the source of the pharmacophore, with Lipitor® grossing over US $13B in sales in 2006.

Although the early compounds are now out of patent, in an excellent example of what can be best described as both pharmacologic and corporate synergies, recently Merck and Schering-Plough were able to obtain FDA approval for a combination drug where Schering-Plough's cholesterol uptake inhibitor ezetimibe (Figure 8.8) was combined with simvastatin under the trade name of Vytorin®. Aside from the unusual situation of two major competing pharmaceutical houses cooperating in production of a new drug, the base structure of ezetimibe is derived from the monobactam nucleus, first described as an antibiotic agent in the early 1980s. The derivation of this cholesterol uptake inhibitor was described by the Schering-Plough chemistry group in 1998.[158] They were designed as acyl coenzyme A:cholesterol acyl transferase(ACAT) inhibitors but were then discovered to inhibit cholesterol uptake not by ACAT inhibition, but rather by inhibition of *NPC1L1* (Niemann-Pick C1-like 1 protein).[159] The review by Burnett and Huff[160] should be consulted for more information on the pharmacology background.

C. Immunosuppressives

1. *Rapamycin and derivatives*

The most important immunosuppressive agent is the fungal secondary metabolite, rapamycin (Figure 8.9), a natural product that has both been highly successful as a drug in its own right and as the basis for a series of potentially important drugs in a variety of disease conditions. This compound was approved as an immunosuppressive drug under the trade name Rapamune® in 1999 though it was first identified as an antifungal drug years earlier. Modification at one site, the C_{43} alcoholic hydroxyl group, has led to three further clinical drugs (everolimus; Figure 8.9, zotarolimus; Figure 8.9, and temsirolimus (CCI-779; Figure 8.9) and one in Phase II clinical trials (AP23573; Figure 8.9). In all cases, modifications were made in one area that avoids both the FKBP-12 and the target of rapamycin (TOR) binding sites, as modifications in other areas would negate the basic biological activity of this molecule.[161]. The recent review by Koehn should be consulted for further details of these and related compounds.[162]

D. Antibiotics

1. *General comments*

Although the pundits will claim that the "Golden Age of Antibiotics" is long past, the necessity for new agents to combat infectious disease of all types is still with us, and with the massive misuse of potent antibiotics, the microbes that are now major causes of diseases of man (and animals) are rapidly exhibiting multiresistant phenotypes. Perhaps nowhere is the effect of multiple resistance phenotype seen more than in the problems that arise in the treatment of tuberculosis. Most if not all clinical strains are resistant to at least two if not more of the commonly used antibiotics, and currently, increasing numbers of patients present with strains

FIGURE 8.10 Natural antiparasitic compounds.

that are resistant to six or more of the antibiotics commonly used for treatment. A recent review by Janin gives an extensive coverage of the drugs in use or in clinical trials for TB and should be consulted for further information.[163]

M. tuberculosis is not the only "problem microbe" with high-level resistance, as shown by the listing from the Infectious Disease Society of America of the following "problem" microbes; methicillin-resistant *S. aureus* (MRSA), vancomycin-resistant *E. faecium* (and early reports of the similar *S. aureus*), extended-spectrum β-lactamase-producing members of the Enterobacteriaceae, and the multiply drug-resistant (MDR) strains of *A. baumannii*, *P. aeruginosa*, and *C. difficile*, with others "waiting in the wings." A fuller discussion of the problem is given in the review by Wright and Sutherland[164] which should be consulted for further information.

Fortunately natural products continue to provide new antibiotics. The review by Lam[165] on natural products in drug discovery, lists seven natural products or derivatives that were approved as antibacterial or antifungal agents in the United States in the time frame 2001–2005. The review by Newman and Cragg,[8] lists 11 antibacterial compounds and three antifungal compounds that are either natural products or derivatives that have been approved by all regulatory authorities in the time frame 1997–2006. A thorough recent review of antibacterial natural products has just been published by von Nussbaum *et al.* and should be consulted for information on such agents.[166]

2. *Avermectin, ivermectin, and doramectin*

The avermectins are a family of broad-spectrum antiparasitic compounds with avermectin A (Figure 8.10) being an example. These compounds were originally sold by Merck (with a slight chemical modification) as veterinary agents (ivermectins) under the name Mectizan® (Figure 8.10). Subsequently, it was discovered that the molecules had excellent activity against some of the West African parasite strains that caused river blindness, and they moved into human medicine and recently (2006) this agent was registered in Japan for the treatment of human scabies infections.

Since these molecules are polyketides, in recent years groups at Pfizer and elsewhere have been working on ways to modify the keto-synthases that would permit novel agents to be expressed by suitable microbial hosts. These efforts have led to the modified agents known collectively as the doramectins which have improved activities (Figure 8.10).[167]

E. Microbial anticancer agents

1. *Epothilones*

Over the last ten or so years, some of the most interesting natural product base structures being considered as agents for clinical trials in cancer chemotherapy, have been the myxobacterial products known collectively as the epothilones. These macrolides, with a mechanism of action similar to that of Taxol®, have been extensively studied, initially by the discoverers in Germany[168] and then by Merck and Kosan in the United States, and also from a chemical perspective, by three major groups, those of Danishefsky at Memorial Sloan-Kettering[169], Nicolaou at the Scripps Institute[138] and Altmann, originally at Schering AG and now at the Eidgenössische Technische Hochschule (ETH) Zurich.[170,171]

Enough material was amassed to begin the evaluation of this class of agents as antitumor drugs, initially by the production of epothilones A and B (Figure 8.11) through a combination of classical fermentation and isolation of

FIGURE 8.11 Epothilone anticancer agents.

the natural products, and then subsequent work on the biosynthetic gene cluster involving the deletion of the terminal P_{450} gene leading to production on a significant scale of epothilones C and D (Figure 8.11) (*cf.* comments in Section III.B.6. above). In addition, total syntheses were performed by a number of groups, including variations on the macrolide ring giving the azaepothilone (ixabepilone or 16-azaepothilone B; Figure 8.11). This compound was approved in late 2007 by the FDA for monotherapy against breast cancer. It is the first non-taxane with tubulin-stabilizing activity to be granted marketing authority. A recent review concludes. "There is a clear need for new agents active against resistant metastatic breast cancer and ixabepilone might be a welcome new compound in this situation."[172]

However, other molecules are still in the race, including some natural products and others that are products of semi-synthesis and in two particular cases, total syntheses. Currently (July 2007) there are at least 16 molecular entities in varying stages of testing from biological testing through to Phase III clinical trials. Of these, four are in clinical trials in addition to ixabepilone. These are composed of epothilone B (patupilone, or EPO-906) in Phase III with Novartis, a slight modification of epothilone B where the thiazole sidechain of epothilone B has an amino group in the 21 position, known as BMS-310705 (Figure 8.11) in Phase I; and a second generation version of epothilone D known as 9,10-didehydroepothilone D or KOS-1584 (Figure 8.11). The latter compound was licenced by Kosan from the Danishefsky group together with epothilone D, but was later selected to replace epothilone D in clinical development; it is currently in Phase I with a projected Phase II trial to commence late in 2007 (Prous Integrity® listing).

From a totally synthetic aspect, another group has introduced a very interesting molecule that is now in Phase II clinical trials with Schering AG. This is the molecule known as ZK-EPO (Figure 8.11), which for a number of years did not have a published structure. In 2006, the full rationale for and synthetic methods employed were published for this molecule, which is a modification of epothilone B with a benzothiazole in place of the thiazole on the Western side (effectively a ring-closure of the pendant thiazole in epothilone B), and the substitution of an allyl group for the methyl on the Eastern side of epothilone B. Although these

are relatively modest changes, the molecule is no longer an MDR substrate, has no effect on confluent normal cells and is as active on resistant phenotypes as on their wild type precursors.[173]

In addition to these molecules in trials there are others that have some interesting substituents from a medicinal chemistry perspective. One is the trifluoromethyl-substituted molecule fludelone[169] (Figure 8.12) and the others have sulfur substituents in the macrolide ring, one being 5-thiaepothilone B (Figure 8.12) and the other 3-thiaepothilone D (Figure 8.12). The sulfur substituted compounds have only been referenced in patents but appear to have activities *in vitro* in the nanomolar to subnanomolar range against selected cell lines. Then last year, in an extension of the concept expressed in the work leading to ixabepilone, where the 16 position was substituted by nitrogen thus giving rise to a formal lactam ring, the Altmann group reported a novel form of "Azathilones" where the epoxide was removed from epothilone B and a ring nitrogen substituted. This ring system is now being further explored by this group utilizing substitution patterns reminiscent of the ZK-EPO molecule, with the most active reported molecule[174] being the *N-tert*-butyloxycarbonyl derivative (Figure 8.12).

Thus just as with the taxanes, which are still actively being worked on close to 15 years after approval (see below), the epothilones, are still exercising the minds of medicinal chemists and will continue to do so, particularly as there are other molecules with similar activity against tubulin such as peloruside, that has structural features similar to the epothilones but comes, at least formally, from a marine invertebrate.

What is potentially exciting is that recently, Khosla's group[175] demonstrated that the biosynthetic pathways that encode the C_1–C_8 portion of the epothilone molecule leading to the formal production of epothilone C, will accept non-natural substrates. This has the potential to lead to "manufacturing" by a combination of precursor-directed biosynthesis coupled to chemical bond formation by standard techniques, of unnatural epothilones whose properties may well be quite different from those previously reported. The C_1–C_8 chain is apparently significant for tubulin binding but changes to the rest of the molecule may "tune" these interactions. However, only time and experimentation will tell if such unnatural molecules will have antitumor activities or other, previously unknown biological interactions.

2. *Other examples*

Rather than discuss other anticancer examples in this chapter, interested readers are referred to a book with data on recent advances in the medicinal chemistry of such agents, including the anthracyclines,[176] the bleomycins,[177] the mitomycins,[178] the endiynes,[179] and the staurosporines.[180]

F. Anticancer agents from plants

The initial studies of natural products as potential anticancer agents were made in the 1950s based on compounds from

FIGURE 8.12 Additional epothilone anticancer agents.

plants, and the vinca alkaloids and the podophyllotoxin analogs etoposide and teniposide were the first fruits of these investigations. Later work led to the very important taxane drugs and the camptothecin analogs.

1. *Paclitaxel*

Paclitaxel (Taxol®; Figure 8.13) is the most exciting plant-derived anticancer drug discovered in recent years. It occurs, along with several key precursors (the baccatins), in the leaves of various *Taxus* species.[181] It was found to act by promoting the assembly of tubulin into microtubules, and the discovery of this activity in 1979 by Schiff and Horwitz[182] was an important milestone in the development of paclitaxel as a drug. After an extended period of development it was finally approved for clinical use against ovarian cancer in 1992 and against breast cancer in 1994. Since then it has become a blockbuster drug, with annual sales of over $1 billion.

The success of paclitaxel spurred an enormous amount of work on the synthesis of analogs, and this work has been summarized in several reviews.[131,183–188] The first analog to be developed is the close chemical relative, docetaxel (Taxotere®; Figure 8.13).[189] The new albumin-bound formulation of paclitaxel known as Abraxane® has also been approved for clinical use and launched in 2005; this formulation offers some important clinical advantages compared with the original Cremophor formulation.[190] The reviews cited above should be consulted for information on new agents in development, such as BMS-184776, BMS-188797, and larotaxel (Figure 8.13).

2. *Other examples*

As in the case of microbially derived anticancer agents (Section IV.E.2.), interested readers can consult the information on podophyllotoxin and derivatives,[191] the vinca alkaloids and derivatives,[192] and camptothecin and derivatives[193] in the recent volume edited by the authors of this chapter.

G. Anticancer agents from marine organisms

The study of marine organisms as sources of anticancer agents only began in the late 1960s, and Yondelis®, the first marine-derived clinical agent has recently been recommended for approval for treatment of soft tissue sarcoma by the EMEA.[194] This section will cover this compound and two additional marine-derived agents which are in advanced clinical trials.

1. *Yondelis®*

The complex alkaloid ecteinascidin 743 (Figure 8.14) was discovered by the late Kenneth Rinehart[195] concomitantly with the Harbor Branch Oceanographic Institution group led by Amy Wright[196] from the colonial tunicate *E. turbinata*. It was found to have a unique mechanism of action, binding to the minor groove of DNA and interfering with cell division, the genetic transcription processes, and DNA repair machinery. The issue of compound supply, always a problem with marine-sourced materials, was solved by the development of a nice semi-synthetic route from the microbial product cyanosafracin B (Figure 8.14).[197] Under the name Yondelis® ecteinascidin 743 has been granted Orphan Drug designation in Europe and the United States, and was recommended for approval by the European Medicine Evaluation Agency (EMEA) in late July, 2007 for the treatment of soft tissue sarcomas (STS).[194] It is also in a Phase III trial in ovarian cancer and in phase II trials for breast, prostate, and pediatric sarcomas.

2. *Other marine agents*

The complex marine polyether halichondrin B (Figure 8.15) was first isolated from a Japanese sponge by Uemura *et al.*,[198] and was subsequently reisolated by Pettit from an *Axinella* species collected in Palau.[199] The compound had excellent bioactivity and showed a pattern of activity

FIGURE 8.13 Taxane anticancer agents.

FIGURE 8.14 Ecteinascidin 743 (Yondelis®) and its semi-synthetic precursor.

FIGURE 8.15 Other marine anticancer agents.

in the NCI 60-cell line screen comparable to the vinca alkaloids and paclitaxel.[199] The compound was only isolated in miniscule yield, however, and its complex structure appeared to make total synthesis impractical as a source for drug development. Fortunately, in the course of synthetic studies on the synthesis of halichondrin B,[200] the group at Eisai Research Institute in the United States, working closely with Kishi's group, discovered that certain macrocyclic ketone analogs of the right hand half of halichondrin B retained all or most of the potency of the parent compound.[201] This key observation was then used as the impetus for the heroic large-scale synthesis of the analog E7389 (Eribulin; Figure 8.15), which is currently in Phase III clinical trials with an NDA filing scheduled for late 2007. The discovery and development of E7389 (Eribulin) has been described in a recent book chapter.[133]

Bryostatin 1 (Figure 8.15) is a complex macrolide natural product originally discovered by Pettit and his collaborators under an early NCI program.[202] It has excellent anticancer activity which is due, at least in part, to its ability to interact with protein kinase C (PKC) isozymes. Compound supply as usual proved to be a major problem, but enough cGMP-grade material could be isolated from wild collections to supply material for clinical trials.[203]

FIGURE 8.16 Natural antimalarial agents and analogs.

Initial clinical results indicated that the drug would be most effective in combination therapy, and several Phase II trials of this nature are in progress.[204]

H. Antimalarial agents

Malaria is a major scourge of mankind, and the discovery of new antimalarial drugs is a worldwide health imperative. The alkaloid quinine was the first effective antimalarial agent to be discovered, and it served mankind well for about 300 years, although resistance to the drug was first noted in 1910. It was largely replaced in the mid 20th century by the synthetic analog chloroquine, but resistance to this drug emerged in 1957, and it is no longer of value in many areas of the world.[205]

The discovery of artemisinin (Figure 8.16) by Chinese scientists in 1971 provided an exciting new natural product lead compound,[206] and artemisinin is now used for the treatment of malaria in many countries. Its unusual endoperoxide bridge is a key to its mechanism of action, which involves complexation with haemin by coordination of the peroxide bridge with iron. This in turn interrupts the detoxification process used by the parasite and generates free radical species which can attack proteins in the parasite.

Many analogs of artemisinin have been prepared in attempts to improve its activity and utility.[207] Two of the most promising of these are the totally synthetic analog OZ277 (Figure 8.16),[208] and the dimeric analog (Figure 8.16). Single doses of the latter compound were shown to cure malaria-infected mice, while corresponding treatments with artemisinin were much less effective.[209]

I. Other natural products

The examples given above are simply a selection of the natural product and natural product analogs that have entered clinical use. There are several recent reviews that cover natural products as drugs and sources of structures.[8,166,169,170,210–215] These references should be consulted for further examples of how natural products have led to novel drugs for a multiplicity of diseases, and for insights into the future potential of natural products in drug discovery.

V. FUTURE DIRECTIONS IN NATURAL PRODUCTS AS DRUGS AND DRUG DESIGN TEMPLATES

A. Introduction

The probability that a directly isolated natural product (e.g. adriamycin or taxanes in the antitumor area) will be the drug for a given disease in the future is relatively low except perhaps in the realm of antibiotics. However, the use of these "base molecules" as structures to be synthesized in a "combinatorial fashion" as leads to both variants of a known active structure, and in the production of novel structures are definitely viable approaches. In a similar fashion, combinatorial biosynthesis can be utilized to produce what are now being called "unnatural natural products" where the biosynthetic machinery of a microbial cell is dissected and the relevant genes are "mixed and matched" followed by expression in a suitable heterologous host. Such compounds may be used in their own right or could be the starting materials for further synthetic modifications. In addition, novel methods of chemical syntheses that have the potential to produce base "natural product" molecules that can be optimized for specific medicinal chemistry purposes are now being reported. That these ideas are not just pipe dreams can be seen in the following examples.

B. Combinatorial chemistry

Combinatorial chemistry is a technique originally developed for the synthesis of large chemical libraries for

high-throughput screening against biological targets, and it has permitted high-throughput approaches to the synthesis of very large libraries (millions) of compounds. Overall, however, this approach to drug discovery over the past decade has been disappointing, with some of the earlier libraries being described as "poorly designed, impractically large, and structurally simplistic,"[136] and "an initial emphasis on creating mixtures of very large numbers of compounds has largely given way in industry to a more measured approach based on arrays of fewer, well-characterized compounds" with "a particularly strong move toward the synthesis of complex natural-product-like compounds – molecules that bear a close structural resemblance to approved natural-product-based drugs."[136] A similar opinion was expressed by Waldmann, who stated "Biological investigation of the million compound speculative combinatorial libraries of the first generation yielded disappointingly low hit rates."[216] He then goes on to cite the example of a library of 76 analogs which was synthesized based on the nakijiquinone C structure and investigated for activity against various kinases. This work uncovered some compounds that inhibit the tyrosine kinase Tie-2 and also yielded a compound that modulates the activity of the VEGF-3 receptor.[217]

Although simple combinatorial chemistry approaches have been disappointing, the use of natural products or natural product-like scaffolds as templates for combinatorial chemistry has been a fruitful approach to drug discovery. The synthesis of natural-product-like libraries is exemplified by the work of the Schreiber group in the synthesis of libraries of over 1,000 compounds bearing significant structural and chiral diversity.[135] As examples of other related work Waldmann *et al.* prepared a pepticinnamin E library by solid-phase synthesis.[218] Solid-phase synthesis of combinatorial libraries of epothilones was used to probe regions of the molecule important to retention or improvement of activity,[138] and combinatorial synthesis of vancomycin dimers has yielded compounds with improved activity against drug-resistant bacteria.[219] The importance of natural products as leads for combinatorial synthetic approaches is embodied in the concept of "privileged structures," originally proposed by Evans *et al.*[139] and then advanced recently by Nicolaou *et al.*[140–142] and Waldmann *et al.*[143,144] Nicolaou[140] stated the underlying thesis as follows: "We were particularly intrigued by the possibility that using scaffolds of natural origin, which presumably have undergone evolutionary selection over time, might confer favorable bioactivities and bioavailabilities to library members." Natural products will thus continue to make contributions to the area of combinatorial chemistry, and will no doubt serve as the foundation of new drugs produced by the synergy of combinatorial chemistry and natural products chemistry. As a very recent example from Waldmann *et al.*, they recently reported the enantiospecific syntheses of a class of simple α,β-unsaturated δ-lactones based on molecules such as leptomycin and pironetin that inhibited the transport of vesicular stomatitis virus to infect host cells.[220]

C. Natural products as design templates

Natural products can also be used as design templates for more traditional chemical syntheses. This approach has been highly successful, as illustrated by some of the examples discussed in the previous sections where a modified natural product became the drug rather than the natural product itself. This section will thus simply give some additional examples to illustrate the power of the approach. Wipf *et al.* prepared some highly modified analogs of the antimitotic natural product curacin A, and found a simpler analog which was more potent than curacin A in inhibiting the assembly of tubulin.[221] In recent work, Crowley and Boger redesigned the vancomycin structure to make it less susceptible to the development of drug resistance;[222] this work provides an interesting parallel to the work from the Nicolaou group mentioned above that used combinatorial chemistry to achieve the same goal.[219]

D. Interactions of microbial sources, genomics, and synthetic chemistry

There have been many significant advances in the knowledge of the microbial and other genomes and the methods that can be used to "mix and match" parts and whole gene clusters with the aim of expressing previously unrecognized metabolites such as bryostatin from as yet uncultured symbionts of *Bugula neritina*.[223] To date, aside from the work with epothilones referred to earlier, where only a terminal enzyme was deleted, combinatorial biosynthetic systems have not, as yet, been used for production of drug entities/precursors, though this will no doubt occur in due course.

However, what is becoming quite evident is that searching older sources for novel agents by utilization of new screens, frequently genetically modified organisms/cell lines, and then coupling these to new synthetic methods, has led to very interesting compounds. In the field of antibiotics from microbial sources, the last few years have seen a veritable explosion of novelty in compounds and screens. As a result of the rapid evolution of genomic sequencing and the ever dropping costs of performing such studies, with the $1,000 genome most likely being achievable in a relatively short time frame, the amount of genomic information is ever increasing. How such information can be utilized is shown below using a small number of relevant examples where novel screens/old sources, or novel sources/novel screens/approaches have led to new structural classes that are only now being investigated.

In 2006, microbial and chemical groups at Merck demonstrated that by screening their older microbial extract libraries against novel screens utilizing antisense technologies,

FIGURE 8.17 New compounds from a variety of approaches.

three entirely new chemical structural classes were identified. The first two, utilizing a screen for *FabF/H* inhibitors (coding for the β-ketoacyl carrier protein synthase I/II) yielded platensimycin (Figure 8.17) from *S. platensis*,[28] platencin (Figure 8.17)[224] and phomallenic acid C (Figure 8.17) from a *Phoma* species.[225] The third, also utilizing antisense technologies but this time directed against ribosomal protein synthesis, specifically ribosomal protein S4, led to the identification of lucensimycin A (Figure 8.17) from *S. lucensis*.[226]

Within a year of the publication of the structure of platensimycin, Nicolaou's group had published both racemic[227] and asymmetric[228] syntheses of this molecule and recently expanded the base structure by constructing the adamantyl derivative of platensimycin (Figure 8.17) with the aim of substituting the more accessible adamantyl substituent in place of the parent caged ketolide.[229] When resolved, the (−)-adamantaplatensimycin (Figure 8.17) exhibited comparable activity to (−)-platensimycin against methicillin resistant *S. aureus* and vancomycin-resistant *Enterococcus faecium*, with the other enantiomers being inactive. This work demonstrates that if a novel structure is isolated from a natural source, there are synthetic chemists just waiting to synthesize and modify these structures, and in ways that will permit scale-up.

As a result of the genomic explosion alluded to earlier, it is now quite evident that the earlier postulates of Zahner[230] were fundamentally correct. What has become evident over the last five to ten or so years, initially from the work of Hopwood, is that the genome of the Streptomycetes and by extension, Actinomycetes in general, contain large numbers of previously unrecognized secondary metabolite clusters.

If these had been expressed under a specific but unrecognized fermentation condition, they were not observed in the screens then used (usually a growth inhibition assay). Perhaps the best current example of this is the work by Banskota *et al*., who investigated the genome of the well known vancomycin producer, *Amycolatopsis orientalis* ATCC 43491 and isolated the novel antibiotic ECO-0501 (Figure 8.17) which was only found by using the genomic sequence to predict the molecular weight and then looking for the molecule directly by HPLC–MS. The compound had a very similar biological profile to vancomycin and was masked by this compound.[231]

Many more examples of the value of this type of investigation are in the literature with two recent reviews[165,232] giving up to date information on the manifold structures that can be found by expression of environmental DNA. This work was pioneered by Handelsman in her studies on zwittermicin production by use of bacterial artificial chromosomes in 1999,[233] which followed on from work on this antibiotic originally published in 1994.

Even the myxobacteria, which are possibly the most prolific organisms in terms of unusual secondary metabolite production, have now yielded to genomic analyses, with two recent reviews being published by Muller's group at the University of the Saarland. The earlier, dealing with the general aspects of genomics on natural product research[234] is complementary to the two mentioned above (Clardy/Lam). The most recent one[235] deals with the major problem in secondary metabolite expression, whether in homologous or heterologous hosts, which is the identification and application of the transcriptional control mechanisms involved. Using genomic analyses, Muller's group has been able to identify and utilize *ChiR*, the gene controlling production of chivosazol, an extremely potent eukaryotic antibiotic.[235]

In conclusion, the examples given in this section are merely a small portion of the immense amount of information that is currently in the literature. It is hoped that this information will lead to novel methods of efficiently generating new agents from many natural sources. Once this has been achieved, an appropriate combination of microbial fermentation, combinatorial biosyntheses or total chemical syntheses will make the original compounds and derivatives available to assist in the generation of new medicinal or other agents for use in the human, veterinary, and agricultural arenas.

VI. SUMMARY

The preceding pages have given just a glimpse of the importance of natural products as both pharmaceutical agents and/or as leads to active molecules. With the advent of novel screening systems related to the explosion of genetic information now becoming available, it will be necessary to rapidly identify novel lead structures. Our belief is that a very significant portion of these will continue to be natural product derived. It should be remembered that Mother Nature has had 3 billion years to refine her chemistry and we are only now scratching the surface of the potential structures that are there.

Due to the relative ease of access to plants, plant-derived materials have been in the majority as far as sources are concerned, with microbial sources being especially important in the antibiotic area. Recent work suggests that marine organisms will play an increasingly important role in the future, especially with the increasing power of organic synthesis to address the supply problems inherent with this source material. In the future, with the advent of genetic techniques that permit the isolation and expression of biosynthetic cassettes, microbes and their marine invertebrate hosts may well be the new frontier for natural products.

Mother Nature has the compounds, it is our job to find and develop them for the good of all.

REFERENCES

1. Borchardt, J. K. The beginnings of drug therapy: Ancient mesopotamian medicine. *Drug News Perspect* 2002, *15*, 187–192.
2. Huang, K. C. *The Pharmacology of Chinese Herbs*, 2nd Edition. CRC Press: Boca Raton, FL, 1999.
3. Kapoor, L. D. *CRC Handbook of Ayurvedic Medicinal Plants*. CRC Press: Boca Raton, FL, 1990.
4. Dev, S. Ancient-modern concordance in ayurvedic plants: some examples. *Environ. Health Persp.* 1999, *107*, 783–789.
5. Johnson, T. *CRC Ethnobotany Desk Reference*. CRC Press: Boca Raton, FL, 1999.
6. Moerman, D. E. *Medicinal Plants of Native America*, Vol. 1. University of Michigan Museum of Anthropology: Ann Arbor, MI, 1986.
7. Farnsworth, N. R., Akerele, R. O., Bingel, A. S., Soejarto, D. D., Guo, Z. Medicinal plants in therapy. *Bull. World Health Org.* 1985, *63*, 965–981.
8. Newman, D. J., Cragg, G. M. Natural products as sources of new drugs over the last 25 years. *J. Nat. Prod.* 2007, *70*, 461–477.
9. Class, S. Pharma overview. *Chem. Eng. News* 2002, *80*(48), 39–49.
10. Rouhi, M. Rediscovering natural products. *Chem. Eng. News* 2003, *81*(41), 77–91.
11. Borman, S. Improving efficiency. *Chem. Eng. News* 2006, *84*(25), 56–78.
12. Borman, S. Organic lab sparks drug discovery. *Chem. Eng. News.* 2002, *80*(40), 23–24.
13. Wilson, R. M., Danishefsky, S. J. Applications of total synthesis to problems in neurodegeneration: Fascinating chemistry along the weay. *Acc. Chem. Res.* 2006, *39*, 539–549.
14. Haslam, E. Secondary metabolism – fact and fiction. *Nat. Prod. Rep.* 1986, *3*, 217–249.
15. Williams, D. H., Stone, M. J., Hauck, P. R., Rahman, S. K. Why are secondary metabolites (natural products) biosynthesized?. *J. Nat. Prod.* 1989, *52*, 1189–1208.
16. Harborne, J. B. Role of secondary metabolites in chemical defence mechanisms in plants. In *Bioactive Compounds from Plants* (Chadwick, D. J., Marsh, J., Eds). John Wiley & Sons: Chichester, 1990, pp. 126–139.
17. Schröeder, F. C., Farmer, J. J., Attygalle, A. B., Smedley, S. R., Eisner, T., Meinwald, J. Combinatorial chemistry in insects: a library of defensive macrocyclic polyamines. *Science* 1998, *281*, 428–431.

18. Bulaj, G., Buczek, O., Goodsell, I., Jiminez, E. C., Kranski, J., Nielsen, J. S., Garrett, J. E., Olivera, B. M. Efficient oxidative folding of conotoxins and the radiation of venomous cone snails. *Proc. Natl. Acad. Sci. USA* 2003, *100*, 14562–14568.
19. Wallace, M. S. Ziconotide: a new nonopioid intrathecal analgesic for the treatment of chronic pain. *Expert Rev. Neurother.* 2006, *6*, 1423–1428.
20. Czaran, T. L., Hoekstra, R. F., Pagie, L. Chemical warfare between microbes promotes biodiversity. *Proc. Natl. Acad. Sci. USA* 2002, *99*, 786–790.
21. Bais, H. P., Vepachedu, R., Gilroy, S., Callaway, R. M., Vivanco, J. M. Allelopathy and exotic plant invasion: from molecules and genes to species interactions. *Science* 2003, *301*, 1377–1380.
22. Fitter, A. Making allelopathy respectable. *Science* 2003, *301*, 1337–1338.
23. Everts, S. Bacterial conversations. *Chem. Eng. News* 2006, *84*(43), 17–26.
24. Rice, S. A., Mcdougald, D., Kumar, N., Kjelleberg, S. The use of quorum-sensing blockers as therapeutic agents for the control of biofilm-associated infections. *Curr. Opin. Investig. Drugs* 2005, *6*, 178–184.
25. Geske, G. D., Wezeman, R. J., Siegel, A. P., Blackwell, H. E. Small molecule inhibitors of bacterial quorum sensing and biofilm formation. *J. Am. Chem. Soc.* 2005, *127*, 12762–12763.
26. Meijer, L. Le cycle de division cellulaire et sa regulation. *Oncologie* 2003, *5*, 311–326.
27. Newman, D. J., Cragg, G. M., Snader, K. M. The influence of natural products upon drug discovery. *Nat. Prod. Rep.* 2000, *17*, 215–234.
28. Wang, J., Soisson, S. M., Young, K., Shoop, W., Kodali, S., Galgoci, A., Painter, R., Parthasarathy, G., Tang, Y. S., Cummings, R., Ha, S., Dorso, K., Motyl, M., Jayasuriya, H., Ondeyka, J., Herath, K., Zhang, C., Hernandez, L., Allocco, J., Basilio, A., Tormo, J. R., Genilloud, O., Vicente, F., Pelaez, F., Colwell, L., Lee, S. H., Michael, B., Felcetto, T., Gill, C., Silver, L. L., Hermes, J. D., Bartizal, K., Barrett, J., Schmatz, D., Becker, J. W., Cully, D., Singh, S. B. Platensimycin is a selective FabF inhibitor with potent antibiotic properties. *Nature* 2006, *441*, 358–361.
29. Tremblay, J.-F. Working the library. Singapore's Merlion builds a business out of its rich library of natural compounds. *Chem. Eng. News* 2006, *84*(10), 44.
30. Downes, D., Laird, S. A., Klein, C., Carney, B. K. Biodiversity prospecting contract. In *Biodiversity Prospecting: Using Genetic Resources for Sustainable Development* (Reid, W. V., Laird, S. A., Meyer, C. A., Gamez, R., Sittenfeld, A., Janzen, D. H., Gollin, M. A., Juma, C., Eds). World Resources Institute: Washington, DC, 1993, pp. 255–287.
31. ten Kate, K., Wells, A. Benefit-sharing case study. The access and benefit-sharing policies of the United States National Cancer Institute: a comparative account of the discovery and development of the drugs calanolide and topotecan. Submission to the Executive Secretary of the Convention on Biological Diversity. The Royal Botanic Gardens: Kew, UK, 1998.
32. ten Kate, K., Laird, S. A. *The commercial use of biodiversity. Access to genetic resources and benefit-sharing*. Earthscan Publications Ltd.: London, UK, 1999.
33. Rosenthal, J. P. Drug discovery, economic development and conservation: the international cooperative biodiversity groups. *Pharm. Biol.* 1999, *37*, 5.
34. Balick, M. J. Ethnobotany and the identification of therapeutic agents from the rainforest. In *Bioactive Compounds from Plants* (Chadwick, D. J., Marsh, J., Eds). John Wiley & Sons: Chichester, 1990, pp. 22–39.
35. Cox, P. A. Ethnopharmacology and the search for new drugs. In *Bioactive Compounds from Plants* (Chadwick, D. J., Marsh, J., Eds). John Wiley & Sons: Chichester, 1990, pp. 40–55.
36. Farnsworth, N. R. The role of ethnopharmacology in drug development. In *Bioactive Compounds from Plants* (Chadwick, D. J., Marsh, J., Eds). John Wiley & Sons: Chichester, 1990, pp. 2–21.
37. Plotkin, M. J. Conservation, ethnobotany and the search for new jungle medicines: pharmacognosy comes of age again. *Pharmacotherapy* 1988, *8*, 257–262.
38. Balandrin, M. F., Kinghorn, A. D., Farnsworth, N. R. In *Human Medicinal Agents from Plants* (Kinghorn, A. D., Balandrin, M. F., Eds), Vol. 534. American Chemical Society: Washington, DC, 1993, pp. 2–12.
39. Raskin, I., Ribnicky, D. M., Komarnytsky, S., Ilic, N., Poulev, A., Borisjuk, N., Brinker, A., Moreno, D. A., Ripoll, C., Yakoby, N., O'Neal, J. M., Cornwell, T., Pastor, I., Fridlander, B. Plants and human health in the twenty-first century. *Trends Biotechnol.* 2002, *20*, 522–531.
40. Newman, D. J., Cragg, G. M. Marine natural products and related compounds in clinical and advanced preclinical trials. *J. Nat. Prod.* 2004, *67*, 1216–1238.
41. Newman, D. J., Hill, R. T. New drugs from marine microbes: the tide is turning. *J. Ind. Microbiol. Biotechnol.* 2006, *33*, 539–544.
42. Diyabalanage, T., Amsler, C. D., McClintock, J. B., Baker, B. J. Palmerolide A, a cytotoxic macrolide from the antarctic tunicate *Synoicum adareanum*. *J. Am. Chem. Soc.* 2006, *128*, 5630–5631.
43. Nicolaou, K. C., Ramakrishna, G., Sun, Y.-P., Banerji, B., Chen, D. Y. K. Synthesis of the originally proposed and revised structures of palmerolide A. *Angew. Chem. Int. Ed.* 2007, *46*, 5896–5900.
44. Perry, N. B., Ettouati, L., Litaudon, M., Blunt, J. W., Munro, M. H. G. Alkaloids from the antarctic sponge *Kirkpatrickia varialosa*. Part 1: variolin B, a new antitumour and antiviral compound. *Tetrahedron* 1994, *50*, 3987–3992.
45. Trimurtulu, G., Faulkner, D. J., Perry, N. B., Ettouati, L., Litaudon, M., Blunt, J. W., Munro, M. H. G., Jameson, G. B. Alkaloids from the antarctic sponge *Kirkpatrickia varialosa*. Part 2: variolin A and *N*(3′)-methyl tetrahydrovariolin B. *Tetrahedron* 1994, *50*, 3993–4000.
46. Ahaidar, A., Fernández, D., Danelón, G., Cuevas, C., Manzanares, I., Albericio, F., Joule, J. A., Álvarez, M. Total syntheses of variolin B and deoxyvariolin B1. *J. Org. Chem.* 2003, *68*, 10020–10029.
47. Poulev, A., O'Neal, J. M., Logendra, S., Pouleva, R. B., Timeva, V., Garvey, A. S., Gleba, D., Jenkins, I. S., Halpern, B. T., Kneer, R., Cragg, G. M., Raskin, I. Elicitation, a new window into plant chemodiversity and phytochemical drug discovery. *J. Med. Chem.* 2003, 2542–2547.
48. McCoy, E., O'Connor, S. E. Directed biosynthesis of alkaloid analogs in the medicinal plant. *Catharanthus roseus*. *J. Am. Chem. Soc.* 2006, *128*, 14276–14277.
49. Pace, N. R. A molecular view of microbial diversity and the biosphere. *Science* 1997, *276*, 734–740.
50. Madigan, M. T., Martinko, J. M., Parker, J. B. *Biology of Microorganisms*, 8th Edition. Prentice-Hall: Upper Saddle River, NJ, 1996.
51. McAlpine, J. B., Bachmann, B. O., Piraee, M., Tremblay, S., Alarco, A.-M., Zazopoulos, E., Farnet, C. M. Microbial genomics as a guide to drug discovery and structural elucidation: ECO-02301, a novel antifungal agent, as an example. *J. Nat. Prod.* 2005, *68*, 493–496.
52. Lautru, S., Deeth, R. J., Bailey, L., Challis, G. M. Discovery of a new peptide natural product by *Streptomyces coelicolor* genome mining. *Nat. Chem. Biol.* 2005, *1*, 265–269.
53. Zengler, K., Toledo, G., Rappe, M., Elkins, J., Mathur, E. J., Short, J. M., Keller, M. Cultivating the uncultured. *Proc. Nat. Acad. Sci. USA* 2002, *99*, 15681–15686.
54. Udwary, D. W., Zeigler, L., Asolkar, R. N., Singan, V., Lapidus, A., Fenical, W., Jensen, P. R., Moore, B. S. Genome sequencing reveals comples secondary metabolome in the marine actinomycete *Salinispora tropica*. *Proc. Natl. Acad. Sci. USA* 2007, *104*, 10376–10381.
55. Gontang, E. A., Fenical, W., Jensen, P. R. Phylogenetic Diversity of Gram-positive bacteria cultured from marine sediments. *Appl. Environ. Microbiol.* 2007, *73*, 3272–3282.
56. Abe, F., Horikoshi, K. The biotechnological potential of piezophiles. *Trends Biotechnol.* 2001, *19*, 102–108.
57. Amato, A. Microbes live near undersea CO_2 lake. *Chem. Eng. News* 2006, *84*(38), 14.

58. Persidis, A. Extremophiles. *Nat. Biotechnol.* 1998, *16*, 593–594.
59. Rossi, M., Ciaramella, M., Cannio, R., Pisani, F. M., Moracci, M., Bartolucci, S. Extremophiles 2002. *J. Bacteriol.* 2003, *185*, 3683–3689.
60. Short, P. L. New Zealand plays to its strengths. *Chem. Eng. News* 2007, *85*(4), 20–21.
61. Cavicchioli, R., Siddiqui, K. S., Andrews, D., Sowers, K. R. Low-temperature extremophiles and their applications. *Curr. Opin. Biotechnol.* 2002, *13*, 253–261.
62. Gomes, J., Steiner, W. *Food Technol. Biotechnol.* 2004, 42.
63. Hoyoux, A., Blaise, V., Collins, T., D'Amico, S., Gratia, E., Huston, A. L., Marx, J. C., Sonan, G., Zeng, Y. X., Feller, G., Gerday, C. Extreme catalysts from low-temperature environments. *J. Biosci. Bioeng.* 2004, *98*, 317–330.
64. Schiraldi, C., De Rosa, M. The production of biocatalysts and biomolecules from extremophiles. *Trends Biotechnol.* 2002, *20*, 515–521.
65. van den Burg, B. Extremophiles as a source for novel enzymes. *Curr. Opin. Microbiol.* 2003, *6*, 213–218.
66. Wiegel, J., Kevbrin, V. V. Alkalithermophiles. *Biochem. Soc. Trans.* 2004, *32*, 193–198.
67. Johnson, D. B., Hallberg, K. B. The microbiology of acidic mine waters. *Res. Microbiol.* 2003, *154*, 466–473.
68. Stierle, A. A., Stierle, D. B., Kemp, K. Novel sesquiterpenoid matrix metalloproteinase-3 inhibitors from an acid mine waste extremophile. *J. Nat. Prod.* 2004, *67*, 1392–1395.
69. Stierle, D. B., Stierle, A. A., Hobbs, D., Stokken, J., Clardy, J. Berkeleydione and berkeleytrione, new bioactive metabolites from an acid mine organism. *Org. Lett.* 2004, *6*, 1049–1052.
70. Gunatilaka, A. A. L. Natural products from plant-associated microorganisms: distribution, structural diversity, bioactivity, and implications of their occurrence. *J. Nat. Prod.* 2006, *69*, 509–526.
71. Strobel, G., Daisy, B., Castillo, U., Harper, J. Natural products from endophytic microorganisms. *J. Nat. Prod.* 2004, *67*, 257–268.
72. Tan, R. X., Zou, W. X. Endophytes: a rich source of functional metabolites. *Nat. Prod. Rep.* 2001, *18*, 448–459.
73. Castillo, U., Harper, J. K., Strobel, G. A., Sears, J., Alesi, K., Ford, E., Lin, J., Hunter, M., Maranta, M., Ge, H., Yaver, D., Jensen, J. B., Porter, H., Robinson, R., Millar, D., Hess, W. M., Condron, M., Teplow, D. Kakadumycins, novel antibiotics from *Streptomyces sp*. NRRL 30566, an endophyte of *Grevillea pteridifolia*. *FEMS Microbiol. Lett.* 2003, *224*, 183–190.
74. Li, J. Y., Harper, J. K., Grant, D. M., Tombe, B. O., Bashyal, B., Hess, W. M., Strobel, G. A. Ambuic acid, a highly functionalized cyclohexenone with antifungal activity from *Pestalotiopsis* spp. and *Monochaetia* sp. *Phytochemistry* 2001, *56*, 463–468.
75. Ezra, D., Castillo, U. F., Strobel, G. A., Hess, W. M., Porter, H., Jensen, J. B., Condron, M. A., Teplow, D. B., Sears, J., Maranta, M., Hunter, M., Weber, B., Yaver, D. Coronamycins, peptide antibiotics produced by a verticillate *Streptomyces* sp. (MSU-2110) endophytic on *Monstera* sp. *Microbiology* 2004, *150*, 785–793.
76. Zhou, G.-K., Wijeratne, E. M. K., Bigelow, D., Pierson, L. S., III, VanEtten, H. D., Gunatilaka, A. A. L. Aspochalasins, I, J, and K: three new cytotoxic cytochalasans of *Aspergillus flavipes* from the rhizosphere of *Ericameria laricifolia* of the Sonoran desert. *J. Nat. Prod.* 2004, *67*, 328–332.
77. Bok, J. W., Hoffmeister, D., Maggio-Hall, L. A., Murillo, R., Glasner, J. D., Keller, N. P. Genomic mining for *Aspergillus* natural products. *Chem. Biol.* 2006, *13*, 31–37.
78. Hoffmeister, D., Keller, N. P. Natural products of filamentous fungi: enzymes, genes, and their regulation. *Nat. Prod. Rep.* 2007, *24*, 393–416.
79. Stierle, A., Strobel, G., Stierle, D. Taxol and taxane production by *Taxomyces andreanae*, an endophytic fungus of Pacific yew. *Science* 1993, *260*, 214–216.
80. Li, J.-Y., Sidhu, R. S., Bollon, A., Strobel, G. A. Stimulation of taxol production in liquid cultures of *Pestalotiopsis microspora*. *Mycolog. Res.* 1998, *102*, 461–464.
81. Puri, S. C., Verma, V., Amna, T., Qazi, G. N., Spiteller, M. An endophytic fungus from *Nothapodytes foetida* that produces camptothecin. *J. Nat. Prod.* 2005, *68*, 1717–1719.
82. Amna, T., Puri, S. C., Verma, V., Sharma, J. P., Khajuria, R. K., Musarrat, J., Spiteller, M., Qazi, G. N. Bioreactor studies on the endophytic fungus *Entrophosphora infrequens* for the production of an anticancer alkaloid camptothecin. *Can. J. Microbiol.* 2006, *52*, 189–196.
83. Eyberger, A. L., Dondapati, R., Porter, J. R. Endophyte fungal isolates from *Podophyllum peltatum* produce podophyllotoxin. *J. Nat. Prod.* 2006, *69*, 1121–1124.
84. Puri, S. C., Nazir, A., Chawla, R., Arora, R., Riyaz-ul-Hasan, S., Amna, T., Ahmed, B., Verma, V., Singh, S., Sagar, R., Sharma, A., Kumar, R., Sharma, R. K., Qazi, G. N. The endophytic fungus *Trametes hirsuta* as a novel alternative source of podophyllotoxin and related tetralin lignans. *J. Biotech.* 2006, *122*, 494–510.
85. Guo, B., Li, H., Zhang, L. Isolation of a fungus producing vinblastine. *J. Yunnan Univ.* 1998, *20*, 214–215.
86. Zhang, L. Q., Guo, B., Li, H., Zeng, S., Shao, H., Gu, S., Wei, R. Preliminary study on the isolation of endophytic fungus of *Catharanthus roseus* and its fermentation to produce products of therapeutic value. *Zhong Cao Yao (Chinese Tradit. Herb. Drugs)* 2000, *31*, 805–807.
87. Yang, X., Zhang, L., Guo, B., Guo, S. Preliminary study of a vincristine-producing endophytic fungus isolated from leaves of *Catharanthus roseus*. *Zhong Cao Yao (Chinese Tradit. Herb. Drugs)* 2004, *35*, 79–81.
88. Rouhi, A. M. Deep-sea harvest. *Chem. Eng. News* 2006, *84*(15), 67–70.
89. Mincer, T. J., Jensen, P. R., Kauffman, C. A., Fenical, W. Widespread and persistent populations of a major new marine *Actinomycete taxon* in ocean sediments. *Appl. Environ. Microbiol.* 2002, *68*, 5005–5011.
90. Feling, R. H., Buchanan, G. O., Mincer, T. J., Kauffman, C. A., Jensen, P. R., Fenical, W. Salinosporamide A: a highly cytotoxic proteasome inhibitor from a novel microbial source, a marine bacterium of the new genus *Salinospora*. *Angew. Chem. Int. Ed.* 2003, *42*, 355–357.
91. Kwon, H. C., Kauffman, C. A., Jensen, P. R., Fenical, W. Marinomycins A–D, antitumor-antibiotics of a new structure class from a marine actinomycete of the recently discovered genus "*Marinispora*". *J. Am. Chem. Soc.* 2006, *128*, 1622–1632.
92. Piel, J. Metabolites from symbiotic bacteria. *Nat. Prod. Rep.* 2004, *21*, 519–538.
93. Yu, T.-W., Floss, H. G. Ansamitocins (maytansinoids). In *Anticancer Agents from Natural Products* (Cragg, G. M., Kingston, D. G. I., Newman, D. J., Eds). Taylor and Francis: Boca Raton, FL, 2005, pp. 321–337.
94. Piel, J., Butzke, D., Fusetani, N., Hui, D., Platzer, M., Wen, G., Matsunaga, S. Exploring the chemistry of uncultivated bacterial symbionts: antitumor polyketides of the pederin family. *J. Nat. Prod.* 2005, *68*, 472–479.
95. Piel, J., Hofer, I., Hui, D. Evidence for a symbiosis island involved in horizontal acquisition of pederin biosynthetic capabilities by the bacterial symbiont of *Paederus fuscipes* beetles. *J. Bacteriol.* 2004, *186*, 1280–1286.
96. Piel, J., Hui, D., Wen, G., Butzke, D., Platzer, M., Fusetani, N., Matsunaga, S. Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the marine sponge *Theonella swinhoei*. *Proc. Nat. Acad. Sci. USA* 2004, *101*, 16222–16227.
97. Partida-Martinez, L. P., Hertweck, C. Pathogenic fungus harbours endosymbiotic bacteria for toxin production. *Nature* 2005, *437*, 884–888.
98. Khosla, C. Natural product biosynthesis: a new interface between enzymology and medicine. *J. Org. Chem.* 2000, *65*, 8127–8133.
99. Staunton, J., Weissman, K. J. Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep.* 2001, *18*, 380–416.
100. Walsh, C. T. Polyketide and nonribosomal peptide antibiotics: modularity and versatility. *Science* 2004, *16*, 1805–1810.

101. Walsh, C. T. The chemical versatility of natural-product assembly lines. *Acc. Chem. Res.* 2007, *40*. DOI: 10.1021/ar7000414.
102. Clardy, J., Walsh, C. T. Lessons from natural molecules. *Nature* 2004, *432*, 829–837.
103. Thomas, M. G., Bixby, K. A., Shen, B. Combinatorial biosynthesis of anticancer natural molecules. In *Anticancer Agents from Natural Products* (Cragg, G. M., Kingston, D. G. I., Newman, D. J., Eds). Taylor and Francis: Boca Raton, FL, 2005, pp. 519–552.
104. Julien, B., Shah, S., Ziemann, R., Goldman, R., Katz, L., Khosla, C. Isolation and characterization of the epothilone biosynthetic gene cluster from *Sorangium cellulosum*. *Gene* 2000, *249*, 153–160.
105. Molnar, I., Schupp, T., Ono, M., Zirkle, R. E., Milnamow, M., Nowak-Thompson, B., Engel, N., Toupet, C., Stratmann, A., Cyr, D. D., Gorlach, J., Mayo, J. M., Hu, A., Goff, S., Schmid, J., Ligon, J. M. The biosynthetic gene cluster for the microtubule-stabilizing agents epothilones A and B from *Sorangium cellulosum* So ce90. *Chem. Biol.* 2000, *7*, 97–109.
106. Lau, J., Frykman, S., Regentin, R., Ou, S., Tsuruta, H., Licari, P. Optimizing the heterologous production of epothilone D in *Myxococcus xanthus*. *Biotechnol. Bioeng.* 2002, *78*, 280–288.
107. Beutler, J. A., McKee, T. C., Fuller, R. W., Tischler, M., Cardellina, J. H., II, McCloud, T. G., Snader, K. M., Boyd, M. R. Frequent occurrence of HIV-inhibitory sulfated polysaccharides in marine invertebrates. *Antiviral Chem. Chemother.* 1993, *3*, 167–172.
108. Cardellina, J. H., II, Munro, M. H. G., Fuller, R. W., Manfredi, K. P., McKee, T. C., Tischler, M., Bokesch, H. R., Gustafson, K. R., Beutler, J. A., Boyd, M. R. A chemical-screening strategy for the dereplication and prioritization of HIV-inhibitory aqueous natural-products extracts. *J. Nat. Prod.* 1993, *56*, 1123–1129.
109. Coombes, J. D. *New Drugs from Natural Sources*. IBC Technical Services: London, 1992.
110. Boyd, M. R., Paull, K. D. Some practical considerations and applications of the national cancer institute *in vitro* anticancer drug discovery screen. *Drug Dev. Res.* 1995, *34*, 91–109.
111. Johnson, R. K., Bartus, H. F., Hofmann, G. A., Bartus, J. O., Mong, S.-M., Faucette, L. F., McCabe, F. L., Chan, J. A., Mirabelli, C. K. Discovery of new DNA-reactive drugs. In *In Vitro and In Vivo Models for Detection of New Antitumor Drugs* (Hanka, L. J., Kondo, T., White, R. J., Eds). University of Tokyo Press: Tokyo, 1986, pp. 15–26.
112. Meyer, B. N., Ferrigni, N. R., Putnam, J. E., Jacobsen, L. B., Nichols, D. E., McLaughlin, J. L. Brine shrimp: a convenient general bioassay for active plant constituents. *Planta Med.* 1982, *45*, 31–34.
113. McLaughlin, J. L. Crown gall tumours on potato discs and brine shrimp lethality: two simple bioassays for higher plant screening and fractionation. In *Methods in Plant Biochemistry: Assays for Bioactivity* (Hostettmann, K., Ed.), Vol. 6. Academic Press: San Diego, CA, 1991, pp. 1–32.
114. Klekota, J., Brauner, E., Roth, F. P., Schreiber, S. L. Using high-throughput screening data to discriminate compounds with single-target effects from those with side effects. *J. Chem. Inf. Model* 2006, *46*, 1549–1562.
115. Liu, B., Li, S., Hu, J. Technological advances in high-throughput screening. *Am. J. Pharmacogen.* 2004, *4*, 263–276.
116. Mullin, R. Drug discovery. *Chem. Eng. News* 2004, *82*(30), 23–32.
117. Shimizu, Y. Bioactive marine natural products with emphasis on handling of water-soluble compounds. *J. Nat. Prod.* 1985, *48*, 223–235.
118. Marston, A., Hostettmann, K. Modern separation methods. *Nat. Prod. Rep.* 1991, *8*, 391–413.
119. Gronquist, M., Meinwald, J., Eisner, T., Schroeder, F. C. Exploring uncharted terrain in nature's structure space using capillary NMR spectroscopy: 13 steroids from 50 fireflies. *J. Am. Chem. Soc.* 2005, *127*, 10810–10811.
120. Hostettmann, K., Wolfender, J.-L., Rodriguez, S. Rapid detection and subsequent isolation of bioactive constituents of crude plant extracts. *Planta Med.* 1997, *63*, 2–10.
121. Jaki, B., Franzblau, S., Pauli, G. F. An NMR method towards the routine chiral determination of natural products. *Phytochem. Anal.* 2004, *15*, 213–219.
122. Konishi, Y., Kiyota, T., Draghici, C., Gao, J. M., Yeboah, F., Acoca, S., Jarussophon, S., Purisima, E. Molecular formula analysis by an MS/MS/MS technique to expedite dereplication of natural products. *Anal. Chem.* 2007, *79*, 1187–1197.
123. Lambert, M., Wolfender, J. L., Staerk, D., Christensen, B., Hostettmann, K., Jaroszewski, J. W. Identification of natural products using HPLC-SPE combined with CapNMR. *Anal. Chem.* 2007, *79*, 727–735.
124. Wolfender, J. L., Queiroz, E. F., Hostettmann, K. Phytochemistry in the microgram domain – a LC-NMR perspective. *Mag. Res. Chem.* 2005, *43*, 697–709.
125. Cordell, G. A. Changing strategies in natural products chemistry. *Phytochemistry* 1995, *40*, 1585–1612.
126. Friesen, J. B., Pauli, G. F. G.U.E.S.S. – a generally useful estimate of solvent systems for CCC. *J. Liq. Chromatogr. Rel. Technol.* 2005, *28*, 2777–2806.
127. Pauli, G. F., Jaki, B. U., Lankin, D. C. Quantitative 1H NMR: development and potential of a method for natural product analysis. *J. Nat. Prod.* 2005, *68*, 138–149.
128. Cragg, G. M., Schepartz, S. A., Suffness, M., Grever, M. R. The taxol supply crisis. New NCI policies for handling the large-scale production of novel natural product anticancer and anti-HIV agents. *J. Nat. Prod.* 1993, *56*, 1657–1668.
129. Fujita, Y. Industrial production of shikonin and berberine. In *Applications of Plant Cell and Tissue Culture* (Bock, G., Marsh, J., Eds). John Wiley & Sons: Chichester, 1988, pp. 157–174.
130. Fowler, M. W., Cresswell, R. C., Stafford, A. M. An economic and technical assessment of the use of plant cell cultures for natural product synthesis on an industrial scale. In *Bioactive Compounds from Plants* (Chadwick, D. J., Marsh, J., Eds). John Wiley & Sons: Chichester, 1990, pp. 228–238.
131. Kingston, D. G. I. Taxol and its analogs. In *Anticancer Agents from Natural Products* (Cragg, G. M., Kingston, D. G. I., Newman, D. J., Eds). CRC Press: Boca Raton, FL, 2005, pp. 89–122.
132. Mickel, S. J., Niederer, D., Daeffler, R., Osmani, A., Kuesters, E., Schmid, E., Schaer, K., Gamboni, R., Chen, W. C., Loeser, E., Kinder, F. R., Jr., Konigsberger, K., Prasad, K., Ramsey, T. M., Repic, O., Wang, R. M., Florence, G., Lyothier, I., Paterson, I. Large-scale synthesis of the anti-cancer marine natural product (+)-discodermolide. Part 5: linkage of fragments C_{1-6} and C_{7-24} and finale. *Org. Process Res. Dev.* 2004, *8*, 122–130.
133. Yu, M. J., Kishi, Y., Littlefield, B. A. Discover of E7389, a fully synthetic macrocyclic ketone analog of halichondrin B. In *Anticancer Agents from Natural Products* (Cragg, G. M., Kingston, D. G. I., Newman, D. J., Eds). Taylor and Francis: Boca Raton, FL, 2005, pp. 241–265.
134. Burke, M. D., Berger, E. M., Schreiber, S. L. Generating diverse skeletons of small molecules combinatorially. *Science* 2003, *302*, 613–618.
135. Burke, M. D., Schreiber, S. L. A planning strategy for diversity-oriented synthesis. *Angew. Chem. Int. Ed.* 2004, *43*, 46–58.
136. Borman, S. The many faces of combinatorial chemistry. *Chem. Eng. News* 2003, *81*(43), 45–56.
137. Borman, S. Rescuing combichem. Diversity-oriented synthesis (DOS) aims to pick up where traditional combinatorial chemistry left off. *Chem. Eng. News* 2004, *82*(40), 32–40.
138. Nicolaou, K. C., Roschangar, F., Vourloumis, D. Chemical biology of epothilones. *Angew. Chem. Int. Ed.* 1998, *37*, 2014–2045.
139. Evans, B. E., Rittle, K. E., Bock, M. G., DiPardo, R. M., Fredinger, R. M., Whitter, W. L., Lundell, G. F., Veber, D. F., Anderson, P. S., Chang, R. S. L., Lotti, V. J., Cerino, D. J., Chen, T. B., Kling, P., Kunkel, K. A., Springer, J. P., Hirshfield, J. Methods for drug discovery: development of potent, selective, orally effective cholecystokinin antagonists. *J. Med. Chem.* 1988, *16*, 2235–2246.

140. Nicolaou, K. C., Pfefferkorn, J. A., Roecker, A. J., Cao, G. Q., Barluenga, S., Mitchell, H. J. Natural product-like combinatorial libraries based on privileged structures. 1. General principles and solid-phase synthesis of benzopyrans. *J. Am. Chem. Soc.* 2000, *122*, 9939–9953.
141. Nicolaou, K. C., Pfefferkorn, J. A., Barluenga, S., Mitchell, H. J., Roecker, A. J., Cao, G.-O. Natural product-like combinatorial libraries based on privileged structures. 3. The "libraries from libraries" principle for diversity enhancement of benzopyran libraries. *J. Am. Chem. Soc.* 2000, *122*, 9968–9976.
142. Nicolaou, K. C., Pfefferkorn, J. A., Mitchell, H. J., Roecker, A. J., Barluenga, S., Cao, G. O., Affleck, R. L., Lillig, J. E. Natural product-like combinatorial libraries based on privileged structures. 2. Construction of a 10,000-membered benzopyran library by directed split-and-pool chemistry using nanokans and optical encoding. *J. Am. Chem. Soc.* 2000, *122*, 9954–9967.
143. Breinbauer, R., Vetter, I. R., Waldmann, H. From protein domains to drug candidates – natural products as guiding principles in the design and synthesis of compound libraries. *Angew. Chem. Int. Ed.* 2001, *41*, 2878–2890.
144. Koch, M. A., Waldmann, H. Protein structure similarity clustering and natural product structure as guiding principles in drug discovery. *Drug Discov. Today* 2005, *10*, 471–483.
145. Ferreira, S. H. A bradykinin-potentiating factor present in the venom of *Bothrops jararaca*. *Br. J. Pharmacol.* 1965, *24*, 163–169.
146. Cushman, D. W., Cheung, H. S., Sabo, E. F., Ondetti, M. A. Design of potent competitive inhibitors of angiotensin-converting enzyme. Carboxyalkanoyl and mercaptoalkanoyl amino acids. *Biochemistry* 1977, 5484–5491.
147. Ondetti, M. A., Rubin, B., Cushman, D. W. Design of specific inhibitors of angiotensin-converting enzyme: new class of orally active hypertensive agents. *Science* 1977, *196*, 441–444.
148. Turk, B. Targeting proteases: successes, failures and future prospects. *Nat. Revs. Drug Discov.* 2006, *5*, 785–799.
149. Endo, A., Kuroda, M., Tsujita, Y., Terahara, A., Tamura, C. Physiologically active compounds. *Ger. Pat. 2524355* 1975.
150. Endo, A. Compactin (ML-236B) and related compounds as potential cholesterol-lowering agents that inhibit HMG-CoA reductase. *J. Med. Chem.* 1985, *28*, 401–405.
151. Brown, A. G., Smale, T. C., King, T. J., Hasenkamp, R., Thompson, R. H. Crystal and molecular structure of compactin, a new antifungal metabolite from *Penicillium brevicompactum*. *J. Chem. Soc. Perkin Trans. 1* 1976, *1*, 1165–1170.
152. Endo, A., Monacolin, K. A new hypocholesterolemic agent produced by a *Monascus* species. *J. Antibiot.* 1979, *32*, 852–854.
153. Endo, A., Monacolin, K. A new hypocholesterolemic agent that specifically inhibits 3-hydroxy-3-methylglutaryl coenzyme a reductase. *J. Antibiot.* 1980, *33*, 334–336.
154. Alberts, A. W., Chen, J., Kuron, G., Hunt, V., Huff, J., Hoffman, C., Rothrock, J., Lopez, M., Joshua, H., Harris, E., Patchett, A., Monaghan, R., Currie, S., Stapley, E., Albers-Shonberg, G., Hensens, O., Hirshfield, J., Hoogsteen, K., Liesch, J., Springer, J. Mevinolin: a highly potent competitive inhibitor of hydroxymethylglutaryl-coenzyme A reductase and a cholesterol lowering agent. *Proc. Natl. Acad. Sci. USA* 1980, *77*, 3957–3961.
155. Monaghan, R. L., Alberts, A. W., Hoffman, C. G., Albers-Schonberg, G. Hypocholesteremic fermentation products and process of preparation. *USP 4231938*, 1980.
156. Alberts, A. W., MacDonald, J. S., Till, A. E., Tobert, J. A. Lovastatin. *Cardiol. Drug Dev.* 1989, *7*, 89–109.
157. Vagelos, P. R. Are prescription drug prices high?. *Science* 1991, *252*, 1080–1084.
158. Rosenblum, S. B., Huynh, T., Afonso, A., Davis, H. R., Jr., Yumibe, N., Clader, J. W., Burnett, D. A. Discovery of 1-(4-fluorophenyl)-(3*R*)-[3-(4-fluorophenyl)-(3*S*)-hydroxypropyl]-(4*S*)-(4-hydroxyphenyl)-2-azetidinone (SCH 58235): a designed, potent, orally active inhibitor of cholesterol absorption. *J. Med. Chem.* 1998, *41*, 973–980.
159. Garcia-Calvo, M., Lisnock, J.-M., Bull, H. G., Hawes, B. E., Burnett, D. A., Braun, M. P., Crona, J. H., Davis, H. R., Jr., Dean, D. C., Detmers, P. A., Graziano, M. P., Hughes, M., MacIntyre, D. E., Ogawa, A., O'Neill, K. A., Iyer, S. P. N., Shevell, D. E., Smith, M. M., Tang, Y. S., Makarewicz, A. M., Ujjainwalla, F., Altmann, S. W., Chapman, K. T., Thornberry, N. A. The target of ezetimibe is Niemann-Pick C1-Like 1 (NPC1L1). *Proc. Natl. Acad. Sci. USA* 2005, *102*, 8132–8137.
160. Burnett, J. R., Huff, M. W. Cholesterol absorption inhibitors as a therapeutic option for hypercholesterolemia. *Expert Opin. Investig. Drugs* 2006, *15*, 1337–1351.
161. Tsang, C. K., Qi, H., Liu, L. F., Zheng, X. F. S. Targeting mammalian target of rapamycin (mTOR) for health and diseases. *Drug Discov. Today* 2007, *27*, 112–124.
162. Koehn, F. E. Therapeutic potential of natural product signal transduction agents. *Curr. Opin. Biotech.* 2006, *17*, 631–637.
163. Janin, Y. L. Antituberculosis drugs: ten years of research. *Bioorg. Med. Chem.* 2007, *15*, 2479–2513.
164. Wright, G. D., Sutherland, A. D. New strategies for combating multidrug-resistant bacteria. *Trends Mol. Med.* 2007, *13*, 260–267.
165. Lam, K. S. New aspects of natural products in drug discovery. *Trends Microbiol.* 2007, *15*, 279–289.
166. von Nussbaum, F., Brands, M., Hinzen, B., Weigand, S., Habich, D. Antibacterial natural products in medicinal chemistry – exodus or revival?. *Angew. Chem. Int. Ed.* 2006, *45*, 5072–5129.
167. Cropp, T. A., Wilson, D. J., Reynolds, K. A. Identification of a cyclohexylcarbonyl CoA biosynthetic gene cluster and application in the production of doramectin. *Nat. Biotechnol.* 2000, *18*, 980–983.
168. Hofle, G., Reichenbach, H. Epothilone, a myxobacterial metabolite with promising antitumor activity. In *Anticancer Agents from Natural Products* (Cragg, G. M., Kingston, D. G. I., Newman, D. J., Eds). Taylor and Francis: Boca Raton, FL, 2005, pp. 413–450.
169. Wilson, R. M., Danishefsky, S. J. Small molecule natural products in the discovery of therapeutic agents: the synthesis connection. *J. Org. Chem.* 2006, *1*, 8329–8351.
170. Altmann, K.-H., Gertsch, J. Anticancer drugs from nature-natural products as a unique source of new microtubule-stabilizing agents. *Nat. Prod. Rep.* 2007, *24*, 327–357.
171. Altmann, K.-H., Pfeifer, B., Arseniyadis, S., Pratt, B. A., Nicolaou, K. C. The chemistry and biology of epothilones – the wheel keeps turning. *ChemMedChem* 2007, *2*, 396–423.
172. Pivot, X., Villanueva, C., Chaigneau, L., Nguyen, T., Demarchi, M., Maurina, T., Stein, U., Borg, C. Ixabepilone, a novel epothilone analog in the treatment of breast cancer. *Expert Opin. Invest. Drugs* 2008, *17,* 593–599.
173. Klar, U., Buchmann, B., Schwede, W., Skuballa, W., Hoffman, J., Lichtner, R. B. Total synthesis and antitumor activity of ZK-EPO: the first fully synthetic epothilone in clinical development. *Angew. Chem. Int. Ed.* 2006, *45*.
174. Feyen, F., Gertsch, J., Wartmann, M., Altmann, K.-H. Design and synthesis of 12-aza-epothilones (azathilones) – "non-natural" natural products with potent anticancer activity. *Angew. Chem. Int. Ed.* 2006, *45*, 5880–5885.
175. Tse, M. L., Watts, R. E., Khosla, C. Substrate tolerance of module 6 of the epothilone synthase. *Biochemistry* 2007, *46*, 3385–3393.
176. Arcamone, F. M. Anthracyclines. In *Anticancer Agents from Natural Products* (Cragg, G. M., Kingston, D. G. I., Newman, D. J., Eds). Taylor and Francis: Boca Raton, FL, 2005, pp. 299–320.
177. Hecht, S. M. Bleomycin group antitumor agents. In *Anticancer Agents from Natural Products* (Cragg, G. M., Kingston, D. G. I., Newman, D. J., Eds). Taylor and Francis: Boca Raton, FL, 2005, pp. 357–381.

178. Remers, W. A. The mitomycins. In *Anticancer Agents from Natural Products* (Cragg, G. M., Kingston, D. G. I., Newman, D. J., Eds). Taylor and Francis: Boca Raton, FL, 2005, pp. 475–497.
179. Hamann, P. R., Upeslacis, J., Borders, D. B. Enediynes. In *Anticancer Agents from Natural Products* (Cragg, G. M., Kingston, D. G. I., Newman, D. J., Eds). Taylor and Francis: Boca Raton, FL, 2005, pp. 451–474.
180. Prudhomme, M. Staurosporines and structurally related indolocarbazoles as antitumor agents. In *Anticancer Agents from Natural Products* (Cragg, G. M., Kingston, D. G. I., Newman, D. J., Eds). Taylor and Francis: Boca Raton, FL, 2005, pp. 499–517.
181. Wani, M. C., Taylor, H. L., Wall, M. E., Coggon, P., McPhail, A. T. Plant antitumor agents. VI. The isolation and structure of taxol, a novel antileukemic and antitumor agent from *Taxus brevifolia*. *J. Am. Chem. Soc.* 1971, *93*, 2325–2327.
182. Schiff, P. B., Fant, J., Horwitz, S. B. Promotion of microtubule assembly *in vitro* by taxol. *Nature* 1979, *277*, 665–667.
183. Wang, X., Itokawa, H., Lee, K.-H. Structure–activity relationships of taxoids. In *Taxus: The Genus Taxus* (Itokawa, H., Lee, K.-H., Eds). Taylor and Francis: London, 2003, pp. 298–386.
184. Cragg, G. M., Newman, D. J. A tale of two tumor targets: topoisomerase I and tubulin. The Wall and Wani contribution to cancer chemotherapy. *J. Nat. Prod.* 2004, *67*, 232–244.
185. Dubois, J. Recent progress in the development of docetaxel and paclitaxel analogues. *Expert Opin. Ther. Patents.* 2006, *16*, 1481–1496.
186. Kingston, D. G. I., Newman, D. J. Taxoids: cancer-fighting compounds from nature. *Curr. Opin. Drug Discov. Dev.* 2007, *10*, 130–144.
187. Geney, R., Chen, J., Ojima, I. Recent advances in the new generation taxane anticancer agents. *Med. Chem.* 2005, *1*, 125–139.
188. Kingston, D. G. I., Jagtap, P. G., Yuan, H., Samala, L. The chemistry of taxol and related taxoids. In *Progress in the Chemistry of Organic Natural Products* (Herz, W., Falk, H., Kirby, G. W., Eds), Vol. 84. Springer: Wien, 2002, pp. 53–225.
189. Guenard, D., Gueritte-Voegelein, F., Potier, P. Taxol and taxotere: discovery, chemistry, and structure–activity relationships. *Acc. Chem. Res.* 1993, *26*, 160–167.
190. Green, M. R., Manikhas, G. M., Orlov, S., Afanasyev, B., Makhson, A. M., Bhar, P., Hawkins, M. J. Abraxane, a novel cremophor-free, albumin-bound particle form of paclitaxel for the treatment of advanced non-small-cell lung cancer. *Ann. Oncol.* 2006, *17*, 1263–1268.
191. Lee, K.-H., Xiao, Z. Podophyllotoxin and analogs. In *Anticancer Agents from Natural Products* (Cragg, G. M., Kingston, D. G. I., Newman, D. J., Eds). CRC Press: Boca Raton, FL, 2005, pp. 71–87.
192. Gueritte, F., Fahy, J. The vinca alkaloids. In *Anticancer Agents from Natural Products* (Cragg, G. M., Kingston, D. G. I., Newman, D. J., Eds). CRC Press: Boca Raton, FL, 2005, pp. 123–135.
193. Rahier, N. J., Thomas, C. J., Hecht, S. M. Camptothecin and its analogs. In *Anticancer Agents from Natural Products* (Cragg, G. M., Kingston, D. G. I., Newman, D. J., Eds). CRC Press: Boca Raton, FL, 2005, pp. 5–21.
194. Anonymous, Yondelis recommended for E.U. approval in soft tissue sarcoma. *DailyDrugNews.com (Daily Essentials)* 2007, *July 23, 2007*.
195. Rinehart, K. L., Holt, T. G., Fregeau, N. L., Stroh, J. G., Keifer, P. A., Sun, F., Li, L. H., Martin, D. G. Ecteinascidins 729, 743, 745, 759A, 759B, and 770: Potent antitumor agents from the Caribbean tunicate *Ecteinascidia turbinate*. *J. Org. Chem.* 1990, *55*, 4512–4515.
196. Wright, A. E., Forleo, D. A., Gunawardana, G. P., Gunasekera, S. P., Koehn, F. E., McConnell, O. J. Antitumor tetrahydroisoquinoline alkaloids from the colonial ascidian *Ecteinascidia turbinate*. *J. Org. Chem.* 1990, *55*, 4508–4512.
197. Cuevas, C., Perez, M., Martin, M., Chicharro, J. L., Fernandez-Rivas, C., Flores, M., Francesch, A., Gallefo, P., Zarzuelo, M., de la Calle, F., Garcia, J., Polanco, C., Rodriguez, I., Manzanares, I. Synthesis of ecteinascidin ET-743 and phthalascidin Pt-650 from cyanosafracin B. *Org. Lett.* 2000, *2*, 2545–2548.
198. Hirata, Y., Uemura, D. Halichondrins – antitumor polyether macrolides from a marine sponge. *Pure Appl. Chem.* 1986, *58*, 701–710.
199. Pettit, G. R., Herald, C. L., Boyd, M. R., Leet, J. E., Dufresne, C., Doubek, D. L., Schmidt, J. M., Cerny, R. L., Hooper, J. N. A., Rutzler, K. C. Antineoplastic agents. 219. Isolation and structure of the cell growth inhibitory constituents from the western Pacific marine sponge *Axinella sp*. *J. Med. Chem.* 1991, *34*, 3339–3340.
200. Choi, H.-W., Demeke, D., Kang, F.-A., Kishi, Y., Nakajima, K., Nowak, P., Wan, Z.-K., Xie, C. Synthetic studies on the marine natural product halichondrins. *Pure Appl. Chem.* 2003, *75*, 1–17.
201. Zheng, W., Seletsky, B. M., Palme, M. H., Lydon, P. J., Singer, L. A., Chase, C. E., Lemelin, C. A., Shen, Y., Davis, H., Tremblay, L., Towle, M. J., Salvato, K. A., Wels, B. F., Aalfs, K. K., Kishi, Y., Littlefield, B. A., Yu, M. J. Macrocyclic ketone analogues of halichondrin B. *Bioorg. Med. Chem. Lett.* 2004, *14*, 5551–5554.
202. Pettit, G. R., Herald, C. L., Doubek, D. L., Herald, D. L., Arnold, E., Clardy, J. Isolation and structure of bryostatin 1. *J. Am. Chem. Soc.* 1982, *104*, 6846–6848.
203. Newman, D. J. Bryostatin – from bryozoan to cancer drug. In *Bryozoans in Space and Time* (Gordon, D. P., Smith, A. M., Grant-Mackie, J. A., Eds). NIWA: Wellington, NZ, 1996, pp. 9–17.
204. Newman, D. J. The bryostatins. In *Anticancer Agents from Natural Products* (Cragg, G. M., Kingston, D. G. I., Newman, D. J., Eds). CRC Press: Boca Raton, FL, 2005.
205. Wongsrichanalai, C., Pickard, A. L., Wernsdorfer, W. H., Meshnick, S. R. Epidemiology of drug-resistant malaria. *Lancet Infect. Diseases* 2002, *2*, 209–218.
206. Klayman, D. L. Qinghaosu (artemisinin): an antimalarial drug from China. *Science* 1985, *228*, 1049–1055.
207. O'Neill, P. M., Posner, G. H. A medicinal chemistry perspective on artemisinin and related endoperoxides. *J. Med. Chem.* 2004, *47*, 2945–2964.
208. Vennerstrom, J. L., Arbe-Barnes, S., Brun, R., Charman, S. A., Chiu, F. C. K., Chollet, J., Dong, Y., Dorn, A., Hunziker, D., Matile, H., McIntosh, K., Padmanilayam, M., Santo Tomas, J., Scheurer, C., Scorneaux, B., Tang, Y., Urwyler, H., Sergio, W., Charman, W. N. Identification of an antimalarial synthetic trioxolane drug development candidate. *Nature* 2004, *430*, 900–904.
209. Posner, G. H., Paik, I.-H., Chang, W., Borstnik, K., Sinishtaq, S., Rosenthal, A. S., Shapiro, T. A. Malaria-infected mice are cured by a single dose of novel artemisinin derivatives. *J. Med. Chem.* 2007, *50*, 2516–2519.
210. Cragg, G. M., Newman, D. J. Plants as a source of anti-cancer and anti-HIV agents. *Ann. Appl. Biol.* 2003, *143*, 127–133.
211. Koehn, F. E., Carter, G. T. The evolving role of natural products in drug discovery. *Nat. Drug Discov. Rev.* 2005, *4*, 206–220.
212. Maison, W. Natural products research: Renaissance with strengthened integration of biology and chemistry. *Angew. Chem. Int. Ed.* 2005, *45*, 3000–3002.
213. Newman, D. J., Cragg, G. M. Advanced preclinical and clinical trials of natural products and related compounds from marine sources. *Curr. Med. Chem.* 2004, *11*, 1693–1713.
214. Newman, D. J., Cragg, G. M., Holbeck, S., Sausville, E. A. Natural products and derivatives as leads to cell cycle pathway targets in cancer chemotherapy. *Curr. Cancer Drug Targ.* 2002, *2*, 279–308.
215. Paterson, I., Anderson, A. E. The renaissance of natural products as drug candidates. *Science* 2005, *310*, 451–453.
216. Waldmann, H. At the crossroads of chemistry and biology. *Bioorg. Med. Chem.* 2003, *11*, 3045–3051.
217. Stahl, P., Kissau, L., Mazitschek, R., Giannis, A., Waldmann, H. Natural Product derived receptor tyrosine kinase inhibitors: identification of IGF1R, tie-2, and VEGFR-3 inhibitors. *Angew. Chem. Int. Ed.* 2002, *41*, 1174–1178.

218. Thutewohl, M., Kissau, L., Popkirova, B., Karaguni, I.-M., Nowak, T., Bate, M., Kuhlmann, J., Müller, O., Waldmann, H. Solid-phase synthesis and biological evaluation of a pepticinnamin E library. *Angew. Chem. Int. Ed.* 2002, *41*, 3616–3620.
219. Nicolaou, K. C., Hughes, R., Cho, S. Y., Winssinger, N., Labischinski, H., Endermann, R. Synthesis and biological evaluation of vancomycin dimers with potent activity against vancomycin-resistant bacteria: target-accelerated combinatorial synthesis. *Chem. Eur. J.* 2001, *7*, 3824–3843.
220. Lessman, T., Leuenberger, M. G., Menninger, S., Lopez-Canet, M., Muller, O., Hummer, S., Bormann, J., Korn, K., Fava, E., Zerial, M., Mayer, T. U., Waldmann, H. Natural product-derived modulators of cell cycle progression and viral entry by enantioselective oxa diels-alder reactions on the solid phase. *Chem. Biol.* 2007, *14*, 443–451.
221. Wipf, P., Reeves, J. T., Balachandran, R., Giuliano, K. A., Hamel, E., Day, B. W. Synthesis and biological evaluation of a focused mixture library of analogues of the antimitotic marine natural product curacin A. *J. Am. Chem. Soc.* 2000, *122*, 9391–9395.
222. Crowley, B. M., Boger, D. L. Total synthesis and evaluation of [psi[CH2NH]Tpg4]vancomycin aglycon: Reengineering vancomycin for Dual D-Ala-D-Ala and D-Ala-D-Lac binding. *J. Am. Chem. Soc.* 2006, *128*, 2885–2892.
223. Sudek, S., Lopanik, N. B., Waggoner, L. E., Hildebrand, M., Anderson, C., Liu, H., Patel, A., Sherman, D. H., Haygood, M. G. Identification of the putative bryostatin polyketide synthase gene cluster from "*Candidatus Endobugula sertula*", the uncultivated microbial symbiont of the marine bryozoan *Bugula neritina*. *J. Nat. Prod.* 2007, *70*, 67–74.
224. Jayasuriya, H., Herath, K. B., Zhang, C., Zink, D. L., Basilio, A., Genilloud, O., Diez, M. T., Vicente, F., Gonzalez, I., Salazar, O., Pelaez, F., Cummings, R., Ha, S., Wang, J., Singh, S. B. Isolation and structure of platencin: a FabH and FabF dual inhibitor with potent broad-spectrum antibiotic activity. *Angew. Chem. Int. Ed.* 2007, 4684–4688.
225. Young, K., Jayasuriya, H., Ondeyka, J. G., Herath, K., Zhang, C., Kodali, S., Galgoci, A., Painter, R., Brown-Driver, V., Yamamoto, R., Silver, L. L., Zheng, Y., Ventura, J. I., Sigmund, J., Ha, S., Basilio, A., Vincente, F., Tormo, J. R., Pelaez, F., Youngman, P., Cully, D., Barett, J. F., Schmatz, D., Singh, S. B., Wang, J. Discovery of FabH/FabF inhibitors from natural products. *Antimicrob. Ag. Chemother.* 2006, *50*, 519–526.
226. Singh, S. B., Zink, D. L., Huber, J., Genilloud, O., Salazar, O., Diez, M. T., Basilio, A., Vincente, F., Byrne, K. M. Discovery of lucensimycins A and B from *Streptomyces lucensis* MA7349 using an antisense strategy. *Org. Lett.* 2006, *8*, 5449–5452.
227. Nicolaou, K. C., Li, D., Edmonds, D. J. Total synthesis of platensimycin. *Angew. Chem. Int. Ed.* 2006, *45*, 7086–7090.
228. Nicolaou, K. C., Edmonds, D. J., Li, A., Tria, G. S. Asymmetric total syntheses of platensimycin. *Angew. Chem. Int. Ed.* 2007, *46*, 3942–3945.
229. Nicolaou, K. C., Lister, T., Denton, R. M., Montero, A., Edmonds, D. J. Adamantaplatensimycin: a bioactive analogue of platensimycin. *Angew. Chem. Int. Ed.* 2007, *46*, 4712–4714.
230. Zahner, H., Drautz, H., Weber, W. Novel approaches to metabolite screening. In *Bioactive Microbial Products: Search and Discovery* (Bu'Lock, J. D., Nisbet, L. J., Winstanley, D. J., Eds). Academic Press: London, 1982, pp. 51–70.
231. Banskota, A. H., Mcalpine, J. B., Sørensen, D., Ibrahim, A., Aouidate, M., Piraee, M., Alarco, A. M., Farnet, C. M., Zazopoulos, E. Genomic analyses lead to novel secondary metabolites. Part 3. ECO-0501, a novel antibacterial of a new class. *J. Antibiot.* 2006, *59*, 533–542.
232. Clardy, J., Fischbach, M. A., Walsh, C. T. New antibiotics from bacterial natural products. *Nat. Biotech.* 2006, *24*, 1541–1550.
233. Rondon, M. R., Raffel, S. J., Goodman, R. M., Handelsman, J. Toward functional genomics in bacteria: Analysis of gene expression in *Escherichia coli* from a bacterial artificial chromosome library of *Bacillus cereus*. *Proc. Natl. Acad. Sci. USA* 1999, *96*, 6451–6455.
234. Bode, H. B., Muller, R. The impact of bacterial genomics on natural product research. *Angew. Chem. Int. Ed.* 2005, *44*, 6828–6846.
235. Rachid, S., Gerth, K., Kochems, I., Muller, R. Deciphering regulatory mechanisms for secondary metabolite production in the myxobacterium *Sorangium cellulosum* So ce56. *Mol. Microbiol.* 2007, *63*, 1783–1796.

Chapter 9

Biology Oriented Synthesis and Diversity Oriented Synthesis in Compound Collection Development

Kamal Kumar, Stefan Wetzel, and Herbert Waldmann

I. INTRODUCTION
II. DIVERSITY ORIENTED SYNTHESIS
A. DOS: Principles
B. DOS of small molecule libraries
C. Applications of DOS libraries
III. BIOLOGY ORIENTED SYNTHESIS
A. Introduction
B. The scaffold tree for structural classification of natural products
C. Protein structure similarity clustering
D. BIOS: The combined application of SCONP and PSSC
E. BIOS: Prospects and future directions
IV. CONCLUSION AND OUTLOOK
REFERENCES

Science discovery is an irrational act. It's an intuition which turns out to be reality at the end of it. I see no difference between a scientist developing a marvelous discovery and an artist making a painting.

C. Rubbia (1934–, Italian physicist)

I. INTRODUCTION

Progress in the biological sciences during the last decades has led to a steep increase in knowledge about basic biological processes as well as the factors leading to their misregulation and ultimately the establishment of disease. In particular, the deciphering of the genomes of various organisms including man has provided the scientific community with a wealth of new data. This new data will provide the basis for identifying interaction points by a chemical–biological approach and offer new opportunities for the development of therapies. In both cases small drug-like molecules are required. In chemical biology research, small molecules are used to perturb biological systems with the goal of gaining insight into biological questions by analyzing the difference between perturbed and non-perturbed state. In medicinal chemistry research the goal is to find hit compounds and to progress them to the lead and development candidate stage in order to modify disease states. Although the routes for optimization of compounds differ in these two approaches, the initial steps are highly similar.

In both cases the chemist faces the question which compounds to synthesize for use in biochemical or biological screens. An answer to this question is not readily given and the selection is even made more complicated by the diverse criteria to be met in the subsequent optimization steps. Given this uncertainty chemical biologists, medicinal and organic chemists often resort to the syntheses of compound collections or libraries to provide a set of candidates for initial screening. However, the range of possible structures to select from is enormous. The total number of molecules with 'drug-like' properties has been estimated to be *ca.* 10^{63}.[1,2] The chemical space covered by all these probable drug-like molecules is so enormous that it cannot be comprehensively or systematically explored by organic synthesis. This dilemma has spurred several different approaches

Copyright © 2008, Elsevier Ltd
All rights reserved.

to cover extended regions of chemical structure space or to identify regions of chemistry space with enhanced likelihood of relevance to chemical biology and medicinal chemistry research.

In this chapter, we address two complimentary approaches to compound collection development namely, Diversity Oriented Synthesis (DOS) and Biology Oriented Synthesis (BIOS). For additional, very well-validated approaches such as fragment-based design and the application of *in silico* methods to develop compound libraries the reader is referred to different chapters in this book and to authoritative reviews.[3–6]

II. DIVERSITY ORIENTED SYNTHESIS

A. DOS: Principles

A target oriented synthesis (TOS) is linear and convergent, is generally planned in a retrosynthetic analysis, and aims to move in the direction of complex target structure from simple substrates. In contrast to TOS, the aim of DOS is the facile preparation of collections of structurally complex and diverse compounds from simple starting materials.[7,8] Therefore DOS needs a different planning strategy which should be in the direction of chemical reactions, that is, from reactants toward products, termed *forward synthetic analysis* (Figure 9.1).

Complexity and diversity are important characteristics of a library or compound collection as the eventual target of a compound in phenotypic screens can be any one of a multitude of diverse and complex proteins inside a cell.[9] Therefore, the overall design of the synthetic pathways in DOS should integrate both complexity and diversity. Structural complexity can, for instance, be generated in DOS through the use of tandem or domino reactions, processes involving a pair of reactions in which the product of the first is a substrate for the second. Another method for the introduction of structural complexity in DOS is the use of conformational restriction. This technique provides an effective means to incorporate macrocyclic rings into the compound design.[10]

Diversity in a collection of compounds can be achieved, for example by, (a) generating different scaffolds or skeletons which could further provide more attachment points (skeletal diversity), (b) incorporating elements of functional group diversity around the given cores structures, (c) generating different stereoisomers to access different binding patterns with protein targets (stereochemical diversity, Figure 9.2). To achieve these goals in DOS, the basic element in forward synthetic planning is the transformation of a collection of substrates into a collection of products by performing a number of common chemical transformations. The key to success in this strategy is the inherent reactivity of common sites available in all the substrates which not only transforms all of them to the products but also ensures their further possible common transformations by generating a new set of common reactive sites.

To achieve maximum efficiency, the synthesis pathways in DOS should be not more than three to five steps. To achieve skeletal complexity in DOS, it is critical to identify and to implement complexity-generating reactions that rapidly and efficiently generate complex molecular skeletons. For example, Schreiber *et al.* used tandem Ugi four component condensation-Diels-Alder reaction to generate complexity (Figure 9.3).

B. DOS of small molecule libraries

DOS takes its origins from combinatorial chemistry efforts that mainly employed increasingly sophisticated organic transformations. These efforts began with a

FIGURE 9.1 Comparison of TOS and DOS. (Reproduced, with permission from The Thomson Corporation from Thomas, G. L., Wyatt, E. E., Spring, D. R. Enriching chemical space with diversity-oriented synthesis. Curr. Opin. Drug Discov. Dev. 2006), *9*(6), 700–712. Copyright 2005, The Thomson Corporation.)

complexity-generating reaction to yield a single molecular skeleton having several attachment points followed by a series of diversity generating appending processes (potentially in split-pool format) to attach all possible combinations of building blocks to this common skeleton. This one-synthesis/one-skeleton approach has proven to be highly general and is capable of generating hundreds, thousands or even millions of distinct small molecules.[12–17] For example, a complexity generating, consecutive transesterification-cycloaddition sequence was used to generate, in one step, the tetracyclic skeleton **B** with potential for functionalization through a series of diversity-generating appending processes (Figure 9.4).[17] A Songashira coupling reaction was then used to append a collection of alkynes to the iodoaryl moiety of **B** and generated a collection of more diverse products **C**. The common lactone moiety **C** was

FIGURE 9.2 Forward synthetic analysis in DOS. (Reproduced with permission from American Chemical Society from Lee, D., Sello, J. K., Schreiber, S.L. Pairwise use of complexity-generating reactions in diversity-oriented organic synthesis. *Org. Lett.* 2000, *2*(5), 709–712. Copyright 2000, American Chemical Society.)

FIGURE 9.3 Ugi reaction for generating complexity and diversity in DOS. Additional complexity was incorporated through a subsequent ring-opening/ring-closing metathesis (RCM) reaction to provide products containing two five-membered and two seven-membered rings. (Figure 3).[11]

FIGURE 9.4 Appending processes in DOS product profiles.[18–22] The fact that altering the relative stereochemistry of a given molecule can drastically change its overall shape, and consequently its biological profile, can be taken as an incentive for diversification.

transformed to a collection of new amide products **D**. Similarly, members of this new collection **D** share a common nucleophilic secondary hydroxyl group, thus making them all substrates for a third appending process, that is, coupling with a collection of carboxylic acid building blocks. This series of products-equals-substrates relationships made it possible to generate the complex arrangement of building blocks found in **E** in a highly efficient manner by using split-pool synthesis.

An important (and intellectually challenging) goal in DOS is to develop efficient synthesis pathways that yield products with diverse displays of chemical information in three-dimensional space. To achieve this goal access to *stereochemical and skeletal diversity* is required.

Stereochemical diversity provides multiple relative orientations of the elements in small molecules that interact with macromolecules. The best way to achieve this diversity is through reactions that proceed with enantio- or diastereoselectivity. Since diversity-generating processes involve the transformation of a collection of substrates into a collection of products, it is critical that the processes used to generate new stereogenic centers are both selective and general. These collective transformations of chiral substrates into products having increased stereochemical diversity require powerful reagents that can overcome any substrate bias and yield the highly diastereoselective product profiles.[18–22] The fact that altering the relative stereochemistry of a given molecule can drastically change its overall shape, and consequently its biological profile, can be taken as an incentive for diversification.

Some examples of stereochemically diverse library generation are shown in Figure 9.5. A novel conformational restriction approach was used by Schreiber *et al.* to favor macrocyclization, *via* strategic placement of ester and amide functionalities in a linear precursor. The macrocyclization also provided further diversity points for structural modifications (Figure 9.5a).[10] Panek *et al*. synthesized 14-, 16- and 22-membered macrodiolides bearing up to six stereogenic centers (Figure 9.5b).[23] Oteras distannoxane transesterification catalysts were employed to effect solution-phase cyclohomodimerization of ω-hydroxyesters. The products were obtained in high yields with limited trimer formation. Again, the products could be diversified further using substrate-controlled stereoselective reactions.

Verdine *et al.* used the concept of stereochemical variation and acyclic stereocontrol to generate non-peptidic ligands for peptide receptors (Figure 9.5c).[24] Inspired by an endogenous ligand for mu opoid receptor, endomorphin-2 (Figure 9.5c(**E**)) a stereodiverse collection of non-peptidic compounds **F**, was generated where the *N*-terminal tripeptide unit of **E** has been replaced by a non-peptidic, stereodiverse unit incorporating a 1,5-enediol moiety. The dense array of stereocenters combined with the rigidifying olefin in **F** were intended to generate geometric diversity.

Recently, Schreiber *et al.* have used macrolides and their linear precursors to probe the relative influences of stereochemical and skeletal diversity upon biological function.[25] A library of 122 macrolides and their 122 linear precursors (Figure 9.5d) was evaluated in 40 different cell-based assays. Statistical analysis of the results revealed

FIGURE 9.5 Stereochemically diverse compound collections.

that active macrocyclic compounds were more likely to exhibit activity in one assay, as opposed to multiple assays, providing a quantitative connection between conformational restriction and biological specificity. Hierarchical clustering of the data also identified stereochemistry as a second dominant factor that influenced global activity patterns.

There have been a number of advances in synthesis directed toward generating *skeletal diversity* in DOS. At present, different reagents are used to transform a common substrate with the potential for diverse reactivity into a collection of products having distinct molecular skeletons.[26] These reagent-based skeletal diversity-generating transformations are, therefore, also referred to as *differentiating processes*.

For example, the Fallis-type[27] triene **A** can be transformed into a collection of products with distinct molecular skeletons by the action of different reagents (Figure 9.6).[28] Treatment of **A** with highly reactive, cyclic disubstituted dienophiles such as ethyl maleimide led to double cycloaddition reactions and yielded unsaturated decalin skeletons functionalized with maleimide-derived building blocks (e.g. **B**). Similarly, treatment with a different reagent, a substituted triazole-3,5-dione, produced the unsaturated tetraazadecalin skeleton **C** through a hetero-Diels-Alder reaction. Treatment of **A** with less reactive dienophiles resulted in single cycloaddition reactions and yielded functionalized cyclohexene derivatives such as **D**. Alternatively, treatment of **A** with halogenated quinones resulted in cycloaddition followed by spontaneous dehydrohalogenation and aromatization to yield benzene derivatives such as **E** (Figure 9.6).

In contrast to appending reactions, processes which create different scaffolds from common intermediates have

FIGURE 9.6 A skeletal diversity approach in small molecules library synthesis.

been used only sparingly to generate skeletal diversity in a combinatorial manner. Doing so requires the identification of such structurally differentiating processes with the products-equals-substrates relationship. Thus, all of the skeletally distinct products of one differentiating process must share a common chemical reactivity that makes them all potential substrates for another differentiating process. This type of forward synthetic planning is challenging and will require non-mutually exclusive approaches to the two, potentially conflicting, goals of maximizing structural diversity and maintaining common reactivity.

C. Applications of DOS libraries

Screening of DOS libraries has yielded important new biological probes which have increased our understanding of various biological processes.[29] These probes have been identified from both drug-like and natural product-like libraries, using a variety of screening techniques[30] ranging from cell-free protein binding, enzyme-linked immunosorbent assays (ELISA) and fluorescence resonance energy transfer (FRET) assays to cell-based reporter gene, cytoblot and phenotypic assays.[31]

In a natural product-based DOS library of small molecules, analogs of dysiherabine (Figure 9.7a), a natural product containing γ–γ-disubstituted glutamate, were discovered as ligands for ionotropic glutamate receptors (iGluRs).[32] Understanding the ligand selectivity of different iGluRs is important because of their role in various central nervous system (CNS) related disorders such as Alzheimer's disease and epilepsy. The availability of ligands for iGluRs may prove invaluable in shedding light on the properties and functions of these receptors.

Uretupamine B (Figure 9.7b) was discovered from a DOS library as function-selective suppressor of the yeast signaling protein Ure2p. This protein regulates cellular responses to the quality of carbon and nitrogen nutrients (for example, glucose versus acetate and ammonium versus proline). Ure2p represses the transcription factors Nil1p and Gln3p, and differential regulation is thought to distinguish carbon- and nitrogen-nutrient-responsive signaling. Thus these two effects cannot be separated using Urep2p knockouts (ure2D), whereas a function-selective small molecule inhibitor can rise to this task.[33]

HR22C16 analogs were discovered as new cell-permeable small molecule inhibitors of cell division which act by targeting Eg5, a molecular-motor protein (Figure 9.7c). A library of HR22C16 analogs synthesized by exploiting solid-phase traceless synthesis also yielded Eg5 inhibitors more potent than HR22C16.[34]

FIGURE 9.7 Small molecules originating from DOS as biological probes.

Schreiber *et al.* reported interesting phenotypes recorded after application of the members of a library based on biaryl containing medium rings.[35] Plants treated with (*P*)-4k (Figure 9.7d) were found to exhibit stunted development that eventually led to noticeable pigment loss (potential inhibition of chlorophyll and/or carotenoid biosynthesis) and death. *Danio rerio* (zebra fish) embryos, when treated with compound **13ab** at 100 nM concentration exhibited delayed development. By the second day postfertilization, all fish exhibited decreased pigmentation, weak hearts, abnormal brains, and misshapen jaws. These experiments indicate that small molecules from the library are cell permeable and capable of interacting directly with intracellular protein targets.

Using the methods of DOS, the Shair group synthesized a 10,000-member library of molecules resembling carpanone, a natural product (Figure 9.7e). A series of molecules that act as vesicular traffic inhibitors by inhibiting exocytosis from the Golgi apparatus was discovered from this compound collection.[36]

Combinatorial chemistry involving DOS principles is a useful method for generating compounds with significant potential in chemical biology and medicinal chemistry research. More importantly, DOS combined with phenotype-based screening has emerged as a powerful tool to study biological systems and has led to the discovery of new bioactive molecules.[37] Continued improvements in library design and in the computational assessment of structural diversity in the starting materials will be conducive to further developments of DOS. The increasing number of biological targets being identified in the postgenomic era will also accelerate drug discovery in academia and pharmaceutical industry.

III. BIOLOGY ORIENTED SYNTHESIS

A. Introduction

As outlined above the chemical structure space accessible to small drug-like molecules is so vast that it cannot be covered by chemical synthesis in a comprehensive and meaningful manner. During evolution nature herself has explored only a tiny fraction of chemical space in the biosynthesis of low molecular weight natural products. The same is true for the evolution of the targets bound and modulated by natural products, for example proteins. It has been estimated that during the evolution of a protein consisting of about 100 amino acids only a tiny fraction of all amino acid combinations could have been biosynthesized.[38] However, in protein evolution, structure is even more conserved than the sequence since similar structures can be formed by very different sequences. Thus the protein structure space explored by nature is limited in size.[39]

Since complementarity between a protein binding site and its ligands is required for binding and thus modulation, the structure space explored by nature during the evolution of small molecules as well as proteins should be highly complementary. The regions in chemical space explored by natural products are certainly not the only regions known to be compatible with protein structure space. Still, number and size of such biologically prevalidated regions can be expected to be limited. The crucial question to be answered therefore is: *How does one identify biologically relevant starting points in chemical structure space leading to the development of compound collection enriched in biological activity*? Various attempts at providing answers to this question have been undertaken from the point of view of both the protein and the ligand. For instance, medicinal chemistry efforts have focused on target-based rational ligand design, for example, the development of compounds based on mechanisms of enzyme catalyzed reactions. Other approaches explored inter-protein relationships based on evolutionary arguments, similarity in sequence or function and comparisons based on shape or electrostatic potential of the binding pocket and more abstract representations, for example, molecular interaction fields.[40] Ligand-based methods include pharmacophore search, shape similarity and *in silico* methods to discriminate between drug-like molecules and others.[41–44]

A new approach termed "Biology Oriented Synthesis" (BIOS) has been developed recently by Waldmann *et al*.[45,46] This approach is based on the structural similarity between small bioactive molecules on the one side and their receptors, that is proteins, on the other side as well as on the complementarity of both. BIOS employs compound classes from biologically relevant regions of chemical space, for example natural product or drug space, to select scaffolds as starting points for the design and synthesis of small focused libraries with limited diversity. In this respect BIOS provides a conceptual alternative to other approaches for compound collection design and synthesis such as DOS to which it is complementary. However, BIOS also links this starting point to a cluster of potential target proteins identified on the basis of structural similarity which introduces the criterion of biological prevalidation (Figure 9.8).

To connect chemical and biology space, BIOS employs two concepts developed earlier by Waldmann *et al*. The first approach which addresses the mapping of chemical space leading to the Scaffold Tree and was introduced in the Structural Classification of Natural Products (SCONP).[47] The Scaffold Tree arises from a hierarchical classification of scaffolds and thus allows charting of and navigation in chemical space. The target world, that is proteins, is addressed by the Protein Structure Similarity Clustering (PSSC) approach which identifies clusters of proteins sharing a similar ligand sensing core and thus are prone to bind similar ligands.

These two approaches, the scaffold tree as applied in SCONP as well as PSSC, are introduced and exemplified by selected applications in the following sections. We also demonstrate first applications of BIOS which successfully combine chemistry and biology.

B. The scaffold tree for structural classification of natural products

Natural product space is diverse both in chemical structure and bioactivity and it only partially overlaps with drug space.[48] Almost half of the drugs introduced to the market over the past 30 years can be classified as natural products or natural product analogs.[49] Thus natural products are one source (importantly not the only source, see below) of validated starting points for the design of focused libraries enriched in biologically activity. Moreover, since often natural product classes target more than one protein in a given organism, *structural information for recognition by proteins is built into these natural products. By the very definition of the pharmaceutical industry the underlying scaffold structures of natural product classes are "privileged structures*."[50] In order to chart natural product space and to generate new insights into its structural and biological diversity, Waldmann *et al.* developed the hierarchical classification of natural products. This classification is based on scaffolds which consist of rings and their linkers but no other aliphatic chains. These chemical entities were then deconstructed in a stepwise manner guided by a set of rules invoking medicinal and synthetic organic chemistry. In their initial approach, the data in the Dictionary of Natural Products (DNP) was used.[51] This very comprehensive database contained *ca*. 190,000 entries in the version 14:2 dating from 2005. These were subjected to a normalization procedure which removed entries without structures and counter ions and normalized charges leading to approximately 170,000 structures (Figure 9.9).

In the next step, stereochemical information was removed because literature indicates that in chemical

FIGURE 9.8 Workflow of the BIOS concept incorporation similarity in the protein and the small molecule world and matching both by biological prevalidation.

grouping techniques, for example clustering, two-dimensional approaches perform as well as three-dimensional approaches.[52–54] Moreover, the stereochemistry of many natural products is often not well defined or unknown and the DNP does not contain any stereo-chemical information at all. The missing stereochemistry has to be addressed at a later stage, most likely during synthesis of a compound collection when most of the stereoisomers of a given natural product scaffold or molecule may be investigated.

In the filtered set of natural product molecules many glycosylated structures were found which exhibited similar aglycons. Since glycosylation in many cases mainly is responsible for modulation of bioactivity, solubility, etc. all molecules were subjected to an *in silico* deglycosylation procedure. The deglycosylated structures were then filtered for all molecules containing rings. The focus on ring systems is justified because many acyclic structures in the DNP are lipids and polypeptides which are usually not of prime interest for the design of small molecule inhibitors. Moreover most small molecules used as inhibitors and drugs contain rings or ring systems that rigidify the structure and reduce the loss of entropy upon binding. From the 149,000 molecules containing rings *ca*. 25,000 scaffolds were extracted. The hierarchical classification itself, which led to the tree, used repeating cycles of scaffold deconstruction and parent–child assignment.[47] All possible parent scaffolds containing one ring less than the larger child scaffold were generated. The final parent in one step was selected by a set of criteria including the following rules:

1. The parent scaffold has to be a substructure of the child scaffold.
2. The parent scaffold has to have fewer rings than the child scaffold.
3. Breaking of ring bonds is forbidden.
4. The parent with the highest number of hetero atoms is chosen.
5. The largest parent scaffold is selected.
6. The more frequent parent scaffold, i.e. the one representing more NPs, is selected.

This procedure led to a unique, hierarchical classification of scaffolds, in which each scaffold in the hierarchy is a well-defined chemical entity, which is a substructure of the original molecules. Thereby more complex scaffolds can be reduced to smaller scaffolds creating branches and several branches merge toward the inner circles of the tree. *This procedure leads to a reduction of molecular complexity from multiple annulated rings at the outer rims to single ring systems at the most inner circle of the scaffold tree* (Figure 9.10).

FIGURE 9.9 Flow-chart of the procedure generating scaffold geneaologies from natural product molecules.

FIGURE 9.10 Tree-like graphical representation of natural product scaffolds. For clarity, only scaffolds that cumulatively represent at least 0.2% of the natural products in the DNP are shown.

The initial procedure described allowed such parents which are present themselves as isolated scaffolds in the dataset. This rendered the tree with "holes" which are intermediate scaffolds, missing in the DNP. A new rule set introduced by Schuffenhauer *et al.*[55] includes modifications to allow for parent scaffolds not contained in the dataset. Thus the dissection of scaffolds becomes independent of the dataset, that is, the parent generated from one molecule will always be the same, irrespective of the dataset used. The resulting "virtual scaffolds," that is, scaffolds not represented in the dataset but generated by the dissection approach, may provide interesting opportunities for organic synthesis. The new set of rules is based mostly on medicinal and synthetic chemistry knowledge and is expanded compared to the criteria used in the initial classification.[47] New rules, for example, keep macrocycles intact, retain unusual structural motifs like spiro- or bridged compounds or conserve aromaticity. Some of the 13 rules in the new rule set are shown in Figure 9.11.

The scaffold tree systematizes structural diversity in an intuitive and chemically meaningful way. In the natural product scaffold tree shown in Figure 9.10 three sectors can be identified: oxygen heterocycles, nitrogen heterocycles and carbocycles. Statistical analysis shows that most of the natural products contain between two and four rings and have a Van der Waal's volume between 100 and 500 $Å^3$ (with a maximum at 250 $Å^3$) which is comparable to the volumes found in the World Drug Index[56] and well in the volume range of known protein binding sites. The actual volume of the scaffolds is often smaller than the volumes of cavities in proteins so that there is enough space for substituents "decorating" the scaffolds.[40]

The natural product scaffold tree offers a new approach to analyze and chart chemical space for applications in the development of compound collections both for chemical biology as well as for medicinal chemistry. One possible application is to chart the chemical space of interest and gain insights into abundant and structurally interesting scaffold families present. These can then be used as templates for the design and synthesis of new focused compound collections. For instance, one structurally and biologically interesting class of compounds incorporates the spiroacetal motif (Figure 9.12**A**). A collection of molecules based on this core motif was synthesized in an enantioselective way using two different methodologies. The key steps in generating stereoselectivity were in the first case stereoselective aldol reactions on solid phase[57,58] and in the second one double intramolecular hetero-Michael additions.[59]

α,β-Unsaturated lactones (Figure 9.12**B**) form another interesting class of natural product scaffolds. Ring-closing metathesis (RCM) coupled with several sequential asymmetric allylations proved to be a viable strategy to access these architectures. Enantioselective allylation on the solid support[60] using different chiral auxillaries was developed and sequential allylations followed by RCM and release from the polymeric support yielded different stereoisomers of natural product analogs in high overall yields.[61,62] Extension of this sequence to all possible combinations gave rise to all eight stereoisomers representative for a natural product with three different stereocenters. This example shows that the allylation synthetic methodology on solid phase is powerful enough to generate all stereoisomers in a library fashion. In another example, a stereochemically flexible asymmetric synthesis of the PP2A inhibitor cytostatin and its analogs was reported.[63,64]

In a different approach the hetero-Diels-Alder reaction of oxygen-substituted dienes with a glyoxylate in the presence of a chiral titanium catalyst yielded the desired dehydrolactones in overall yields of 10–40% and with enantiomer- and diastereomer ratios >90%.[65] This compound collection contained new chemotypes which influence the organization of the cytoskeleton leading to phenotypic changes in spindle formation and defects in chromosome alignment. Alternatively, a hetero-Diels-Alder reaction of a resin bound aldehyde with Danishefsky diene-type compounds generated the lactones in high yield and with very high enantiomeric excess using chiral catalysts. The dehydropyrones were further modified on solid phase to yield tetrahydropyran compound collection (Figure 9.12**C**).[66]

Further, a compound collection based on the scaffold of the natural product Furanodictin (Figure 9.12**D**) revealed a previously uncharacterized inhibitor class for the protein tyrosine phosphatases PTP1B and Shp-2.[45] A compound collection which embodies the underlying scaffold structure of alkaloid cytisine and related natural products (Figure 9.12**E**) revealed the first inhibitor class of the vascular endothelial protein tyrosine phosphatase (VEPTP) at a hit rate of 1.57% as well as a completely new class of inhibitors for protein tyrosine phosphatase-1B and the phosphatase Shp-2. These two enzymes are targets for the treatment of the metabolic syndrome and diabetes as well as cancer respectively. In these cases the hit rates were *ca.* 0.3 to 0.4%. The screens revealed selective inhibitors for the proteins.[45]

As further example for a nitrogen containing heterocyclic scaffold from the scaffold tree of natural products a small collection of 40 indolactams (Figure 9.12**F**) was synthesized on solid phase.[67] Biochemical evaluation of these potential ligands of the protein kinase C (PKC) family yielded a selective ligand for PKC delta. To synthesize more complex indole derivatives, indoloquinolizidine alkaloids were targeted (Figure 9.12**G**) and about 500 indoloquinolizidines were synthesized using efficient solid-phase strategy.[45] The indole scaffold itself occurs in many natural products and is also frequently present in synthetic drugs. Using different synthetic methodologies on solid phase, about 400 indole derivatives were synthesized[68–69] which delivered a new class of protein tyrosine kinase inhibitors as well as new ligands for a multi-drug resistance protein.

FIGURE 9.11 A selection of important rules guiding the most recent scaffold tree generating algorithm.

Access to libraries of complex and diverse small molecules demands advancements in synthetic methodologies. There has been a continuous improvement in the existing synthetic methodologies which can be applied to library synthesis in addition to the development of newer and sophisticated reactions for this purpose. Natural product derived and inspired compound collections in general are accessible by application of current synthetic methodology.

FIGURE 9.12 Structural motifs identified in the natural product scaffold tree and used as templates for natural product-based compound collections.

This impression is enforced and supported by various further reports covering additional natural product classes. For a comprehensive overview the reader is referred to additional review articles.[70–73]

C. Protein structure similarity clustering

Proteins are modular entities and many of them consist of several different modules, so called domains. Although there is no clear-cut definition for the term "domain" it may be described as an autonomous folding unit consisting of one single peptide chain.[74] The term "fold" describes the spatial arrangement of secondary structure elements, that is α-helices and β-sheets, relative to each other. Modern structural biology combined with bioinformatics analyses of genome-based data have revealed that protein folds are well conserved in nature and during evolution.[75–78] The SCOP (Structural Classification of Proteins) database predicts about 1,000 folds corresponding to 28,000 entries in the Protein Data Bank (PDB).[79–81] Different fold comparison methods disagree on the total number of folds, and depending on the algorithms used, the estimates of the overall number of folds range from 1,000 to 10,000, that is, there are only very few considering the theoretical number of possibilities. Another feature of the conservatism in protein domain folding is that the distribution of folds is highly non-homogeneous with some folds occurring abundantly and some rarely.[82–87] It has been proposed that a majority of protein domains can be attributed to *ca.* 1,000 most commonly observed folds. Based on this structural conservatism in protein architecture PSSC was developed as a guiding principle for the selection of biologically validated starting points for compound library development.[88–94]

In the protein world structural conservatism and diversity are combined on two different levels: conservatism in the more macroscopic, that is, the structural level and diversity on the microscopic level, that is, the individual amino acid sequence. The fold defines the scaffold of the protein, that is, the 3D structure of the amino acid backbone, as well as the shape and size of the active site and the spatial orientation of the catalytic residues. The individual amino acid side chains forming the active site and its catalytic residues determine the molecular interactions between the protein and the ligand. The same fold can be assembled by amino acid sequences with only as little as a few percent sequence similarity. Thus both, fold and sequence, determine together the binding properties of any protein and enable the vast number of specific functions to be carried out by a limited number of fold types.[95–98]

The analysis of the natural product chemical space described in the previous section revealed a similar dualism in the small molecule world. The number of natural product classes with particular scaffolds was found to be limited.[99] However, any given scaffold type can be decorated with a large number of diverse substituents at different positions theoretically enumerating a large number of possible molecules. Analogous to the protein fold, the scaffolds define the frameworks of the protein ligands while the individual substituents decorating the scaffolds define the molecular interaction between the individual ligand and its target.

This analogy between the protein and the small molecule world led Waldmann *et al.* to the suggestion that there may be complementarity between the proteins and their small molecule ligands at the scaffold level, similar to the complementarity on the atomic level when interaction patterns of individual proteins and their ligands are compared

FIGURE 9.13 Structural conservatism and diversity in proteins and their natural ligands.

(Figure 9.13). This approach may then allow a prospective prediction of small molecule scaffold types most suitable for ligand design. These ligands are targeted at a group of structurally similar proteins by a prediction based exclusively on structure and nature's conservatism during evolution.

However, the individual side chains decorating the small molecule and the protein backbone determine the interactions on the molecular level and thus modulate the binding of a ligand to a protein. Even with similar backbone structures, proteins can display very diverse interaction patterns due to different amino acid sequences. Most likely, a particular natural product will not bind to all structurally similar binding sites found in different proteins. One can imagine an extreme example of two nearly identical binding sites, one of which carries a positively charged residue and the other one a negatively charged amino acid side chain at a similar position in their binding pocket. A negatively charged ligand would in the first case be bound tightly through a salt bridge and in the other case be repelled from the pocket. Consequently, *it is necessary to generate sufficient chemical diversity to match biological diversity in the quest for biologically active molecules/ligands for proteins*.

In the PSSC procedure itself, initially the full structure of a protein of interest was subjected to search for structural similarity using the Dali (FSSP) and Combinatorial Extension (CE) algorithms.[100–102] The searches were performed across the entire PDB and yielded lists of structurally similar proteins ordered by decreasing similarity (Figure 9.14). The entries which were deemed interesting

FIGURE 9.14 Bioinformatics procedure for the identification of PSSCs.

according to different criteria, for example, pharmaceutical relevance or low sequence similarity, were then inspected manually. For this step, ligand sensing cores, that is, the subfolds around the binding sites, were manually isolated and aligned. The cores that showed sufficient similarity in their 3D structures were assigned to a protein structure similarity cluster. Known ligand types for one individual cluster member are regarded as complementary to this subfold and thus used as a template for the development of compound collections targeting all cluster members.

This approach was first applied to the natural product dysidiolide and a known target, the Cdc25A phosphatase[103] which regulates cell-cycle progression and is a target in cancer research.[104–106] PSSC with Cdc25A as a template yielded proteins with similar ligand sensing cores including acetylcholine esterase (AChE) with an RMSD of 2.74 Å over 49 aligned residues and a sequence identity of 8.2%. AChE is involved in signaling in synapses and is a target in the treatment of Alzheimer's disease.[107] 11β-Hydroxysteroid dehydrogenase (11βHSD) was identified as further cluster member with an RMSD of 4.13 Å to Cdc25A over 80 aligned residues and a sequence identity of only 5%. 11βHSDs are involved in the regulation of gene transcription and are targets in the treatment of diabetes.[108–111]

Based on this analysis the hydroxybutenolide part of dysidiolide was used as the guiding structure for the synthesis of a compound collection targeting the two other proteins.[89] The choice of the hydroxybutenolide moiety was based on the assumption that it would be the critical motif for phosphatase inhibition (Figure 9.15). Screening of 150 synthesized compounds revealed 3 hits for AChE with IC_{50} values of 1.7–6.9 μM. Assessment of inhibition of 11βHSD revealed 7 hits with IC_{50} values from 2.4 to 10 μM. For one compound a pronounced selectivity in inhibition of the 11βHSD type 1 over the type 2 isoenzyme was observed. Closer analysis revealed that despite the limited similarity in the ligand sensing cores the catalytic residues in Cdc25A, 11βHSD and AChE are located at similar positions. Thus this approach may also be a viable method to identify and define unknown super sites in proteins.

Another example from a retrospective analysis of literature data using the PSSC concept deals with the development of ligands for the farnesoid X receptor. Selective ligands for this receptor have been found in a 10,000-membered combinatorial library based on a benzopyran core structure synthesized by Nicolaou *et al.* (Figure 16).[112–114] These ligands were used in a chemical genetics approach to unravel the function of the farnesoid X receptor in lipid metabolism.[115]

FIGURE 9.15 Application of the PSSC concept for de novo compound library design. Superimposition of the catalytic cores of AChE (blue) & Cdc25A (green) and Cdc25A (red) & 11βHSD1 (blue). Analogs of the naturally occurring Cdc25A inhibitor Dysdiolide screened for binding to the PSSC member enzymes Cdc25A, AChE and 11βHSD1/2 (IC_{50} values are given).

The farnesoid X receptor is a member of the class of nuclear hormone receptors, which have key roles in development and homeostasis, as well as in many diseases like obesity, diabetes and cancer.[116,117] The ligand-binding domain of the farnesoid X receptor was found to be structurally similar to the ones of the estrogen receptor β (ERβ)[118] and the peroxisome proliferation-activated receptor γ (PPARγ).[119] Thus all three receptors can be assigned to one protein structure similarity cluster. They exhibit a similar fold pattern (Figure 9.16a); albeit sequence similarities below 20%. The natural product genistein (Figure 9.16b) which contains a benzopyran scaffold, is an active inhibitor for both the ERβ and PPARγ receptor.[120] The drug troglitazone (Figure 9.16b) modulates the activity of the PPARγ receptor.[121] A PSSC analysis would have suggested the benzopyran scaffold as template for a library and would thus have predicted the discovery of farnesoid X inhibitors in this compound class. The benzopyran library synthesized by Nicolaou *et al*. also yielded ligands for the other members of this PSSC cluster (Figure 9.16b, synthetic ligands). This example further supports the application of PSSC in library design on the quest for new protein inhibitors.

The use of PSSC for compound library development provides new opportunities and a clear alternative to approaches based, for example, on mechanism or evolutionary relationships. For instance, the grouping of Cdc25A, AChE and the 11βHSD is non-obvious and could hardly have been achieved by sequence- or function-based methods. Thus, PSSC is a new method for grouping proteins which opens new routes to small molecule inhibitor design. Moreover, it also provides a cluster of proteins to test the compounds against which may help to identify cross-inhibition and thereby potentially also reasons for side effects of drugs at a very early stage of development.

D. BIOS: The combined application of SCONP and PSSC

SCONP and PSSC themselves are cheminformatics- and bioinformatics-based approaches for the design of biologically prevalidated compound collections which aim to improve the probability for successful discovery of small molecule ligands and inhibitors. The combined use

FIGURE 9.16 (a) Superposed X-ray structures of the ligand-binding domains of ERβ, PPARγ and (farnesoid X receptor) FXR, each with bound ligand. ERβ with genistein (**I**, blue), PPARγ with rosiglitazone (red), FXR with **V** (yellow); (b) Natural, non-natural and synthetic ligands for ERβ, PPARγ and FXR receptors

of SCONP and PSSC, however, offers a new route toward biologically relevant compound classes. For example, the natural product tree may be employed in attempts to simplify structures of known inhibitors while retaining their basic biological activity. In the scaffold tree diagram this step is achieved by moving inward along the branches (brachiation) to reach regions of the tree populated by less complex scaffolds. While such structural simplification has been tried many times before with varying success, the scaffold tree offers a reduced set of simplifications preselected on the basis of biological relevance. This may lead to situations where it may not be possible to choose the obvious retrosynthetic disconnection but rather the solution suggested by the scaffold tree and therefore by evolution. In a first application example, it could be shown that the structure of a natural product assigned to the carbocycle branch of the tree could be simplified retaining its biological activity. Glycyrrhetinic acid, a ligand of the enzyme 11βHSD was chosen as the structurally complex template. Placement of glycyrrhetinic acid into the tree leads to reduction to its scaffold and stepwise simplification from the pentacyclic scaffold to compounds with two ring scaffolds. The final decision of which compound collection to make first was based on a second line of argument, in this case PSSC. Dehydrodecalines were revealed by sequential brachiation along the branches of the natural product tree as shown in Figure 9.17.

The precise structure of the dehydrodecalin collection to be synthesized was chosen based on the PSSC analysis described before. Glycyrrhetinic acid is a known ligand of 11βHSD1 which, in turn, is part of a PSSC cluster with Cdc25A and AChE. Dysidiolide[122] contains a particular dehydrodecalin and Cdc25A, one of its known binding partners, is also part of the cluster which renders the Dysidiolide core motif a viable starting point for the design of library targeting 11βHSD. Based on this rationale, a collection of *ca.* 500 dehydrodecalines was synthesized using an asymmetric Robinson annulation as the key transformation.[47] Biochemical testing of 162 compounds from this collection for inhibitory activity on 11βHSD yielded 30 inhibitors with IC_{50} values below 10 μM; 4 of them (Figure 9.18) even showed IC_{50} between 310 and 740 nM. The most potent of these ligands also proved to be active in cellular assays.

Thus, BIOS may allow for the identification of structurally simpler starting points for library design while leading to high hit rates and yielding successful inhibitors. In an attempt to further explore this approach, a second example was investigated. Structurally complex alkaloids – yohimbine and ajmalicine – were identified as inhibitors of the protein phosphatase Cdc25A. Structural simplification of the yohimbine and ajmalicine scaffold led from the pentacyclic via the tetracyclic to tricyclic indole-based scaffolds (Figure 9.19). Investigation of indoloquinolizidines as potential inhibitors of Cdc25A revealed two compounds

FIGURE 9.17 Strategic use of PSSC and SCONP; the scaffold of glycyrrhetinic acid is analyzed according to SCONP rules leading to an octahydronaphtalene scaffold, which is a substructure of dysidiolide.

FIGURE 9.18 A combined application of PSSC and SCONP in the development of inhibitors of 11β-HSD1.

FIGURE 9.19 Brachiation along the indole branch of the natural product tree.

with IC_{50} values comparable to the pentacyclic natural product.[45,46] Extending this screen to other phosphatases allowed the identification of structurally new inhibitors of the *Mycobacterium tuberculosis* protein tyrosine phosphatase B from the same library. Structurally simpler compounds with three- and two-membered rings yielded one inhibitor of Cdc25A with an IC_{50} value similar to the value recorded for the pentacyclic molecule. The compound collection also contained inhibitors of protein tyrosine phosphatase 1B and MPTPB including eight inhibitors with submicromolar IC_{50} values.

These examples indicate that the combined use of SCONP and PSSC as proposed in BIOS may provide a viable strategy for the structural simplification of complex natural products while retaining most of their biological activity. Generality cannot be claimed on the basis of these two examples and even proof of principle will require additional successful examples. Moreover, the brachiation approach cannot be expected to work for all branches and the level of simplification that is possible may vary significantly. The extent to which a structure can be simplified, that is, the hierarchy level at which to stop, is difficult to predict and may require a second criterion. However, in principle, it has been shown that the use of SCONP and PSSC provides an advantageous opportunity for the design of structurally simplified protein ligands based on complex natural product structure.

E. BIOS: Prospects and future directions

PSSC and the scaffold tree are guiding principles which evolved from the analysis of nature's choice of structures in proteins as well as in small molecules. These concepts combine the biologically prevalidated parts of both worlds and thus provide starting points for library design as well as the choice of corresponding target clusters for biochemical screening. The combination of protein space and small molecule chemical space by merging PSSC and the scaffold tree into BIOS provides new opportunities for compound design and evaluation. BIOS enhances the chance of finding modulators of protein function and their corresponding targets, a challenge which lies at the heart of chemical biology.

The BIOS concept is based on biological prevalidation rather than on occurrence in nature. While occurrence in nature can be regarded as biological prevalidation *per se* it is definitely not an exclusive criterion. A large fraction of known biologically active molecules and current drugs and drug candidates originate from decades of highly successful medicinal chemistry research and combinatorial chemistry efforts. Additional compound classes with known biological effects including published screening hits, pesticides, herbicides or food ingredients made using the methods of modern organic synthesis can also be regarded as biologically prevalidated. Including these classes in the scaffold tree approach will allow access to larger regions of chemical space and may pave a way to the discovery of new small molecule inhibitors. Thus, the natural product tree will have to be extended to a tree of all known biologically relevant small molecules.

The PSSC approach on BIOS can be enhanced in applicability as well as in its scope. It has been shown that besides protein structures determined by crystallography or NMR methods and homology models of good quality can also be successfully employed.[89] Still, the static structures forming the basis of the analysis limit the specificity of the PSSC approach. Protein structures in their physiological environment are dynamic and conformational changes in the protein, for example induced fit, are known to modulate the binding properties and thus influence the analysis. Therefore, the omission of the dynamic component in protein crystal structures can lead to false positive and false negative results. One approach that introduces protein flexibility to the PSSC approach, as described by Charette *et al*. is to apply molecular dynamics to the protein structure of interest and subsequently use a selection of possible conformations as templates the structure similarity search.[123] The authors used molecular dynamics calculations to produce a number of conformations which discovered proteins known to bind similar ligands but previously undetected in the subsequent similarity search. Thus this method is an important and meaningful extension of the PSSC procedure.

IV. CONCLUSION AND OUTLOOK

The recent upsurge in the synthesis of natural product inspired compound collections and DOS around complex molecular architectures has successfully benefited from the structural information encoded in nature. Whereas DOS takes into consideration the structural aspects of the small molecule world, BIOS revolves around a common axis in the evolution of the worlds of proteins and small molecules and tries to analyze and explore their natural interfaces. Both approaches for compound collection development are compared in Table 9.1. Although at some points both approaches seem to be almost identical; large differences remain. DOS, for example, generally involves making large and diverse libraries to explore larger parts of chemical structure space. In contrast, BIOS prefers small focused libraries exploring a well-defined part of chemical space previously identified as of interest to the particular problem. The structural complementarity between proteins and their ligands forms the basis for one of the key hypotheses of BIOS, the "similar proteins bind similar ligands" postulate.

Developments in synthetic methodology like solid-phase synthesis, split-pool synthesis, etc. have helped considerably in streamlining the initial stages of the discovery

TABLE 9.1 Comparison of DOS and BIOS

Diversity oriented synthesis (DOS)	Biology oriented synthesis (BIOS)
• Aims to achieve maximum diversity around a core structure in chemical space selected because of its chemical accessibility. DOS also aims at diversity in the core structure if possible.	• Aims to achieve focused diversity around a starting point in the biologically relevant chemical space. These starting points are often natural product core structures.
• Diversity being the core criterion, the synthesis is planned generally by forward synthetic analysis.	• To synthetically access the core structure of validated small molecule, e.g., a natural product scaffold, retrosynthetic analysis is important. However, for adding diversity, forward synthetic analysis is also helpful.
• Synthetic routes generally involve established powerful methodologies like high yielding, stereoselective and functional group tolerant reactions.	• Known synthetic methods are used in the synthesis of the core structure and attachment of substituents, new synthetic methodologies to access diastereo- and enantiopure complex molecules in library formats are in demand.
• Building blocks should be commercially available in great variety.	• Building blocks are commercially available or have to be synthesized, often in solution.
• Members of a library can be used for varying target proteins due to their large diversity.	• Compound collections are focused and thus may provide ligands only for a limited number of related proteins.
• Generally large size.	• Focused small sized libraries.

processes including library synthesis and hit optimization. Still, the synthesis of natural product-like structures involving more complexity, diversity and stereochemistry demands new and better asymmetric tools to be added to the toolbox of organic synthesis which should also be applicable to compound library synthesis.[124] Research successfully addressing these challenges will definitely broaden the chemical space exploited by today's libraries toward more complex and natural product-like structures.

The exciting potential of the two complementary approaches DOS and BIOS to facilitate advances in chemical biology research and drug discovery will be realized to a greater extend in the coming years as the scientific community continues to explore the interactions between chemistry and biology.

REFERENCES

1. Dobson, C. M. Chemical Space and Biology. *Nature* 2004, *432*, 824–828.
2. Bohacek, R. S., McMartin, C., Guida, W. C. The art and practice of structure-based drug designing: a molecular modeling perspective. *Med. Res. Rev.* 1996, *16*, 3–50.
3. Augen, J. The evolving role of information technology in the drug discovery process. *Drug Discov. Today* 2002, *7*, 315–323.
4. Blundell, T. L., Jhoti, H., Abell, C. High-throughput crystallography for lead discovery in drug design. *Nat. Rev. Drug Discov.* 2002, *1*, 45–54.
5. Carr, R., Hann, M. The right road to drug discovery? *Mod. Drug Discov.* 2002, *5*, 45–48.
6. Patani, G. A., LaVoie, E. J. Bioisosterism: a rational approach in drug design. *Chem. Rev.* 1996, *96*, 3147–3176.
7. Burke, M. D., Schreiber, S. L. A planning strategy for diversity-oriented synthesis. *Angew. Chem. Int. Ed.* 2004, *43*, 46–58.
8. Spring, D. R. Diversity-oriented synthesis: a challenge for synthetic chemists. *Org. Biomol. Chem.* 2003, *1*, 3867–3870.
9. Mayer, T. U., Kapoor, T. M., Haggarty, S. J., King, R. W., Schreiber, S. L., Mitchison, T. J. Small molecule inhibitor of mitotic spindle bipolarity identified in a phenotype-based screen. *Science* 1999, *286*, 971–974.
10. Lee, D., Sello, J. K., Schreiber, S. L. A Strategy for macrocyclic ring closure and functionalization aimed toward split-pool syntheses. *J. Am. Chem. Soc.* 1999, *121*, 10648–10649.
11. Lee, D., Sello, J. K., Schreiber, S. L. Pairwise use of complexity generating reactions in diversity-oriented organic synthesis. *Org. Lett.* 2000, *2*, 709–712.
12. DeWitt, S. H., Kiely, J. S., Stankovic, C. J., Schroeder, M. C., Cody, D. M., Pavia, M. R. Diversomers: an approach to nonpeptide, nonoligomeric chemical diversity. *Proc. Natl. Acad. Sci. USA* 1993, *90*, 6909–6913.
13. Pelish, H. E., Westwood, N. J., Feng, Y., Kirchhausen, T., Shair, M. D. Use of biomimetic diversity-oriented synthesis to discover galanthamine-like molecules with biological properties beyond those of the natural product. *J. Am. Chem. Soc.* 2001, *123*, 6740–6741.
14. Dolle, R. E., Nelson, K. H., Jr. Comprehensive survey of combinatorial library synthesis: 1998. *J. Comb. Chem.* 1999, *1*, 235–282.
15. Dolle, R. E. Comprehensive survey of combinatorial library synthesis: 1999. *J. Comb. Chem.* 2000, *2*, 383–433.
16. Dolle, R. E. Comprehensive survey of combinatorial library synthesis: 2000. *J. Comb. Chem.* 2001, *3*, 477–517. (Since 1998, the comprehensive surveys of combinatorial library synthesis published by Dolle *et al*. are valuable source of information covering essentially all efforts in small molecule library synthesis)
17. Tan, D. S., Foley, M. A., Shair, M. D., Schreiber, S. L. Stereoselective synthesis of over two million compounds having structural features both reminiscent of natural products and compatible with miniaturized cell-based assays. *J. Am. Chem. Soc.* 1998, *120*, 8565–8566.
18. Masamune, S., Choy, W., Petersen, J. S., Sita, L. R. Double stereodifferentiation and a new strategy for stereocontrol in organic syntheses. *Angew. Chem. Int. Ed.* 1985, *24*, 1–30.

19. Masamune, S. Recent advances in stereochemical control: multiple asymmetric induction. *Heterocycles* 1984, *21*, 107–136.
20. Ko, S. Y., Lee, A. W. M., Masamune, S., Reed III, L. A., Sharpless, K. B., Walker, F. J. Total synthesis of the l-hexoses. *Science* 1983, *220*, 949–951.
21. Harrison, B. A., Verdine, G. L. The synthesis of an exhaustively stereodiversified library of *cis*-1,5 enediols by silyl-tethered ring-closing metathesis. *Org. Lett.* 2001, *3*, 2157–2159.
22. Gierasch, T. M., Shi, Z., Verdine, G. L. Extensively stereodiversified scaffolds for use in diversity-oriented library synthesis. *Org. Lett.* 2003, *5*, 621–624.
23. Su, Q., Beeler, A. B., Lobkovsky, E., Porco, J. A., Panek, J. S. Stereochemical diversity through cyclodimerization: synthesis of polyketide-like macrolides. *Org. Lett.* 2003, *5*, 2149–2152.
24. Harrison, B. A., Gierasch, T. M., Neilan, C., Pasternak, G. W., Verdine, G. L. High-affinity mu opoid receptor ligands discovered by the screening of an exhaustively stereodiversified library of 1,5-enediols. *J. Am. Chem. Soc.* 2002, *124*, 13352–13353.
25. Kim, Y.-K., Arai, M. A., Arai, T., Lamenzo, J. O., Dean, E. F., III, Patterson, N., Clemons, P. A., Schreiber, S. L. Relationship of stereochemical and skeletal diversity of small molecules to cellular measurement space. *J. Am. Chem. Soc.* 2004, *126*, 14740–14745.
26. Crews, C. M., Mohan, R. Small-molecule inhibitors of the cell cycle. *Curr. Opin. Chem. Bio.* 2000, *4*, 47–53.
27. Woo, S., Squires, N., Fallis, A. G. Indium-mediated γ-pentadienylation of aldehydes and ketones: cross-conjugated trienes for diene-transmissive cycloadditions. *Org. Lett.* 1999, *1*, 573–575.
28. Kwon, O., Park, S. B., Schreiber, S. L. Skeletal diversity via a branched pathway: efficient synthesis of 29,400 discrete, polycyclic compounds and their arraying into stock solutions. *J. Am. Chem. Soc.* 2002, *124*, 13402–13404.
29. Potuzak, J. S., Moilanen, S. B., Tan, D. S. Discovery and applications of small molecule probes for studying biological processes. *Biotechnol. Genet. Eng. Rev.* 2004, *21*, 11–78.
30. Walters, W. P., Namchuk, M. A guide to drug discovery: designing screens: how to make your hits a hit. *Nature Rev. Drug Discov.* 2003, *2*, 259–266.
31. Stockwell, B. R., Haggarty, S. J., Schreiber, S. L. High-throughput screening of small molecules in miniaturized mammalian cell-based assays involving post-translational modifications. *Chem. Biol.* 1999, *6*, 71–83.
32. Cohen, J. L., Limon, A., Miledi, R., Chamberlin, A. R. Design, synthesis, and biological evaluation of a scaffold for iGluR ligands based on the structure of (–)-dysiherbaine. *Bioorg. Med. Chem. Lett.* 2006, *16*, 2189–2194.
33. Kuruvilla, F. G., Shamji, A. F., Sternson, S. M., Hergenrother, P. J., Schreiber, S. L. Dissecting glucose signalling with diversity-oriented synthesis and small-molecule microarrays. *Nature* 2002, *416*, 653–657.
34. Hotha, S., Yarrow, J. C., Yang, J. G., Garrett, S., Renduchintala, K. V., Mayer, T. U., Kapoor, T. M. HR22C16: a potent small-molecule probe for the dynamics of cell division. *Angew. Chem. Int. Ed.* 2003, *42*, 2379–2382.
35. Spring, D. R., Krishnan, S., Blackwell, H. E., Schreiber, S. L. Diversity-oriented synthesis of biaryl-containing medium rings using a one bead/one stock solution platform. *J. Am. Chem. Soc.* 2002, *124*, 1354–1363.
36. Goess, B. C., Hannoush, R. N., Chan, L. K., Kirchhausen, T., Shair, M. D. Synthesis of a 10,000-membered library of molecules resembling carpanone and discovery of vesicular traffic inhibitors. *J. Am. Chem. Soc.* 2006, *128*, 5391–5403.
37. Miertus, S., Fassina, G., Seneci, P. F. Concepts of combinatorial chemistry and combinatorial technologies. *Chem. Listy* 2000, *94*, 1104–1110.
38. Smith, J. M. Natural selection and the concept of a protein space. *Nature* 1970, *225*, 563–564.
39. Russell, R. B., Sasieni, P. D., Sternberg, M. J. E. Supersites within superfolds. Binding site similarity in the absence of homology. *J. Mol. Biol.* 1998, *282*, 903–918.
40. Schmitt, S., Kuhn, D., Klebe, G. A new method to detect related function among proteins independent of sequence and fold homology. *J. Mol. Biol.* 2002, *323*, 387–406.
41. Rush, T. S., Grant, J. A., Mosyak, L., Nicholls, A. A shape-based 3-D scaffold hopping method and its application to a bacterial protein-protein interaction. *J. Med. Chem.* 2005, *48*, 1489–1495.
42. Shah, A. V., Walters, W. P., Murcko, M. A. Can we learn to distinguish between drug-like and nondrug-like molecules? *J. Med. Chem.* 1998, *41*, 3314–3324.
43. Walters, W. P., Shah, A. V., Murcko, M. A. Recognizing molecules with drug-like properties. *Curr. Opin. Chem. Biol.* 1999, *3*, 384–387.
44. Langer, T., Hoffmann, R. D. *Pharmacophores and Pharmacophore Searches in the Series of Methods and Principles in Medicinal Chemistry*, 32. Wiley-VCH: Weinheim, 2006. published by R. Mannhold, H. Kubinyi, G. Folkers, 1. edition.
45. Noeren-Mueller, A., Reis-Correa, I., Prinz, H., Rosenbaum, C., Saxena, K., Schwalbe, H. J., Vestweber, D., Cagna, G., Schunk, S., Schwarz, O., Schiewe, H., Waldmann, H. Discovery of protein phosphatase inhibitor classes by biology-oriented synthesis. *Proc. Nat. Acad. Sci. USA* 2006, *103*, 10606–10611.
46. Reis-Correa, I., Noeren-Mueller, A., Ambrosi, H. D., Jakupovic, S., Saxena, K., Schwalbe, H., Kaiser, M., Waldmann, H. Identification of inhibitors for mycobacterial protein tyrosine phosphatase B (MptpB) by biology-oriented synthesis (BIOS). *Chem. Asian J.* 2007, *2*, 1109–1126.
47. Koch, M. A., Schuffenhauer, A., Scheck, M., Wetzel, S., Casaulta, M., Odermatt, A., Ertl, P., Waldmann, H. Charting biologically relevant chemical space: a structural classification of natural products (SCONP). *Proc. Natl. Acad. Sci. USA* 2005, *102*, 17272–17277.
48. Feher, M., Schmidt, J. M. Property distributions: differences between drugs, natural products, and molecules from combinatorial chemistry. *J. Chem. Inf. Comput. Sci.* 2003, *43*, 218–227.
49. Newman, D. J., Cragg, G. M. Natural products as sources of new drugs over the last 25 years. *J. Nat. Prod.* 2007, *70*, 461–477.
50. Evans, B. E., Rittle, K. E., Bock, M. G., DiPardo, R. M., Freidinger, R. M., Whitter, W. L., Lundell, G. F., Veber, D. F., Anderson, P. S., Chang, R. S. LV., Lotti, J., Cerino, D. J., Chen, T. B., Kling, P. J., Kunkel, K. A., Springer, J. P., Hirshfieldt, J. Methods for drug discovery: development of potent, selective, orally effective cholecystokinin antagonists. *J. Med. Chem.* 1988, *31*, 2235–2246.
51. *Dictionary of Natural Products* (Chapman & Hall/CRC Informa, London), Version 14:2, 2005.
52. Brown, R. D., Martin, Y. C. Use of structure-activity data to compare structure-based clustering methods and descriptors for use in compound selection. *J. Chem. Inf. Comp. Sci.* 1996, *36*, 572–584.
53. Matter, H., Poetter, T. Comparing 3D pharmacophore triplets and 2D fingerprints for selecting diverse compound subsets. *J. Chem. Inf. Comp. Sci.* 1999, *39*, 1211–1225.
54. Robert, P. S., Simon, K. K. Why do we need so many chemical similarity search methods? *Drug Discov. Today* 2002, *7*, 903–911.
55. Schuffenhauer, A., Ertl, P., Roggo, S., Wetzel, S., Koch, M. A., Waldmann, H. The scaffold tree-visualization of the scaffold universe by hierarchical scaffold classification. *J. Chem. Inf. Model* 2007, *47*, 47–58.
56. *Thomson Scientific*, Derwent World Drug Index, London, UK.
57. Barun, O., Sommer, S., Waldmann, H. Asymmetric solid-phase synthesis of 6,6-spiroketals. *Angew. Chem. Int. Ed.* 2004, *43*, 3195–3199.
58. Barun, O., Kumar, K., Sommer, S., Langerak, A., Mayer, T. U., Mueller, O., Waldmann, H. Natural product-guided synthesis of a spiroacetal collection reveals modulators of tubulin cytoskeleton integrity. *Eur. J. Org. Chem.* 2005, 4773–4788.
59. Sommer, S., Waldmann, H. Solid phase synthesis of a spiro[5.5]ketal library. *Chem. Commun.* 2005, 5684–5686.

60. Mamane, V., Garcia, A. B., Umarye, J. D., Lessmann, T., Sommer, S., Waldmann, H. Stereoselective allylation of aldehydes on solid support and its application in biology-oriented synthesis (BIOS). *Tetrahedron* 2007, *63*, 5754–5767.
61. Garcia, A. B., Lessmann, T., Umarye, J. D., Mamane, V., Sommer, S., Waldmann, H. Stereocomplementary synthesis of a natural product-derived compound collection on a solid phase. *Chem. Commun.* 2006, 3868–3870.
62. Umarye, J. D., Lessmann, T., Garcia, A. B., Mamane, V., Sommer, S., Waldmann, H. Biology-oriented synthesis of stereochemically diverse natural-product-derived compound collections by iterative allylations on a solid support. *Chem. Eur. J.* 2007, *13*, 3305–3319.
63. Bialy, L., Waldmann, H. Synthesis of the protein phosphatase 2A inhibitor (4S,5S,6S,10S,11S,12S)-cytostatin. *Angew. Chem. Int. Ed.* 2002, *41*, 1748–1751.
64. Bialy, L., Waldmann, H. Total synthesis and biological evaluation of the protein phosphatase 2A inhibitor cytostatin and analogues. *Chem. Eur. J.* 2004, *10*, 2759–2780.
65. Lessmann, T., Leuenberger, M. G., Menninger, S., Lopez-Canet, M., Müller, O., Huemmer, S., Bormann, J., Korn, K., Fava, E., Zerial, M., Mayer, T. U., Waldmann, H. Natural product-derived modulators of cell cycle progression and viral entry by enantioselective oxa diels-alder reactions on the solid phase. *Chem. Biol.* 2007, *14*, 443–451.
66. Sanz, M. A., Voigt, T., Waldmann, H. Enantioselective catalysis on the solid phase: synthesis of natural product-derived tetrahydropyrans employing the enantioselective oxa-Diels-Alder reaction. *Adv. Synth. Cat.* 2006, *348*, 1511–1515.
67. Meseguer, B., Alonso-Díaz, D., Griebenow, N., Herget, T., Waldmann, H. Natural product synthesis on polymeric supports-synthesis and biological evaluation of an indolactam library. *Angew. Chem. Int. Ed.* 1999, *38*, 2902–2906.
68. Rosenbaum, C., Baumhof, P., Mazitschek, R., Müller, O., Giannis, A., Waldmann, H. Synthesis and biological evaluation of an indomethacin library reveals a new class of angiogenesis-related kinase inhibitors. *Angew. Chem. Int. Ed.* 2004, *43*, 224–228.
69. Rosenbaum, C., Katzka, C., Marzinzik, A., Waldmann, H. Traceless Fischer indole synthesis on the solid phase. *Chem. Commun.* 2003, 1822–1823.
70. Ganesan, A. Recent developments in combinatorial organic synthesis. *Drug Discov. Today* 2002, *7*, 47–55.
71. Arya, P., Baek, M.-G. Natural-product-like chiral derivatives by solid phase synthesis. *Curr. Opin. Chem. Biol.* 2001, *5*, 292–301.
72. Wessjohann, L. A. Synthesis of natural-product-based compound libraries. *Curr. Opin. Chem. Biol.* 2000, *4*, 303–309.
73. Mentel, M., Breinbauer, R. Combinatorial solid phase natural product chemistry. *Top. Curr. Chem.* 2007, *278*, 209–241.
74. Whitford, D. *Proteins, Structure and Function*, 1st edition. John Wiley & Sons, 2005 pp. 56–57.
75. Grant, A., Lee, D., Orengo, C. Progress towards mapping the universe of protein folds. *Genome Biol.* 2004, *5*, 107.
76. Koonin, E. V., Wolf, Y. I., Karev, G. P. The structure of the protein universe and genome evolution. *Nature* 2002, *420*, 218–223.
77. Leonov, H., Mitchell, J. S. B., Arkin, I. T. Monte Carlo estimation of the number of possible protein folds: effects of sampling bias and folds distributions. *Proteins* 2003, *51*, 352–359.
78. Coulson, A. F. W., Moult, J. A unifold, mesofold, and superfold model of protein fold use. *Proteins* 2002, *46*, 61–71.
79. Murzin, A. G., Brenner, S. E., Hubbard, T., Chothia, C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 1995, *247*, 536–540.
80. Andreeva, A., Howorth, D., Brenner, S. E., Hubbard, T. J. P., Chothia, C., Murzin, A. G. SCOP database in 2004: refinements integrate structure and sequence family data. *Nucl. Acids Res.* 2004, *32*, D226–D229.
81. Statistics taken from the SCOP website, (http://scop.mrclmb.cam.ac.uk/scop/count.htm#scop-1.71).
82. Grishin, N. V. Fold change in evolution of protein structures. *J. Struc. Biol.* 2001, *134*, 167–185.
83. Ponting, C. P., Schultz, J., Copley, R. R., Andrade, M. A., Bork, P. Evolution of domain families. *Adv. Protein Chem.* 2000, *54*, 185–244.
84. Apic, G., Gough, J., Teichmann, S. A. An insight into domain combinations. *Bioinformatics* 2001, *17*(Suppl. 1), S83–S89. (Oxford, England).
85. Chothia, C., Gough, J., Vogel, C., Teichmann, S. A. Evolution of the protein repertoire. *Science* 2003, *300*, 1701–1703.
86. Liu, J., Rost, B. Domains, motifs and clusters in the protein universe. *Curr. Opin. Chem. Biol.* 2003, *7*, 5–11.
87. Lee, D., Grant, A., Buchan, D., Orengo, C. A structural perspective on genome evolution. *Curr. Opin. Struc. Biol.* 2003, *13*, 359–369.
88. Koch, M. A., Breinbauer, R., Waldmann, H. Protein structure similarity as guiding principle for combinatorial library design. *Biol. Chem.* 2003, *384*, 1265–1272.
89. Koch, M. A., Wittenberg, L.-O., Basu, S., Jeyaraj, D. A., Gourzoulidou, E., Reinecke, K., Odermatt, A., Waldmann, H. Compound library development guided by protein structure similarity clustering and natural product structure. *Proc. Natl. Acad. Sci. USA* 2004, *101*, 16721–16726.
90. Koch, M. A., Waldmann, H. Natural product-derived compounds libraries and protein structure similarity as guiding principles for the discovery of drug candidates. In *Chemogenomics in Drug Discovery: A Medicinal Chemistry Perspective* (Kubinyi, H., Müller, G., Eds). Wiley-VCH, 2004, pp. 377–403.
91. Koch, M. A., Waldmann, H. Protein structure similarity clustering and natural product structure as guiding principles in drug discovery. *Drug Discov. Today* 2005, *10*, 471–483.
92. Balamurugan, R., Dekker, F. J., Waldmann, H. Design of compound libraries based on natural product scaffolds and protein structure similarity clustering (PSSC). *Mol. BioSyst.* 2005, *1*, 36–45.
93. Dekker, F. J., Koch, M. A., Waldmann, H. Protein structure similarity clustering (PSSC) and natural product structure as inspiration sources for drug development and chemical genomics. *Curr. Opin. Chem. Bio.* 2005, *9*, 232–239.
94. Dekker, F. J., Wetzel, S., Waldmann, H. Natural product scaffolds and protein structure similarity clustering (PSSC) as inspiration sources for compound library design in chemogenomics and drug development. *Chemogenomics* 2006, 59–84.
95. Stark, A., Shkumatov, A., Russell, R. B. Finding functional sites in structural genomics. *Proteins* 2004, *12*, 1405–1412.
96. Russell, R. B., Sasieni, P. D., Sternberg, M. J. E. Supersites within superfolds. Binding site similarity in the absence of homology. *J. Mol. Biol.* 1998, *282*, 903–918.
97. Jones, S., Thornton, J. M. Searching for functional sites in protein structures. *Curr. Opin. Chem. Biol.* 2004, *8*, 3–7.
98. Anantharaman, V., Aravind, L., Koonin, E. V. Emergence of diverse biochemical activities in evolutionarily conserved structural scaffolds of proteins. *Curr. Opin. Chem. Biol.* 2003, *7*, 12–20.
99. Lamb, S. S., Wright, G. D. Accessorizing natural products: adding to nature's toolbox. *Proc. Natl. Acad. Sci. USA* 2005, *102*, 519–520.
100. Holm, L., Sander, C. Dali/FSSP classification of three-dimensional protein folds. *Nucl. Acids Res.* 1997, *25*, 231–234.
101. Shindyalov, I. N., Bourne, P. E. Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Engin.* 1998, *11*, 739–747.
102. Shindyalov, I. N., Bourne, P. E. A database and tools for 3-D protein structure comparison and alignment using the combinatorial extension (CE) algorithm. *Nucl. Acids Res.* 2001, *29*, 228–229.
103. Gunasekera, S. P., McCarthy, P. J., Kelly-Borges, M., Lobkovsky, E., Clardy, J. Dysidiolide: a novel protein phosphatase inhibitor from the Caribbean sponge Dysidea etheria de Laubenfels. *J. Am. Chem. Soc.* 1996, *118*, 8759–8760.

104. Lyon, M. A., Ducruet, A. P., Wipf, P., Lazo, J. Dual-specificity phosphatases as targets for antineoplastic agents. *Nature Rev. Drug Discov.* 2002, *1*, 961–976.
105. Fauman, E. B., Cogswell, J. P., Lovejoy, B., Rocque, W. J., Holmes, W., Montana, V. G., Piwnica-Worms, H., Rink, M. J., Saper, M. A. Crystal structure of the catalytic domain of the human cell cycle control phosphatase, Cdc25A. *Cell* 1998, *93*, 617–625.
106. Bialy, L., Waldmann, H. Inhibitors of protein tyrosine phosphatases: next-generation drugs?. *Angew. Chem. Int. Ed.* 2005, *44*, 3814–3839.
107. Racchi, M., Mazzucchelli, M., Porrello, E., Lanni, C., Govoni, S. Acetyl-cholinesterase inhibitors: novel activities of old molecules. *Pharma. Res.* 2004, *50*, 441–451.
108. Walker, B. R., Seckl, J. R. 11-β-Hydroxysteroiddehydrogenase type 1 as a novel therapeutic target in metabolic and neurodegenerative disease. *Expert Opin. Therap. Targets* 2003, *7*, 771–783.
109. Chrousos, G. P. Is 11-beta-hydroxysteroid dehydrogenase type 1 a good therapeutic target for blockade of glucocorticoid actions? *Proc. Natl. Acad. Sci. USA* 2004, *101*, 6329–6330.
110. Ross, S. A., Gulve, E. A., Wang, M. Chemistry and biochemistry of type 2 diabetes. *Chem. Rev.* 2004, *104*, 1255–1282.
111. Sandeep, T. C., Yau, J. L. W., MacLullich, A. M. J., Noble, J., Deary, I. J., Walker, B. R., Seckl, J. R. 11-β-Hydroxysteroiddehyd rogenase inhibition improves cognitive function in healthy elderly men and type 2 diabetics. *Proc. Natl. Acad. Sci. USA* 2004, *101*, 6734–6739.
112. Nicolaou, K. C., Evans, R. M., Roecker, A. J., Hughes, R., Downes, M., Pfefferkorn, J. A. Discovery and optimization of non-steroidal FXR agonists from natural product-like libraries. *Org. Biomol. Chem.* 2003, *1*, 908–920.
113. Nicolaou, K. C., Pfefferkorn, J. A., Roecker, A. J., Cao, G. Q., Barluenga, S., Mitchell, H. J. Natural product-like combinatorial libraries based on privileged structures. 1. General principles and solid-phase synthesis of benzopyrans. *J. Am. Chem. Soc.* 2000, *122*, 9939–9953.
114. Nicolaou, K. C., Pfefferkorn, J. A., Mitchell, H. J., Roecker, A. J., Barluenga, S., Cao, G.-Q., Affleck, R. L., Lillig, J. E. Natural product-like combinatorial libraries based on privileged structures. 2. Construction of a 10,000-membered benzopyran library by directed split-and-pool chemistry using NanoKans and optical encoding. *J. Am. Chem. Soc.* 2000, *122*, 9954–9967.
115. Downes, M., Verdecia, M. A., Roecker, A. J., Hughes, R., Hogenesch, J. B., Kast-Woelbern, H. R., Bowman, M. E., Ferrer, J.-L., Anisfeld, A. M., Edwards, P. A., Rosenfeld, J. M., Alvarez, J. G. A., Noel, J. P., Nicolaou, K. C., Evans, R. M. A chemical, genetic, and structural analysis of the nuclear bile acid receptor FXR. *Mol. Cell* 2003, *11*, 1079–1092.
116. Robinson-Rechavi, M., Escriva, G. H., Laudet, V. The nuclear receptor superfamily. *J. Cell Sci.* 2003, *116*, 585–586.
117. Gronemeyer, H., Gustafsson, J. A., Laudet, V. Principles for modulation of the nuclear receptor superfamily. *Nature Rev. Drug Discov.* 2004, *3*, 950–964.
118. Jordan, V. C. Selective estrogen receptor modulation: concept and consequences in cancer. *Cancer Cell* 2004, *5*, 207–213.
119. Alarcon de la Lastra, C., Sanchez-Fidalgo, S., Villegas, I., Motilva, V. New pharmacological perspectives and therapeutic potential of PPAR-gamma agonists. *Curr. Pharma. Design* 2004, *10*, 3505–3524.
120. Dang, Z. C., Audinot, V., Papapoulos, S. E., Boutin, J. A., Loewik, C. W. G. M. Peroxisome proliferator-activated receptor γ(PPARγ) as a molecular target for the soy phytoestrogen genistein. *J. Biol. Chem.* 2003, *278*, 962–967.
121. Van Gaal, L., Scheen, A. J. Are all glitazones the same? *Diabetes/Metab. Res. Rev.* 2002, *18*(Suppl. 2), S1–S4.
122. Gunasekera, S. P., McCarthy, P. J., Kelly-Borges, M., Lobkovsky, E., Clardy, J. Dysidiolide: a novel protein phosphatase inhibitor from the Caribbean sponge Dysidea etheria de Laubenfels. *J. Am. Chem. Soc.* 1996, *118*, 8759–8760.
123. Charette, B. D., MacDonald, R. G., Wetzel, S., Berkowitz, D. B., Waldmann, H. Protein structure similarity clustering: dynamic treatment of PDB structures facilitates clustering. *Angew. Chem. Int. Ed.* 2006, *45*, 7766–7770.
124. Leßmann, T., Waldmann, H. Enantioselective synthesis on the solid phase. *Chem. Commun.* 2006, 3380–3389.

Chapter 10

In Silico Screening: Hit Finding from Database Mining

Thierry Langer and Sharon D. Bryant

I. INTRODUCTION
A. Chemoinformatics in drug discovery
B. What is the difference between a hit and a lead structure?
C. Data mining using chemoinformatics
II. REPRESENTATION OF CHEMICAL STRUCTURES
A. Structural keys and 1D fingerprints
B. Topological descriptors
C. 3D descriptors
D. Further descriptors
III. DATA MINING METHODS
IV. DATABASE SEARCHES
A. Distance and similarity searches
B. 2D database searches
C. 3D database searches
V. APPLICATIONS
A. Ligand-based *in silico* screening
B. Structure-based *in silico* screening
C. Assessing affinity profiles using parallel *in silico* screening
D. Example: Parallel pharmacophore-based virtual screening
VI. CONCLUSION AND FUTURE DIRECTIONS
REFERENCES

Today's scientists have substituted mathematics for experiments, and they wander off through equation after equation, and eventually build a structure which has no relation to reality.

Nikola Tesla (1857–1943), Modern Mechanics and Inventions, July 1934

I. INTRODUCTION

While a majority of drugs present on today's market has been developed by intelligently following serendipitous results,[1] there has been ongoing effort devoted to the implementation of rational approaches in the area of drug discovery and development. In this context, up to five decades ago medicinal chemists were following hypothetical activity models when synthesizing new compounds and biological activity was assessed using experiments with animals or at best with isolated organs. The output of compounds was therefore limited by the speed of biological tests. Twenty years later, techniques using computer-aided molecular modeling emerged, promoted by recent advances in gene technology, and the possibility to produce proteins of interest and elucidate their 3D structure. This boosted the development of *in vitro* models for receptor binding and enzyme inhibition. Suddenly the number of compounds that could be assessed for their bioactivity was drastically enhanced. In this time, the synthesis of new compounds became the time-limiting step and soon thereafter new technologies emerged based on the idea of high-throughput combinatorial synthesis for obtaining large numbers of novel compounds. The possibility of synthesizing combinatorial libraries containing hundreds of thousands of compounds and the technology to assess the capability of such compounds to bind to proteins of interest created an unprecedented hype for this approach. Unfortunately, by making the size of the haystack bigger, the

Copyright © 2008, Elsevier Ltd
All rights reserved.

chance is not higher to find the needle, as stated by Lahana[2] in a paper dating from 1999 and in fact, the number of interesting drug candidates that has emerged by the application of such an approach is disappointingly low.[3] When analyzing the situation, it turned out that the majority of compounds designed either by computer-aided methods (e.g. structure-based design) or obtained from combinatorial chemistry exhibited inappropriate ADMET (absorption, distribution, metabolism, excretion, and toxicity) properties which later was a major cause for attrition in preclinical studies or clinical trials.

Nevertheless, in the past 50 years, enormous progress was achieved in various disciplines of pharmaceutical research; one of the highlights was the elucidation of the human genome. Thus, the number of possible targets used now in pharmacotherapy has been estimated from a few hundred to several thousand that could be considered as "druggable."[4] The enormous amount of data generated by modern methods used in drug discovery and development, including genomics, proteomics, bioinformatics, metabolomics, combinatorial chemistry, and ultrahigh-throughput screening (μHTS) requires powerful and efficient data mining methods. In this context, electronic or *in silico* screening gained influence on the validation of targets, hit finding, hit to lead, expansion, lead optimization, and especially the prediction of suitable ADMET and off-target effect profiles. In the present chapter, a brief overview is presented on the application of *in silico* screening and database mining for the efficient identification of hits and the rational expansion to leads in drug design, together with concise insight into the methods used for this purpose.

A. Chemoinformatics in drug discovery

The fact that there is still an absence of a precise definition that is generally accepted for the term chemoinformatics is not an indication for low importance of this area of science. It is true that chemoinformatics has gained enormous interest since it has become indispensable in drug discovery and development. According to Brown,[5] "Chemoinformatics is the mixing of those information resources to transform data into information and information into knowledge for the intended purpose of making better decisions faster in the area of drug lead identification and organization." There are many reviews,[6–9] and a couple of books[10,11] available that focus on this topic. One specific aspect of the new emphasis of this discipline is the sheer magnitude of chemical information that needs to be processed. When thinking about the fact that Chemical Abstract Services add approximately a million of new compounds to its database annually for which large amounts of property data are available, and that industrial groups generate hundreds of thousands to millions of compounds on a regular basis through combinatorial chemistry, the importance of chemoinformatics becomes easily understandable. The pharmaceutical industry has been facing a dramatic loss of innovation in the last decade, bringing a decreasing number of new chemical entities (NCE) to the market. This number has continuously dropped from around 60 per year in the 1980s to about 35 in the beginning of the new millennium and costs for developing and bringing a new drug to the market have been estimated to be around $800 million to a billion. Certainly, regulatory pressure has become higher than ever before, however, the high attrition rate cannot only be attributed to regulation issues. There are published studies dealing with the reasons for attrition of drug candidates in different development phases.[12,13] Clearly, one of the reasons for failure in the early development stages can be attributed to unfavorable physicochemical properties, resulting in undesirable pharmacokinetic behaviors and suboptimal ADMET profiles. In this regard, chemoinformatics methods can help through the prediction of pharmacological properties and aid the proper selection of compounds for follow-up, which is a central task in early drug development process. The drug discovery and development process comprise the following steps:

1. target identification
2. target validation
3. lead finding (including *in silico* and *in vitro* screening compound libraries for hit finding, as well as the design and the synthesis of compound series)
4. lead optimization (acceptable ADME profile, no toxicity and mutagenicity)
5. preclinical studies
6. clinical studies (phase I, II, and III)
7. regulatory approval.

In this chapter, we will focus on the use of chemoinformatics methods involved in Step (3) of this entire process.

B. What is the difference between a hit and a lead structure?

Valler and Green have defined the term lead structure as "a representative of a compound series with sufficient potential (as measured by potency, selectivity, pharmacokinetics, physicochemical properties, absence of toxicity and novelty) to progress to a full drug development programme."[14] An optimized lead structure, or in short "lead" is therefore by definition a compound that has been studied extensively, and has been modified starting from appropriate hit structures. Thus it is clear that a lead will generally not directly result from a screening campaign, irrespective of whether it is performed *in silico* or *in vitro*, or combined, as recently termed *in combo*.[15] Compounds emerging from such screening experiments that are bioactive for a target are called "hits." Their selectivity and affinity for other targets should be determined in follow-up experiments. Usually, a hit is found to exhibit an affinity in the low micromolar range and

can be turned into a potent and selective lead using classical medicinal chemistry approaches, as described elsewhere in this book. For hit finding, *in silico* approaches have been found to be extremely useful. When prescreening a compound collection using a virtual approach, the number of compounds that needs to be screened experimentally can be reduced by a factor of several orders of magnitude. The term "enrichment" describes the hit rate obtained from a virtual screening procedure followed by experimental screening. While in a typical random screening campaign, the hit rate is found to be approximately between 0.025% and 0.1%,[16] hit rates for *in combo* approaches in exceptional cases have been reported to be up to higher than 50%.[17,18]

C. Data mining using chemoinformatics

One of the most important issues in drug research is the establishment of a sound relationship between a chemical structure and its biological activity. Tools that are necessary to navigate through the massive amounts of data gathered from synthesis and screening and to extract relevant information have been developed and are widely used in the pharmaceutical industry nowadays. Terms and expressions closely associated with this area of data mining research are "data warehousing," "knowledge acquisition," "knowledge discovery," "data harvesting," "fuzzy modeling," "machine learning," "web farming," etc. A definition of the term data mining by Fayyad[19] describes it as "nontrivial extraction of implicit, previously unknown and potentially useful information from data, or the search for relationships and global patterns that exist in databases." For extracting useful information from huge quantities of data and gaining knowledge from this information, deep analysis and exploration have to be performed. This can be done only by automatic or at least semi-automatic methods. A typical data mining process is divided into several consecutive steps: (i) selection, (ii) preprocessing, (iii) transformation, (iv) interpretation, and (v) evaluation. Visualization of data plays an important role, especially in the latter two steps. However, the most crucial part remains the correct representation of chemical structures, together with their efficient storage in databases that should be able to store and rapidly process several millions up to billions of compounds. In the following section, we will focus on the use of chemoinformatics for selectively retrieving bioactive compounds from large collections and discuss the most common methods used for this so-called *in silico* screening task.

II. REPRESENTATION OF CHEMICAL STRUCTURES

Retrieving bioactive hits from compound databases requires the analysis of the relationship between the structure of compounds and their biological activity, that is, their binding affinity to a certain target. In view of the large amounts of data to be processed and analyzed, it seems advisable to use a hierarchical representation of the chemical structures. This can start using 1D information, such as 1D fingerprints, continue with topological descriptors, such as graph-based descriptors or 2D autocorrelation, and finally end with 3D structures, even considering their multiconformational behaviors and the properties attributed to their entire surfaces or to special areas thereof. It is clear that dealing with 1D descriptors is much more efficient than with 2D, and even 3D. However, a correct structure–activity relationship can often be obtained only when using the correct 3D representations of molecules.

The simplest representation of a molecular structure is the linear notation converting the connection matrix of a molecule (when interpreted in the classical way as a model consisting of atoms and bonds connecting them) into a sequence of alphanumeric symbols using a set of rules. The most widespread method used for linear 1D representation of molecules is the "simplified molecular input line system" (SMILES).[20,21] Another prominent example and among the most popular within the first attempts to derive a linear notation for molecular structures is the Wiswesser Line Notation (WLN).[22] It is straightforward to search for structures in a database by string matching. A prerequisite for successful retrieval of compounds, however, is that unique SMILES or WLN strings are used. The term "Canonical SMILES" refers to the version of the SMILES specification that includes rules for ensuring that each distinct chemical molecule has a single unique SMILES representation. Recently, the IUPAC has introduced the International Chemical Identifier (InChI)[23] as a standard for linear formula representation. It represents the digital equivalent of the IUPAC name for any particular covalent compound. Chemical structures are expressed in InChI in terms of five layers of information – connectivity, tautomeric, isotopic, stereochemical, and electronic (e.g. see the SMILES, WLN, and InChI notations of (2E)-3-cyclohexyl-2-[(R)-hydroxy(phenyl)methyl]acrylonitrile given in information Box 10.1). For encoding substructure search queries, the "SMILES arbitrary target specification" (SMARTS)[24] can be used. They allow retrieving a particular pattern (subgraph) in a molecule (graph), which is one of the most important tasks for chemoinformatics-based data mining. Substructure search is used virtually in every application that employs a digital representation of a molecule, including depiction (to highlight a particular functional group), drug design (searching a database for similar structures and activity), analytical chemistry (looking for previously characterized structures and comparing their data to that of an unknown), and a number of other tasks.

A. Structural keys and 1D fingerprints

Structural keys are used in order to transform structural information of different molecules into normalized

BOX 10.1 SMILES, WLN, and InChI Notation of (2E)-3-cyclohexyl-2-[(R)-hydroxy(phenyl)methyl]acrylonitrile

Notation Type*	
SMILES	c2ccccc2[C@@H](O)/C(=C/C1CCCCC1)C#N
Wiswesser Line Notation	RYQ&YCN&U1L6TJ
InChI	1/C16H19NO/c17-12-15(11-13-7-3-1-4-8-13)16(18)14-9-5-2-6-10-14/h2,5-6,9-11,13,16,18H,1,3-4,7-8H2/b15-11+/t16-/m1

*taken from http://www.oci.unizh.ch/edu/lectures/material/DBC/LinNot/LinNot.shtml

bitstrings. They describe the chemical composition and eventually structural motifs of molecules represented as a Boolean array: The presence of a certain structural feature in a molecule or substructure of a molecule is indicated by a bit that is set to 1 (true), absence of such a feature is indicated by a bit that is set to zero (false). In such an array, bits may encode particular functional groups (such as a carboxylic group, an amide linkage, a benzene ring, etc.), structural elements (e.g. a substituted cyclohexane), or at least *n* occurrences of a particular element (e.g. an oxygen atom). Alternatively, a structural key can be defined as an array of integers, with the elements of this array containing the frequency of how often specific features occur in the compound. Thus, similarities among pairs of molecules can be determined easily and expressed as similarity coefficients (e.g. Tanimoto index).[25] When a database search is performed using such an approach, the structural key of the query molecule or substructure is compared with the stored structural keys of all database entries. This necessitates that each array element in the key has to be defined initially, implicating that the key is inflexible and may become extremely long. Search speed across the whole database is influenced by the choice and the number of patterns included in the key. Using short keys will allow for fast operation, while long keys slow down searching. On the other hand, searching with short keys may result in retrieving a lot of structures that are of no interest.

Molecular 1D fingerprints were introduced to overcome the inherent problems and shortcomings of structural keys. Such a fingerprint is also a Boolean array, however, in contrast to a structural key, the meaning of any particular bit is not predefined. Initially, all bits of a fingerprint with a fixed size are set to zero. In the following step, a list of patterns is generated for each atom, each pair of adjacent atoms and the bonds connecting them, and for each group of atoms joined by longer pathways. A hash coding algorithm is used to assign a unique set of bits (typically 4 or 5 bits per pattern) to each pattern along the fingerprint. The set of bits obtained in this way is added to the fingerprint with a logical OR. When assuming that a pattern represents a substructure of a given molecule, each bit on the pattern's fingerprint will be set in the molecule's fingerprint.

Databases can be searched by simple Boolean operations using both structural keys and fingerprints. The latter have a higher information density than structural keys without losing specificity. Hence, database searches using fingerprints instead of structural keys are more efficient.

Several commercial providers of chemoinformatic tools have issued systems for calculating molecular fingerprints. Commonly used formats are those used in the molecular structure database systems developed by MDL (ISIS and MACCS)[26,27] and the Daylight[28] fingerprints.

B. Topological descriptors

A large number of descriptors based on molecular topology have been published in the last 50 years. Among the first ones in this context was Wiener's topological index.[29–31]. This index (or "Wiener path number") is calculated as the sum of all topological distances wherein the hydrogen atoms can be omitted from the molecular graph. Further topological descriptors in this category are the Randic connectivity index;[32] the connectivity indexes described by Kier and Hall,[33] as well as their electrotopological-state index (or *E*-state index),[34] and eigenvalue-based descriptors like Burden eigenvalues.[35] Charge-based indexes have also been introduced as topological descriptors.[36] All of these descriptors have been used in a large number of different studies for establishing structure–activity relationships within series of congeners. Before discussing in more detail the most relevant 2D topological descriptors for *in silico* screening, a detailed review by Estrada on recent advances on the role of topological indices in drug discovery research is recommended[37] for those that are interested in this topic.

The following section gives an overview of feature trees and 2D autocorrelation vectors, the two most important graph-based topological descriptors used for virtual screening.

1. *Feature trees*

The structural diagram of a molecule can be interpreted as a mathematical graph. Each atom therein is represented by a node in the graph, and accordingly, the bonds are represented

by the edges. The search for substructures[38] or for the maximum common substructure of a set of molecules[39,40] can be performed by using algorithms developed in graph theory.

In 1998, Rarey and Dixon introduced the concept of feature trees as molecular descriptors.[41] This approach is based on a similarity value for two molecules that can be calculated starting from molecular profiles and rough mapping: Molecules are represented by a so-called feature tree. Within this molecular graph, the nodes are fragments of the molecule, while the atoms belonging to one node are connected in the graph. A node consists of at least of one atom, and rings are collapsed to single node. Edges in the feature tree connect two nodes having atoms in common, or having atoms connected in the molecular graph.[42]

Within this framework, it is possible to represent molecules with feature trees at various levels of resolution. Obviously, the maximum simplification of a molecule would be its representation as a feature tree with a single node. On the other hand, the highest level of representation would be a feature tree with each acyclic atom forming a node. Due to the hierarchical nature of such a representation, feature trees of all levels of resolution can be derived from the highest hierarchical level. A subtree is replaced by a single node representing the union of the atom sets of the nodes belonging to this subtree. In addition to the topological information contained in such a tree, properties such as chemical or steric features are additionally assigned to nodes. The latter consists of the number of atoms in the fragment and an estimate for the van der Waals volume. In the case that atoms belong to several fragments, only the relevant fractional amount is taken into account. The chemical features are stored in an array and denote the fragment's ability to form interactions with complementary groups. Atom type profiles consider the number of carbon, nitrogen, oxygen, phosphorus, sulfur, fluorine, chlorine, bromine, iodine, or other non-hydrogen atoms, as well as their different hybridization states. The interaction profile contained in the FlexX environment[43] comprises hydrogen bond donors and acceptors, aromatic ring atoms and centers, and a hydrophobic interaction. When comparing molecules, for a pair of feature values, a similarity value in the range of 0 (dissimilar) to 1 (identical) is calculated. For comparison of two feature trees, the trees have to be matched to each other and a weighted average of the similarity values of all matches within the two feature trees is calculated. This approach has been shown to be useful in lead finding and optimization, analysis of HTS, and in other general virtual screening applications.[44–47]

2. 2D *autocorrelation vectors*

In 1980, Moreau and Broto[204] introduced an autocorrelation function (1) for transforming the information from a molecule's structural diagram into a representation with fixed numbers of components. In the equation (10.1), $A(d)$ is the component of the autocorrelation vector for the topological distance d. N represents the number of atoms, and the topological distance between atoms i and j is denoted d_{ij} (the number of bonds for the shortest path in the structure diagram). Properties of atoms i and j are referred to as p_i and p_j, respectively. The value of the autocorrelation function $A(d)$ for a defined topological distance d can be calculated from the summation over all products of a property p of atoms i and j having the required distance d. Transformation molecular structure information using this procedure is of high interest when considering that both statistical methods and artificial neural networks need a fixed number of descriptors for the analysis of a set of molecules, independent of their size and number of atoms.

$$A(d) = \sum_{j=i}^{N} \sum_{i=1}^{N} \delta(d_{ij} - d) p_j p_i \qquad \delta = \begin{cases} 1 & \forall \quad d_{ij} = d \\ 0 & \forall \quad d_{ij} \neq d \end{cases} \tag{10.1}$$

The application of artificial neural networks using such autocorrelation vectors for characterizing molecules has been studied and further developed by Gasteiger and colleagues, and the interested reader is referred to Ref. [47] for more detailed information. They made available the program package ADRIANA. *Code*[48] in which a range of physicochemical properties such as partial atomic charges[49] or measures of the polarizability[50] can be calculated and transferred into the appropriate artificial neural network for further processing and analysis.

C. 3D descriptors

1. 3D *structure generation*

Physical, chemical, and biological properties are related to the 3D structure of a molecule. Sources for experimental determination of 3D structure information are essentially X-ray crystallography, electron and neutron diffraction, and NMR spectroscopy. There are several databases containing 3D structure information. The commercially available Cambridge Structural Database (CSD)[51] currently comprises about 400,000 experimentally determined molecular structures of small organic molecules and coordination compounds. The publicly available Protein Databank (PDB)[52] contains approximately 45,000 entries of large biopolymers, mainly proteins and molecules of DNA or RNA. These numbers of experimentally determined 3D structures are low compared to the overall known number of organic and inorganic compounds (by now more than 32 M).[53] In addition, the virtual space of compounds that might be synthesized within a drug discovery project still has to be investigated. While the exact prediction of 3D structures of proteins and RNA still remains challenging, for small organic molecules – the typical application case for drug compounds – there are

several algorithms that are well suited for calculating 3D molecular structure models and molecular properties with accuracy. The quantum or molecular mechanics calculations work well, however, they are currently too slow to process millions of compounds in a reasonable period of time. Additionally, most of these methods require a 3D structure as starting point and therefore, automatic methods for transformation of 2D connectivity information into 3D models are required. For a detailed review about this subject, the reader is referred to Ref. 54.

Four widely used programs for the generation of 3D structures that work excellently are CORINA,[55–58] CONCORD,[59–61] OMEGA,[62] and CATALYST.[63] Studies about the performance of the latter two have been published.[64–67] In addition to 3D structure generation, conformational sampling is necessary in order to reflect the flexibility of molecules binding to receptors. Different approaches starting with systematic grid search,[68] expanding to stochastic methods,[69,70] application of genetic algorithms,[71,72] and finally simulation methods such as molecular dynamics,[73] Monte Carlo,[74] or simulated annealing[75] address this problem. For high-speed conformational generation of diverse and user-controlled conformational ensembles starting from a given 3D model, the programs ROTATE,[76] OMEGA,[62] CATALYST[63] and CAESAR[77] are available.

2. 3D *autocorrelation*, 3D MoRSE *code, and radial distribution function code*

In contrast to the topological autocorrelation vectors in the 3D autocorrelation vector, the spatial distance between atoms is used for calculation. Hence, using 3D autocorrelation vectors, it is possible to distinguish between different conformations of a molecule. The calculation of autocorrelation vectors of surface properties[78] is similar to equation (10.2):

$$A(d) = \frac{1}{L}\sum_{x} p(x) \cdot p(x+d) \tag{10.2}$$

with the distance d within the interval $d_l < d \leq d_u$, a certain property $p(x)$ at a point x on the molecular surface, and the number of distance intervals L. The component of the autocorrelation vector for a certain distance d within the interval $d_l < d \leq d_u$ is the sum of the product of the property $p(x)$ at a point x on the molecular surface, with the same property $p(x + d)$ at a certain distance d normalized by the number of distance intervals L. All pairs of points on the surface are considered only once.

Another code for representation of the 3D structure of a molecule with a fixed number of variables irrespective of the number of atoms in the molecule (3D MoRSE code) has been proposed by Soltzberg and Wilkins.[79] This molecular description is based on methods used in the interpretation of electron diffraction data. The approach has been used successfully for both the simulation of infrared spectra and the classification of a dataset of 31 corticosteroids for which binding affinity data to the corticosteroid binding globulin (CBG) receptor was available.[80]

The radial distribution function code (RDF code) is closely related to the 3D-MoRSE code and it is calculated by equation (10.3):

$$g(r) = f\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} p_i p_j e^{-B(r-r_{ij})^2} \tag{10.3}$$

where f is a scaling factor, N is the number of atoms in the molecule, p_i and p_j are properties of the atoms i and j, B a smoothing parameter, and r_{ij} the distance between the atoms i and j. $g(r)$ is computed at a number of discrete points with defined intervals.[81,82] B can be regarded as a temperature factor that defines the movement of the atoms. By including characteristic atomic properties p_i and p_j the RDF code can easily be adapted to the requirements of information to be represented in different application cases. For studying a set of ligands, for example, in a drug discovery process, it may be useful to utilize properties describing the atomic partial charges or their capability to act as hydrogen bond donors or acceptors, respectively. The dimension and length of the RDF code is independent of the number of atoms and the size of a molecule. Moreover, it is unambiguous regarding the 3D arrangement of the atoms, and invariant against rotation and translation of the entire structure.

3. 3D *pharmacophore descriptors*

In the past decade, several software programs that rely on the concept of chemical feature-based pharmacophore models have exerted an increasing influence on rational drug design. According to IUPAC's definition verbalized by Wermuth,[83] "A pharmacophore is the ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target structure and to trigger (or to block) its biological response. A pharmacophore does not represent a real molecule or a real association of functional groups, but a purely abstract concept that accounts for the common molecular interaction capacities of a group of compounds toward their target structure. The pharmacophore can be considered the largest common denominator shared by a set of active molecules. This definition discards a misuse often found in the medicinal chemistry literature, which consists of naming as pharmacophores simple chemical functionalities such as guanidines, sulfonamides or dihydroimidazoles (formerly imidazolines), or typical structural skeletons such as flavones, phenothiazines, prostaglandins or steroids."

In this pharmacophoric context, rather than comparing molecular structures or substructures to each other, the binding pattern of a ligand to its binding site is characterized by location and tolerance constraints in 3D space encoding for different kinds of interactions. Those include vectors

FIGURE 10.1 Pharmacophore model representation derived from the cyclin-dependent kinase 2 (CDK2) inhibitor 3-{[(2,2-Dioxido-1,3-dihydro-2-benzothien-5-yl)amino]-methylene}-5-(1,3-oxazol-5-yl)-1,3-dihydro-2H-indol-2-one in its binding site,[96] PDB entry 1ke7 processed using LigandScout;[95] (a) Automatically generated 3D pharmacophore model (green arrow: H-bond donor; red arrow: H-bond acceptor; yellow sphere: hydrophobic region). (b) Projection showing the automatically generated schematic 2D pharmacophore model depiction together with the interacting amino acids in the binding pocket.

for H-bonds, planes for aromatic Pi stacking, or spheres for hydrophobic or electrostatic interactions. Each interaction feature represents the region in 3D space that should be occupied by a certain atom or group of atoms capable of the kind of physicochemical interaction specified by the feature type. This kind of generalization is highly effective for database mining, since in organic molecules, different structural motifs can express a similar chemical behavior and therefore the same biological effect. While in the majority of the cases no experimental information on either the biological conformation of the ligand or the target protein are currently available, the so-called ligand-based pharmacophore approach is able to provide essential information for medicinal chemists. Several successful applications within this subject have been reviewed.[84–87] In these articles, the authors outline the theoretical background as well as describe several significant studies including 3D pharmacophore-based database search strategies. Programs for pharmacophore modeling are CATALYST,[63] PHASE,[88,89] MOE,[90] GALAHAD,[91–93] and LIGANDSCOUT[94,95] (Figure 10.1).

D. Further descriptors

In addition to the descriptors mentioned above, grid-based methods have been frequently used in the field of quantitative structure–activity relationships (QSAR).[97] In one of the most successful approaches, the comparative molecular field analysis (CoMFA),[98] the molecule is placed in a box and the interaction energy values between this molecule and another small molecule (the probe) are calculated for an orthogonal grid of points. Thus, a grid map is obtained that characterizes the molecular shape and charge distribution. Additional properties, such as hydrogen bonding capability, hydrophobicity, etc. can be assessed by choosing appropriate molecules as a probe. After appropriate alignment of a set of ligands that bind to the same receptor, a CoMFA allows determination and visualization of molecular interaction regions involved in ligand-target binding. Using statistical methods that are able to deal with the problem of unsymmetrical data distribution such as partial least squares (PLS) regression,[99] correlations are searched between the CoMFA descriptors and biological activity. A more detailed overview about 3D-QSAR methods is given in Chapter 29.

Further commonly used molecular descriptors are for example, eigenvalue-based descriptors including 3D information, such as BCUT descriptors (Burden, CAS, University of Texas eigenvalues),[100,101] EVA,[102,103] and WHIM descriptors (weighted holistic invariant molecular descriptors).[104,105] For more information about molecular descriptors, the reader is referred to a handbook by Todeschini and Consonni[106] and a recent review by Xue and Bajorath.[107]

III. DATA MINING METHODS

In the data mining process, one uses the combination of methods and tools from three areas: Statistics, machine

learning, and databases. The most important methods from statistics and machine learning are listed below:

Statistical methods

- Correlation analysis
- Factor analysis
- Multilinear regression analysis (MLRA)
- Principal component analysis (PCA)
- Partial least squares regression (PLS)
- Principal component regression (PCR).

Machine learning

- Decision-tree learning or rule induction (ID3, C5)
- K-nearest neighbor analysis (KNN)
- Hierarchical and non-hierarchical clustering (k-means, Ward, Jarvis-Patrick)
- Artificial neural networks (feed-forward neural networks, self-organizing neural networks, counterpropagation neural networks, Bayesian neural networks)
- Genetic algorithms
- Support vector machines.

The learning algorithm can be performed in supervised or unsupervised mode. In order to avoid bias, it is advisable to start data mining investigations of datasets with unsupervised learning methods.

IV. DATABASE SEARCHES

In order to perform a search for molecular structures in a database, the structural data have to be present in a searchable format. In the second section of this chapter, we have discussed the "Representation of chemical structures." It is important to mention that the encoding scheme can be a reversible or an irreversible representation. Only reversible or two-way representation schemes such as the systematic nomenclature, SMILES notation, connectivity tables, etc. allow conversion of the original structure diagram into the representation and vice versa. Using irreversible one-way encoding it is impossible to regenerate the structural graph from the representation. Such a representation is only useful, for example, for indexing the database for search speed enhancement.

Performance is an important factor when searching compound databases. Therefore, in addition to smart indexing, reducing the number of structures in a first step has become common practice in virtual screening workflows. Such prescreening can be performed using simple 1D descriptors, for example, molecular weight, number of heavy atoms, number of rotatable bonds, estimated log P, number of hydrogen bond donors and acceptors, etc. Combination of such filters have been proposed, such as the "rule of five" from Lipinski *et al.*[108] This rule is applied for removing compounds likely to exhibit poor absorption and permeation properties (oral bioavailability). Other such simple filters used for prescreening account for drug-[109–111] or lead-likeness,[112,113] appropriate ADMET profile,[114–117] or favorable predicted properties concerning receptor binding.[118]

Chemical structural databases range in addition to the aforementioned CSD and PDB databases from the comprehensive collection of small organic molecules in the CAS registry file (Chemical Abstracts Services)[53] and the Beilstein/Crossfire database,[119] both commercially available, to the publicly available PubChem database,[120] emerging as a component of NIH's Molecular Libraries Roadmap Initiative. According to the principle that similar structures should exhibit similar properties, (known as "similarity property principle") high ranked structures in similarity-based searches are supposed to exhibit similar physicochemical and biological properties to that of the target structure. Consequently, similarity searches play an important role in database searches related to drug design. On the other hand, when relying only on topological similarity, scaffold hopping might become a difficult task. Since ligand binding to a receptor is driven by interaction of functional groups at the ligand's solvent-accessible surface, approaches that focus on the detection of such interaction regions are likely to perform better than simple similarity searches using structure or substructure matching. In the following section, a brief overview on database search methods is given.

A. Distance and similarity searches

Essentially, the similarity between compounds is estimated in terms of a distance measure between two different objects, described by vectors. Scaling of the variables is advisable if they do not have comparable magnitude. The most prominent distance measures are the Euclidean distance, the average Euclidian distance, and the Manhattan distance. A comprehensive overview of methods for chemical similarity searching has been published by Willet *et al.*[121] Except for similarity searches among compounds in databases, such similarity measures are frequently applied in the design and analysis of combinatorial libraries.[122–126]

B. 2D database searches

Foundations of graph theory and the algorithms required for structure and substructure searching are reported in a review by Hopkinson.[127] When discussing 2D database searching, four different general cases can be distinguished: (i) Exact 2D structure search, (ii) exact 2D substructure search, (iii) R-group and Markush searching, and (iv) 2D similarity searching. While in case (i), all entries in the database that match exactly and completely to a unique target query structure have to be retrieved, in case (ii) all

entries in the database that contain the user-defined partial structure need to be retrieved. While the exact 2D structure search is usually only useful when trying to retrieve the same compound, for example, from different providers, exact 2D substructure searching is an application scenario in virtual screening procedures aimed at finding closely related compounds to a target hit. Typically, this kind of searching proceeds in two steps; after screening the database based on 2D chemical fragment indices such as molecular fingerprints, all records that passed are compared by atom-by-atom mapping with the query substructure. In R-group and Markush searching, the substructure is represented as a partial structure with a substitution pattern and a list of substituents. This approach is usually used in particular for chemical patent searching. Retrieval of compounds by R-groups and Markush searching results is done in the same way as substructure searching. The 2D similarity searching, as the fourth method in this context, is the most commonly used. The motivation for performing such searches follows from the aforementioned similarity property principle: All compounds exhibiting a similarity above a certain predefined threshold will be retrieved from the database by 2D similarity searching.[128–130] Mostly, similarity measures are based on molecular fingerprints, as mentioned in the Section II.A.

C. 3D database searches

In contrast to the aforementioned 2D database search methods, the structural graph of the molecules does not account for a match of two molecules in 3D searching; 3D database searching falls into the categories of pharmacophore-based, and shape- or volume-based methods, respectively. While pharmacophore-based queries give topological and geometric constraints for the search, the latter uses a query built from either the ligand shape or the accessible volume of the binding site. Combination of both query types has proven to be extremely useful in yielding high rates of actives when performing virtual screening campaigns.[131] While the simplest definition of a pharmacophore can be a combination of three points with 3D location and 1D property constraints,[132] in modern applications the feature-based pharmacophore approaches allow much more specific definitions of such models. They, in fact, represent specific binding modes of ligands to their targets. Sometimes it is not straightforward to derive only one pharmacophore model for different families of chemotypes, especially when there is no information about the 3D structure of the binding site. When building pharmacophore models using the ligand-based approach, it is generally assumed that all ligands exhibit the same binding mode. Unfortunately, this assumption often proves to be wrong and models generated under these conditions will likely fail when used as queries for 3D database searching. Therefore, it is advisable to build more than one model for different classes of ligands. In the case where the 3D structure of the target is determined, the structure-based pharmacophore model building procedure is straightforward. Tools, such as the program LIGANDSCOUT[94,95] allow one to rapidly build feature-based 3D pharmacophore models and analyze similarity among large number of them using novel and efficient pharmacophore alignment techniques.[133] A comprehensive overview on pharmacophore perception and searching is given in two books edited by Güner[134] and Langer and Hoffmann,[135] respectively.

Docking procedures, on the other hand, take the 3D shape of a ligand and a receptor and their complementarities into account.[136–146] There are many different docking programs available, and review about recent advances in docking and scoring, and the performance of the most important docking tools is given in papers by Krovat *et al.*[147] and Warren *et al.*[148] When comparing docking-based approaches to 3D database searching using pharmacophore models with respect to search speed, there is a clear advantage using the latter one. Therefore, it seems advisable to use pharmacophore-based 3D search methods as a prefiltering step before compute-intensive docking procedures.

V. APPLICATIONS

A. Ligand-based *in silico* screening

According to the methods outlined above, a diverse range of ligand-based virtual screening methods exists. Their degree of sophistication, and thus their ultimate computational cost, depend highly on the type of structural information being used.[42] In all cases, generally significant enrichments over random selection of molecules in databases are obtained. After the searching procedure, the top scoring molecules can be prioritized for experimental testing. As described before, similarity searches are the less expensive side and accordingly, all the molecules in a particular database can be scored by similarity to one or multiple bioactive ligands and then ranked to reflect decreasing probability of being active. One of the most successful approaches in this respect is the usage of self-organizing artificial neural networks that can generate projections of large data sets defined in high-dimensional space. The resulting self-organizing maps can be used in many applications in the drug discovery process, such as to analyze combinatorial libraries for their similarity or diversity and to select descriptors for structure–activity relationships.[149,150] Teckentrup *et al.*[151] have used this methodology for analyzing a set of 5,513 compounds containing 185 as identified hits (hit rate 3.4%).

In contrast to topological approaches, methods based on geometrical representations of molecular structures can be used instead. Among them, flexible superimposition of molecules onto one or multiple conformations of a reference bioactive ligand is a well-established methodology

in virtual screening.[152] More sophisticated methods like pharmacophore-based searching can be used if the level of information on the bioactivity of ligands is already higher and if some knowledge is available about structure–activity relationships. Although initially slow to gain an industrial foothold, pharmacophore approaches have subsequently been applied to many therapeutic targets for the virtual screening of compound databases.[153–155] Successful applications of the use of pharmacophores in virtual screening include the identification of hits for a variety of targets such as protein kinase C,[156] farnesyltransferase,[157] HIV integrase,[158,159] endothelial differentiation gene receptor antagonists,[160] urotensin antagonists,[161] CCR5 antagonists,[162] and mesangial cell proliferation inhibitor discovery,[163] endothelin receptor antagonists,[164] and sigma receptor ligands[165] to mention a few. Pharmacophores have also been generated for numerous ADME/Tox-related proteins.[166] These efforts suggest that pharmacophore-based *in silico* searching approaches have considerable versatility and applicability to be used with difficult biological targets.

B. Structure-based *in silico* screening

Structure-based virtual screening of 3D molecular databases has been successfully applied to find new hits in structure-based design.[167–169] Krier *et al.*[170] optimized phosphodiesterase 4 inhibitors using molecular docking of a combinatorial library. Barreca *et al.*[171] used structure-based 3D pharmacophore generation and multiconformational database searching together with molecular docking in order to identify new scaffolds of HIV-1 non-nucleoside reverse transcriptase inhibitors (NNRTIs). Rollinger *et al.*[172] used a similar approach to discover acetylcholine esterase inhibitors from a 3D database of natural compounds. Wang and colleagues[173] discovered a novel and potent dopamine transporter inhibitor by 3D-database pharmacophore searching and subsequent chemical modification as well as inhibitors of *Tritrichomonas fetus* hypoxanthine-guanine-xanthine phosphoribosyltransferase (HGXPRT).[174] Zhang *et al.*[175] identified several potent and selective non-peptidic protein tyrosine phosphatase 1B (PTB1B) inhibitors using the DOCK methodology. Other successful structure-based virtual screening campaigns have been described for the estrogen receptor b,[176] the angiotensin converting enzyme-2 (ACE2),[177] influenza virus neuraminidase,[178] peroxisome proliferators-activated receptor (PPAR) ligands,[179] dipeptidyl peptidase IV inhibitors,[180] and for HMG CoA reductase inhibitors,[181] to mention a few. Even in the absence of experimentally determined 3D structures of target proteins such approaches can be useful: Homology models in these cases can substitute experimentally determined models and subsequent virtual screening experiments yield satisfactory results, as demonstrated recently both by Evers and Klabunde[182] and by Triballeau *et al.*[183]

C. Assessing affinity profiles using parallel *in silico* screening

In a series of recent reviews, Ekins, Mestres, Testa, and Rognan[184–186] report the current and latest trends in virtual screening, now aiming at *in silico* modeling of chemogenomics and of effects related to polypharmacology. One of the first attempts to predict activity profiles has been described by Poroikov *et al.*[187] The PASS approach was based on the analysis of structure–activity relationships for a training set of molecules consisting of about 35,000 biologically active compounds extracted from the literature. In another pioneering work, relationships between the chemical structures of 48 compounds and their pharmacological profile against a set of more than 70 receptors, transporters, and channels relevant to a central nervous system (CNS)-oriented project were analyzed.[188] Along the same line, a biospectra similarity analysis was performed by clustering a set of 1,567 drugs for which percent inhibition values determined at single high ligand concentration was available for a set of 92 assays.[189] In yet another study, the MDL Drug Data Report (MDDR) database was used as a source of ligands annotated to the four major target classes, namely, enzymes, G-protein-coupled receptors (GPCRs), nuclear receptors, and ligand-gated ion channels.[190] All these studies clearly indicate that research in virtual screening should no longer be focused on single targets but rather on entire families of related proteins or on full metabolic pathways. Techniques are nowadays in our hand that allow for rapid and efficient *in silico* profiling of compounds, even before they are synthesized. In the next section, an example is given for the use of multiple structure-based pharmacophore models built for various viral targets for *in silico* bio-affinity profile assessment.

D. Example: Parallel pharmacophore-based virtual screening

The relationship between the structural properties of ligands, especially when described with pharmacophore descriptors, and their biological activity has been demonstrated to be useful for drug discovery applications.[191] When there are experimentally determined 3D structure of the targets available, pharmacophore models are highly reliable for retrieving bioactive compounds. Thus, in a first proof-of-principle study that was carried out with the software tools LIGANDSCOUT[95] and CATALYST,[63] a model workflow for the application of parallel pharmacophore-based virtual screening on a set of 50 structure-based pharmacophore models built for various viral targets and 100 antiviral compounds were elaborated.[192, 193] The latter were screened against all pharmacophore models in order to determine if their real biological targets could be correctly predicted *via* an enrichment of corresponding pharmacophores matching these ligands.

TABLE 10.1 Target Proteins Used for Structure-Based Pharmacophore Modeling

Target protein	Disease	Function	Mechanism of inhibition
HIV protease	HIV infection, AIDS	Cleavage of gag and gag-pol precursor polyproteins into mature, structural and functional viral proteins	Inhibition at active site
HIV reverse transcriptase	HIV infection, AIDS	Synthesis of a double-stranded DNA from virus RNA for integration into host chromosomal DNA and transcription to viral genomic and messenger RNA	Inhibition at allosteric site
Influenza virus neuraminidase	Influenza	Viral envelope glycoprotein involved in viral release, cleavage of sialic acid residues of new virus particles and host membranes	Inhibition at active site
Human rhinovirus coat protein	Common cold	Attachment to host cell receptor, viral entry, and uncoating	Binding in hydrophobic pocket (capsid stabilization)
Hepatitis C virus RNA polymerase	Hepatitis C	Viral replication, transcription of genomic RNA	Inhibition at three different allosteric sites

1. *Profiling antiviral compounds*

a. Target proteins used in the study

This study involved the selection of several ligand binding sites belonging to five viral target proteins which were represented in the screening setup by sets of chemical feature-based pharmacophore models. In order to be selected for our application example, a target had to fulfill certain criteria. Since pharmacophore hypotheses were derived in a structure-based approach, the existence of a sufficient number of complexes from the PDB[52] was a requirement. Further, we focused on proteins, whose inhibition offers therapeutic strategies in the combat of highly relevant viral diseases. These include human immunodeficiency virus (HIV) infection, influenza, common cold, and hepatitis C. To increase the applicability of these conclusions, the study aimed to provide diversity concerning the nature of the proteins as well as inhibitory mechanisms. The macromolecular targets used are listed in Table 10.1.

b. Pharmacophore model generation

For each of the selected target proteins, a set of 10 pharmacophore models was generated based on receptor–inhibitor complexes from the PDB.[52] The three allosteric sites of hepatitis C virus (HCV) polymerase were represented by three, five, and two models, respectively. Structure-based pharmacophore model generation was performed with the software LIGANDSCOUT using the default settings and the standard workflow and thereafter exported into the CATALYST data format for 3D multiconformational database searching.

c. Structure preparation and parallel screening procedure

Compound structure models of these 100 inhibitors were generated by a standard procedure, starting from SMILES code, followed by 3D structure generation within the CATALYST software. In a subsequent step, 3D structure minimization was performed, followed by conformational model generation using the following parameter settings for building the 3D multiconformer database: A maximum of 250 conformers, the *best* conformer generation algorithm, and an energy threshold of 20 kcal above the calculated lowest energy conformation. All 100 antiviral compounds were then screened against the 50 structure-based pharmacophore models using both the *fast flexible search* algorithm within CATALYST. The results were analyzed in order to determine which molecule was retrieved as active by which models. For each of the compounds a pharmacophore hit list was generated, that is a set of hypotheses by which this particular ligand is retrieved (Figure 10.2). A correct model represents the target for which the ligand is a known inhibitor. An incorrect model means that the pharmacophore hypothesis was built for another target. For evaluation, the percentage of correct and incorrect models in the pharmacophore hit list was calculated. Furthermore, it was interesting to explore

FIGURE 10.2 Heat map obtained from parallel virtual screening experiment. Rows: Antiviral compounds; Columns: Pharmacophore models. Green signal: compound is found with model built from correct target; red signal: compound is found with a model from another target.

to what extent the available correct models were retrieved and which false target was most extensively identified. The relation between the last two parameters was found critical for activity profile prediction accuracy. Additionally, the pharmacophore models were validated in a similar manner inspecting the hit map vertically: The percentage of known active versus inactive compounds in the obtained hit list plays a vital role for pharmacophore quality.

d. Results

The interpretation of the results obtained in this parallel screening study is straightforward. Visualization of hits (true ones and false positives) in a heat map (Figure 10.2) where the pharmacophore models are represented in columns and the ligands in rows: green boxes indicate correct retrieval of a compound by a model for the correct target, while red

boxes indicate retrieval of a compound with a model from another target, where the molecule is presumably inactive. A critical issue, however, in this context must be noted. It was assumed that one molecule displays activity only at one particular target and is inactive at all the others – an assumption which has not been tested and verified in biological assays. Implicitly, it should be kept in mind that what is referred to as an incorrect or false model or a misleading prediction, might in fact be a correct one and only indicate a new and so far unknown activity spectrum of a compound. The activity profile obtained with the *fast flexible search* algorithm in CATALYST was convincing. While a majority of inhibitors (89%) was predicted with the correct activity profiles, the prediction was wrong in only 8% of the cases. For three of the compounds no mapping hypothesis was found and therefore no clues concerning their biological activities were gained, which is indicative of a problem occurring in the conformational analysis procedure.

2. Profiling HIV protease inhibitors

As an extension of the previous example, in this section, we describe the use of parallel pharmacophore-based *in silico* screening in order to model selectivity of ligands. A set of HIV protease inhibitor pharmacophore models was used to screen four datasets of compounds, with known affinities to different aspartate proteases.[194] A set of known HIV protease inhibitors, a group of other protease inhibitors, and two sets of inactive compounds – one derived from a literature reference, the other one randomly extracted from a large virtual drug-like library were selected and submitted to the virtual screening procedure as outlined above. After screening with the HIV protease inhibitor pharmacophore collection, the activity profiles obtained were analyzed and the expected trends were found: (i) higher enrichment for known active ligands, followed (ii) by the related protease inhibitors, and (iii) lowest retrieval rates for the inactive structures.

VI. CONCLUSION AND FUTURE DIRECTIONS

Over the years, the pharmaceutical industry has learned to accept that *in silico* screening methods indeed can be an efficient complement to HTS[195,196] to the point that they have undoubtedly become an integral part of today's hit identification and lead generation process.[197–199] The vast amount of data produced in pharmaceutical research enforces the use of database and data mining techniques. In contrast to technology-driven HTS, virtual screening is a knowledge-driven approach that requires structural information either on bioactive ligands for the target of interest (ligand-based virtual screening) or on the target itself (target-based virtual screening). Comparative studies have suggested that information about a target obtained from known bioactive ligands is as valuable as knowledge of the target structures for identifying novel bioactive scaffolds through virtual screening.[200] Therefore, the final choice for a method to use will ultimately depend on the type and amount of information available without *a priori* having a large impact on performance of pharmacophore selectivity screening. If virtual ligand screening extended QSAR via the addition of chemical dimension, recent trends in virtual affinity profiling are adding another biological dimension. A wave of new methods that are capable of estimating the pharmacological profile of molecules on multiple targets have been recently reported. These methods promise to have a strong influence in drug discovery as a means for detecting potential side effects of compounds due to off-target affinities during the optimization process.[201] Notwithstanding, it should be recognized that the current expansion of these methods is mainly a consequence of the important progress experienced by some coordinated initiatives dedicated to data collection, classification and storage, both in extracting pharmacological data for ligands as well as in gathering structural information for proteins.[202] Integrated systems that are suitable for optimizing new compounds in parallel for their potency, selectivity, and ADMET profile will be available soon, and the reliability of *in silico* models will be improved. One should never forget, however, that the greater diversity in a dataset, the more difficult is the generation of predictive structure-activity models. Models developed based on compounds representing only a subspace of the chemical universe usually have low predictivity for compounds beyond its boundaries. Having this in mind, one might be more successful when starting a project with optimizing compounds that have already successfully been used in other target areas.[203]

REFERENCES

1. Kubinyi, H. Chance favors the prepared mind. From serendipity to rational drug design. *J. Recept. Signal Transduct. Res.* 1999, *19*, 15–39.
2. Lahana, R. How many hits from HTS?. *Drug Discov. Today* 1999, *4*, 447–448.
3. Kubinyi, H. The changing landscape in drug discovery. In *Computational Approaches to Structure Based Drug Design* (Stroud, R. M., Ed.). Royal Society of Chemistry: London, 2007, pp. 24–45.
4. Drews, J. Drug discovery: a historical perspective. *Science* 2000, *287*, 1960–1964.
5. Brown, F. Chemoinformatics: What is it and how does it impact drug discovery?. *Annu. Rep. Med. Chem.* 1998, *33*, 375–384.
6. Richie, T. Chemoinformatics; manipulating chemical information to facilitate decision-making in drug discovery. *Drug Discov. Today* 2001, *6*, 813–814.
7. Johnson, D. E., Blower, P. E., Jr., Myatt, G. J., Wolfgang, G. H. I. Chem-tox informatics: data mining using a medicinal chemistry building block approach. *Curr. Opin. Drug Discov. Dev.* 2001, *4*, 92–101.
8. Oprea, T. I., Gottfries, J. Chemography: the art of navigating in chemical space. *J. Combin. Chem.* 2001, *3*, 157–166.
9. Oprea, T. I. Chemical space navigation in lead discovery. *Curr. Opin. Chem. Biol.* 2002, *6*, 384–389.

10. Leach, A. R., Gillet, V. J. *An Introduction to Chemoinformatics*. Kluwer Academic Publishers: Dordrecht, Netherlands, 2003.
11. Oprea, T. I. Chemoinformatics in drug discovery. In *Methods and Principles in Medicinal Chemistry* (Mannhold, R., Kubinyi, H., Folkers, G., Eds), Vol. 23. Wiley-VCH: Weinheim, 2004.
12. Kola, I., Landis, J. Can the pharmaceutical industry reduce attrition rates?. *Nat. Rev. Drug Discov.* 2004, *3*, 711–715.
13. Schuster, D., Laggner, C., Langer, T. Why drugs fail – A study on side effects in new chemical entities. *Curr. Pharmaceut. Des.* 2005, *11*, 3545–3559.
14. Valler, M. J., Green, D. Diversity screening versus focused screening in drug discovery. *Drug Discov. Today* 2000, *5*, 286–293.
15. Van de Waterbeemd, H. Which *in vitro* screens guide the prediction of oral absorption and volume of distribution?. *Basic Clin. Pharmacol. Toxicol.* 2005, *96*, 162–166.
16. Oprea, T. I. Current trends in lead discovery: Are we looking for the appropriate properties?. *J. Comput.-Aided Mol. Des.* 2002, *16*, 325–334.
17. Langer, T., Krovat, E.-M. Chemical feature-based pharmacophores and virtual library screening for discovery of new leads. *Curr. Opin. Drug Discov. Dev.* 2003, *6*, 370–376.
18. Güner, O., Clement, O., Kurogi, Y. Pharmacophore modeling and three dimensional database searching for drug design using catalyst: recent advances. *Curr. Med. Chem.* 2004, *11*, 2991–3005.
19. Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P. From data mining to knowledge discovery. In *Advances in Knowledge Discovery and Data Mining* (Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., Uthurusamy, R., Eds). AAAI Press: Menlo Park, CA, 1996, pp. 1–37.
20. Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inform. Comput. Sci.* 1988, *28*, 31–36.
21. Weininger, D., Weininger, A., Weininger, J. L. SMILES. 2. Algorithm for generation of unique SMILES notation. *J. Chem. Inform. Comput. Sci.* 1989, *29*, 97–101.
22. Wiswesser, W. J. Historical development of chemical notations. *J. Chem. Inform. Comput. Sci.* 1985, *25*, 258–263.
23. Stein, S. E., Heller, S. R., Tchekhovski, D. An open standard for chemical structure representation: the IUPAC chemical identifier. *Proceedings of the 2003 International Chemical Information Conference (Nimes, France), Infonortics*, 2003, pp. 131–143.
24. SMARTS, a language created for matching substructures of chemical files created by Daylight Chemical Information Systems, Inc. alongside the SMILES format.
25. Tanimoto, T. T. IBM Internal Report November 17 1957.
26. Elsevier MDL, 2440 Camino Ramon, Suite 300, San Ramon, CA 94583, USA. http://www.mdl.com
27. Durant, J. L., Leland, B. A., Henry, D. R., Nourse, J. G. Reoptimization of MDL keys for use in drug discovery. *J. Chem. Inform. Comput. Sci.* 2002, *42*, 1273–1280.
28. Daylight Chemical Information Systems, Inc, 7401 Los Altos, Suite 360, Mission Viejo, CA 92691, USA. http://www.daylight.com
29. Wiener, H. Structural determination of paraffin boiling points. *J. Am. Chem. Soc.* 1947, *69*, 17–20.
30. Wiener, H. Correlation of heat of isomerization, and differences in heats of vaporization of isomers, among the paraffin hydrocarbons. *J. Am. Chem. Soc.* 1947, *69*, 2636–2638.
31. Wiener, H. Influence of interatomic forces of paraffin boiling points. *J. Chem. Phys.* 1947, *15*, 766.
32. Randic, M. On characterization of molecular branching. *J. Am. Chem. Soc.* 1975, *97*, 6609–6615.
33. Kier, L. B., Hall, L. H. The nature of structure-activity relationships and their relation to molecular connectivity. *Eur. J. Med. Chem.* 1977, *12*, 307–312.
34. Kier, L. B., Hall, L. H. An electrotopological-state index for atoms in molecules. *Pharmaceut. Res.* 1990, *7*, 801–807.
35. Burden, F. R. Molecular identification numbers for substructure searches. *J. Chem. Inform. Comput. Sci.* 1989, *29*, 225–228.
36. Gálvez, J., Garcia, R., Salabert, M. T., Soler, R. Charge indexes. New topological descriptors. *J. Chem. Inform. Comput. Sci.* 1994, *34*, 520–525.
37. Estrada, E., Uriarte, E. Recent advances on the role of topological indices in drug discovery. *Curr. Med. Chem.* 2001, *8*, 1573–1588.
38. Hopkinson, G. A. Structure and substructure searching. In *Encyclopedia of Computational Chemistry* (Von Ragué Schleyer, P., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds), Vol. 4. John Wiley: New York, 1998, pp. 2764–2771.
39. Bron, C., Kerbosch, J. A. G. M. Algorithm 457: Finding all cliques of an undirected graph. *Commun. ACM* 1973, *16*, 575–577.
40. Carraghan, R., Pardalos, P. M. An exact algorithm for the maximum clique problem. *Oper. Res. Lett.* 1990, *9*, 375–382.
41. Rarey, M., Dixon, J. S. Feature trees: a new molecular similarity measure based on tree matching. *J. Comput.-Aided Mol. Des.* 1998, *12*, 471–490.
42. Rarey, M., Fricker, P., Hindle, S., Metz, G., Rummey, C., Zimmermann, M. Feature trees: theory and applications from large-scale virtual screening to data analysis. In *Pharmacophores and Pharmacophore Searches* (Langer, T., Hoffmann, R. D., Eds). Wiley-VCH: Weinheim, 2006, pp. 81–116.
43. FlexX as well as FTrees are available from BioSolveIT GmbH, An der Ziegelei 75, 53757 Sankt Augustin, Germany (www.biosolveit.com).
44. Baringhaus, K.-H., Hessler, G. Fast similarity searching and screening hit analysis. *Drug Discov. Today Technol.* 2004, *1*, 197–202.
45. Lengauer, T., Lemmen, C., Rarey, M., Zimmermann, M. Novel technologies for virtual screening. *Drug Discov. Today* 2004, *9*, 27–34.
46. Briem, H., Lessel, U. F. *In vitro* and *in silico* affinity fingerprints: finding similarities beyond structural classes. *Perspect. Drug Discov. Des.* 2000, *20*, 231–244.
47. Zupan, J., Gasteiger, J. *Neural Networks in Chemistry*, 2nd Edition. Wiley-VCH: Weinheim, 1999.
48. ADRIANA.*Code* is available from Molecular Networks GmbH, Henkestraße 91, 91052 Erlangen, Germany (http://www.molecular-networks.com)
49. Gasteiger, J., Marsili, M. Iterative partial equalization of orbital electronegativity – A rapid access to atomic charges. *Tetrahedron* 1980, *36*, 3219–3228.
50. Gasteiger, J., Hutchings, M. G. Quantification of effective polarisability. Applications to studies of X-ray photoelectron spectroscopy and alkylamine protonation. *J. Chem. Soc. Perkin.* 1984, *2*, 559–564.
51. The CSD is available from the Cambridge Crystallographic Data Center, 12 Union Road, Cambridge, CB2 1EZ, UK (http://www.ccdc.cam.uk)
52. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., Bourne, P. E. The protein data bank. *Nucleic Acids Res.* 2000, *28*, 235–242. The Protein Database is available on the web (http://www.rcsb.org/pdb)
53. Chemical Abstract Services, 2540 Olentangy River Road, PO Box 3012, Columbus, OH 43202, USA. (http://www.cas.org/cgi-bin/cas/regreport.pl)
54. Sadowski, J. Three-dimensional structure generation: automation. In *Encyclopedia of Computational Chemistry* (Von Ragué Schleyer, P., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds), Vol. 5. John Wiley: New York, 1998, pp. 2976–2988.
55. Sadowski, J., Gasteiger, J. From atoms to bonds to three-dimensional atomic coordinates: automatic model builders. *Chem. Rev.* 1993, *93*, 2567–2581.
56. Sadowski, J., Gasteiger, J., Klebe, G. Comparison of automatic three-dimensional model builders using 639 X-ray structures. *J. Chem. Inform. Comput. Sci.* 1994, *34*, 1000–1008.

57. Gasteiger, J., Rudolph, C., Sadowski, J. Automatic generation of 3D-atomic coordinates for organic molecules. *Tetrahedron Comput. Meth.* 1990, *3*, 537–547.
58. CORINA is available from Molecular Networks GmbH, Henkestraße 91, 91052 Erlangen, Germany (http://www.molecular-networks.com)
59. Pearlman, R. S. Rapid generation of high quality approximate 3D molecular structures. *Chem. Des. Autom. News* 1987, *2*, 1–7.
60. Pearlman, R. S. 3D molecular structures: generation and use in 3D-searching. In *3D QSAR in Drug Design: Theory, Methods, and Applications* (Kubinyi, H., Ed.). ESCOM Science Publishers: Leiden, 1993, pp. 21–58.
61. CONCORD is available from Tripos, 1699 South Hanley Road, St. Louis, MO 63144-2319, USA (http://www.tripos.com)
62. OMEGA is available from OpenEye Scientific Software, 9 Bisbee Court, Suite D Santa Fe, NM 87508, USA (http://www.eyesopen.com)
63. CATALYST is available from Accelrys, Inc. 10188 Telesis Court, Suite 100 San Diego, CA 92121, USA (http://www.accelrys.com)
64. Boström, J. Reproducing the conformations of protein-bound ligands: a critical evaluation of several popular conformational searching tools. *J. Comput.-Aided Mol. Des.* 2001, *15*, 1137–1152.
65. Boström, J., Greenwood, J. R., Gottfries, J. Assessing the performance of OMEGA with respect to retrieving bioactive conformations. *J. Mol. Graph. Model.* 2003, *21*, 449–462.
66. Kirchmair, J., Laggner, C., Wolber, G., Langer, T. Comparative analysis of protein-bound ligand conformations with respect to catalyst's conformational space subsampling algorithms. *J. Chem. Inform. Model.* 2005, *45*, 422–430.
67. Kirchmair, J., Wolber, G., Laggner, C., Langer, T. Comparative performance assessment of the conformational model generators omega and catalyst: a large-scale survey on the retrieval of protein-bound ligand conformations. *J. Chem. Inform. Model.* 2006, *46*, 1848–1861.
68. Leach, A. R. A survey of methods for searching the conformational space of small and medium-sized molecules. In *Reviews in Computational Chemistry* (Lipkowitz, K. B., Boyd, D. B., Eds), Vol. 2. VCH Verlagsgesellschaft: New York, 1991, pp. 1–55.
69. Saunders, M. Stochastic exploration of molecular mechanics energy surfaces. Hunting for the global minimum. *J. Am. Chem. Soc.* 1987, *109*, 3150–3152.
70. Saunders, M. Stochastic search for the conformations of bicyclic hydrocarbons. *J. Comput. Chem.* 1989, *10*, 203–208.
71. McGarrah, D. B., Judson, R. S. Analysis of the genetic algorithm method of molecular conformation determination. *J. Comput. Chem.* 1993, *14*, 1385–1395.
72. Judson, R. S., Jaeger, E. P., Treasurywala, A. M., Peterson, M. L. Conformational searching methods. II. Genetic algorithm approach. *J. Comput. Chem.* 1993, *14*, 1407–1414.
73. Lybrand, T. P. Computer simulations of biomolecular systems using molecular dynamics and free energy perturbation methods. In *Reviews in Computational Chemistry* (Lipkowitz, K. B., Boyd, D. B., Eds), Vol. 1. VCH Verlagsgesellschaft: New York, 1990, pp. 295–320.
74. Chang, G., Guida, W. C., Still, W. C. An internal coordinate Monte Carlo method for searching conformational space. *J. Am. Chem. Soc.* 1989, *111*, 4379–4386.
75. Schönberger, H., Schwab, C. H., Hirsch, A., Gasteiger, J. Molecular modelling of fullerene dendrimers. *J. Mol. Model.* 2000, *6*, 379–395.
76. ROTATE is available from Molecular Networks GmbH, Henkestraße 91, 91052 Erlangen, Germany (http://www.molecular-networks.com)
77. Li, J., Ehlers, T., Sutter, J., Varma-O'Brien, S., Kirchmair, J. CAESAR: a new conformer generation algorithm based on recursive buildup and local rotational symmetry consideration. *J. Chem. Inform. Model.* 2007. ASAP Article 10.1021/ci700136x S1549-9596(70)00136-5 Web Release August 11, 2007
78. Wagener, M., Sadowski, J., Gasteiger, J. Autocorrelation of molecular surface properties for modelling corticosteroid binding globulin and cytosolic Ah receptor activity by neural networks. *J. Am. Chem. Soc.* 1995, *117*, 7769–7775.
79. Soltzberg, L. J., Wilkins, C. L. Molecular transforms: a potential tool for structure-activity studies. *J. Am. Chem. Soc.* 1977, *99*, 439–443.
80. Schuur, J. H., Selzer, P., Gasteiger, J. The coding of the three-dimensional structure of molecules by molecular transforms and its applications to structure-spectra correlations and studies of biological activity. *J. Chem. Inform. Comput. Sci.* 1996, *36*, 334–344.
81. Gasteiger, J., Schuur, J., Selzer, P., Steinhauer, L., Steinhauer, V. Finding the 3D structure of a molecule in its IR spectrum. *Fresen. J. Anal. Chem.* 1997, *359*, 50–55.
82. Hemmer, M. C., Steinhauer, V., Gasteiger, J. Deriving the 3D structure of organic molecules from their infrared spectra. *Vibrat. Spectrosc.* 1999, *19*, 151–154.
83. Wermuth, C.-G., Ganellin, C. R., Lindberg, P., Mitscher, L. A. Glossary of terms used in medicinal chemistry (IUPAC recommendations 1997). *Annu. Rep. Med. Chem.* 1998, *33*, 385–395.
84. Kurogi, Y., Güner, O. F. Pharmacophore modeling and three-dimensional database searching for drug design using catalyst. *Curr. Med. Chem.* 2001, *8*, 1035–1055.
85. Langer, T., Krovat, E.-M. Chemical feature-based pharmacophores and virtual library screening for discovery of new leads. *Curr. Opin. Drug Discov. Dev.* 2003, *6*, 370–376.
86. Langer, T., Hoffmann, R. D. Pharmacophore modeling: applications in drug discovery. *Exp. Opin. Drug Discov.* 2006, *1*, 261–267.
87. Manetti, F., Botta, M., Tafi, A. Application of pharmacophore models in medicinal chemistry. In *Pharmacophores and Pharmacophore Searches* (Langer, T., Hoffmann, R. D., Eds). Wiley-VCH: Weinheim, 2006, pp. 253–282.
88. Dixon, S. L., Smondyrev, A. M., Rao, S. R. PHASE: a novel approach to pharmacophore modeling and 3D database searching. *Chem. Biol. Drug Des.* 2006, *67*, 370–372.
89. PHASE is available from Schrödinger Inc., 120 W. 45th St, 29th Floor, New York, NY 10036, USA (http://www.schrodinger.com)
90. MOE is available from the Chemical Computing Group, Suite 910–1010 Sherbrooke St. W, Montreal, Canada H3A 2R7 (http://www.chemcomp.com)
91. Richmond, N. J., Willett, P., Clark, R. D. Alignment of three-dimensional molecules using an image recognition algorithm. *J. Mol. Graph. Model.* 2004, *23*, 199–209.
92. Cottrell, S. J., Gillet, V. J., Taylor, R., Wilton, D. J. Generation of multiple pharmacophore hypotheses using multiobjective optimisation techniques. *J. Comput.-Aided Mol. Des.* 2004, *18*, 665–682.
93. GALAHAD is available from Tripos, 1699 South Hanley Road, St. Louis, MO 63144-2319, USA (http://www.tripos.com)
94. Wolber, G., Langer, T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J. Chem. Inform. Model.* 2005, *45*, 160–169.
95. LIGANDSCOUT is available from Inte:Ligand GmbH, Mariahilferstrasse 74b/11, 1070 Vienna, Austria (http://www.inteligand.com)
96. Bramson, H. N., Corona, J., Davis, S. T., Dickerson, S. H., Edelstein, M., Frye, S. V., Gampe, R. T., Jr, Harris, P. A., Hassell, A., Holmes, W. D., Hunter, R. N., Lackey, K. E., Lovejoy, B., Luzzio, M. J., Montana, V., Rocque, W. J., Rusnak, D., Shewchuk, L., Veal, J. M., Walker, D. H., Kuyper, L. F. Oxindole-based inhibitors of cyclin-dependent kinase 2 (CDK2): design, synthesis, enzymatic activities, and X-ray crystallographic analysis. *J. Med. Chem.* 2001, *44*, 4339–4358.
97. Goodford, P. J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* 1985, *28*, 849–857.
98. Cramer, R. D., III, Patterson, D. E., Bunce, J. D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* 1988, *110*, 5959–5967.
99. Wold, S., Ruhe, A., Wold, H., Dunn, W. J. The covariance problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM J. Sci. Stat. Comput.* 1984, *5*, 735–743.

100. Pearlman, R. S., Smith, K. M. Novel software tools for chemical diversity. *Perspect. Drug Discov. Des.* 1998, *9*, 339–353.
101. Pearlman, R. S., Smith, K. M. J. Metric validation and the receptor-relevant subspace concept. *J. Chem. Inform. Comput. Sci.* 1999, *39*, 28–35.
102. Ferguson, A. M., Heritage, T. W., Jonathon, P., Pack, S. E., Phillips, L., Rogan, J., Snaith, P. J. EVA: a new theoretically based molecular descriptor for use in QSAR/QSPR analysis. *J. Comput.-Aided Mol. Des.* 1997, *11*, 143–152.
103. Turner, D. B., Willet, P., Ferguson, A. M., Heritage, T. W. Evaluation of a novel infrared range vibration-based descriptor (EVA) for QSAR studies. 1. General applications. *J. Comput.-Aided Mol. Des.* 1997, *11*, 409–422.
104. Todeschini, R., Lasagni, M., Marengo, E. New molecular descriptors for 2D- and 3D-structures. *Theory. J. Chemo.* 1994, *8*, 263–273.
105. Todeschini, R., Gramatica, P. 3D-modelling and prediction by WHIM descriptors. Part 5. Theory development and chemical meaning of WHIM descriptors. *Quant. Struct.-Act. Relat.* 1997, *16*, 113–119.
106. Todeschini, R., Consonni, V. The handbook to molecular descriptors. In *Methods and Principles in Medicinal Chemistry* (Mannhold, R., Kubinyi, H., Timmermann, H., Eds), Vol. 11. Wiley-VCH: Weinheim, 2000.
107. Xue, L., Bajorath, J. Molecular descriptors in chemoinformatics, computational combinatorial chemistry, and virtual screening. *Combin. Chem. High Throughput Screening* 2000, *3*, 363–372.
108. Lipinski, C. A., Lombardo, F., Dominy, B. W., Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 1997, *23*, 3–25.
109. Clark, D. E., Pickett, S. D. Computational methods for the prediction of 'drug-likeness'. *Drug Discov. Today* 2000, *5*, 49–58.
110. Ajay, A., Walters, W. P., Murcko, M. A. Can we learn to distinguish between 'drug-like' and 'non drug-like' molecules?. *J. Med. Chem.* 1998, *41*, 3314–3324.
111. Blake, J. F. Chemoinformatics – predicting the physicochemical properties of 'drug-like' molecules. *Curr. Opin. Biotechnol.* 2000, *11*, 104–107.
112. Teague, S. J., Davis, A. M., Leeson, P. D., Oprea, T. The design of leadlike combinatorial libraries. *Angew. Chem. Int. Ed. Engl.* 1999, *38*, 3743–3747.
113. Oprea, T. I., Davis, A. M., Teague, S. J., Leeson, P. D. Is there a difference between leads and drugs? A historical perspective. *J. Chem. Inform. Comput. Sci.* 2001, *41*, 1308–1315.
114. Li, A. P., Segall, M. Early ADME/Tox studies and *in silico* screening. *Drug Discov. Today* 2002, *7*, 25–27.
115. Li, A. P. Screening for human ADME/Tox drug properties in drug discovery. *Drug Discov. Today* 2001, *6*, 357–366.
116. Thompson, T. N. Early ADME in support of drug discovery: the role of metabolic stability studies. *Curr. Drug Metabol.* 2000, *1*, 215–241.
117. Keserü, G. M., Molnár, L. METAPRINT: a metabolic fingerprint. Application to cassette design for high-throughput ADME screening. *J. Chem. Inform. Comput. Sci.* 2002, *42*, 437–444.
118. Andrews, P. R., Craik, D. J., Martin, J. L. Functional group contributions to drug-receptor interactions. *J. Med. Chem.* 1984, *27*, 1648–1657.
119. The Beilstein Crossfire database is available from Elsevier MDL, 2440 Camino Ramon, Suite 300, San Ramon, CA 94583, USA (http://www.mdl.com)
120. The PubChem database contains more than 10 M chemical structures and is publicly available on the web (http://pubchem.ncbi.nlm.nih.gov)
121. Willet, P., Barnard, J. M., Downs, M. Chemical similarity searching. *J. Chem. Inform. Comput. Sci.* 1998, *38*, 983–996.
122. Agrafiotis, D. K. Diversity of chemical libraries. In *Encyclopedia of Computational Chemistry* (Von Ragué Schleyer, P., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds), Vol. 1. John Wiley: New York, 1998, pp. 742–761.
123. Agrafiotis, D. K. Multiobjective optimization of combinatorial libraries. *IBM J. Res. Dev.* 2001, *45*, 545–566.
124. Lewis, R. A., Pickett, S. D., Clark, D. E. Computer-aided molecular diversity analysis and combinatorial library design. In *Reviews in Computational Chemistry* (Lipkowitz, D. B., Boyd, D. B., Eds), Vol. 16. Wiley-VCH: New York, 2000, pp. 1–51.
125. Warr, W. A. Combinatorial chemistry and molecular diversity. An overview. *J. Chem. Inform. Comput. Sci.* 1997, *37*, 134–140.
126. Blaney, J. M., Martin, E. J. Computational approaches for combinatorial library design and molecular diversity analysis. *Curr. Opin. Chem. Biol.* 1997, *1*, 54–59.
127. Hopkinson, G. A. Structure and substructure searching. In *Encyclopedia of Computational Chemistry* (Von Ragué Schleyer, P., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds), Vol. 4. John Wiley: New York, 1998, pp. 2764–2771.
128. Willet, P. Similarity and clustering techniques in chemical information systems. In *Chemometric Series* (Bawden, D., Ed.), Vol. 18. Research Studies Press: Letchworth, UK, 1987, pp. 1–254.
129. Willet, P. Structural similarity measures for database setting. In *Encyclopedia of Computational Chemistry* (Von Ragué Schleyer, P., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds), Vol. 4. John Wiley: New York, 1998, pp. 2748–2756.
130. Willet, P. Chemoinformatics – similarity and diversity in chemical libraries. *Curr. Opin. Biotechnol.* 2000, *11*, 85–88.
131. Hoffmann, R. D., Meddeb, S., Langer, T. Use of 3D pharmacophore models in 3D database searching. In *Computational Medicinal Chemistry and Drug Discovery* (Tollenaere, J., De Winter, H., Langenaeker, W., Bultinck, P., Eds). Dekker, Inc.: New York, 2003, pp. 461–482.
132. Güner, O. F., Hughes, D. W., Dumont, L. M. An integrated approach to three-dimensional information management with MACCS-3D. *J. Chem. Inform. Comput. Sci.* 19991, *31*, 408–414.
133. Wolber, G., Dornhofer, A. A., Langer, T. Efficient overlay of small organic molecules using 3D pharmacophores. *J. Comput.-Aided Mol. Des.* 2006, *20*, 773–788.
134. Güner, O. F. (Ed.). *Pharmacophore Perception, Development, and Use in Drug Design*. International University Line: La Jolla, CA, 2000.
135. Langer, T., Hoffmann, R. D. Pharmacophores and pharmacophore searches. In *Methods and Principles in Medicinal Chemistry* (Mannhold, R., Kubinyi, H., Folkers, G., Eds), Vol. 32. Wiley-VCH: Weinheim, 2006.
136. Kuntz, I. D., Blaney, J. M., Oatley, S. J., Langridge, R., Ferrin, T. E. A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.* 1982, *161*, 269–288.
137. Jones, G., Willet, P. Docking small-molecules ligands into active sites. *Curr. Opin. Biotechnol.* 1995, *6*, 652–656.
138. Jones, G., Willet, P., Glen, R. C., Leach, A. R., Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* 1997, *267*, 727–748.
139. Shoichet, B. K., Bodian, D. L., Kuntz, I. D. Molecular docking using shape descriptors. *J. Comput. Chem.* 1992, *13*, 380–397.
140. Meng, E. C., Shoichet, B. K., Kuntz, I. D. Automated docking with grid-based energy evaluation. *J. Comput. Chem.* 1992, *13*, 505–524.
141. Sun, Y., Ewing, T. J. A., Skilman, A. G., Kuntz, I. D. CombiDOCK: structure-based combinatorial docking and library design. *J. Comput.-Aided Mol. Des.* 1998, *20*, 773–788.
142. Rarey, M., Kramer, B., Lengauer, T., Klebe, G. A fast flexible docking method using an incremental construction algorithm. *J. Mol. Biol.* 1996, *261*, 470–489.
143. Rarey, M., Kramer, B., Lengauer, T. Multiple automatic base selection: protein-ligand docking based on incremental construction without manual intervention. *J. Comput.-Aided Mol. Des.* 1997, *11*, 369–384.

144. Yang, J.-M., Kao, C.-Y. Flexible ligand docking using a robust evolutionary algorithm. *J. Comput. Chem.* 2000, *21*, 988–998.
145. Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., Olson, A. J. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* 1998, *19*, 1639–1662.
146. Steindl, T. M., Langer, T. Docking versus pharmacophore model generation: a comparison of high-throughput virtual screening strategies for the search of human rhinovirus coat protein inhibitors. *QSAR Combin. Sci.* 2005, *24*, 470–479.
147. Krovat, E.-M., Steindl, T., Langer, T. Recent advances in docking and scoring. *Curr. Comput.-Aided Drug Des.* 2005, *1*, 93–102.
148. Warren, G. L., Andrews, C. W., Capelli, A.-M., Clarke, b., LaLonde, J., Lambert, M. H., Lindvall, M., Nevins, N., Semus, S. F., Senger, S., Tedesco, G., Wall, I. D., Woolven, J. M., Peishoff, C. E., Head, M. S. A critical assessment of docking programs and scoring functions. *J. Med. Chem.* 2006, *49*, 5912–5931.
149. Bauknecht, H., Zell, A., Bayer, H., Levi, P., Wagener, M., Sadowski, J., Gasteiger, J. Locating biologically active compounds in medium-sized heterogeneous datasets by topological autocorrelation vectors: dopamine and benzodiazepine antagonists. *J. Chem. Inform. Comput. Sci.* 1996, *36*, 1205–1213.
150. Sadowski, J., Wagener, M., Gasteiger, J. Assessing similarity and diversity of combinatorial libraries by special autocorrelation function and neural networks. *Angew. Chem. Int. Ed. Engl.* 1995, *24*, 2674–2677.
151. Teckentrup, A., Briem, H., Gasteiger, J. Mining high-throughput screening data of combinatorial libraries: development of a filter to distinguish hits from nonhits. *J. Chem. Inform. Comput. Sci.* 2004, *44*, 626–634.
152. Mestres, J., Veeneman, G. H. Identification of 'latent hits' in compound screening collections. *J. Med. Chem.* 2003, *46*, 3441–3444.
153. Sprague, P. W. Automated chemical hypothesis generation and database searching with Catalyst. *Perspect. Drug Discov. Des.* 1995, *3*, 1–20.
154. Barnum, D., Greene, J., Smellie, A., Sprague, P. W. Identification of common functional configurations among molecules. *J. Chem. Inform. Comput. Sci.* 1996, *36*, 563–571.
155. Sprague, P. W., Hoffman, R. CATALYST pharmacophore models and their utility as queries for searching 3D databases. In *Computer-Assisted Lead Finding and Optimization* (Van de Waterbeemd, H., Testa, B., Folkers, G., Eds). Verlag Helvetica Chimica Acta: Basel, 1997, pp. 225–240.
156. Wang, S., Zaharevitz, D. W., Sharma, R., Marquez, V. E., Lewin, N. E., Du, L., Blumberg, P. M., Milne, G. W. A. The discovery of novel, structurally diverse protein kinase C agonists through computer 3D-database pharmacophore search. Molecular modeling studies. *J. Med. Chem.* 1994, *37*, 4479–4489.
157. Kaminski, J. J., Rane, D. F., Snow, M. E., Weber, L., Rothofsky, M. L., Anderson, S. D. S. L. L. Identification of novel farnesyl protein transferase inhibitors using three-dimensional searching methods. *J. Med. Chem.* 1997, *40*, 4103–4112.
158. Nicklaus, M. C., Neamati, N., Hong, H., Mazumder, A., Sunder, S., Chen, J., Milne, G. W. A., Pommier, Y. HIV-1 integrase pharmacophore: discovery of inhibitors through three-dimensional database searching. *J. Med. Chem.* 1997, *40*, 920–929.
159. Carlson, H. A., Masukawa, K. M., Rubins, K., Bushman, F. D., Jorgenson, W. L., Lins, R. D., Briggs, J. M., McCammon, J. A. Developing a dynamic pharmacophore model for HIV-1 integrase. *J. Med. Chem.* 2000, *43*, 2100–2114.
160. Koide, Y., Hasegawa, T., Takahashi, A., Endo, A., Mochizuki, N., Nakagawa, M., Nishida, A. Development of novel EDG3 antagonists using a 3D database search and their structure–activity relationships. *J. Med. Chem.* 2002, *45*, 4629–4638.
161. Flohr, S., Kurz, M., Kostenis, E., Brkovich, A., Fournier, A., Klabunde, T. Identification of nonpeptidic urotensin II receptor antagonists by virtual screening based on a pharmacophore model derived from structure–activity relationships and nuclear magnetic resonance studies on urotensin II. *J. Med. Chem.* 2002, *45*, 1799–1805.
162. Debnath, A. K. Generation of predictive pharmacophore models for CCR5 antagonists: study with piperidine- and piperazine-based compounds as a new class of HIV-1 entry inhibitors. *J. Med. Chem.* 2003, *46*, 4501–4515.
163. Kurogi, Y., Miyata, K., Okamura, T., Hashimoto, K., Tsutsumi, K., Nasu, M., Moriyasu, M. Discovery of novel mesangial cell proliferation inhibitors using a three-dimensional database searching method. *J. Med. Chem.* 2001, *44*, 2304–2307.
164. Funk, O. F., Kettmann, V., Drimal, J., Langer, T. Chemical function based pharmacophore generation of endothelin-A selective receptor antagonists. *J. Med. Chem.* 2004, *47*, 2750–2760.
165. Laggner, C., Schieferer, C., Fiechtner, B., Poles, G., Hoffmann, R. D., Glossmann, H., Langer, T., Moebius, F. F. Discovery of high-affinity ligands of sigma-1 receptor, ERG2, and emopamil binding protein by pharmacophore modeling and virtual screening. *J. Med. Chem.* 2004, *48*, 4754–4764.
166. Ekins, S., Swaan, P. W. Development of computational models for enzymes, transporters, channels and receptors relevant to ADME/TOX. *Rev. Comput. Chem.* 2004, *20*, 333–415.
167. Klebe, G. (Ed.) Virtual screening: an alternative or complement to high throughput screening, *Perspect. Drug Discov. Des.* 2000, *20*, 1–187.
168. Bissantz, C., Folkers, G., Rognan, D. Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations. *J. Med. Chem.* 2000, *43*, 4759–4767.
169. Böhm, H.-J., Schneider, G. Virtual screening for bioactive molecules. In *Methods and Principles in Medicinal Chemistry* (Mannhold, R., Kubinyi, H., Timmermann, H., Eds), Vol. 10. Wiley-VCH: Weinheim, 2000, pp. 1–307.
170. Krier, M., de Araujo-Junior, J. X., Schmitt, M., Duranton, J., Justiano-Basaran, H., Lugnier, C., Bourguignon, J.-J., Rognan, D. Design of small-sized libraries by combinatorial assembly of linkers and functional groups to a given scaffold: application to the structure-based optimization of a phosphodiesterase 4 inhibitor. *J. Med. Chem.* 2005, *48*, 3816–3822.
171. Barreca, M. L., De Luca, L., Iraci, N., Rao, A., Ferro, S., Maga, G., Chimirri, A. Structure-based pharmacophore identification of new chemical scaffolds as non-nucleoside reverse transcriptase inhibitors. *J. Chem. Inform. Model.* 2007, *47*, 557–562.
172. Rollinger, J. M., Hornick, A., Langer, T., Stuppner, H., Prast, H. Acetylcholinesterase inhibitory activity of scopolin and scopoletin discovered by virtual screening of natural products. *J. Med. Chem.* 2004, *47*, 6248–6254.
173. Wang, S., Sakamuri, S., Enyedy, I. J., Kozikowski, A. P., Deschaux, O., Bandyopadhyay, B. C., Tella, S. R., Zaman, W. A., Johnson, K. M. Discovery of a novel dopamine transporter inhibitor, 4-hydroxy-1-methyl-4-(4-methylphenyl)-3-piperidyl 4-methylphenyl ketone, as a potential cocaine antagonist through 3D-database pharmacophore searching. Molecular modeling, structure-activity relationships, and behavioral pharmacological studies. *J. Med. Chem.* 2000, *43*, 351–360.
174. Aronov, A. M., Munagala, N. R., Ortiz de Montellano, P. R., Kuntz, I. D., Wang, C. Rational design of selective submicromolar *Tritrichomonas foetus* hypoxanthine-guanine-xanthine phosphoribosyltransferase. *Biochemistry* 2000, *39*, 4684–4691.
175. Sarmiento, M., Wu, L., Keng, Y.-F., Song, L., Luo, Z., Huang, Z., Wu, G.-Z., Yuan, A. K., Zhang, Z.-Y. Structure-based discovery of small molecule inhibitors targeted to protein tyrosine phosphatase 1B. *J. Med. Chem.* 2000, *43*, 146–155.
176. Zhao, L., Brinton, R. D. Structure-based virtual screening for plant-based ER-selective ligands as potential preventative therapy against age-related neurodegenerative diseases. *J. Med. Chem.* 2005, *48*, 3463–3466.

177. Rella, M., Rushworth, C. A., Guy, J. L., Turner, A. J., Langer, T., Jackson, R. M. Structure-based pharmacophore design and virtual screening for novel angiotensin converting enzyme 2 inhibitors. *J. Chem. Inform. Model.* 2006, *46*, 708–716.
178. Steindl, T., Langer, T. Influenza virus neuraminidase inhibitors: generation and comparison of structure-based and common feature pharmacophore hypotheses and their application in virtual screening. *J. Chem. Inform. Comput. Sci.* 2004, *44*, 1849–1856.
179. Lu, I. L., Huang, C.-F., Peng, Y.-F., Lin, Y.-T., Hsieh, H.-P., Chen, C.-T., Lien, T.-W., Lee, H.-J., Mahindroo, N., Prakash, E., Yueh, A., Chen, H.-Y., Goparaju, C. M. V., Chen, X., Liao, C.-C., Chao, Y.-S., Hsu, J. T.-A., Wu, S.-Y. Structure-based drug design of a novel family of PPAR partial agonists: virtual screening, X-ray crystallography, and *in vitro*/*in vivo* biological activities. *J. Med. Chem.* 2006, *49*, 2703–2712.
180. Ward, R. A., Perkins, T. D. J., Stafford, J. Structure-based virtual screening for low molecular weight chemical starting points for dipeptidyl peptidase IV inhibitors. *J. Med. Chem.* 2005, *48*, 6991–6996.
181. Zhang, Q. Y., Wan, J., Xu, X., Yang, G. F., Ren, Y. L., Liu, J. J., Wang, H., Guo, Y. Structure-based rational quest for potential novel inhibitors of human HMG-CoA reductase by combining CoMFA 3D QSAR modeling and virtual screening. *J. Combin. Chem.* 2007, *9*, 131–138.
182. Evers, A., Klabunde, T. Structure-based drug discovery using GPCR homology modeling: successful virtual screening for antagonists of the alpha1A adrenergic receptor. *J. Med. Chem.* 2005, *48*, 1088–1097.
183. Triballeau, N., Acher, F., Brabet, I., Pin, J.-P., Bertrand, H.-O. Virtual screening workflow development guided by the "receiver operating characteristic" curve approach. Application to high-throughput docking on metabotropic glutamate receptor subtype 4. *J. Med. Chem.* 2005, *48*, 2534–2547.
184. Ekins, S., Mestres, J., Testa, B. *In silico* pharmacology for drug discovery: methods for virtual ligand screening and profiling. *Br. J. Pharmacol.* 2007, *152*, 9–20.
185. Ekins, S., Mestres, J., Testa, B. *In silico* pharmacology for drug discovery: applications to targets and beyond. *Br. J. Pharmacol.* 2007, *152*, 21–37.
186. Rognan, D. Chemogenomic approaches to rational drug design. *Br. J. Pharmacol.* 2007, *152*, 38–52.
187. Poroikov, V. V., Filimonov, D. A., Borodina, Y. V., Lagunin, A. A., Kos, A. Robustness of biological activity spectra predicting by computer program PASS for noncongeneric sets of chemical compounds. *J. Chem. Inform. Comput. Sci.* 2000, *40*, 1349–1355.
188. Poulain, R., Horvath, D., Bonnet, B., Eckhoff, C., Chapelain, B., Bodinier, M. C., Déprez, B. From hit to lead. Analyzing structure–profile relationships. *J. Med. Chem.* 2001, *44*, 3391–3401.
189. Fliri, A. F., Loging, W. T., Thadeio, P. F., Volkmann, R. A. Biospectra analysis: model proteome characterizations for linking molecular structure and biological response. *J. Med. Chem.* 2005, *48*, 6918–6925.
190. Schuffenhauer, A., Zimmermann, J., Stoop, R., van der Vyver, J. J., Lecchini, S., Jacoby, E. An ontology for pharmaceutical ligands and its application for *in silico* screening and library design. *J. Chem. Inform. Comput. Sci.* 2002, *42*, 947–955.
191. Schuffenhauer, A., Jacoby, E. Annotating and mining the ligand-target chemogenomics knowledge space. *Drug Discov. Today Biosilico* 2004, *2*, 190–200.
192. Steindl, T. M., Schuster, D., Laggner, C., Langer, T. Parallel Screening: a novel concept in pharmacophore modeling and virtual screening. *J. Chem. Inform. Model.* 2006, *46*, 2146–2157.
193. Steindl, T. M., Schuster, D., Wolber, G., Laggner, C., Langer, T. High-throughput structure-based pharmacophore modelling as a basis for successful parallel virtual screening. *J. Comput.-Aided Mol. Des.* 2006, *20*, 703–715.
194. Steindl, T. M., Schuster, D., Laggner, C., Chuang, K., Hoffmann, R. D., Langer, T. Parallel screening and activity profiling with HIV protease inhibitor pharmacophore models. *J. Chem. Inform. Model.* 2007, *47*, 563–571.
195. Stahura, F. L., Bajorath, J. Virtual screening methods that complement HTS. *Combin. Chem. High Throughput Screening* 2004, *7*, 259–269.
196. Klebe, G. Virtual ligand screening: strategies, perspectives and limitations. *Drug. Discov. Today* 2006, *11*, 580–594.
197. Bajorath, J. Integration of virtual and high-throughput screening. *Nat. Rev. Drug Discov.* 2002, *1*, 882–894.
198. Shoichet, B. K. Virtual screening of chemical libraries. *Nature* 2004, *432*, 862–865.
199. Bleicher, K. H., Böhm, H.-J., Müller, K., Alanine, A. I. Hit and lead generation: beyond high-throughput screening. *Nat. Rev. Drug Discov.* 2003, *2*, 369–378.
200. Zhang, Q., Muegge, I. Scaffold hopping through virtual screening using 2D and 3D similarity descriptors: ranking, voting, and consensus scoring. *J. Med. Chem.* 2006, *49*, 1536–1548.
201. Paolini, G. V., Shapland, R. H., van Hoorn, W. P., Mason, J. S., Hopkins, A. L. Global mapping of pharmacological space. *Nat. Biotechnol.* 2006, *24*, 805–815.
202. Mestres, J. Computational chemogenomics approaches to systematic knowledge-based drug discovery. *Curr. Opin. Drug Discov. Dev.* 2004, *7*, 304–313.
203. Wermuth, C.-G. The "Sosa" approach: an alternative to high-throughput screening. *Med. Chem. Res.* 2001, *10*, 431–439.
204. Moreau, G. and Broto, P. Autocorrelation of molecular strucutres: Application to SAR studies. *Nouv. J. Chem.* 1980, *4*, 757–764.

Chapter 11

Fragment-Based Drug Discovery

Bennett T. Farmer and Allen B. Reitz

I. LIGAND–PROTEIN INTERACTIONS: FIRST PRINCIPLES
A. Binding energy as the sum of the parts
B. Historical development
C. Ligand efficiency
II. STATUS OF LATE 1990s DRUG DISCOVERY IN THE PHARMACEUTICAL INDUSTRY
III. WHAT IS FBDD?
A. Concept and overview
B. Differences between FBDD and HTS/HTL approaches
C. The role of the medicinal chemist in FBDD
IV. CREATION AND ANALYSIS OF FBDD LIBRARIES
A. General evaluation and analysis
B. Computational approaches
C. Use of primary data: sprouting and merging to create secondary libraries
V. NUCLEAR MAGNETIC RESONANCE
A. 1D (ligand-based) screening
B. Example
C. 2D (protein-based) screening
VI. X-RAY CRYSTALLOGRAPHY
A. General principles and limitations
B. Examples
VII. OTHER BIOPHYSICAL AND BIOCHEMICAL SCREENING METHODS
A. Substrate activity screening
B. *In situ* click chemistry
C. SPR spectroscopy
D. SAR by mass spectroscopy
VIII. METHODS FOR FRAGMENT HIT FOLLOW-UP
A. How to best reduce false positives (NMR, MS) and false negatives (X-ray)
B. Isothermal and differential titration calorimetry and further secondary analysis
IX. TRENDS FOR THE FUTURE
REFERENCES

Everything should be made as simple as possible, but not simpler

Albert Einstein

I. LIGAND–PROTEIN INTERACTIONS: FIRST PRINCIPLES

When Paul Ehrlich conceived of a drug as a magic bullet interacting as a key in a lock, he was not far from the reality of ligand–protein interactions. Although target proteins are flexible and can adopt one or more of a manifold of induced conformations, binding sites on proteins have evolved to recognize a limited number of endogenous modulators and substrates and to exclude others.

A. Binding energy as the sum of the parts

The Gibbs free energy (ΔG) of multivalent ligand–protein binding is the sum of the energies involving each of the substructures or fragments that comprise the ligand.[1,2] To the extent that we understand ligand–protein interactions on as small a scale as possible, we can design inhibitors and modulators from first principles more easily than the hit or miss approach implicit in high-throughput screening (HTS; see Figure 11.1). In this example, fragments consist of diverse, low molecular weight compounds. Typical screening libraries can be considered as the combination of fragments in various ways, but these can only sample a small fraction of the total possible diversity space for their larger molecular weight range. As ligands become larger and more complex, the probability of finding useful information for any randomly selected compound is diminishingly small.[3] When hits emerge from HTS, typically in the 1–10 μM potency range, some of the fragments contained

Copyright © 2008, Elsevier Ltd
All rights reserved.

FIGURE 11.1 FBDD: general principles.

within these hits will not be and others may be optimized for interaction with the protein target. Importantly, in cases where it is possible to screen the fragments separately in advance, one can design more efficient ligands *a priori* by further synthetic manipulation such as by linking different fragments together. Therefore, less complex and smaller molecules are better starting points early on in the drug discovery process. This is the essence of fragment-based drug discovery (FBDD) – determining which molecular substructures or fragments interact with targets of interest and determining how they bind, and then using that information to obtain drugs for therapy.[4–10] FBDD represents a paradigm shift in our thinking of how to approach the lead generation process in drug discovery, and is an attempt to get more information rapidly while doing the same amount of work overall.

An important caveat to FBDD is that fragments will not always orient in the same way individually as when combined together in an optimized structure. For example, β-lactamase inhibitor **1** was conceptually deconstructed into fragments **2**–**5**, which were prepared and separately evaluated for both functional activity and binding mode by ligand–protein X-ray crystallography.[11] Fragments **2**–**4** were found to bind to the protein in very different orientations when compared to the binding mode of **1** (Figure 11.2b). In fact, **2** and **3** induced a conformational change in the protein itself and mapped into a previously unidentified tunnel carboxylate site (Figure 11.2a). Fragment **2** was also seen to adopt a second binding mode which along with **4** bound into a different new distal carboxylate recognition site (Figure 11.2c). Ligand efficiencies, a measure of the relative efficiency of biological activity per heavy atom count (see below), of **2**–**5** were lower than that of **1**. Larger fragment **5**, which resembles **1** to a greater degree than **2**–**4**, was found to bind in the same location on the protein as the thiophenecarboxylic acid portion of **1** (Figure 11.3). One explanation for the disparate binding modes of **2**–**4** when compared to **1** is that these fragments were indeed not positioned in the best possible way in **1**, which has only a 1 μM K_i. In any event, the extra insight provided by the alternative orientations of **2**–**4** determined separately could be used in the design of more potent inhibitors.

B. Historical development

Abbott Laboratories pioneered the introduction of FBDD into the drug discovery process. Abbott introduced structure–activity relationship (SAR) by nuclear magnetic resonance (NMR)[12] in the mid-1990s followed by high-throughput X-ray crystallography[13] in 2000. Both of these techniques have gained increasing acceptance as a variety of success stories have emerged in which novel leads have been discovered and developed more rapidly using FBDD than with traditional HTS.[5] Several biotechnology companies have been founded based on core technologies in the area of FBDD.

It is difficult to compare the success rates in FBDD between laboratories because of the different methods and procedures that are employed. Out of 18 examples discussed in a review of the incremental increases in biological activities obtained by using an X-ray crystallography fragment-based approach, the biotechnology company SGX reported that 14 showed >10-fold and 9 had >100-fold improvements in IC_{50} values.[14] Recently, it was reported that Abbott has examined 23 proteins by the SAR by NMR method with a 0–0.9% hit rate for various samplings of a *ca.* 10,000 member fragment library.[15] They were able to identify potent (<300 nM) inhibitors using subsequent iterations of synthesis and testing for 10 of the 23. From these examples, it appears that the overall success rate per target for FBDD is *ca.* 50%.

In this chapter, we review FBDD on the historical and operational level. There are many different techniques which are now classified as fragment based, and we attempt to explain how they can be applied on a case-by-case basis.

FIGURE 11.2 Binding modes in β-lactamase of (a) **1** (white), along with **2** (cyan) and **3** (yellow) in a tunnel carboxylate site, (b) overlay of **1**–**4**, and (c) **1**, along with **2** and **4** (yellow) in a distal carboxylate binding site. *Source*: Reprinted with permission from Ref. [11] (Nature Publishing Group).

FIGURE 11.3 Binding of **1** and **5** (purple) in β-lactamase. *Source*: Reprinted with permission from Ref. [11] (Nature Publishing Group).

C. Ligand efficiency

One powerful tool frequently used to rank hits coming from any screening campaign is ligand efficiency (LE).[16]

$$\text{Ligand Efficiency (LE)} = -\frac{(\text{pki or pIC}_{50})}{\#\text{ of heavy (non-hydrogen) atoms}} \tag{11.1}$$

LE normalizes activity relative to size with values of >0.3 as generally favorable. A 500 MW hit from a screening effort with a 100 nM K_i may actually be less useful as a starting point for lead optimization than a 2 μM hit with a 250 MW. Deconstruction of a lead or drug into fragments and examination of the LEs can provide insight as to which fragment(s) deserve the greatest synthetic effort. Once a high LE has been achieved, LE and potency can be sacrificed to some extent as other important ADME (absorption, distribution, metabolism and excretion) and selectivity issues are addressed.

II. STATUS OF LATE 1990s DRUG DISCOVERY IN THE PHARMACEUTICAL INDUSTRY

The late 1990s witnessed a maturing of HTS technology and a realization that specialized hit-to-lead (HTL) chemistry and informatics expertise are required to maximally leverage the volumes of data produced by this technology.[17,18] The late 1990s are also known as the time when the benefits of combinatorial chemistry, the rapid synthesis of large numbers of compounds, were finally balanced against its limitations, a lack of real diversity and drug-likeness in those compounds. General combinatorial libraries had been the mainstream, with no easy way to assess their impact. Large libraries had

been preferred because they could be made more efficiently on a per compound basis. Diversity had been the watch word, but the desire for a few large libraries, with their associated synthetic constraints, led to highly focused diversity rather than true diversity.[19] Non-polar functionality seemed to be preferred, mainly because of the ease with which such reagents adhered to the synthetic constraints. The result was many large, hydrophobic compounds, whose overall characteristics were not at all drug-like and which could not serve as good lead generation starting points given the very nature of the lead optimization (LO) process.

In hindsight, it seems quite obvious that HTS can be only as good as the compounds that it screens. The screening of large numbers of compounds did not yield the expected return in NCE's.[20,21] Slowly, the industry began to understand the significance of drug-likeness and equally to recognize that many compounds in the current screening collections were simply not drug-like. So began an effort to improve the drug-likeness of the compound screening collections in order to boost the success of the HTS technology in drug discovery.

In the meantime, a seminal article published by Fesik *et al.*,[12] in which they coined the phrase "SAR by NMR," pointed out that, drug-likeness notwithstanding, current compound screening collections would always remain woefully inadequate in covering pharmaceutically relevant chemical space.[5] Hann *et al.* codified this statement by relating the complexity of a compound to the probability that such a compound would experience a binding event.[3] In short, because the compounds in the current screening collections were both highly complex and rather hydrophobic, being derived mainly from large combinatorial libraries and many previous LO campaigns, HTS provided only a limited success in identifying novel hits that could be progressed into novel lead candidates and then into marketed drugs.[22]

The SAR by NMR approach is to screen small, highly polar fragments, as opposed to large, drug-like compounds; and create lead candidates through the linking of multiple, independently optimized fragments. An example of this approach is the development of potent inhibitor **6** against stromelysin where NMR screening has identified that **7** and **8** bind to stromelysin in two distinct, neighboring locations, namely the catalytic site and the S1' pocket, respectively.[23]

The rationale behind this screening approach is simple. Less complex and smaller fragments are more likely to find at least one suitable location to bind in a protein target. In addition, screening 1,000 fragments against three independent binding subpockets, and then sampling 10 different linkers to connect fragments found to bind in each of these three subpockets, is analogous to screening 10^{11} drug-like compounds by HTS. Hence, the screening of fragments and the incremental construction of fragment hits into lead candidates, a set of processes now collectively referred to as FBDD, also provides a more tractable solution to the diversity problem that even today plagues the historical, large compound collections, despite their markedly improved drug-likeness.

III. WHAT IS FBDD?

Many HTS hits that are sufficiently potent to be actively considered by HTL chemistry and the vast majority of lead candidates destined for LO represent rather complex molecular structures comprised of multiple, interconnected ring systems onto which any number of substituent pharmacophores are grafted. These structures are complex because of their size, extent of conformational freedom and array of diverse chemical substitution. A careful analysis of such structures reveals that they can be deconstructed into a set of denuded chemical building blocks, such as heterocyclic and phenyl rings, and a set of linkers comprised of amide, urea, ketone and methylene functionalities. From these basic building blocks and linkers, one can envision constructing smaller, less complex molecular structures that present only a limited number of both pharmacophores and degrees of conformational freedom. These structures are called fragments.

A. Concept and overview

A fragment may be thought of as that molecular unit which occupies a single pocket or subpocket in an enzyme such as the S1 pocket of Factor VIIa, with sufficient potency and directionality such that the binding event can be measured and the binding mode characterized by preferably a single, defined orientation.[24] A fragment is typically defined by the *Astex Rule of 3*™ in which the MW is ≤300, and the number of hydrogen-bond acceptors, hydrogen-bond donors and rotatable bonds is each ≤3.[25] An analysis of fragments in terms of Similog keys, which are related to 3-point pharmacophores, has suggested that a minimum of 13–20 such keys must be present in order for a binding event to be detectable by any one of the myriad techniques currently available.[26] In short, a fragment must exhibit some complexity in its interaction with the target protein. One challenge in fragment selection is therefore to balance

the two competing attributes of ensuring sufficient complexity while limiting size.

The goal of FBDD is to identify one or more fragments sets, with each set comprising fragments that bind predominantly to a unique pocket or region of the desired binding site on the target protein; and then to select the most optimum fragment either from each such set, eventually to be linked together based on their spatial proximity to form a sufficiently potent compound,[4,27] or from a single such set, eventually to be grown into a sufficiently potent compound.[8,25] While this goal is simple in concept, the process whereby it can be reasonably achieved is more complex.

The process of FBDD as illustrated in Figure 11.4 starts with the selection of a library of fragments, which is the subject of Section IV. These fragments must then be screened against the target protein to detect a binding event or, in the case of an enzyme, an inhibitory event. Fragments that produce a binding or inhibitory event are termed fragment hits. The technique most commonly used to detect a fragment binding event is NMR,[28,29] although surface plasmon resonance (SPR),[24,30] mass spectrometry (MS),[31,32] and X-ray crystallography[13,33] have also been used with success. Detecting a fragment inhibitory event is accomplished by increasing the concentration of the fragment in the functional assay, typically to 0.2–1.0 mM, depending on the expected solubility within a particular fragment library.[12,34]

Once a fragment has been detected to bind to the target protein, the magnitude of its binding must be established. Functional inhibition is routinely determined on initial fragment hits, although it is generally understood that weak inhibition (>10 μM) can often be variable due to artifacts such as aggregation. These early functional data so obtained provide both an additional means to evaluate the relative merits of individual fragments and a baseline against which to gauge improvement through subsequent iterations of screening and synthesis. The quality of binding is assessed based on stoichiometry, dose dependence, location and orientation of each such fragment hit. The Hill coefficient from the dose–response curve in the functional assay and the maximum SPR signal are often good indicators of whether a fragment hit is engaged in multiple binding events. However, a fragment hit that is so engaged may still exhibit a component of binding that involves one or more desired pockets or regions on the target protein. Competitive displacement experiments, using a potent inhibitor known to occlude the desired pockets or regions, are an efficient means to determine whether this is indeed the case or not.[35]

Although competitive displacement experiments can limit the area where the fragment hit binds, this area usually spans multiple pockets given the tendency of potent inhibitors to engage these multiple pockets on a target protein. NMR and X-ray crystallography are the only general techniques that can both sufficiently and reliably limit the binding location of the fragment hit and subsequently determine its binding orientation. While NMR can achieve these two outcomes in a step-wise fashion, X-ray crystallography

FIGURE 11.4 FBDD workflow and estimated timelines.

provides both at once. In many instances, a particular fragment hit may be ill suited for either NMR or X-ray to determine its binding orientation, either because of some known or suspected molecular or physicochemical deficiency such as poor solubility. In those cases, a series of fragment analogs may need to be acquired and subjected to the same NMR or X-ray analyses before the respective fragment hit is abandoned.

A fragment hit that exhibits a high-quality binding or inhibitory event will necessarily occupy the desired pocket or region in a single preferred orientation. At this point, the medicinal chemist must evaluate whether the orientation of that particular fragment hit is such that reasonable synthetic opportunities exist to grow into a desired neighboring pocket or region, or to link onto a different fragment hit that already occupies that neighboring pocket or region. As these two paradigms have been treated in increasingly sophisticated ways by computational *de novo* design programs during the past 10–15 years, FBDD can be viewed as the experimental equivalent to this more well-established computational strategy.[19] As such, the lessons previously learned in applying such computational strategies, and the very programs themselves, may benefit the medicinal chemists tasked with evolving one or more fragment hits into a potent lead candidate.

B. Differences between FBDD and HTS/HTL approaches

The fundamental difference between the FBDD and HTS/HTL approaches is that FBDD seeks to incrementally construct a lead candidate based on the identification of suitably oriented fragments within the overall binding site on the target protein, whereas the HTS/HTL approach seeks to identify one or more reasonably mature chemical starting points whose structural and physicochemical characteristics are such that HTL chemistry can optimize them in a timely fashion to one or more lead candidates both by exploring alternatives to existing substituents and by gently probing new substitution sites. This fundamental difference brings with it a host of other differences, some related to process and some related to risk. While the differences in process are perhaps more readily recognized, the differences in risk may be less obvious.

The risk in the HTS/HTL approach is that after screening 1 million or so drug-like compounds, no tractable chemical starting points can be identified. The upside to this risk is that reaching such a decision point is both deterministic and efficient, and therefore highly manageable. This statement presumes that enabling NMR or X-ray crystallography for structure determination, while perhaps highly desirable, is not a required component of the HTS/HTL approach, and that any decision point encountered by this approach will be addressed regardless of the status of such structural enablement.

In the case of HTS, enablement means having a suitable assay system. Producing the assay system, while not trivial, has become an almost industrialized process, greatly facilitated by the plethora of available, validated assay types and formats and, in the case of *in vitro* functional assays, by the extremely low protein purity that is often tolerated (see below). In addition, only limited chemistry resources are typically expended to determine whether the chemical starting points that emerge from HTS are indeed tractable.

The downside to the risk in the HTS/HTL approach is that a potentially valuable therapeutic target may be terminated as a drug discovery project because of a lack of progressible chemical matter. Such processes often have a high rate of both false positives and negatives, arising necessarily from the need to balance speed and efficiency with accuracy and success. The associated downside can be minimized by prosecuting a greater number of less validated targets, for which the eventually necessary target validation not only is far more demanding of time and resources than the HTS/HTL campaign but also would benefit significantly for a suitable lead candidate to emerge either through systematic design or serendipity.

The risk in the FBDD approach is entirely different, mainly because of the significant resources and time that must be devoted *upfront* to enabling the technology and because of the predictable need to invest significant chemistry resources to determine whether any fragment chemical matter can even be progressed further. Important milestones to achieve include desired potency and selectivity, favorable ADME properties, and preliminary *in vivo* analysis. The upside to this risk is that a very innovative and unique chemical lead series may emerge that can provide a competitive advantage. Furthermore, a truly negative outcome from a FBDD campaign, especially when taken in concert with a negative outcome from an HTS/HTL campaign, more strongly indicates that the particular target may be nondruggable.[36] The downside to this risk is that a significant and unpredictable level of resources may be required to drive an FBDD campaign to a favorable outcome, or a termination decision.

In the case of FBDD, structural enablement means having either suitable crystals for X-ray or a suitably isotopically labeled protein sample for NMR. Obtaining suitable crystals for X-ray can take anywhere from 4 (literature precedent) to 18 (novel) months. These times presume that protein production must be developed in-house. In cases where high-quality proteins can be readily purchased, we have seen on rare occasions this time to structural enablement cut down to approximately 2 months.

The crystals to be used for X-ray in an FBDD campaign must present a higher diffraction quality than is typically

required by conventional structure-based drug design. A minimum resolution of <2.5 Å, but more realistically on the order of 2 Å, is required to reliably orient a fragment that is bound to the target protein (see below). Such stringency arises because fragments are deliberately chosen to be small in size, and therefore, in contrast to most HTS hits and LO compounds, typically present a rather compact and only minimally asymmetric 3D structure. In contrast, reliable orientation of an HTS hit or LO compound can often be achieved even at a resolution of 3.2 Å.

Previously, prosecuting a greater number of less validated targets has been promulgated as a way to minimize the downside risk to the HTS/HTL approach. In contrast, the downside risk to the FBDD approach can be minimized by prosecuting a smaller number of targets that have achieved at least strong pre-clinical therapeutic validation. Such targets virtually guarantee that any quality LO compound will proceed into development, justifying the significant upfront resources required by the FBDD approach. Such targets are also typically pursued at many pharmaceutical companies and have many associated patents describing a sundry of inhibitor chemotypes and scaffolds, typically derived either directly from HTS/HTL campaigns or from patent-busting operations on previous HTS/HTL-derived compounds. Under these circumstances, the downside risk inherent to the FBDD approach is further reduced because the very strength of FBDD lies in its proven ability to span effectively a wider chemical space, which plays to chemical novelty; in its natural selection of more polar chemical starting points, thereby avoiding at least the initial lipophilicity of many HTS hits that may later on become associated with specific ADME liabilities;[37] and in the very way that chemical elaboration of initial fragment hits is aggressively probed both functionally and structurally to insure that such elaborations are as ligand-efficient as possible.[38]

C. The role of the medicinal chemist in FBDD

The fundamental difference between FBDD and HTS/HTL also means that the medicinal chemist plays a different role in each approach. In the HTS/HTL approach, the medicinal chemist is tasked to determine, often with limited chemistry resources, whether any HTS hit represents chemical matter that can be progressed in a timely fashion into LO. Much data must be analyzed and synthetically challenging analogs are generally avoided. The HTL chemist is part informatician, medicinal and synthetic chemist.

In the FBDD approach, the medicinal chemist plays the role of a combined synthetic and structural chemist. The emphasis on informatics is greatly reduced because there is less data overall and most of it, such as from NMR or X-ray crystal structures, is visually analyzed, typically being complemented only by functional assay data on just the target itself. The emphasis on pure medicinal chemistry is also reduced, especially in the early stages of Fragment-based Hit To Lead (FHTL), because the overriding initial task in FHTL is to select the appropriate fragment hits to be elaborated by the appropriate synthetic methods to establish that potency, and subsequently selectivity can be increased in an efficient manner.

As each selected fragment hit is expected to be diverse, the synthetic pathways for elaborating each such hit have a high probability of being unique. Early SAR analysis based on lower molecular weight fragments can often be done by the selective purchase of further analogs from chemistry library providers – "chemistry by commerce." The more synthetically adept the FBDD chemist is, the more rapidly any chosen set of fragment hits can be chemically triaged. In this regard, it is important to note that there may be no substitute for such chemical triaging, and that it must be done at risk because no *a priori* guarantee exists that any chosen fragment hit can in fact be efficiently elaborated to achieve the desired potency, much less selectivity.

IV. CREATION AND ANALYSIS OF FBDD LIBRARIES

As previously stated, fragments are defined by certain physical criteria, such as MW ≤ 300 with hydrogen-bond donors ≤3 and acceptors ≤3. FBDD libraries can potentially cover more diversity space in the lower MW range than can similar libraries that have a larger MW average. Further, FBDD libraries are easier to accumulate because many of the members can be purchased from commercial vendors. The key is to construct the libraries optimally for the specific questions addressed, or the techniques used to evaluate and interact with the protein targets being considered. Fragments in libraries are often designed with a synthetic handle in the form of suitable functionality for further synthetic elaboration if required.

A. General evaluation and analysis

Typical fragment libraries consist of 500–10,000 members, and are thus generally much smaller than libraries used in traditional HTS campaigns. Even with this lower number, greater diversity of structure and topology for the fragment MW range can often be achieved relative to the larger libraries. The potency of fragments upon screening is often weak such as in the mM range. However, it has been shown that the greatest chance of success is found when the fragments already bear at least some measurable activity since a linear relationship frequently exists between MW and binding affinity.[38]

Fragment libraries are typically selected with specific attributes in mind. For example, certain techniques require

covalent attachment to either protein targets or solid support which necessitates the incorporation of suitable functionality for that purpose, such as the sulfhydryl group. Some libraries for X-ray crystallography are designed to incorporate heavy atom substitution such as with bromine in order to solve the X-ray structures more rapidly. Many techniques such as those for X-ray and NMR entail clustering of fragments into mixtures or cocktails, followed by deconvolution if activity is seen. The clustering algorithms can be based on either similarity or diversity. Mixtures based on similarity can give a pharmacophore map right from the start if more than one compound is found to interact, and those based on diversity can be easier to deconvolute from the primary data. It is also common to create fragment libraries which bear representative substructures known to have favorable drug-like properties or are found to a high degree in currently marketed drugs.[39]

Chemical novelty is not as important a criterion for inclusion in a primary fragment library because unique composition of matter desirable for patent protection is expected to be added later during the sprouting and merging HTL chemistry phase. There are several vendors of fragment libraries who accumulate fragments from various sources and then cluster, package and resell them.

B. Computational approaches

Several different computational methods have been developed to prescreen fragment libraries as a way to select members for further study and consideration. Grand canonical Monte Carlo simulations can provide useful information including the discovery of novel binding sites, and can serve as the starting point for further chemistry synthesis validation and SAR development.[40,41] The fragment libraries selected for computational prescreening typically have a very low MW average and can thus cover a large fraction of the possible diversity space.

C. Use of primary data: sprouting and merging to create secondary libraries

After the original discovery of weakly active fragments, the challenge for the medicinal chemist is to take this information and generate novel leads suitable for development into clinical candidates. One of the ways in which potency can be improved is by sprouting, growing or fragment evolution, which is SAR development based on intuition, commercial reagent library availability, knowledge of the active site from structural analyses, and insight from the literature. Additionally, if multiple fragment binding sites are determined, these can be leveraged in order to increase potency. There are several advantages in starting small fragments with regard to MW: synthesis for the early SAR work is often easier to perform with fewer chemical steps than when working with larger hits from traditional HTS campaigns, and many related and relevant compounds can be purchased. In additional, SAR that is performed on weakly active smaller fragments can often be translated into useful information that can be applied to that same region of the molecule when incorporated into the larger leads later.

V. NUCLEAR MAGNETIC RESONANCE

As previously stated, NMR is the technique most represented in the literature as being used to detect a fragment binding event, for both historical and technical reasons.

A. 1D (ligand-based) screening

Ligand-based NMR techniques are, as the name suggests, based on detecting the ligand signal in a way that can be modulated by the binding of that ligand to the protein target. To this end, a key difference, relevant to NMR, between a small ligand and a large protein is the efficiency with which magnetization can be transferred between protons. Cross-relaxation is the mechanism for such magnetization transfer, and because of the $1/r^6$ dependence in its rate, is restricted to proton pairs whose average separation distance is typically <5 Å, implying a sustained, close contact.[42] The cross-relaxation rate is also a strong, monotonic function of molecular weight, becoming approximately linear therein for globular proteins >3–5 kDa.[43,44]

Saturation transfer difference (STD) NMR is probably the most common ligand-based NMR technique used to screen the binding of fragments. STD-NMR takes advantage of the enhanced cross-relaxation rate in the large protein system, and of the fact that protons on the bound-state ligand become part of this larger proton–proton relay network.[34,45] Saturation of the protein methyl resonances is therefore quickly propagated through this relay network to the bound-state ligand, leading to a saturation of the ligand proton resonances. To this end, protein methyl resonances (-1 to 0.5 ppm) are chosen because most ligands of pharmaceutical interest have few, if any, protons that resonate in that region. In addition, methyl groups serve as relaxation sinks in proteins and, as such, are efficient propagators of proton saturation throughout the protein and hence bound-state ligand, and being at the terminus of amino acid side-chains, which protrude into the interior of most binding pockets, are most likely to be proximal to the bound-state ligand.[46]

The STD-NMR experiment is carried out such that ligand saturation, effected when the ligand is bound to the protein, is actually measured on the free ligand, not the bound-state ligand. This is achieved by requiring that the

ligand-to-protein ratio is on the order of 40–100, and that the ligand binding kinetics are such that $k_{off} >> 1/T_{1free}$, where $1/T_{1free}$ is the rate at which net magnetization recovers from saturation in the free ligand, and k_{off} is sufficiently large to insure that a sufficient concentration of saturated free ligand is established during the saturation period (on the order of 1–2 s) to enable the detection of the actual saturation level. This constraint on k_{off} is easily met for fragments. $1/T_{1free}$ for small ligands, such as fragments, is typically on the order of 0.1–0.3 s^{-1}. Given that k_{on} for fragment binding is assumed to be diffusion controlled at , $10^8 M^{-1}s^{-1}$,[45] and that the associated K_d values typically range from 10 to 10 mM, k_{off} values can be estimated to fall in the range of 10^3–$10^6 s^{-1}$.

The sensitivity of the STD-NMR experiment is a function of both experimental parameters and specific characteristics of the protein. In terms of experimental parameters, there is an optimum ligand-to-protein ratio for a given ligand k_{off}. As k_{off} decreases, the optimum ligand-to-protein ratio also decreases (and vice versa). The optimum ligand-to-protein ratio is a balance between a high-enough ligand concentration to maintain a sufficient occupancy level of the protein binding site, only during the time when the ligand can actually be saturated, and a low-enough ligand concentration not to dilute beyond detection the percentage of saturated free ligand (relative to unsaturated free ligand) that builds up during the saturation period.

The absolute protein concentration required for the STD-NMR experiment is inversely related to the intra-protein cross-relaxation rate, which, for a globular protein, is directly proportional to its molecular weight. For increasingly larger proteins with concomitantly larger cross-relaxation rates, saturation is more efficiently transferred to the ligand during its bound-state lifetime ($1/k_{off}$). Therefore, the same integrated ligand saturation is achieved at increasingly lower protein concentrations, that is, lower occupancy levels of the protein binding site. For monomeric proteins on the order of 20 kDa, an absolute protein concentration of 5–10 μM has been found to be sufficient. Concomitantly, absolute ligand concentrations range from 100–400 μM for fragments. In our NMR work, we currently screen at a fragment concentration of 300 μM.

There are other specific characteristics of the protein, namely the amino acid composition of the targeted binding site, that also impact the sensitivity of the STD-NMR experiment. It is well known in the NMR literature that methyl groups are the most efficient in transferring magnetization via cross-relaxation.[46] Therefore, the more methyl groups that are proximal (<5Å) to the bound-state ligand, the more efficiently that ligand will become saturated. If the ligand binding site were such that only amide and Hα protons are proximal, the sensitivity of the STD-NMR experiment would be decreased. The flexibility of the bound-state ligand also plays a role in the absolute magnitude of the protein-to-ligand cross-relaxation rate, and hence the ultimate sensitivity achievable in the STD-NMR experiment. Increased ligand flexibility, which may arise either due to the way the ligand interacts with the protein or to the particular geometry of the binding site (e.g., highly concave versus flat and open), decreases the protein-to-ligand cross-relaxation rate, and therefore decreases the STD-NMR sensitivity. This decrease in sensitivity may be overcome at least in part by using a higher protein concentration.

The STD-NMR signal requires that the protein remains competent to bind the ligand. Therefore, STD-NMR does not suffer from false positives resulting from the precipitation or denaturation of the protein due to the high ligand concentrations associated with fragment-based screening. In addition, one typically collects a straight 1D proton spectrum prior to collecting the STD-NMR spectrum. The straight 1D proton spectrum should show a clear, free-state ligand signal whose intensity is directly proportional to the concentration of that ligand. Hence, one can also easily assess the integrity of the ligand before, or even after, the STD-NMR binding measurement. These two attributes of STD-NMR significantly increase the robustness of this technique in detecting true, weakly binding ligands to protein targets.

The STD-NMR signal represents the aggregate signal from bound-state ligand at one or more unique sites on the protein. There is no way to know the number of unique ligand binding sites, the relative occupancy of each such binding site, the number of discrete binding orientations per site, or the binding affinities of each such orientation that are reflected in this aggregate signal. By using competitive displacement STD-NMR,[47] in which a known, potent reference inhibitor is added to the binary ligand–protein solution, one may assess the relative percentage of the aggregate STD-NMR signal that arises from ligand binding to a single, desired binding site on the protein. Ideally, one would like to see at least 50% of the aggregate signal derive from the desired binding site. Even then, however, one does not know whether that population of bound-state ligand is characterized by a single, well-defined orientation or by multiple orientations. This can be assessed only either by X-ray crystallography or by the detection of intermolecular NOEs that are consistent with one predominant ligand orientation.

B. Example

STD-NMR was used to screen a set of 34 potential binders to the S1 pocket of human Factor Xa.[48] Fragments **9**–**12** were observed to produce the strongest STD signals.

TPAM (**13**), whose bezamidine group is known to bind in the S1 pocket of Factor Xa and whose tolylsulfonamide group in the S4 pocket,[49] has been used to probe the binding region for both benzamidine (**9**) and the three fragments **10**–**12**.

TPAM clearly displaced **9** as well as **11**. However, TPAM only weakly displaced **10**, and did not displace **12** at all. Therefore, despite the fact that **12** presented the highest binding affinity, it was not pursued further because it did not bind in the desired region on Factor Xa. Given that TPAM bridges two key subpockets in Factor Xa (S1 and S4), the competitive displacement experiments with TPAM could not further localize the binding of the remaining two novel fragments to any one subpocket.

C. 2D (protein-based) screening

Protein-based NMR techniques are more powerful than ligand-based techniques in characterizing ligand binding, but are much less widely applicable because these protein-based techniques require both the protein be either ^{15}N or ^{13}C isotopically labeled and that a protein so labeled present a high-resolution 2D ^{1}H–^{15}N or ^{1}H–^{13}C heteronuclear correlation spectrum. In cases where both these requirements are met, the binding of a ligand can be monitored based on its bound-state perturbation to the chemical shifts of specific protein resonances (proton and heteronuclear). The specific protein resonances whose chemical shifts are thus perturbed serve to localize the ligand binding site, thereby eliminating the need to use competitive displacement experiments to this end, as would be required for all ligand-based NMR screening methods.

The sensitivity of this approach depends on the maximum net chemical-shift perturbation that a specific ligand induces upon binding to the protein. This net perturbation is a sum of both the perturbation induced by the ligand itself upon binding and of that induced by changes in protein conformation upon ligand binding. As such, one could envision situations where the net perturbation is close to zero, even though a ligand has bound to a protein binding site with a normally sufficient level of occupancy. Variations on this theme may decrease the overall sensitivity of protein-based NMR to weakly binding fragments, especially those fragments that do not contain moieties capable of producing ring-current shifts, for example, saturated ring systems or alkyl functionalities.

The observed chemical-shift perturbation to the protein resonances can also be highly sensitive to the conformation of the bound-state ligand, so sensitive in fact that proton chemical shifts have been used directly to refine the bound-state ligand conformation in an NMR co-structure.[50] If a ligand adopts several different conformations, even if substantially similar in overall appearance, each conformation may induce a significantly different chemical-shift perturbation in one or more protein resonances. The net perturbation to each protein resonance is a sum of the perturbations from each ligand conformation. In short, how sensitive protein-based NMR is to each weakly binding fragment is quite difficult to predict, and there is a non-negligible risk that not all fragments that bind will be detected. However, this limitation in protein-based NMR may also serve a useful purpose by minimizing the detection of false positives due to non-stoichiometric, non-specific binding – a fact that clearly differentiates protein-based from ligand-based NMR methods.

VI. X-RAY CRYSTALLOGRAPHY

The use of X-ray crystallography in FBDD has become more prevalent, fostered by the development of techniques to more rapidly find conditions for protein crystallization and robotic methods for obtaining crystal structures. Having a ligand–protein crystal structure provides structural information about the mode of binding and gives considerable insight into the chemistry optimization process.[51]

A. General principles and limitations

The two primary modalities by which crystallography is performed in FBDD are soaking and co-crystallization. In soaking, crystals of generally a very small amount of protein (μg amounts) are obtained in solution and either mixtures or single compounds are added at high concentration (25–200 mM).[52] After a period of 1–24 h, the samples are cooled and stored prior to data collection. The hit rate found

in typical examples when using X-ray crystallography in FBDD is *ca.* 0.5–10%, and the most valuable information is obtained when multiple and unanticipated binding sites are identified. When protein crystals are not robust enough in terms of forming apo crystals without any added substrate or inhibitor, then fragments can be evaluated for their ability to co-crystallize with the protein and a known ligand. The greatest limitation for the use of X-ray crystallography in FBDD is that only *ca.* 20–30% of soluble proteins are amenable to crystallization, and membrane-bound targets such as G-protein-coupled receptors (GPCRs) and ion channels are excluded from screening in this way. One of the confounding factors that can impede crystallization is that proteins can display heterogeneous glycosylation resulting in multiple crystal forms. The resource and personnel commitment to high-throughput X-ray crystallography is substantial, often requiring collaborations with specialized academic research groups and biotechnology companies.

B. Examples

Fragments **14** and **15** were identified as binding to the ATP-binding site hinge region of non-phosphorylated p38aMAP kinase.[53] Although both of these fragments displayed relatively weak biological activity, they served as starting points for structure-based drug design programs. From the insight gained by inspection of the initial structure and further iterative crystal structure determinations, lead **16** was obtained from **14,** and lead **17** was obtained from **15**. In addition, **16** was found to selectively inhibit p38α relative to five other kinases (>10 nM IC_{50}s).

In another case, a core fragment library of 20,000 members was prescreened for activity in a high-throughput scintillation proximity assay against the phosphodiesterase family members PDE1B, PDE2A, PDE4D, PDE5A and PDE7B.[54] The 316 fragments that showed >30% inhibition at 200 for three or more PDEs in the screening panel were then co-crystallized with PDE4D and PDE4B. From this study, 269 fragments co-crystallized, and 107 crystal structures were determined. Pyrazole **18** was selected for further study because of its mode of binding between two hydrophobic amino acid residues ($F372^{4D}$ and $I336^{4D}$). After the chemical synthesis of only 21 analogs, the potency had been improved by 4,000 fold to give *N*-aryl pyrazole **19**.

In another example, four fragments (0.12–1.34 nM ED_{50}s) were identified that inhibited nucleoside 2-deoxyribosyltranserase from *Trypanosoma brucei*.[55] In this study, 304 fragments were evaluated in 31 mixtures or cocktails. Short soaks of *ca.* 10 s were sufficient for incorporation of even hydrophobic ligands into the active site groove of this enzyme.

VII. OTHER BIOPHYSICAL AND BIOCHEMICAL SCREENING METHODS

The value of screening fragments of low molecular weight and high LE is somewhat intuitive. However, the key has been to identify and exploit screening technologies that can detect the weak activity or binding that is generally inherent in the fragment-based approach. In addition to NMR and X-ray crystallography, there are now several other approaches that are being adapted for FBDD.

A. Substrate activity screening

This is a fragment screening strategy in which substrates for a particular target protein are identified and then optimized rapidly.[56,57] For example, substrates for the cysteinyl protease cathepsin S bearing a fluorogenic group were optimized for

cleavage.[58] The hydrolyzed amide bond serves as a structural anchor as to where in the protein the substrates are binding, so that enhancement of substrate activity determines what structural features improve the binding interaction off of the C-terminal side. After the substrate activity is optimized, then the scissile bond is modified to obtain an inhibitor such as by the use of bioisosteric or reactive functionalities. For cathepsin S, a 15 nM K_i inhibitor was identified that was selective (>1,000-fold) relative to the other cysteinyl proteases cathepsins B, L and K.

B. *In situ* click chemistry

Several reports have appeared in which fragment-like monomers are allowed to react together in the presence of target proteins.[59] In certain cases, the rate of formation of a few possible products from a potentially large number is enhanced because of the appropriate orientation of the fragments mediated by the protein, and the resulting compounds can be potent inhibitors or modulators because they are pre-selected to be good binders. This approach has been effectively applied to the [3 + 2]-triazole forming reaction of azides with acetylenes, also known as click chemistry because the reactive partners appear to snap together when properly oriented.[60–62] In one example, acetylene **20** (60 μM concentration) was incubated with a library of 24 azide-bearing fragments such as **21** (400 μM concentration) in the presence of bovine carbonic anhydrase II.[63] A set of 12 reagent combinations specifically formed a triazole in the presence of the enzyme, including **22** formed by reaction of **20** and **21**.

C. SPR spectroscopy

SPR is used to monitor the concentration of proteins at the surface of a solid support by measuring changes in refractive index. Libraries of functionalized fragments differing by virtue of the spacer length and composition as well as the linker functionality itself have been conjugated to solid support, followed by exposure to target proteins of interest. In this way, early SAR trends relative to fragment binding have been determined providing new insight for inhibitors of enzymes such as thrombin.[64] Isothermal calorimetry has also been used to evaluate the binding of fragments with protein targets and has mapped out the structural determinants for co-factor recognition of a model dehydrogenase, *Escherichia coli* ketopantoate reductase.[65]

D. SAR by mass spectroscopy

In some cases, it is possible to observe ligand–protein complexes of fragment libraries by electrospray ionization mass spectrometry.[66] This method was used to identify novel hits that functionally inhibit bacterial protein synthesis by binding to a subdomain of the 23S rRNA subunit.[33]

Covalently attached fragments can be detected when they are bound onto proteins. In one variant, libraries of sulfhydryl-group (-SH) containing fragments are incubated with protein targets which are mutated to incorporate Cys residues near the active site.[59,67] Those that demonstrate affinity for the protein may orient in such a way as to form a covalent disulfide bond with neighboring cysteine residues which are then detected by MS analysis. Traditional SAR sprouting and merging approaches can then be employed to prepare a further generation of sulfhydryl-containing substrates, and then the sulfhydryl group may be removed entirely to prepare inhibitors and modulators. This approach has the limitation that the sulfhydryl group may itself perturb the binding, and there may be important binding sites for which there are no cysteines nearby for linking and detection.

VIII. METHODS FOR FRAGMENT HIT FOLLOW-UP

A. How to best reduce false positives (NMR, MS) and false negatives (X-ray)

Ligand-based NMR suffers from non-specific (i.e. non-stoichiometric) fragment binders, which further competition displacement studies attempt to filter out. However, not all reference inhibitors used in such studies will bind in all relevant pockets that may be productively occupied by a fragment. Although protein-based NMR is much less sensitive to such false positives in fragment-based screening, this technique is severely limited in the particular proteins to which it can be practically applied. Functional assays suffer from the effects of both high fragment concentration (0.2–1 mM) in the presence of nM protein concentrations and from the inability to determine protein integrity after the assay is complete.

In-house experience would suggest that the intersection of fragment hits from a ligand-based NMR and those from

a functional assay screen can produce a set of fragments with a rate of false positives acceptable to immediate characterization by X-ray crystallography. However, fragment hits identified by only one screening methodology have been observed (see below) to present a significantly higher rate of false positives and may therefore benefit from further biophysical characterization prior to crystallography.

This further characterization of fragment binding may take the form of assessing dose-dependency, estimating the stoichiometry of binding, and estimating the potency of binding. Ligand-based NMR is not well suited to such further characterization of fragment binding. However, protein-based NMR, SPR, and differential scanning calorimetry (DSC) are potentially capable of providing information on one or more aspects of this further characterization.

B. Isothermal and differential titration calorimetry and further secondary analysis

SPR uses only unlabeled protein, but requires that the protein retain binding competency upon being tethered to the SPR surface. DSC uses only unlabeled protein, but may suffer in reliably quantifying weakly binding interactions (see below). Protein-based NMR requires labeled protein and a high-resolution 2D ^{1}H–^{15}N or ^{1}H–^{13}C correlation spectrum. Both DSC and protein-based NMR methods require substantially more protein than SPR.

All three methods can measure the dose-dependency of fragment binding and could therefore in principle yield an estimation of binding potency. For protein-based NMR, such measurements are extremely time consuming, and have an implicit sensitivity that is maximally constrained by the extent of net protein chemical-shift perturbation induced by 100% fragment occupancy. For DSC, the overall sensitivity of the measurement is often such that even establishing binding at a single fragment concentration is difficult.

In general, SPR appears to be the most universally applicable technique for establishing dose-dependent binding or a dose–responsive binding curve. The implicit sensitivity is governed by changes in molecular weight upon fragment binding, and is therefore expected to be more predictable from fragment to fragment for a given protein. Only small amounts of unlabeled protein are required, typically <1 mg for 100 fragments.

In terms of the stoichiometry of fragment binding, the most important requirement is that a sufficient level of fragment binding occurs at a single protein site; how many additional protein sites present some level of occupation is largely irrelevant. Therefore, only approaches that can discriminate between different populations of either bound fragments or fragment binding sites can provide such information. Protein-based NMR is one such approach. Competitive displacement ligand-based NMR is another approach. However, neither SPR nor DSC can provide such discrimination. While it is true that the total signal intensity measured in an SPR experiment can give some insight as to whether fragment binding is single or multiple event, this information provides no insight into either the total number of binding events or the relative proportion of each such event – the latter of which is the most critical to know.

While the above discussion is focused mainly on approaches to triage a set of primary fragment hits down to those select fragments most likely to yield an X-ray co-structure, there is the added complication that further optimization of any such fragment, whether it yields an X-ray co-structure or not, is dependent on being able to readily measure some parameter that is highly correlated with the binding potency of that fragment. An interesting point to realize is that a comparison of two different fragment co-structures, with a given protein, does not easily, if at all, allow one to infer the relative binding potency of these two fragments. Therefore, developing fragment SAR based solely on co-structures is not currently possible.

Once binding potencies $<10\,\mu M$ have been achieved, the ability to develop fragment SAR is more likely achievable as this is the regime in which most, if not all, HTS hits fall. The challenge lies in developing SAR for those select fragments whose binding potency remains $>10\,\mu M$. In cases where fragment binding produces inhibition and where the functional assay is amenable to fragment concentrations up to 1 mM, one may readily obtain reasonable estimates of fragment IC_{50} values which are <0.5–1 mM. In cases where such select fragments yield only an approximate value in a functional assay or where the particular fragment binding mode does not yet lead to functional inhibition, developing SAR for fragment binding remains a significant challenge. A worthwhile strategy then may be to design and accumulate secondary fragment libraries building upon the insight obtained during the screening of the primary fragments, and test these using the original experimental paradigm. The secondary fragment libraries are expected to display a higher hit rate and may drive the analysis forward for the preparation of tertiary libraries and more traditional SAR development.

IX. TRENDS FOR THE FUTURE

The immediate need in FBDD is to achieve a sustainable delivery of high-quality lead candidates that can be more quickly and reliably progressed into the clinical development pipeline. Meeting this need is no longer a question of technology, but rather of deployment. Correctly deploying FBDD starts with the question "What targets shall we address by FDBB?". If FBDD is used to address only those targets recalcitrant to HTS/HTL, then FBDD will remain a niche approach and consequently not be developed to its full potential.

The nature of the binding site on the chosen protein target plays a key role in determining the success rate of an FBDD campaign. Shallow, surface exposed pockets may bind fragments, but may not do so with sufficient occupancy or with a preferred binding conformation to yield a suitable co-structure. Highly hydrophobic pockets, in which there are no obviously conserved waters, may also prove problematic given that most fragments are quite hydrophilic and that hydrophobic interactions are non-directional in nature. Binding sites, in which key subpockets of interaction must be induced, may not yield to the incremental constructionist approach of FBDD. Finally, a single-pocket binding site with one large, diffuse pocket is much less preferred to a multi-pocket binding site with several smaller, more constrained subpockets.

In a lead generation setting, FBDD is best deployed on targets where the need for structure-based support is high even for the HTS/HTL approach; where the barrier to structural enablement is lower, and overcoming this barrier, more predictable; where the HTS hit set, although providing insight into the chemical make-up of potential lead candidates, may not itself contain a hit that can be progressed to a suitable lead candidate; and where the target is at least validated in a pre-clinical model.

In an LO setting, FBDD can be successfully deployed in three ways. The first is by identifying replacement moieties for an existing LO series where the moiety to be replaced is associated with a specific liability. Because many liabilities are associated with lipophilicity, FBDD, through the very physicochemical criteria by which fragments are selected, offers an approach uniquely qualified to address such liabilities. The second way is by identifying any additional pockets for small-molecule interaction on the target protein that are proximal to the binding site of the current lead series. The third way is by generating a novel lead series, either from first principles or from a progressive replacement of various substructures within an existing lead series. While the latter is nothing more than lead generation in an LO setting, the LO setting itself provides a critical difference in that targets that have made it to LO are in general more highly validated, at least up to the pre-clinical stage.

Initial fragment libraries are often made up of commercially available compounds. While this would present an IP issue where these LO-like compounds, that their are fragments are expected to render the issue much less severe because the final FBDD lead candidate will undoubtedly be comprised of another fragment, additional moieties, and several linkers, wherein IP will be garnered based on the unique way in which these constituents are ultimately arranged. In addition, novel fragment core structures can be derived through a scaffold-hopping exercise that exploits the superposition of pharmacophoric interactions made by not just one, but rather multiple commercially available fragments. With time, it is possible that fragment libraries will become increasingly populated with proprietary fragments based on combichem library synthesis, novel fragments synthesized in previous FHTL campaigns; and those produced by custom synthesis.

The size of fragment libraries should also be expected to grow, but at a rate commensurate with the ability to screen them in a reasonable period of time. To the extent that a functional assay can tolerate a high concentration of ligand, such an assay may be used to prioritize fragment hits for further screening. If the resulting hits are further assessed by biophysical assays, for example, NMR, that establish a physical interaction between the fragment hit and the target protein, the ultimate outcome should be similar to that produced if the fragment hits were initially derived from a biophysical screen and then further assessed by a functional assay. A front-end functional assay can leverage the automation and expertise in HTS laboratories even if the assay itself is not fully HTS compatible. Furthermore, functional assays require only a small amount of protein, even to screen 1 million compounds. In such a scenario, fragment libraries can be easily imagined to grow to well over 100,000 fragments.

A major process bottleneck in FBDD occurs at the point when fragment co-structures are required. High-throughput crystallography is a powerful, emerging tool in the lead generation process, but requires substantial time and manual effort in growing, soaking, freezing and evaluating crystals. The maximum sustainable throughput in our laboratories on those steps is not much more than 60 fragments per month per crystallization scientist given the current workflow and other associated tasks.

An additional challenge for FBDD is one of informatics, not only in storing all of the assay and structural data, but also in extracting the most useful information from these data toward selecting the best fragments for elaboration (or linking) and deciding on the best elaboration (or linking) strategies. Very few computational chemists have had the occasion to look at hundreds of co-structures, and an integrated computational tool to automatically extract the necessary information does not yet exist commercially.

FBDD offers new ways to think about lead generation in drug discovery and the enabling tools to put these into practice. In the future, the relative fraction of programs that are conducted using this approach is expected to increase as are the success stories that emerge. In general, as a starting point in the HTL and LO processes, smaller and simpler is better and provides more possibilities for both subsequent SAR development and the identification of leads suitable for pre-clinical development.

REFERENCES

1. Jencks, W. P. On the attribution and additivity of binding energies. *Prod. Natl. Acad. Sci. USA* 1981, *78*, 4046–4050.
2. Nakamura, C. E., Abeles, R. H. Mode of interaction of β-hydroxy-β-methylglutaryl coenzyme A reductase with strong binding inhibitors:

compactin and related compounds. *Biochemistry* 1985, *24*, 1364–1376.
3. Hann, M. M., Leach, A. R., Harper, G. Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inf. Comput. Sci.* 2001, *41*, 856–864.
4. Jahnke, W., Flörsheimer, A., Blommers, M. J. J., Paris, C. G., Heim, J., Nalin, C. M., Perez, L. B. Second-site NMR screening and linker design. *Curr. Top. Med. Chem.* 2003, *3*, 69–80.
5. Hajduk, P. J., Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat. Rev. Drug Discov.* 2007, *6*, 211–219.
6. Bartoli, S., Fincham, C. I., Fattori, D. The fragment-approach: an update. *Drug Discov. Today: Technologies* 2006, *3*, 425–431.
7. Zartler, E. R., Shapiro, M. J. Fragonomics: fragment-based drug discovery. *Curr. Opin. Chem. Biol.* 2005, *9*, 366–370.
8. Rees, D. C., Cosgreve, M., Murray, C. W., Carr, R. Fragment-based lead discovery. *Nat. Rev. Drug. Discov.* 2004, *3*, 660–672.
9. Erlanson, D. A., McDowell, R. S., O'Brien, T. Fragment-based drug discovery. *J. Med. Chem.* 2007, *47*, 3463–3482.
10. Jahnke, W., Erlanson, D. A. *Fragment-based approaches in drug discovery*. Wiley-VCH Verlag GmbH & Co. KgaA: Weinheim, 2006. pp. 1–369.
11. Babaoglu, K., Shoichet, B. K. Deconstructing fragment-based inhibitor discovery. *Nat. Chem. Biol.* 2006, *2*, 720–723.
12. Shuker, S. B., Hajduk, P. J., Meadows, R. P., Fesik, S. W. Discovering high-affinity ligands for proteins: SAR by NMR. *Science* 1996, *274*, 1531–1534.
13. Nienaber, V. L., Richardson, P. L., Koighofer, V., Bouska, J. J., Giranda, V. L., Greer, J. Discovering novel ligands for macromolecules using X-ray crystallographic screening. *Nat. Biotechnol.* 2000, *18*, 1105–1108.
14. Blaney, J., Neinaber, V., Burley S. K., (Jahnke, W., Erlanson, D. A., Eds) *Fragment-Based Approaches in Drug Discovery*. Wiley-VCH Verlag GmbH & Co. KgaA: Weinheim, 2006, pp. 215–248.
15. Sen, D. S., (Jahnke, W., Erlanson, D. A., Eds) *Fragment-Based Approaches in Drug Discovery*. Wiley-VCH Verlag GmbH & Co. KgaA: Weinheim, 2006, pp. 149–180.
16. Hopkins, A. L., Groom, C. R., Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discov. Today.* 2004, *9*, 430–431.
17. Davis, A. M., Keeling, D. J., Steele, J., Tomkinson, N. P., Tinker, A. C. Components of successful lead generation. *Curr. Top. Med. Chem.* 2005, *5*, 421–439.
18. Michne, W. Hit-to-lead chemistry: a key element in new lead generation. *Pharmaceutical News* 1996, *3*, 19–21.
19. Honma, T. Recent advances in *de novo* design strategy for practical lead identification. *Med. Res. Rev.* 2003, *23*, 606–632.
20. Brown, D., Superti-Furga, G. Rediscovering the sweet spot in drug discovery. *Drug Discov. Today* 2003, *8*, 1067–1077.
21. Hird, N. Isn't combinatorial chemistry just chemistry?. *Drug Discov. Today* 2000, *5*, 307–308.
22. Verdonk, M. L., Hartshorn, M. J. Structure-guided fragment screening for lead discovery. *Curr. Opin. Drug Discov. Dev.* 2004, *7*, 404–410.
23. Hajduk, P. J., Sheppard, G., Nettesheim, D. G., Olejniczak, E. T., Shuker, S. B., Meadows, R. P., Steinman, D. H., Carrera, G. M., Jr., Marcotte, P. A., Severin, J., Walter, K., Smith, H., Gubbins, E., Simmer, R., Holzman, T. F., Morgan, D. W., Davidsen, S. K., Summers, J. B., Fesik, S. W. Discovery of potent nonpeptide inhibitors of stromelysin using SAR by NMR. *J. Am. Chem. Soc.* 1997, *119*, 5818–5827.
24. Dickopf, S., Frank, M., Junker, H.-D., Maier, S., Metz, G., Ottleben, G., Rau, H., Schellhaas, N., Schmidt, K., Sekul, R., Vanier, C., Vetter, D., Czech, J., Lorenz, M., Matter, H., Schudok, M., Schreuder, H., Will, D. W., Nestler, H. P. Custom chemical microarray production and affinity fingerprinting for the S1 pocket of factor VIIa. *Anal. Biochem.* 2004, *335*, 50–57.
25. Gill, A., Cleasby, A., Jhoti, H. The discovery of novel protein kinase inhibitors by using fragment-based high-throughput X-ray crystallography. *Chem Bio Chem* 2005, *6*, 506–512.
26. Schuffenhauer, A., Ruedisser, S., Marzinzik, A. L., Jahnke, W., Blommers, M., Selzer, P., Jacoby, E. Library design for fragment-based screening. *Curr. Top. Med. Chem.* 2005, *5*, 751–762.
27. Oblak, M., Grdadolnik, S. G., Kotnik, M., Jerala, R., Filipic, M., Solmajer, T. In silico fragment-based discovery of indolin-2-one analogues as potent DNA gyrase inhibitors. *Bioorg. Med. Chem. Lett.* 2005, *15*, 5207–5210.
28. Roberts, G. C. K. Applications of NMR in drug discovery. *Drug Discov. Today* 2000, *5*, 230–240.
29. Fejzo, J., Lepre, C., Xie, X. Application of NMR screening in drug discovery. *Curr. Top. Med. Chem.* 2003, *3*, 81–97.
30. Gestwicki, J. E., Hsieh, H. V., Pitner, J. B. Using receptor conformational change to detect low molecular weight analytes by surface plasmon resonance. *Anal. Chem.* 2001, *73*, 5732–5737.
31. Moy, F. J., Haraki, K., Mobilio, D., Walker, G., Powers, R., Tabei, K., Tong, H., Siegel, M. M. MS/NMR: a structure-based approach for discovering protein ligands and for drug design by coupling size exclusion chromatography, mass spectrometry, and nuclear magnetic resonance spectroscopy. *Anal. Chem.* 2001, *73*, 571–581.
32. Swayze, E. E., Jefferson, E. A., Sannes-Lowery, K. A., Blyn, L. B., Risen, L. M., Arakawa, S., Osgood, S. S., Hofstadler, S. A., Griffey, R. H. SAR by MS: a ligand based technique for drug lead discovery against structured RNA targets. *J. Med. Chem.* 2002, *45*, 3816–3819.
33. Carr, R., Congreve, M., Murray, C., Rees, D. Fragment-based lead discovery: leads by design. *Drug Discov. Today* 2005, *10*, 987–992.
34. Boehm, H.-J., Boehringer, M., Bur, D., Gmuender, H., Huber, W., Klaus, W., Kostrewa, D., Kuehne, H., Luebbers, T., Meunier-Keller, N., Mueller, F. Novel inhibitors of DNA gyrase: 3D structure based biased needle screening, hit validation by biophysical methods, and 3D guided optimization, a promising alternative to random screening. *J. Med. Chem.* 2000, *43*, 2663–2674.
35. Schade, M. NMR fragment screening: advantages and applications. *IDrugs* 2006, *9*, 110–113.
36. Cheng, A. C., Coleman, R. G., Smyth, K. T., Cao, Q., Soulard, P., Caffrey, D. R., Salzberg, A. C., Huang, E. S. Structure-based maximal affinity model predicts small-molecule druggability. *Nat. Biotechnol.* 2007, *25*, 71–75.
37. Riley, R. The potential pharmacological and toxicological impact of P450 screening. *Curr. Opin. Drug Discov. Dev.* 2001, *4*, 45–54.
38. Hajduk, P. J. Fragment-based drug design: how big is too big?. *J. Med. Chem.* 2006, *49*, 6972–6976.
39. Siegel, M. G., Vieth, M. Drugs in other drugs: a new look at drugs as fragments. *Drug Discov. Today* 2007, *12*, 71–79.
40. Clark, M., Guarnieri, F., Shkurko, I., Wiseman, J. Grand canonical Monte Carlo simulation of ligand–protein binding. *J. Chem. Inf. Mod.* 2006, *46*, 231–242.
41. Moore, W. R., Jr. Maximizing discovery efficiency with a computationally driven fragment approach. *Curr. Opin. Drug Discov. Dev.* 2005, *8*, 355–364.
42. Macura, S., Ernst, R. R. Elucidation of cross relaxation in liquids by two-dimensional NMR spectroscopy. *Mol. Phys.* 1980, *41*, 95–117.
43. Schurr, J. M., Babcock, H. P., Fujimoto, B. S. A test of the model free formulas: effects of anisotropic rotational diffusion and dimerization. *J. Magn. Reson. Ser. B* 1994, *105*, 211–224.
44. Van Holde, K. E. Physical biochemistry. *Foundations of Modern Biochemistry Series*. Prentice-Hall, Inc.: Englewood Cliffs, New Jersey, 1971.
45. Mayer, M., Meyer, B. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angew. Chem. Int. Ed.* 1999, *38*, 1784–1788.
46. Kalk, A., Berendsen, H. J. C. Proton magnetic relaxation and spin diffusion in proteins. *J. Magn. Reson.* 1976, *24*, 343–366.
47. Davis, D. G., Perlman, M. E., London, R. E. Direct measurements of the dissociation-rate constant for inhibitor-enzyme complexes via the $T_1\rho$ and T_2 (CPMG) methods. *J. Magn. Reson. Ser. B.* 1994, *104*, 266–275.

48. Fielding, L., Fletcher, D., Rutherford, S., Kaur, J., Mestres, J. Exploring the active site of human factor Xa protein by NMR screening of small molecule probes. *Org. Biomol. Chem.* 2003, *1*, 4235–4241.
49. Gabriel, B., Stubbs, M. T., Bergner, A., Hauptmann, J., Bode, W., Stürzebecher, J., Moroder, L. Design of benzamidine-type inhibitors of factor Xa. *J. Med. Chem.* 1998, *41*, 4240–4250.
50. Wittekind, M., Mapelli, C., Lee, V., Goldfarb, V., Friedrichs, M. S., Meyers, C. A., Mueller, L. Solution structure of the Grb2 N-terminal SH3 domain complexed with a ten residue peptide derived from SOS: direct refinement against NOEs, J-couplings and ^{1}H and ^{13}C chemical shifts. *J. Mol. Biol.* 1997, *267*, 933–952.
51. Rondeau, J.-M., Schreuder, H. Protein crystallography and drug discovery. In *The Practice of Medicinal Chemistry* (Wermuth, C. A., Ed.), 2nd Edition. Elsevier Academic Press: New York, 2003, pp. 417–444.
52. Hartshorn, M. J., Murray, C. W., Cleasby, A., Frederickson, M., Tickle, I. J., Jhoti, H. Fragment-based lead discovery using X-ray crystallography. *J. Med. Chem.* 2005, *48*, 403–413.
53. Gill, A. L., Frederickson, M., Cleasby, A., Woodhead, S. J., Carr, M. G., Woodhead, A. J., Walker, M. T., Congreve, M. S., Devine, L. A., Tisi, D., O'Reilly, M., Seavers, L. C. A., Davis, D. J., Curry, J., Anthony, R., Padova, A., Murray, C. W., Carr, R. A. E., Jhoti, H. Identification of p38a kinase inhibitors using fragment-based lead generation. *J. Med. Chem.* 2005, 48, 414–426.
54. Card, G. L., Blasdel, L., England, B. P., Zhang, C., Suzuki, Y., Gillette, S., Fong, D., Ibrahim, P. N., Artis, D. R., Bollag, G., Milburn, M. V., Kim, S.-H., Schlessinger, J., Zhang, K. Y. L. A family of phosphodiesterase inhibitors discovered by cocrystallography and scaffold-based drug design. *Nat. Biotechnol.* 2005, *23*, 201–207.
55. Bosch, J., Robien, M. A., Mehlin, C., Boni, E., Riechers, A., Buckner, F. S., Van Voorhis, W. C., Myler, P. J., Worthey, E. A., DeTitta, G., Luft, J. R., Lauricella, A., Gulde, S., Anderson, L. A., Kalyuzhniy, O., Neely, H. M., Ross, J., Earnest, T. N., Soltis, M., Schoenfeld, L., Zucker, F., Merritt, E. A., Fan, E., Verlinde, C. L. M. J., Hol, W. G. J. Using fragment cocktail crystallography to assist inhibitor design of *Trypanosoma brucei* nucleoside 2-deoxyribosyltransferase. *J. Med. Chem.* 2006, *49*, 5939–5946.
56. Wood, W. J. L., Patterson, A. W., Tsuruoka, H., Jain, R. K., Ellman, J. A. Substrate activity screening: a fragment-based method for the rapid identification of nonpeptidic protease inhibitors. *J. Am. Chem. Soc.* 2005, *127*, 15521–15527.
57. Salisbury, C. M., Ellman, J. A. Rapid identification of potent nonpeptidic serine protease inhibitors. *ChemBioChem* 2006, *7*, 1034–1037.
58. Patterson, A. W., Wood, W. J. L., Hornsby, M., Lesley, S., Spraggon, G., Ellman, J. A. Identification of selective, nonpeptidic nitrile inhibitors of cathepsin S using the substrate activity screening method. *J. Med. Chem.* 2006, *49*, 6298–6307.
59. Erlanson, D. A., Wells, J. A., Braisted, A. C. Tethering: fragment-based drug discovery. *Annu. Rev. Biophys. Biomol. Struct.* 2004, *33*, 199–223.
60. Whiting, M., Muldoon, J., Lin, Y.-C., Silverman, S. M., Lindstrom, W., Olson, A. J., Kolb, H. C., Finn, M. G., Sharpless, K. B., Elder, J. H., Fokin, V. V. Inhibitors of HIV-1 protease by using *in situ* click chemistry. *Angew. Chem. Int. Ed.* 2006, *45*, 1435–1439.
61. Manetsch, R., Krasinski, A., Radic, Z., Raushel, J., Taylor, P., Sharpless, K. B., Kolb, H. C. In situ click chemistry: enzyme inhibitors made to their own specifications. *J. Am. Chem. Soc.* 2004, *126*, 12809–12818.
62. Lewis, W. G., Green, L. G., Grynszpan, F., Radic, Z., Carlier, P. R., Taylor, P., Finn, M. G., Sharpless, K. B. Click chemistry *in situ*: acetylcholinesterase as a reaction vessel for the selective assembly of a femtomolar inhibitor from an array of building blocks. *Angew. Chem. Int. Ed.* 2002, *41*, 1053–1057.
63. Mocharla, V. P., Colasson, B., Lee, L. V., Roper, S., Sharpless, K. B., Wong, C.-H., Kolb, H. C. In situ click chemistry: enzyme-generated inhibitors of carbonic anhydrase II. *Angew. Chem. Int. Ed.* 2005, *44*, 116–120.
64. Neumann, T., Junker, H.-D., Keil, O., Burkert, K., Ottleben, H., Gamer, J., Sekul, R., Deppe, H., Feurer, A., Tomandl, D., Metz, G. Discovery of thrombin inhibitor fragments from chemical microarray screening. *Lett. Drug Design Discov.* 2005, *2*, 590–594.
65. Ciulli, A., Williams, G., Smith, A. G., Blundell, T. L., Abell, C. Probing hot spots at protein–ligand binding sites: a fragment-based approach using biophysical methods. *J. Med. Chem.* 2006, *49*, 4992–5000.
66. Hofstadler, S. A., Sannes-Lowery, K. A. Applications of ESI-MS in drug discovery: interrogation of noncovalent complexes. *Nat. Rev. Drug Discov.* 2006, *5*, 585–595.
67. Oslob, J. D., Erlanson, D. A. Tethering in early target assessment. *Drug Discov. Today: Targets* 2004, *3*, 143–150.

Chapter 12

Lead-Likeness and Drug-Likeness

Alex Polinsky

I. INTRODUCTION
II. ASSESSING "DRUG-LIKENESS"
A. Avoiding known threats
B. Mimicking known drugs
C. Direct property prediction
III. SELECTING BETTER LEADS: "LEAD-LIKENESS"
A. What is a "high-quality" lead compound?
B. Designing "lead-like" libraries for biochemical screening
IV. CONCLUSION
REFERENCE

Design is not just what it looks like and feels like. Design is how it works.

Steve Jobs

I. INTRODUCTION

Since the 1970s, several major changes in drug discovery paradigms have occurred, driven both by the pressure to increase the productivity of pharmaceutical research and by our expanding knowledge of the underlying biology of diseases. Before that time, the drug discovery process often started with the identification of endogenous active compounds or natural products, followed by chemical modifications to optimize their characteristics in *in vivo* models. Although suffering from very low throughput, this approach had an advantage of working well inside of what would later be defined as the "drug-like" chemical space: using *in vivo* models as the main optimization driver automatically steered chemists toward bioavailable and non-toxic structures.

With explosive growth of our understanding of molecular biology and biochemistry in the 1980s, target-based discovery became the main approach, leading to the emergence of *in vitro* assays that had a much higher throughput. As a result, many more compounds could be tested, culminating in the late 1980s–early 1990s with the development of high-throughput screening (HTS) technology and a race to assemble large screening collections. It is this shift to target-based *in vitro* assays, as well as screening compounds from sources other than historical medicinal chemistry collections, that helped establish the two-step process that is still most widely used: first using *in vitro* screening to find compounds – leads – that bind to and inhibit or activate a biological target, then optimizing their pharmacological properties, such as selectivity, pharmacokinetic (PK) and safety, by changing their chemical structure. The development of high-speed chemistry methods in the 1990s further expanded the chemical diversity of the screened collections, leading in turn to further improvements in screening throughput. A wealth of accumulated experience with this process by the end of the 1990s led chemists to ask questions about how the choices they make in selecting leads and – later – selecting candidates for clinical trials can influence the ultimate success of the entire process of launching a drug. Lipinski *et al.*[1] was the first to introduce the concept of "drug-likeness" and suggest his famous Rule of 5 (Ro5), stating that compounds outside its boundaries have low probability of being orally bioavailable.

The genomics "revolution" of the early 2000s focused biologists' efforts on developing high-throughput methods for the identification of disease-relevant targets but did not change chemists' two-step paradigm. Moreover, further analysis of the industry record of converting leads to drugs has led to the definition of a concept of "lead-likeness"[2] which has helped better rationalize lead selection criteria and develop lead discovery approaches other than HTS, such as, for example, fragment-based screening, described elsewhere in this section. Emerging concepts of systems biology may over time lead to studying the effect of drugs on an entire

Copyright © 2008, Elsevier Ltd
All rights reserved.

biological network of pathways and provide better ways to predict physiological properties of chemical compounds based directly on *in vitro* experiments, thus making concepts of "drug-likeness" or "lead-likeness" less relevant. Until then, however, these concepts will continue to serve as useful guides for chemists searching for new drugs.

One could think about "drug-likeness" as a strategic guide – small molecule drug candidates falling within the "drug-like" chemical space (as defined by past industry experience) have a better chance (but not a guarantee!) of surviving clinical trials and becoming a drug. "Lead-likeness," on the other hand, is a tactical guide for selecting starting points for chemical optimization that offer the best chance to deliver "drug-like" candidates at the end of drug discovery programs.

II. ASSESSING "DRUG-LIKENESS"

"Drug-likeness" literally means similarity to known drugs. Considering the diversity of mechanisms of action and properties the known drugs have, and a variety of methods by which these drugs are administered, it is not surprising that "drug-likeness" is difficult to define precisely and that there are multiple approaches to its assessment.[3–5] But first of all a question arises as to why anybody would want to come up with formal algorithms of assessing "drug-likeness."

When working on a specific drug discovery project, medicinal chemists are guided by property measurements and experience to achieve a desired drug property profile. Since not all properties can be measured during lead optimization (e.g. human PK and toxicity), even compounds that fit the desired property profile often fail in development stages. It is this uncertainty that drives our efforts to understand what structural features or molecular properties constitute the liabilities that might cause candidates to fail. For this reason, "drug-likeness" is most commonly thought of in terms of those liabilities, and it is assessed to alert chemists about them.

The ability to assess "drug-likeness" is even more important in the design of libraries made using high-speed chemistry. If library design is driven primarily by the ease of synthesis and structural diversity of an array of compounds, it is likely that significant portions of the library will have undesirable properties, for example, high molecular weight or low solubility. For libraries containing many thousands of compounds, computational methods for assessing "drug-likeness" need to be applied to filter out molecules with potential liabilities before resources are spent on synthesis, characterization and HTS.

A. Avoiding known threats

One approach to define a space of "drug-like" molecules is to do so by elimination. There are molecular features that in industry experience are known to cause toxicity, lack of bioavailability or other undesirable effects. Based on these features, a set of "filters" can be defined and applied to libraries of molecules – real or virtual – to eliminate the ones possessing those features.

1. *Exclusion of known undesirable functionalities*

Structural alerts have long been used to detect compounds with potential mutagenicity and carcinogenicity.[6] The following groups are defined as the causative substructure of one or more chemical carcinogens: arylamine function, ring epoxides, alkane sulfonate groups, arylnitro functions, azo groups, ring *N*-oxides and NMe_2 groups, methylols and aliphatic aldehydes, vinyl groups attached to aromatic rings, aziridines, nitrogen mustards and chloramines, benzyl halides, alkylnitrosamines and alkylurethanes. It should be noted that analysis based on these structural alerts is oversensitive: most carcinogens possess these features, yet there are many benign, non-carcinogenic molecules that have these same functional groups. This is why the exclusion lists applied in practice are subject to corporate or personal choice, especially when general libraries intended for lead discovery are designed.

2. *Rule of 5*

In order to be bioavailable, a drug molecule must be transported through biological membranes. Therefore, molecular properties that correlate with poor membrane permeability (in the absence of active transport) can be used to filter out undesirable molecules.

The "Rule-of-Five" (Ro5)[1] is one of the best known and widely accepted "drug-likeness" filters. The Ro5 was derived from an analysis of 2,245 drug candidates from World Drug Index (WDI)[7] with assigned United States Adopted Name (USAN) or International Proprietary Name (INN) that are believed to have reached Phase II trials. It states that a compound violating any two of the following rules is likely to be poorly absorbed:

1. molecular weight less than 500 Da;
2. number of hydrogen bond donors (OH or NH groups) equal or less than 5;
3. number of hydrogen bond acceptors less than 10;
4. calculated Log *P* less than 5.0 (by *C* log *P*) or 4.15 (by *M* log *P*).

The Ro5 is intended to filter out molecules with potential absorption problems and it does so very well using simple terms that are easily implemented and used by medicinal chemists. This rule should not be expected, however, to discriminate well between drugs and non-drugs. Neither can it provide a quantitative estimate of compound absorption or take into account active transport mechanisms.

In using Ro5, one should also bear in mind that Log *P* characterizes the lipophilicity of a molecule in its neutral state. Yet the vast majority of drugs – 95% – are ionizable. It was

recently suggested[8] that using log D, a pH dependent version of log P, is more appropriate now that *in silico* methods for predicting pK_a and log D are available. In particular, since drugs are absorbed in the small intestine at pH ~ 5.5, log D at this pH (log $D_{5.5}$) should be used as a parameter in Ro5. When applied to several databases of known drugs, over 5% more molecules passed the Ro5 lipophilicity criterion using log $D_{5.5}$ as a descriptor compared to log P. Interestingly, an average difference of ~3 log units between log P and log $D_{5.5}$ was observed for drugs that failed log $P < 5$ criterion.

3.. Other rule-based filters

A consensus definition of a "drug-like" molecule has been derived[9] from the analysis of the CMC database[10] by defining qualifying (covering 80% of drug molecules) ranges of calculated physical properties such as molecular weight, log P, molar refractivity, and number of atoms:

- molecular weight between 160 and 480, average 357;
- calculated log P between –0.4 and 5.6, average 2.52;
- molar refractivity between 40 and 130, average 97;
- total number of atoms between 20 and 70, average 48.

In addition to these property ranges, the consensus definition requires molecules to have a combination of several of the following functional groups: a benzene ring, a heterocyclic ring (both aliphatic and aromatic), an aliphatic amine (preferably tertiary), a carboxamide group, an alcoholic hydroxyl group, a carboxy ester, and a keto group.

Another study[11] analyzed property distribution in drug databases and found that 70% of drugs are found within the following ranges of properties:

- number of H-bond donors between 0 and 2;
- number of H-bond acceptors between 2 and 9;
- number of rotatable bonds between 2 and 8;
- number of rings between 1 and 4.

An important parameter that strongly correlates with membrane permeability is polar surface area (PSA) – a sum of van der Waals surface areas of polar atoms (oxygens and nitrogens).[12,13] Multiple studies[12,14] have found an upper PSA threshold value of ~140–150 Å^2 for oral absorption. Similarly, for blood–brain barrier penetration, the upper PSA threshold is ~90 Å^2 (Figure 12.1). Recently, fast algorithms for PSA calculation have been suggested that do not need a

FIGURE 12.1 Distribution of the polar surface area (Å) for drugs that have reached at least Phase II efficacy studies. Panel A: 776 orally administered CNS drugs. Panel B: 1,590 orally administered non-CNS drugs.[12]

3D structure and use molecular topology instead (topological PSA, or TPSA).[15] Using these algorithms, a PSA filter can be applied to large collections of real or virtual compounds to eliminate the ones with high PSA.

An even more refined PSA-based filter[16] that also takes into account molecular flexibility is based on the analysis of oral bioavailability measurements in rats for over 1,100 drug candidates. The following two criteria were suggested for achieving bioavailability in rats above 20–40%:

- PSA equal or less than 140 $Å^2$ (or 12 or fewer H-bond donors and acceptors);
- ten or fewer rotatable bonds.

The first criterion is related to the desolvation energy of polar groups that is required for permeation through membrane. The second criterion probably reflects the entropic cost of conformational changes required to present an appropriate exterior to the hydrocarbon interior of the membrane.[16]

It is noteworthy that these two criteria worked well independently of molecular weight, a property that has been perceived as very important: consider a 500 Da cutoff that is commonly used to define the "drug-like" molecular weight range. The perception that this property is important probably arises from the apparent dependence of oral bioavailability on properties that correlate with molecular weight, which certainly include the number of rotatable bonds and PSA or H-bond donor count. As such, the value of a hard molecular weight cutoff may need to be reexamined.

A Pharmacophore Point Filter[17] contains a set of rules based on medicinal chemistry experience and on the observation that non-drug molecules are often underfunctionalized. Fundamental to this system is evaluating molecules for the occurrence of Pharmacophore Points, functional groups that potentially provide key interaction with the receptor or enzyme. The following rules constitute this filter:

1. Molecules with less than 2 Pharmacophore Points are not "drug-like."
2. Molecules with more than 7 Pharmacophore Points are not "drug-like."
3. The basic set of Pharmacophore Points includes: amine, amide, alcohol, ketone, sulfone, sulfonamide, carboxylic acid, carbamate, guanidine, amidine, urea, and ester.
4. Pharmacophore Points are fused and counted as one if their heteroatoms are not separated by more than one carbon atom.
5. Intracyclic amines that occur in the same ring (e.g. piperazine) are counted as one Pharmacophore Point.
6. Primary, secondary and tertiary amines are considered Pharmacophore Points, but not pyrrole, indole, thiazole, isoxazole, other azoles and diazines.
7. Compounds with more than one carboxylic acid are not "drug-like."
8. Compounds without a ring structure are not "drug-like."

When applied to the MDDR[18] and CMC databases, this filter classified almost 70% of molecules as "drug-like," while in the Advanced Chemistry Development (ACD) database only 36% of molecules were found to be "drug-like." The advantage of this filter is that it provides a detailed structural reason for classifying molecules as drugs or non-drugs.

The filters described above or others described in the literature[19,20] should not be used indiscriminately. It is critical to consider the context in which they are applied. For example, different property filters should be considered when looking for an oral central nervous system (CNS) agent for chronic use and for a chemotherapeutic agent administered intravenously for a short period of time.

B. Mimicking known drugs

The introduction of desirable features found in known drugs into the structures of library compounds is an alternative to the use of negative filters.

1. *Privileged structures*

The concept of privileged structures has been introduced[21] to describe select structural types capable of binding to multiple, unrelated classes of receptors or enzymes with high affinity. Privileged structures are typically constrained, heterocyclic multi-ring systems capable of orienting varied substituent patterns in a well-defined 3D space.[22] In addition to known drugs, privileged structures can be derived from natural products.[23] Examples of privileged structures are shown in Figure 12.2.

Using privileged structures as scaffolds is a powerful approach to make "drug-like" libraries.[24] One has to keep in mind, though, that having these scaffolds does not guarantee that all compounds in the library will be bioavailable and non-toxic. Proper choice of substituents is required to

FIGURE 12.2 Examples of "privileged" structures.

FIGURE 12.3 Topological frameworks for compounds in the CMC database (numbers indicate frequency of occurrence).[27] Drug molecules are first reduced to a combination of rings and linkers forming the framework of the molecule, then atom and bond type identities are removed.

avoid structural alerts and keep physicochemical properties such as PSA or solubility within the desired range.

In recent years, the meaning of the term "privileged structures" has been expanded to include all structural fragments, not just scaffolds that are frequently found in drugs. Sets of such fragments are offered commercially as building blocks for libraries.[25,26]

2. *Extraction of privileged substructures from drug databases*

Databases of known drugs can be examined to identify any preferred substructures or combinations thereof that could be used in the design of "drug-like" libraries. An evaluation of 5,120 structures from the CMC database identified a diverse set of 1,179 different topological frameworks.[27] Only 32 frameworks, however, were required to describe 50% of the drugs in the set (Figure 12.3). This indicates that a surprisingly small number of common privileged "shapes" can be used to design a variety of drug molecules with a wide variety of physical properties, binding to different classes of receptors. A six-member ring is the most commonly used framework, with 23 of the above-mentioned 32 frameworks containing at least two linked or fused six-member rings. Privileged frameworks tend to be relatively rigid, with only 3 out of 32 having more than five

rotatable bonds. Only 6% of analyzed drugs had an acyclic framework.

An analysis of side chains in drug molecules, performed by the same authors[28] showed that there are only 1,246 different side chains among molecules in the CMC database. The average number of side chains among the molecules is 4, and the average number of heavy atoms in each side chain is 2, although 66% of all side chains have only one heavy atom. Similar analysis of the same database[9] found that the most abundant functional groups are tertiary amines, alcohols, and carboxamides.

3. *Artificial Intelligence-based classifiers discriminating drugs from non-drugs*

Utilizing privileged structures to mimic known drugs has the advantage of providing chemists with very specific suggestions about what molecular fragments to include in their synthesis. However, this concept does not cover all molecular properties that might make a molecule "drug-like." Because of the general vagueness of the term "drug-like," it is difficult to even define what those properties or combinations of properties are.

Meanwhile, there are several public and commercial drug databases that lend themselves to examination for possible property patterns intrinsic to drugs.[7,10,18,29,30] Several pattern recognition approaches have been tried[31–36] in which computer is presented with examples of drugs (molecules from drug databases) and non-drugs (typically, ACD database cleaned of reactive molecules) and is trained to discriminate between them. After the training, these algorithms recognized 80–90% of molecules in drug databases correctly as drugs.

A concern intrinsic to all these computational classifiers is that they merely reflect the characteristics of existing drugs, and that using these classifiers as filters would impede the discovery of novel structural classes of drugs in the future. Furthermore, since these classifiers are trained on specific, carefully processed databases, the results could be strongly database-dependent.

C. Direct property prediction

Prediction of general "drug-likeness" based on chemical structure is an ambitious goal considering how vague the concept of "drug-likeness" is. A large number of factors determine whether a compound can become a drug or not, and only some of them are structure-dependent. This is why most general "drug-likeness" classifiers derived from analyzing databases of known drugs cannot be expected to be precise and unambiguous. When large general libraries are designed for lead generation, these classifiers are instrumental in filtering out compounds with liabilities from huge virtual libraries. On the other hand, at the lead optimization stage of drug discovery projects, when specific issues need to be addressed for specific chemical series, more reliable predictions are desirable.

The issues chemists deal with during lead optimization are the same as those that contribute to general "drug-likeness": the need to reduce toxicity or to improve solubility, oral absorption, blood–brain barrier penetration, and metabolic stability. Computational methods that can predict directly these properties of compounds would allow for the design of individual compounds and libraries that are instrumental in solving these specific needs.

1. *Solubility*

Solubility plays an important role at all stages of drug discovery. Insoluble compounds cause problems in screening. Compounds with low solubility are poorly absorbed in the intestine. Predicting solubility is a challenge because it is a complex phenomenon. It depends on compound lipophilicity, the number of H-bonds that can be formed with the solvent, the ability to form an intramolecular H-bond, the ionization state of functional groups, and the properties of the crystalline form, in particular, the nature and energy of crystal packing.

Of a particular interest are methods that do not require any measured parameters (e.g. melting point) as an input. In one example of such an approach,[37] standard regression statistics have been applied to a series of whole-molecule descriptors that have a straightforward physical interpretation (as opposed to fragment-based descriptors): Lennard-Jones interaction energy, solvent accessible surface area, number of H-bond donors and acceptors, and a count of amino and nitro groups. QikProp,[38] a commercially available program implements this algorithm to calculate solubility and other physicochemical properties. The reader is referred to a more comprehensive account of solubility prediction methods[39] for further details.

2. *Oral bioavailability*

Oral bioavailability, a fraction of the oral dose that reaches blood circulation, is a multi-mechanism phenomenon, as well. It depends on the drug's solubility, chemical and metabolic stability, and membrane permeability. Most compounds are absorbed through passive diffusion across the membrane, while some are transferred by a carrier or by transporter molecules residing in the membrane. For example, amino acids and sugars are actively transported by specific transporters, including peptide transporters, organic cation transporters, and ATP-binding cassette (ABC) transporter proteins. On the other hand, some efflux proteins, especially P-gp, localized in the apical or basolateral cell membranes pump drugs out from the cell into extracellular fluids.

The key factors determining membrane permeability (for passive transport) include PSA and lipophilicity (log P), discussed earlier in this chapter. Many statistical models for predicting permeability have been suggested and reviewed in detail.[40]

3. Blood–brain barrier penetration

The determinants of blood–brain barrier penetration are similar to the determinants of membrane permeability. They include lipophilicity (log P), H-bonding capacity, ionization profile, size, and flexibility. An example of a simple quantitative structure–activity relationship (QSAR) equation to calculate the ratio of the steady state concentration of the drug molecule in the brain and in the blood have been described[41] (for a comprehensive review of the *in silico* methods see[42]):

$$\log \frac{C_{\text{brain}}}{C_{\text{blood}}} = -0.0148 * \text{PSA} + 0.152 * C \log P + 0.139$$

4. Metabolic stability

Currently, there are no reliable methods for the quantitative prediction of metabolic stability. However, over the years researchers have accumulated significant knowledge of what chemical groups might be metabolically labile and what the resulting metabolites are. Two software packages based on knowledge bases are available – MetabolExpert[43] and META.[44] Both provide alerts about potential metabolic pathways but do not indicate how probable those pathways are.

Understanding the structural requirements of specific metabolic enzymes (e.g. various isoforms of P450) is a very active area of research, and in the future it might provide quantitative predictive models of metabolic stability.

5. Toxicity

The main causes of chemical toxicity are side reactions of drug molecules with DNA or proteins, as well as interference with enzymatic systems. Although toxicity information is often fragmented and not standardized, there are more than a dozen public or commercial toxicity databases available that cover cancerogenicity, genetic toxicity/mutagenicity, target organ toxicity, reproductive toxicity, developmental toxicity, and immunologic toxicity data.[45]

Computational approaches to predict toxicity are based on expert systems containing empirical or computationally derived rules, typically encoded via structural fragments that potentially could cause toxicity.[46,47]

III. SELECTING BETTER LEADS: "LEAD-LIKENESS"

A. What is a "high-quality" lead compound?

1. Smaller and less lipophilic

Although few hoped to find drugs directly from HTS, it seemed logical to limit the set of screened compounds to the "drug-like" space so that it would be easier to stay within that space during lead optimization. The various "drug-likeness" assessment tools and filters described above have been developed and applied to both historical compound collections and to combinatorial libraries with the intention of obtaining "high quality" lead compounds after screening. This approach makes an implicit assumption: that any "drug-like" molecule is in fact a "high-quality" lead. It was observed that, in many projects, the structure of a drug candidate sent to clinic was very close to the structure of the original lead,[48] so it made sense to start with leads that most closely resembled a drug.

Certainly, a "drug-like" lead is more desirable than a compound that is carcinogenic or insoluble or too polar. However, it does not necessarily follow that "drug-like" lead compounds are indeed the best starting points for optimization. Several comparative studies[2,49,50] conclusively demonstrate that, on average, lead optimization tends to increase molecular weight, lipophilicity, solubility, and molecular complexity. This happens because, in the absence of direct structural information (e.g. X-ray structure), adding more non-directional hydrophobic interactions are more likely to improve potency than forming a new, highly direction-sensitive hydrogen bond that does not increase the size of the molecule. Another possible reason is that changing the bulk properties of a molecule is often used as a means of tuning its PK properties. Thus if a lead compound is already at the boundary of the drug-like space in terms of complexity and physicochemical properties, chemists will be more constrained in what they can do during the optimization phase.

2. Potency alone is a poor predictor of lead quality: ligand efficiency

Intuitively, one might be concerned that smaller molecules might not have enough interacting surface to bring about the high affinity interaction that is typically required to produce a biological effect. Kuntz[51] showed that maximal free-energy contribution per non-hydrogen atom is approximately −1.5 kcal/mol. Using a simple equation that expresses the free energy of ligand binding through its dissociation constant:

$$\Delta G = -RT \ln K_{\text{d}}$$

one can see that subnanomolar binding can theoretically be achieved with as few as 8–9 atoms and there are drugs that bind very tightly in spite of their very small size (allopurinol, 10 atoms, pK_d 9.17; guanobenz, 14 atoms, pK_d 9.02; aminoclonidine, 15 atoms, pK_d 9.32). Many endogenous ligands are also quite small yet bind with high affinity to their receptors (acetylcholine, 10 atoms, pK_d 8.14; dopamine, 11 atoms, pK_d 8.65; serotonin, 13 atoms, pK_d 8.30). In these small molecules, all atoms are exposed on the molecular surface and most contribute to the interaction with the complementary receptor surface. In other words, these molecules "use" each of their atoms very efficiently for binding to the receptor.

It is the value of binding free energy per non-hydrogen atom (or "ligand efficiency," $\Delta g = \Delta G/N_{\text{nonhydrogen atoms}}$[52]), not potency, that should be used as the measure of lead compound quality. Consider a situation where one needs to choose between two screening hits. One has 40 non-hydrogen atoms (MW = ~530[52]) and binding affinity of 100 nM and the other has 22 non-hydrogen atoms (MW ~ 290) and binding affinity of 1 μM. Although it is tempting to select a more potent hit, let's first calculate their ligand efficiencies. Using the formula for ΔG introduced above, ΔG of the 100 nM hit is −9.6 kcal/mol and its ligand efficiency, Δg, is −0.24 kcal/mol, while for the second hit ΔG is −7.2 kcal/mol and Δg is −0.33 kcal/mol. A less potent hit turns out to be a more efficient ligand and is therefore a better starting point for lead optimization. The molecular weight of the first hit is already slightly above the "drug-like" boundary and its optimization would require figuring out which parts of the molecule are not essential (have low Δg) and replacing them with different functional groups with better Δg. Such improvements are certainly possible in many cases, but could be difficult and time consuming. Starting with the second hit, on the other hand, would allow a medicinal chemist to build the molecule up with functional groups of total MW up to ~200 and still stay within "drug-like" boundaries.

Most compounds medicinal chemists work with have between 20 and 50 non-hydrogen atoms, and their ligand efficiency falls far below the theoretical limit of −1.5 kcal/mol per non-hydrogen atom.[51] It is instructive to analyze why this might be the case.

One reason is purely sterical and is related to the 3D structure of a molecule itself. As the size of the molecule grows, more and more atoms get buried in the bulk of the molecule and are not exposed to its surface. They are thus unable to interact with the receptor and cannot contribute to binding energy.

The second reason has to do with how well the ligand surface properties (such as shape, charge distribution and lipophilicity) match those of the receptor or enzyme active site. Some non-hydrogen atoms might be at the ligand's surface but match poorly and thus not contribute to binding energy. Using a very simple model of ligand and receptor interaction, Hann *et al.*[2] demonstrated that the probability of a good ligand–receptor match quickly decreases as molecular complexity increases.

The concept of ligand efficiency can be extended to include other molecular properties relevant for lead optimization when one would like to control these properties while retaining potency. Other ligand efficiency indices can be calculated by normalizing binding free energy (or binding constants) by log *P*, serum protein binding affinity, or PSA.[53]

B. Designing "lead-like" libraries for biochemical screening

Ligand efficiency indices can only be calculated in the context of a particular target, after screening is completed and binding affinities measured, so they cannot be used in selecting specific molecules into the screening library. Therefore, various "lead-likeness" filters have been proposed, based on analyzing the industry track record of converting leads of various types into drugs. In addition to physicochemical properties such as molecular weight or lipophilicity, these filters consider chemical classes that historically have failed to deliver drugs from screening leads as well as the amenability of a lead to chemical analoging.

1. *Rule of* 3

The first "lead-likeness" filter was suggested by Teague *et al.*[49] who proposed to include molecules with MW between 100 and 350 and $C\log P$ between 1 and 3 as opposed to the Ro5 for "drug-like" molecules (a similar "Rule of 3" filter was later proposed in the context of designing libraries for fragment-based screening[54]). Teague *et al.* also pointed out that in order for molecules this small to achieve a level of affinity measurable in HTS (10–20 μM), it is beneficial to have functional groups capable of strong binding with little MW penalty, for example, basic groups that are charged at physiological pH. Both small size and strongly interacting groups make high ligand efficiency more likely for the members of a "lead-like" library.

Special attention has to be paid when following up on "weak" HTS hits in order to separate them from the numerous false positives usually found at this low level of potency. Physical methods such as mass spectroscopy (MS), nuclear magnetic resonance (NMR) and X-ray crystallography, which are capable of detecting binding at much higher concentrations allow chemists to work with even smaller molecular fragments that bind to their protein targets even more weakly (see chapter 11).

There is another reason to favor smaller molecules. Screening a library is an experiment in which chemical space is sampled with regard to the ability to bind to the target of interest. The better a given library samples chemical space, the higher the probability of finding all possible

FIGURE 12.4 Reactive functional groups responsible for *in vitro* false positives (X = F, Cl, Br, I, tosyl, mesyl, etc.; R = alkyl, aryl, heteroalkyl, heteroaryl, etc.). These reactive functional groups are generally prone to decomposition under hydrolytic conditions (i.e. aqueous Na_2CO_3/methanol). They are reactive toward protein and biological nucleophiles (e.g. glutathione, dithiothreitol), and they exhibit poor stability in serum.[58]

different lead compounds. The number of possible chemical structures grows exponentially with the square of the number of heavy atoms. The size of the chemical space of "drug-like" (MW ~ 500) molecules ($\sim 10^{60}$)[55] is many orders of magnitude larger than the size of current screening libraries ($\sim 10^{6-7}$), so the sampling of chemical space in such libraries is extremely sparse. At lower molecular weights (MW < 300), compounds that are accessible for screening (e.g. from commercial sources or from proprietary collections) represent a much better sampling of chemical space.[56]

2. *Avoiding chemical classes with a known high failure rate in drug discovery and development*

Many chemical functionalities have been identified as undesirable in a lead compound.[57,58] Compounds that have reactive electrophilic functional groups can form covalent bonds with proteins and will generate false positives in screening or may appear to be toxic. Such compounds should be excluded from the screening library. An example of an exclusion list of reactive groups is shown in Figure 12.4. Hydroxamates and thiol chelators could also cause artifacts in biochemical assays.

Another type of undesirable compounds to be excluded from screening libraries is "frequent hitters."[59] These compounds are detected as hits in many different biological assays covering a wide range of targets, which can happen for two main reasons: (i) the activity of the compound is not specific for the target (promiscuous compounds); (ii) the compound perturbs the assay or detection method, for example, colored or fluorescent molecules. In both cases, such molecules are usually poor starting points for lead optimization programs and can cost money and time without delivering any benefits. Computational filters have been developed to identify "frequent hitters."[59] Interestingly, the application of such filters identified some known drugs as frequent hitters, probably because these drugs might act as artifact-generating agents in biochemical screens against other targets. Polyphenols, polyionics and compounds with extended electron conjugation that are known to form aggregates in aqueous solution are represented in this group.

3. *Ease of chemical analoging*

Typical HTS screen generates multiple screening hits. To select the right lead compound, it is essential to determine if the chemical series around these screening hits has a well-defined SAR. If so, it is an indication that the lead can be built upon to add structural elements required for better potency, selectivity or ADME (absorption, distribution, metabolism, excretion) properties. Methods of parallel chemistry can be applied to a variety of scaffolds and functional groups in order to speed up the hit follow-up process. Enriching "lead-like" libraries with compounds that contain functional groups (linkers) amenable to reactions optimized for parallel chemistry becomes highly desirable. This approach is not universally applicable because some of these groups are reactive (e.g. aldehydes) and others, like amino groups or carboxylic groups could be essential for binding to the receptor and using them for derivatization would interfere with binding.

IV. CONCLUSION

Medicinal chemists can chose an infinite number of paths in chemistry space in their journey from an idea to a drug. Both drug-likeness and lead-likeness concepts steer chemists toward those paths that, based on industry experience, have higher probability of success. It is important to remember that as any other statistics-based rule, rules of drug- and lead-likeness are not absolute, and there are exceptions, Lipitor being among the most notable examples. What drives a chemist's choices at all phases of drug discovery and development is measured data about the liabilities of each particular compound or a series of compounds. Finding an optimization path that leads to a compound with fewest liabilities ultimately determines the success of a program.

The rationale for enriching screening libraries with lead-like compounds is very strong. Does it mean that we should stop screening historical collections of "drug-like" molecules? That would have been a tremendous waste of valuable chemical matter. Ideally, screening should produce a variety of potential leads – both "lead-like" and "drug-like." A combination of multiple factors, some related to the chemical structure and some not, will point to the best starting point for further optimization. Chemical optimization strategy, of course, will be different for "drug-like" versus "lead-like" compounds, but winning strategies can usually be found in both cases.

REFERENCES

1. Lipinski, C. A., Lombardo, F., Dominy, B. W., Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 1997, *23*, 3–25.
2. Hann, M. M., Leach, A. R., Harper, G. Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inform. Comput. Sci.* 2001, *41*, 856–864.
3. Matter, H., Baringhaus, K. H., Naumann, T., Klabunde, T., Pirard, B. Computational approaches towards the rational design of drug-like compound libraries. *Combin. Chem. High Throughput Screening* 2001, *4*, 453–475.
4. Walters, W. P. Ajay, Murcko, M. A. Recognizing molecules with drug-like properties. *Curr. Opin. Chem. Biol.* 1999, *3*, 384–387.
5. Muegge, I. Selection criteria for drug-like compounds. *Med. Res. Rev.* 2003, *23*, 302–321.
6. Ashby, J. Fundamental structural alerts to potential carcinogenicity or noncarcinogenicity. *Environ. Mutagen.* 1985, *7*, 919–921.
7. World Drug Index, Derwent Information: Alexandria, VA www.derwent.com.
8. Bhal, S. K., Kassam, K., Peirson, I. G., Pearl, G. M. The Rule of Five revisited: applying log *D* in place of log *P* in drug-likeness filters. *Mol. Pharm.* 2007, *4*, 556–560.
9. Ghose, A. K., Viswanadhan, V. N., Wendoloski, J. J. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of know drug databases. *J. Combin. Chem.* 1999, *1*, 55–68.
10. Comprehensive Medicinal Chemistry Database, 94.1; Distributed by MDL Information Systems: San Leandro, CA.
11. Oprea, T. I. Property distribution of drug-related chemical databases. *J. Comput.-Aided Mol. Des.* 2000, *14*, 251–264.
12. Palm, K., Stenberg, P., Luthman, K., Artursson, P. Polar molecular surface properties predict the intestinal absorption of drugs in humans. *Pharmacol. Res.* 1997, *14*, 568–571.
13. Clark, D. E. Rapid calculation of polar molecular surface area and its application to the prediction of transport phenomena. 1. Prediction of intestinal absorption. *J. Pharm. Sci.* 1999, *88*, 807–814.
14. Kelder, J., Grootenhuis, P. D. J., Bayada, D. M., Delbressine, L. P. C., Ploemen, J. P. Polar molecular surface as a dominating determinant for oral absorption and brain penetration of drugs. *Pharmacol. Res.* 1999, *16*, 1514–1519.
15. Ertl, P., Rohde, B., Selzer, P. Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *J. Med. Chem.* 2000, *43*, 3714–3717.
16. Veber, D. F., Johnson, S. R., Cheng, H. Y., Smith, B. R., Ward, K. W., Kopple, K. D. Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* 2002, *45*, 2615–2623.
17. Muegge, I., Heald, S. L., Brittelli, D. Simple selection criteria for drug-like chemical matter. *J. Med. Chem.* 2001, *44*, 1841–1846.
18. MACCS-II Drug Data Report Database, Distributed by MDL Information Systems, Inc., San Leandro, CA.
19. Wang, J., Ramnarayan, K. Toward designing drug-like libraries: a novel computational approach for prediction of drug feasibility of compounds. *J. Combin. Chem.* 1999, *1*, 524–533.
20. Xu, J., Stevenson, J. Drug-like index: a new approach to measure drug-like compounds and their diversity. *J. Chem. Inform. Comput. Sci.* 2000, *40*, 1177–1187.
21. Evans, B. E., Rittle, K. E., Bock, M. G., DiPardo, R. M., Freidinger, R. M., Whitter, W. L., Lundell, G. F., Veber, D. F., Anderson, P. S., Chang, R. S. L., Lotti, V. J., Cerino, D. J., Chem, T. B., Kling, P. J., Kunkel, K. A., Springer, J. P., Hirshfield, J. Methods for drug discovery: development of potent, selective, orally effective cholecystokinin antagonists. *J. Med. Chem.* 1988, *31*, 2235–2246.
22. Mason, J. S., Morize, I., Menard, P. R., Cheney, D. L., Hulme, C., Labaudiniere, R. F. New 4-point pharmacophore method for molecular similarity and diversity applications: overview of the method and applications, including a novel approach to the design of combinatorial libraries containing privileged substructures. *J. Med. Chem.* 1999, *42*, 3251–3264.

23. Nicolaou, K. C., Pfefferkorn, J. A., Roecker, A. J., Cao, G. Q., Barluenga, S., Mitchell, H. J. Natural product-like combinatorial libraries based on privileged structures. 1. General principles and solid-phase synthesis of benzopyrans. *J. Am. Chem. Soc.* 2000, *122*, 9939–9953.
24. DeSimone, R. W., Currie, K. S., Mitchell, S. A., Darrow, J. W., Pippin, D. A. Privileged structures: applications in drug discovery. *Combin. Chem. High Throughput Screening* 2004, *7*, 473–493.
25. PHARMABlock® Proprietary Combinatorial Building Blocks, ChemBridge Corporation: San Diego, CA, www.chembridge.com.
26. Optimer Building Blocks, Array Biopharma: Bolder, CO, http://www.arraybiopharma.com.
27. Bemis, G. W., Murcko, M. A. The properties of known drugs. 1. Molecular frameworks. *J. Med. Chem.* 1996, *39*, 2887–2893.
28. Bemis, G. W., Murcko, M. A. Properties of known drugs. 2. Side chains. *J. Med. Chem.* 1999, *42*, 5095–5099.
29. WOMBAT 2006.1, Distributed by Sunset Molecular Discovery LLC: Santa Fe, NM.
30. The PubChem Database, Available online at the National Center for Biotechnology Information, http://pubchem.ncbi.nlm.nih.gov/.
31. Walters, W. P. Ajay, Murcko, M. A. Can we learn to distinguish between drug-like and nondrug-like molecules? *J. Med. Chem.* 1998, *41*, 3314–3324.
32. Sadowski, J., Kubinyi, H. A scoring scheme for discriminating between drugs and nondrugs. *J. Med. Chem.* 1998, *41*, 3325–3329.
33. Frimurer, T. M., Bywater, R., Naerum, L., Lauritsen, L., Brunak, S. Improving the odds in discriminating drug-like from non drug-like compounds. *J. Chem. Inform. Comput. Sci.* 2000, *40*, 1315–1324.
34. Gillet, V. J., Willett, P., Bradshaw, J. Identification of biological activity profiles using substructural analysis and genetic algorithms. *J. Chem. Inform. Comput. Sci.* 1998, *38*, 165–179.
35. Wagener, M., van Geerestein, V. J. Potential drugs and nondrugs: prediction and identification of important structural features. *J. Chem. Inform. Comput. Sci.* 2000, *40*, 280–292.
36. Muller, K. R., Ratsch, G., Sonnenburg, S., Mika, S., Grimm, M., Heinrich, N. Classifying "drug-likeness" with kernel-based learning methods. *J. Chem. Inform. Model*. 2005, *45*, 249–253.
37. Jorgensen, W. L., Duffy, E. M. Prediction of drug solubility from Monte Carlo simulations. *Bioorg. Med. Chem. Lett.* 2000, *10*, 1155–1158.
38. QikProp, Schrodinger, Inc., Portland, OR. www.schrodinger.com.
39. Balakin, K. V., Savchuk, N. P., Tetko, I. V. *In Silico* approaches to prediction of aqueous and DMSO solubility of drug-like compounds: trends, problems and solutions. *Curr. Med. Chem.* 2006, *13*, 223–241.
40. Hou, T., Wang, J., Zhang, W., Wang, W., Xu, X. Recent advances in computational prediction of drug absorption and permeability in drug discovery. *Curr. Med. Chem.* 2006, *13*, 2653–2667.
41. Clark, D. E. Rapid calculation of polar molecular surface area and its application to the prediction of transport phenomena. 2. Prediction of blood–brain barrier penetration. *J. Pharm. Sci.* 1999, *88*, 815–821.
42. Ecker, G. F., N., C. R. *In Silico* prediction models for blood–brain barrier permeation. *Curr. Med. Chem.* 2004, *11*, 1617–1628.
43. MetabolExpert, CompuDrug: South San Francisco, CA. www.compudrug.com.
44. META, Charles River Laboratories: Wilmington, MA. www.criver.com.
45. Yang, C., Benz, R. D., Cheeseman, M. A. Landscape of current toxicity databases and database standards. *Curr. Opin. Drug. Discov. Dev.* 2006, *9*, 124–133.
46. Mohan, C. G., Gandhi, T., Garg, D., Shinde, R. Computer-assisted methods in chemical toxicity prediction. *Mini-Rev. Med. Chem.* 2007, *7*, 499–507.
47. Richard, A. M. Future of toxicology-predictive toxicology: an expanded view of chemical toxicity. *Chem. Res. Toxicol.* 2006, *19*, 1257–1262.
48. Proudfoot, J. R. Drugs, leads, and drug-likeness: an analysis of some recently launched drugs. *Bioorg. Med. Chem. Lett.* 2002, *12*, 1647–1650.
49. Teague, S. J., Davis, A. M., Leeson, P. D., Oprea, T. The design of leadlike combinatorial libraries. *Angew. Chem. Int. Ed.* 1999, *38*, 3743–3748.
50. Oprea, T. I., Allu, T. K., Fara, D. C., Rad, R. F., Ostopovichi, L., Bologa, C. G. Lead-like, drug-like or pub-like: How different are they? *J. Comput.-Aided Mol. Des.* 2007, *21*, 113–119.
51. Kuntz, I. D., Chen, K., Sharp, K. A., Kollman, P. A. The maximal affinity of ligands. *Proc. Natl. Acad. Sci. USA*, 1999, *96*, 9997–10002.
52. Hopkins, A. L., Groom, C. R., Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discov. Today*. 2004, *9*, 430–431.
53. Leach, A. R., Hann, M. M., Burrows, J. N., Griffen, E. J. Fragment screening: an introduction. *Mol. BioSyst.* 2006, *2*, 429–446.
54. Congreve, M., Carr, R., Murray, C. W., Jhoti, H. A Rule of Three for fragment-based lead discovery?. *Drug Discov. Today*. 2003, *8*, 876–877.
55. Bohacek, R. S., Martin, C., Guida, W. C. The art and practice of structure-based drug design: a molecular modeling perspective. *Med. Res. Rev.* 1996, *16*, 3–50.
56. Hann, M. M., Oprea, T. I. Pursuing the leadlikeness concept in pharmaceutical research. *Curr. Opin. Chem. Biol.* 2004, *8*, 255–263.
57. Rishton, G. M. Reactive compounds and *in vitro* false positives in HTS. *Drug Discov. Today*. 1997, *2*, 382–384.
58. Rishton, G. M. Non-leadlikeness and leadlikeness in biochemical screening. *Drug Discov. Today*. 2003, *8*, 86–96.
59. Roche, O., Schneider, P., Zuegge, J., Guba, W., Kansy, M., Alanine, A., Bleicher, K., Danel, F., Gutknecht, E. M., Rogers-Evans, M., Neidhart, W., Stalder, H., Dillon, M., Sjogren, E., Fotouhi, N., Gillespie, P., Goodnow, R., Harris, W., Jones, P., Taniguchi, M., Tsujii, S., von der Saal, W., Zimmermann, G., Schneider, G. Development of a virtual screening method for identification of frequent hitters in compound libraries. *J. Med. Chem.* 2002, *45*, 137–142.

Chapter 13

Web Alert: Using the Internet for Medicinal Chemistry

David Cavalla

I. INTRODUCTION
II. BLOGS
III. WIKIS
A. RSS information feeds
IV. COMPOUND INFORMATION
A. Chemspider
B. The NIH Roadmap and PubChem
C. ChemBank
V. BIOLOGICAL PROPERTIES OF COMPOUNDS
A. Prediction of biochemical properties
B. Molecular datasets
C. Information on metabolic properties
VI. DRUG INFORMATION
A. DrugBank
VII. PHYSICAL CHEMICAL INFORMATION
VIII. PREDICTION AND CALCULATION OF MOLECULAR PROPERTIES
IX. CHEMICAL SUPPLIERS
X. CHEMICAL SYNTHESIS
XI. CHEMICAL SOFTWARE PROGRAMS
A. Chemical drawing and viewing software
B. Various chemoinformatics software
C. Datasets for virtual screening
XII. ANALYSIS
XIII. CHEMICAL PUBLICATIONS
A. Journals
B. Open Access
C. Theses
XIV. PATENT INFORMATION
A. Japanese patents
XV. TOXICOLOGY
XVI. METASITES AND TECHNOLOGY SERVICE PROVIDER DATABASES

Imagine a world in which every single person on the planet is given free access to the sum of all human knowledge. That's what we're doing.

Jimmy Wales, founder of Wikipedia

I. INTRODUCTION

The internet has undergone further substantial change since the 2nd edition of Practice of Medicinal Chemistry, both in the continued growth in the internet, and the availability of additional resources. This article can only describe the situation as it currently stands and give some predictions as to the future. It is in the nature of such reviews that they can never be complete and up-to-date, moreover they deteriorate rapidly with time. Further more, the resources on the internet are now too vast for a comprehensive article in the space available; a view has had to be taken to include only what can be considered to be the most significant and likely to last.

This review is intended to provide freely available resources for the tools that medicinal chemists use generally in the work they do, which necessarily involves a variety of tasks, from drug design to chemical synthesis. Sites for prediction of physical activity are just as relevant as those for prediction of biological activity and patents.

While there has been a substantial increase in the number of commercial sites for chemistry since the last edition, there has not been the predicted shift in balance away from freely available resources. If anything there has been a surge in open resources, most notably in the open-access publishing movement (see later). There has also been the advent of certain forms of internet use which were not evident, or very substantially smaller in impact and not predicted to grow in the way they have done over the last 4 years. These include the introduction of the blog and the rise and rise of the wiki.

Copyright © 2008, Elsevier Ltd
All rights reserved.

II. BLOGS

The blog, or web log, has rapidly risen to a position of great importance in news and other forms of media. However, it has been relatively more slowly introduced into science in general. There are no quality controls in the blog medium; however, some gems are to be found. Some of the more well-known generalist sites are given in Table 13.1.

There is a "hyperblog" site at http://wiki.cubic.uni-koeln.de/pg. This site, among other topics, features a section on popular stories and areas of common interest. There is a section dealing with blogs on various aspects of chemistry from analytical to chemometrics/bioinformatics to pharmacology and genetic modification.

In addition, Table 13.2 shows the sites with specialization on medicinal and pharmaceutical chemistry.

TABLE 13.1

Title	URL	Comments
A Zephyr in Time	http://echiral.blogspot.com/search/label/Chemistry	A set of general articles on chemistry.
Chemical Blogspace	http://wiki.cubic.uni-koeln.de/cb/index.php	Chemical blogspace collects data from tens of scientific chemistry blogs and presents it in one place.
Chemical Forums Blog	http://blog.chemicalforums.com/	Tied to the chemical forums, a place for open discussion related to chemistry.
Chemistry World Blog	http://prospect.rsc.org/blogs/cw/	Blog from the UK Royal Society of Chemistry (http://www.rsc.org).
Culture of Chemistry	http://cultureofchemistry.blogspot.com/	Authored by Michelle Francl, Professor at Bryn Mawr College.
Manufacturing Chemistry	http://pharmamanufacturing.wordpress.com/about	Administered by Pharmaceutical Manufacturing's editor in chief, Agnes Shanley (see http://www.PharmaManufacturing.com).
Science Quick Picks	http://www.pontotriplo.org/quickpicks/	This site deals with popular chemistry but also other branches of science, education, and technology.
SciScoop	http://www.sciscoop.com	General science blog site and forum.
The Curious Wavefunction	http://ashutoshchemist.blogspot.com/	Miscellaneous thoughts, facts, and tidbits, recent and past, about chemistry, science, and society.
The Skeptical Chymist	http://blogs.nature.com/thescepticalchymist/	The Sceptical Chymist is a blog by the editors of *Nature* and the Research journals

TABLE 13.2

Title	URL	Comments
Medicinal Chemistry: In the Pipeline	http://www.corante.com/pipeline/	The author Derek Lowe has worked for several major pharmaceutical companies, on drug discovery projects against schizophrenia, Alzheimer's, diabetes, osteoporosis, and other diseases.
One in Ten Thousand	http://walkerma.wordpress.com/	General discussion on medicines and pharmaceutical chemistry.
The Half Decent Pharmaceutical Chemistry Blog	http://the-half-decent-pharmaceutical-chemistry-blog.chemblogs.org/	A blog on pharmacology and pharmaceutical chemistry.
The Molecule of the Day	http://scienceblogs.com/moleculeoftheday/	Focused on medicines in the context of real life.
Totally Medicinal	http://totallymedicinal.wordpress.com/	A blog looking at the world of pharmaceutical chemistry.
Kinasepro	http://kinasepro.wordpress.com/	Excellent site on kinase medicinal chemistry.

Further examples of blogs related specifically to synthetic chemistry are detailed in the section on Chemical Synthesis.

A report has been written on the effect of blog sites on the Pharmaceutical Industry (http://www.pharmaceutical-business-review.com/research.asp?guid=BFHC0707). As blogs become more widespread, they will influence the way in which healthcare is delivered, and peer-to-peer health advice will become another means by which information is delivered to patients.

III. WIKIS

From software to encyclopedias, collaborative projects are one of the most evidently disruptive applications of the internet, posing multiple challenges to conventional business models. Wikipedia (http://www.wikipedia.org) is well established, even though its first page went online only in 2001. It is a generalist information source but its scope and depth exceeds many specialist alternatives. The word is a composite of "wiki" being the Hawaiian word for quick and encyclopedia, of which it is now the world's biggest. An interesting article on the origins, implementation, and phenomenal growth of Wikipedia was published in the The Atlantic in 2006 (http://www.theatlantic.com/doc/200609/wikipedia).

It is the collaborative (and indeed co-operative) nature of wikis that has enabled the rapid growth of Wikipedia; by comparison with the first *Oxford English Dictionary* which took 78 years before the first product was published in 1928 (http://www.oed.com/), Wikipedia is 6 years old and has more than 1.7 million articles in its English language edition, growing by nearly 2,000 a day. In 2006, a report in *Nature* (http://www.nature.com) compared Encyclopedia Britannica and Wikipedia science articles and suggested that the former are usually only marginally more accurate than the latter.

As an example, the entry for "angiotensin II antagonists" (http://en.wikipedia.org/wiki/Angiotensin_II_receptor_antagonist) nicely leads to a list of seven members of the group. While incomplete as a list, each entry contains a graphical structure, an IUPAC name, CAS number, Pubchem link, bioavailability, protein binding, metabolism, half life, and so on. In the main section, the articles include information on regulatory status, dosing frequency, therapeutic indications, and side effects.

Other wiki-based information resources are more candidly scientific in focus, but nascent in coverage. Biocrawler.com (http://www.biocrawler.com) claims to be "a [scientific] encyclopedia written collaboratively." It is divided into sections on biotechnology (http://www.biocrawler.com/biocorp/index.php?words=biotech) and drug discovery (http://www.biocrawler.com/biocorp/index.php?words=drug) companies, as well as videos, images, and a directory. The site currently offers a much smaller collection of entries than Wikipedia, and moreover, less well developed information even in those entries that exist. "Mitochondrion" for instance garners a much more impressive set of information in the generalist Wikipedia than in Biocrawler.

The Chemical Information Sources Wiki (http://cheminfo.informatics.indiana.edu/cicc/cis/index.php/Main_Page) is a guide to the many sources of reference materials available for those with questions related to chemistry. The site includes information on primary, secondary, and tertiary publication sources, chemical information databases, physical property information, chemical patent searching, and molecular visualization tools and sites. The material is based on an undergraduate course offered for many years in the Indiana University Department of Chemistry by Gary Wiggins.

A. RSS information feeds

Nowadays web information is often delivered by means other than the simple (static) web browser. RSS feeds (a new format for notifying new content at a web site), which stands for really simple syndication, and is a new way of getting news in general, has recently been introduced for the *American Chemical Society* (http://www.acs.org) journals (see http://pubs.acs.org/alerts/rss/index.html). The information is not fed directly into a browser, but into a news aggregator such as FeedDemon (http://www.feeddemon.com).

IV. COMPOUND INFORMATION

A. Chemspider

There is a growing number of compound databases, and the dominance of Chemical Abstracts (http://www.cas.org) is thus challenged, however the mere diversity of these databases poses its own difficulty. ChemSpider (http://www.chemspider.com/) is a new chemistry search engine built with the intention of aggregating and indexing chemical structures and their associated information into a single searchable repository and making it available to everybody, at no charge. Some properties have been added to each of the chemical structures within the database such as structure identifiers like SMILES, InChI, IUPAC and Index names as well as many physicochemical properties. In addition, ChemSpider provides access to a series of property prediction algorithms.

ChemSpider currently searches over 14 million compounds in multiple chemical structure databases. These include databases of curated literature data, chemical vendor catalogs, molecular properties, environmental data, toxicity data, analytical data, and so on. ChemSpider intends to aggregate into a single database all chemical structures available within open access and commercial databases and to provide the necessary pointers from the ChemSpider search engine to the information of interest. This service will allow users to either access the data immediately via open-access links or have the information necessary to continue their searches into commercially available systems.

Two blogs support the system, one for the science, politics and vision behind ChemSpider (http://www.chemspider.com/blog) and another for incremental changes in functionality at http://www.chemspider.com/news.

B. The NIH Roadmap and PubChem

The NIH (National Institutes of Health) "Roadmap" was launched in October 2004 (http://nihroadmap.nih.gov/initiatives.asp) to encompass five themes, namely building blocks, biological pathways, and networks, molecular libraries and molecular imaging, structural biology, bioinformatics and computational biology and nanomedicine. PubChem (http://pubchem.ncbi.nlm.nih.gov/) is the child of the chemoinformatics initiative, developing a "new and comprehensive database" of chemical structures together with their biological activities. The information from the screening centers (together with publicly available information) will be housed in PubChem, which will also feature new algorithms for computational chemistry and virtual screening. Already, both Nature Chemical Biology (http://www.nature.com/nchembio/) and NMRShiftDB (http://www.nmrshiftdb.org) are available through PubChem, which also provides links to Medical Subject Annotations and PubMed biomedical literature citations. The database, growing rapidly, now has 17,000,000 substances and 500 bioassays in its collection. PubChem provides a limited set of structure properties, selected to be relevant for typical drug design applications. Presently, it is possible to do chemical similarity searches based on SMILES.

PubChem is intended to develop new bioassays and perform massive high-throughput screening experiments on a large number of compounds, resulting in a very large public store of biological activity data associated with chemical structures. The structure database will ultimately contain full catalogs of major suppliers of screening compounds, as well as the structures from other public databases (NCI, NIAID, NIST), and it will provide extensive linkouts to original data. Examples for productive queries in the PubChem system can be found at http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pcsubstance. PubChem also claims it is already the largest freely accessible chemical structure store. If its proposed developments are delivered it could become very useful indeed.

Tensions have arisen between the "not-for-profit" American Chemical Society (ACS; http://www.acs.org) and the NIH's PubChem, and also with Google's Google Scholar. The ACS have been concerned about the scale and freely available nature of PubChem, and have also claimed that the literature-search function of Google Scholar infringed upon its own SciFinder Scholar trademark (http://battellemedia.com/archives/001116.php). These disputes can be seen part of a wider, revolutionary change in the publishing climate, due to the rising importance of Open Access (see later).

C. ChemBank

ChemBank (http://chembank.broad.harvard.edu/) is an initiative of Broad Institute Chemical Biology Program (BCB), and sponsored by the National Cancer Institute's Initiative for Chemical Genetics (ICG; http://deainfo.nci.nih.gov/ICG.htm). ChemBank was developed by the informatics group at the Harvard Institute of Chemistry and Cell Biology and utilizes software toolkits supplied by Daylight Chemical Information Systems (http://www.daylight.com).

ChemBank is a freely available collection of data about small molecules and resources for studying their properties, especially their effects on biology. Currently, *ChemBank* stores information on hundreds of thousands of small molecules and hundreds of biomedically relevant assays. The database can be searched by chemical name or activity, by substructure (SMILES string input), or for structure similarity (SMILES string input). Searches can be limited to subsets of available compounds, defined as natural products, known drugs, FDA-approved drugs, commercially available compounds, orally available compounds and primary metabolites amongst other categories. *ChemBank* stores an increasingly varied set of cell measurements derived from, among other biological objects, cell lines treated with small molecules. It is possible to pick an assay, and then view both the details of the screen and/or the data from the assay (http://chembank.broad.harvard.edu/screens/screen_finder.html). There is an additional option to enable viewing of the chemical structure employed in the assay, and even to enable export of spreadsheet files into Microsoft Excel of comma-separated value format.

ChemIDplus (http://chem.sis.nlm.nih.gov/chemidplus) is a search engine which allows retrieval of about 380,000 chemical substance files. The structure-searchable database may include structure (263,000 structures available), official name, systematic name, other names, classification code (therapeutic use), molecular formula, STN locator code, and CAS registry number. Compounds are also searchable by toxicity data and physical properties.

ChemFinder (http://chemfinder.camsoft.com/) is a very large and specific chemical substances search engine, which provides basic information about chemicals such as physical property data and 2D chemical structures. Obvious spelling errors and invalid CAS registry numbers are corrected. Chemicals and pharmaceuticals can be searched by chemical name, CAS registry number, molecular formula or molecular weight. About 75,000 compounds are registered to date.

V. BIOLOGICAL PROPERTIES OF COMPOUNDS

The NCI DIS 3D database (http://dtp.nci.nih.gov/docs/3d_database/dis3d.html) is a collection of 3D structures for over 400,000 compounds which was built and is maintained by the Developmental Therapeutics Program Division of Cancer

Treatment, National Cancer Institute (http://www.nci.nih.gov). While the information stored therein is only a connection table of atomic linkages, it can be interpreted by computer software to provide a 3D structure for each entry. This can then be cross checked against available biological pharmacophores, representing the preferred 3D arrangement for certain biological activities. Drugs that match the pharmacophore could have similar biological activity, but have very different patterns of atomic connections. This approach has been used to search for certain novel protein kinase C (PKC) agonists (http://dtp.nci.nih.gov/docs/3d_database/pharms/pkcsearch.html), using a pharmacophore derived from phorbol. A similar approach has been used to find new ligands for HIV protease, HIV integrase and HIV reverse transcriptase (http://dtp.nci.nih.gov/docs/3d_database/pharms/ncisearches.html.

The Reciprocal Net project, run by the Indiana University Molecular Structure Center (http://www.iumsc.indiana.edu/) is a distributed, open, extensible digital collection of molecular structures together with software tools for visualizing, interacting with, and rendering printable images of the contents, and their automated conversion into standard formats which can be globally shared. The contents of the collection come principally from structures contributed by participating crystallography laboratories, which include universities from the Midwest, the East and West coasts of the US, the UK and Australia. Reciprocal Net's common molecules include a section on therapeutic compounds (http://www.reciprocalnet.org/edumodules/commonmolecules/biochemical/list.html#therapeutic).

One of the most complete and useful resources for CNS pharmacologists is the detailed ligand-receptor chart from http://www.neurotransmitter.net/neurosignaling.html. This table includes over 100 signaling molecules, including a wide variety of neurotransmitters, neuromodulators, neuropeptides, neurosteroids, and neuroactive hormones. The list does not include growth factors, cytokines, and intracellular second messengers. In almost all cases, the substances listed are known to or very likely to affect neurons in the human brain. There are links to gene sequence from the OMIM site (Online Mendelian Inheritance in Man; http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM). This database is a catalog of human genes and genetic disorders authored and edited by Dr. Victor A. McKusick and his colleagues at Johns Hopkins and elsewhere, and developed for the World Wide Web by NCBI, the National Center for Biotechnology Information (http://www.ncbi.nih.gov). There is also a substantial range of reviews referenced for the various neurological agents.

A. Prediction of biochemical properties

FlexX (http://www.biosolveit.de/flexx) is a fast, robust, and highly configurable (FlexX-able) computer program for predicting protein–ligand interactions. Its main application is the prediction of binding. For instance, FlexX predicts the geometry of the protein–ligand complex for a protein with known 3D structure and a small ligand molecule, and estimates the binding affinity. The speed of the calculation permits operation in a virtual high-throughput screening (vHTS) mode: FlexX is capable of screening a database consisting of ~100,000 compounds in about 8 hours on a 30-node cluster. One of the recent features is a new module called PERMUTE (http://www.biosolveit.de/Permute), which protonates molecules and generates tautomers. An evaluation license for FlexX is valid for approximately 6 weeks free of charge and provides access to the full functionality of the software. After the evaluation period the software must be purchased.

AutoDock (http://autodock.scripps.edu/) which is designed to predict how small molecules, such as substrates or drug candidates, bind to a receptor of known 3D structure. AutoDock has applications in X-ray crystallography, lead optimization, structure-based design, combinatorial chemistry, protein–protein docking and chemical mechanism studies

GRID (http://www.moldiscovery.com/soft_grid.php) is a computational procedure for determining energetically favorable binding sites on molecules of known structure. It may be used to study individual molecules such as drugs, molecular arrays such as membranes or crystals, and macromolecules such as proteins, nucleic acids, glycoproteins or polysaccharides. Several different molecules can be processed one after the other.

Lastly in this section, a new improvement to the JME molecular editor (see above, http://www.molinspiration.com/cgi-bin/properties) permits prediction of bioactivity.

B. Molecular datasets

There is a list of free molecular datasets to correlate chemical structure and biological properties, incorporating information on QSAR, QSPR, toxicity, metabolism, permeability, etc. available at http://www.cheminformatics.org/. Downloadable structures are available for 29 out of 31 of the datasets. (The two datasets that are not freely available are restricted due to license reasons, because they are taken from the MDL Drug Data Report (MDDR) database: interested parties can gain access through the MDL web site at http://www.mdl.com/products/knowledge/drug_data_report/.) The Cheminformatics site includes the information in a tabular format, with links to the chemical datasets, in structure-data format, and to the peer-reviewed articles, accessible through a Document Object Identifier ("DOI") linkup. The articles are free to subscribers only – all others must pay a copyright fee.

There is a lot of information here, in a ready-to-use format. For example, blood–brain barrier penetration data is available on a training set of 57 compounds and a data set of 13 more; long-term animal carcinogenicity results are available for over 1,400 compounds, drawn from the Carcinogenic Potency Database (CPDB), an initiative of the Lawrence

Berkeley Laboratory (Berkeley, California); and pharmacological data are available on a wide range of receptors

C. Information on metabolic properties

For metabolism data, there is a superb database in the University of Minnesota Biocatalysis/Biodegradation Database at http://umbbd.msi.umn.edu/. The database includes a search capability for compound, enzyme, or micro-organism name; chemical formula; CAS registry number; or EC (enzyme classification) code. It also has lists of reaction pathways, enzymes, micro-organism entries, and organic functional groups. It specifically includes a large number of reactions of naphthalene 1,2-dioxygenase and of toluene dioxygenase. There is a paper describing the database, which was published in Nucleic Acids Research in January 2001, which can also be downloaded in full text or in .pdf format from the site.

PharmGKB (http://www.pharmgkb.org/) is an integrated resource about how variation in human genes leads to variation in our response to drugs. The database contains genetic and clinical information about people who have participated in research studies at various medical centers. Genomic data, molecular and cellular phenotype data, and clinical phenotype data are accepted from the scientific community at large. These data are then organized and the relationships between genes and drugs are then categorized into clinical outcome, pharmacodynamics and drug responses, pharmacokinetics, molecular and cellular functional assays and genotype. The database itself has been created by Stanford University in a nationwide effort funded by the US National Institutes of Health.

The site refers to the interesting set of tools available at http://www.drug-interactions.com/, which is located in the Indiana University Department of Medicine. This site includes the Cytochrome P450 Drug Interaction Table, a text-based list of drugs which are known to have interactions with cytochrome P450. A quick way to find a specific drug on this page is to use your web browser's Search feature: press Ctrl-F and type all or part of the drug name. The drugs themselves are linked to entries in RxList (http://www.rxlist.com) and to pre-composed search routines on PubMed (http://www.ncbi.nlm.nih.gov). The site is additionally categorized into compounds which are known substrates, inhibitors and inducers of a particular p450 subtype. There is an abbreviated table used for clinical purposes at http://medicine.iupui.edu/flockhart/clinlist.htm. Overall, this site is simply but clearly and excellently put together.

VI. DRUG INFORMATION

A. DrugBank

The University of Alberta supported by Genome Alberta and Genome Canada, has introduced the freely available online resource DrugBank, which contains detailed chemical, pharmaceutical, medical and molecular biological information on more than 3,000 drug targets and 4,100 approved or experimental drugs products (http://redpoll.pharmacy.ualberta.ca/drugbank/index.html).

DrugBank brings the latest data from the Human Genome Project together with detailed chemical information about drugs and drug products. It provides more than 80 data fields for each drug, including brand names, chemical structures, protein and DNA sequences, and links to relevant internet sites, prescription information, and detailed patient information.

The database contains nearly 4,300 drug entries including >1,000 FDA-approved small molecule drugs, 113 FDA-approved biotech (protein/peptide) drugs, 62 nutraceuticals and >3,000 experimental drugs. Additionally, more than 6,000 protein (i.e. drug target) sequences are linked to these drug entries.

Users may query DrugBank through a simple text query, for general text queries of the entire textual component of the database; they may browse for a tabular synopsis of database content, such as for instance for compounds grouped by their indication; and they may draw a structure (using a ChemSketch applet or SMILES string) a chemical compound to search for chemicals similar or identical to the query compound. Finally there is also a facility to conduct BLASTP (protein) sequence searches of the 15,000 sequences contained in DrugBank. Both single and multiple sequence (i.e. whole proteome) BLAST queries are supported. A relational query search tool allows users to select or search over various combinations of subfields.

While the FDA has a very good searchable web site of approved drugs at http://www.fda.gov/cder/ob/ (and FDA-approved biologics and other biopharmaceutical products are at http://www.biopharma.com) this is not structure searchable and does not contain information on compounds in development. More complete database products, like PJB Publications' Pharmaprojects® (http://www.pjbpubs.co.uk), Prous' Ensemble (http://www.dailydrugnews.com) and Current Drugs' IDdb (http://www.current-drugs.com) are only available for a substantial price.

There are multiple other sources of information on marketed compounds, similar to that which is conventionally available in pharmacopoeias, indeed the names of these sites often reflects that connection. The Internet Drug Index (http://www.rxlist.com) is a prescription drug database, which provides good basic information about products on the market, searchable by keyword, brand, or interaction. RxList is a trove of pharmaceutical knowledge with more than 4,500 medications on file, a pharmaceutical discussion board, and an online dictionary of medical jargon. It provides useful basic information about conventional drugs and a handful of herbal remedies as well, in the form of drug FAQs (frequently asked questions) and patient monographs.

Another source is the electronic Medicines Compendium (eMC; http://emc.vhn.net/), with electronic versions of data sheets and Summaries of Product Characteristics (SPCs, sometimes also called SmPCs to differentiate them from Supplementary Patent Certificates) for medicines. It provides the same information as that contained in the latest edition of the Compendium of Data Sheets and SPCs, which covers thousands of medicines licensed in the UK. As an ongoing process, the eMC is also incorporating the SPCs of several thousand other medicines approved by the licensing authorities.

The eMC ultimately aims to provide information on every licensed prescription, pharmacy and general sale medicine in the UK, including generics. As well as SPCs, the eMC will eventually include all Patient Information Leaflets (PILs), and will also be enhanced with dynamic updating and online links to complementary sources of medicines information.

VII. PHYSICAL CHEMICAL INFORMATION

A comprehensive practical guide for determination of physical properties has been produced by Ben Wagner at the State University of New York at Buffalo (http://www.che.utoledo.edu/findmatprop.pdf).

ChemFinder (http://www.chemfinder.com/) is one of the most important sites for physical property information. This site provides the structure, synonyms, CAS registry number, and up to nine physical properties directly for each compound (melting point (m.p.), boiling point (b.p.), refractive index, evaporation rate, flash point, density, vapor density, vapor pressure, and water solubility). Other information such as the physical description and odor detection limits are also given, when available. ChemFinder also acts as a metasearch engine, searching over 350 web sites and displaying direct links to these sites. The links are arranged into several categories including biochemistry, health, MSDS, physical properties, regulations, structures, and usage.

The MatWeb site (http://www.matweb.com/index.asp?ckck=1) is different, since it deals mostly with materials, instead of individual chemical substances. The free sites, while offering significant amounts of data, do not compare with the information available from Beilstein's Crossfire product, which of course is commercially priced, either in terms of number of compounds or in terms of number of properties for each "hit."

Syracuse Research Corporation (SRC) (http://www.syrres.com/esc/physdemo.htm) offers commercial online searches of a number of physical property databases, including online logP measurements (octanol–water partition coefficient), environmental fate for over 25,000 chemicals.

There is a good discussion of the theory and application of various kinds of solubility parameters at http://palimpsest.stanford.edu/byauth/burke/solpar/. There are a belwindering array of solubility parameters, such as Kauri-Butanol number, solubility grade, aromatic character, aniline cloud point, wax number, heptane number, and Hildebrand solubility parameter, among others. The Hildebrand solubility parameter (http://palimpsest.stanford.edu/byauth/burke/solpar/solpar2.html), the most widely applicable of all the systems, includes such variations as the Hildebrand number, hydrogen bonding value, Hansen parameter, and fractional parameter, to name a few. They are directly related to the heat of vaporization.

Information specifically on solvents is to be found at Solv-DB (http://solvdb.ncms.org/solvdb.htm). This site, sponsored by the National Center for Manufacturing Sciences (NCMS), gives information on eight different parameters including solvent name, CAS registry number, molecular formula, and chemical category for over 200 solvents. Nine different properties are range searchable including flash point, vapor pressure, density, and surface tension. Up to 33 more properties can be displayed for each solvent. Results can be sorted by solvent name or any of the nine range-searchable properties. Extensive information is provided for each solvent with display of health, safety, regulatory, and environmental fate data.

The ChemExper Chemical Directory (http://www.chemexper.com/) is also listed below as a resource for searching available chemicals from various supplier catalogs. Links are provided to the supplier's web site and to MSDS. Only the basic properties are directly provided: density, m.p., b.p., and flash point. However, links to the full text of the MSDS will usually provide some additional properties.

The NIST Chemistry webBook (http://webbook.nist.gov/chemistry/) from the National Institute of Standards and Technology (formerly the National Bureau of Standards), lists up to 45 thermochemical, thermophysical, and ion energetics properties which are available for over 40,000 compounds.

Finally, the Organic Compounds Database (http://www.colby.edu/chemistry/cmp/cmp.html), maintained at Colby College, features a database of 2,483 compounds compiled by Harry M. Bell of Virginia Tech. Though only a few common properties are provided, the search screen allows the selection of a wide variety of parameters including property values, element counts, and the presence or absence of certain broad structural entities such as amines or hydroxyl groups. Unfortunately, retrieval sets are limited to twenty compounds, though the search engine does report the total number of hit compounds.

VIII. PREDICTION AND CALCULATION OF MOLECULAR PROPERTIES

Molecular property prediction is becoming a useful tool in the generation of libraries of "beautiful" molecules, or molecules with the correct parameters to be useful drug candidates. Used in a more focused way, drug design and lead optimization benefits from an ability to predict physical properties such as lipophilicity and solubility, as well as physical molecular properties such as polar molecular

surface area (PSA). Methods for prediction of the latter are outlined in an old publication by David Clark at http://www.documentarea.com/qsar/davclark.pdf. PSA, along with others such as charge, polarizability, molecular surface area and numbers of H-bond donor–acceptors are predicted by plugins for ChemAxon (http://www.chemaxon.com/marvin/chemaxon/marvin/help/calculator-plugins.html). Alternatively, a program for the calculation of PSA directly from SMILES input, which claims to be 2–3 orders of magnitude faster than other methods, called tpsa.c is available from http://www.daylight.com/meetings/emug00/Ertl/.

A nice way of visualizing five drug-like properties in a single graph is the PK RADAR presentation (http://www.rsc-modelling.org/CEAtoDD/RitchieRadar.ppt). Alternatively, Vlaavis is a free visualization tool in which each circle represents a single compound and each slice of the "pie" represents a normalized response to a particular assay or property; originally a structure-activity relationship (SAR) tool, it is available from http://www.vlaaivis.com/index.htm.

In addition to experimental information, SRC (referred to above) have also developed software to predict physical properties, such as the Estimation Program Interface (EOI) Suite (http://www.epa.gov/oppt/exposure/docs/episuite.htm) which was developed for the US EPA (Environmental Protection Agency). By entering a single SMILES notation as the search key, results from 10 separate programs are displayed. These are shown in Table 13.3. The program contains a SMILES notation database searchable by CAS registry numbers. By entering a registry number, the SMILES notation is automatically retrieved and entered into the search box.

Another useful sites in this regard is ChemExper, which (see below) in addition to resources for searching available chemicals and their physical properties, hosts the OSIRIS Property Explorer (http://www.chemexper.com/tools/propertyExplorer/main.html) for calculation/prediction of a compound's physical parameters. OSIRIS calculates various drug-relevant properties using a user-drawn structure. Prediction results are valued and color coded. Properties with high risks of undesired effects like mutagenicity or a poor intestinal absorption are shown in red, whereas a green color indicates likelihood of conforming to drug-like behavior. As the user is building the molecule, the cLogP and solubility characteristics are being calculated. The kinds of toxicological and safety issues that are predicted include mutagenicity, tumorigenicity, reproductive effects and irritancy. The algorithms used to calculate these properties are described in some detail – for instance the toxicity risk assessment is explained at http://www.chemexper.com/tools/propertyExplorer/tox.html. A substructure search process determines the occurrence frequency of any fragment (core and constructed fragments) within all compounds of that toxicity class. Similar explanations follow the fragment-based drug-likeness score (http://www.chemexper.com/tools/propertyExplorer/druglikeness.html) and the overall drug-likeness score (http://www.chemexper.com/tools/propertyExplorer/drugScore.html). The OSIRIS Property Explorer is an integral part of Actelion's (http://www.actelion.com) in-house substance registration system.

The prediction process relies on a pre-computed set of structural fragments that give rise to toxicity alerts in case they are encountered in the structure currently drawn. These fragment lists and toxicities (e.g. mutagenicity) were drawn from the RTECS database. RTECS, the Registry of Toxic Effects of Chemical Substances, aims to "list … all known toxic substances … and the concentrations at which … toxicity is known to occur;" currently there are over 133,000 such substance listed at http://www.ntis.gov/products/types/databases/rtecs.asp?loc=4-4-3.

The Interactive Laboratory (I-Lab: http://www.acdlabs.com/ilab) is a commercial product (but with a free demonstration version) available from Advanced Chemistry Development (ACD). It provides online computation of molecular physical properties for LogP, pKa, LogD, and aqueous solubility. I-Lab also includes database searching of ACD's compilations of Spectra and Physical Properties. The ACD/logP calculator (http://www.acdlabs.com/download/logp.html), now offered as freeware has been compared with competitive products at: http://www.acdlabs.com/products/phys_chem_lab/logp/competit.html. It is claimed to calculate an accurate logP derived from an internal ACD/LogP database containing over 5,000 experimental LogP values.

An interactive web service for the calculation of molecular properties relevant to drug design and QSAR has been established at the Molinspiration Cheminformatics web site http://www.molinspiration.com/cgi-bin/properties. Properties calculated include logP, PSA, and Lipinski Rule of five parameters. Soon also a drug-likeness index will be available. Molinspiration is offering this as a free service to internet chemistry community for up to 100 determinations per month. In addition to single calculations, the site also offers the possibility to search in the web molecular database by substructure, structure similarity or pharmacophore similarity (http://www.molinspiration.com/cgi-bin/search).

TABLE 13.3

Aquatic toxicity (LD50, LC50)	Henry's law constant
Aqueous hydrolysis rates	m.p., b.p., and vapor pressure
Atmospheric oxidation rates	Octanol–water partition coefficient
Bioconcentration factor (BCF)	Soil sorption coefficient (Koc)
Biodegradation probability	Water solubility

Source: Properties available from the EOI Suite *at http://www.epa.gov/oppt/exposure/docs/episuite.htm.*

Previously, the calculation of logP could be performed from SMILES strings at the Daylight site (http://www.daylight.com/daycgi/clogp), however, there was no facility for calculation based on molecular structure.

IX. CHEMICAL SUPPLIERS

There is currently a very large amount of information on available chemicals on the web. This information is relevant for both laboratory-scale synthesis and for larger scale preparations, however, it is more easily searched for laboratory synthesis. A useful reference is to the list of the web sites of online searchable chemicals and suppliers; for example http://www.mdpi.org/forum.htm#chemicals offers a range of options, which are also accessible through its European mirror site at http://www.unibas.ch/mdpi/forum.htm#chemicals

Examples of sites for online searching of available chemicals are provided in Table 13.4.

The site http://www.chemexper.com also allows access to Expereact™ WEB, a laboratory management program that helps to keep stock control, order products, add reactions (electronic laboratory journal), export all the information to a word processor, etc. Another site providing software for inventory management is ChemSW, at http://www.chemsw.com. Products include the CIS Inventory system Pro 2000, and a digital "MSDS digital filing cabinet" (very useful for managing data sheets as they go out of date), as well as more conventional chemical drawing and molecular modeling programs.

Many suppliers offer database searching capabilities themselves. Large companies such as Sigma-Aldrich have managed to offer a complete searchable database of their products, by name, structure and CAS number (http://www.sigma-aldrich.com). They also feature online ordering via a secure interface. The smaller suppliers have been later in arriving at an online database with searching and secure ordering.

There are some commercial databases products in addition to the sources listed above, such as ChemSources International (http://www.chemsources.com/csintl.htm), which includes the products of more than 8,000 chemical companies worldwide. The Chemical Section lists approximately 275,000 chemical compounds and provides contact data necessary for making direct inquiries to each chemical firm. The product is available at a cost of approximately $1,100 for a single user CD-ROM license.

TABLE 13.4

Web site address	Comments
http://www.buyersguidechem.de/	Excellent site with wide variety of chemicals, no prices; useful for bulk and for MSDS; a directory of over 100,000 chemicals.
http://www.chemexper.be/	Excellent search capability on a wide variety of research chemicals, and information that includes the exact chemical name as well as formula, melting point and other physical properties. Searching can be conducted by CAS number, molecular formula, substructure, name and a range of other terms. Hot links allow the user to directly go to the individual supplier.
http://www.molmall.org	MolMall features the Rare Chemical Samples ExchangeCenter. Compounds are made available from small samples provided by individual researchers. Full structure search or substructure searches are permitted on the web site, as well as for the name of the submitter and several other very useful searching functions. Links to Molecules MolBank (http://www.molbank.org) papers if the compounds are published there. There are plans to allow the sample submitters to add additional information to the data sheet, such as the literature where the compound was published.
http://www.icis.com/Search/default.aspx	Fairly wide variety of chemicals but no prices.
http://www.chemindustry.com/apps/chemicals	ChemIndustry.com site enables the user to enter a product name and then search a database of web sites related to various chemical suppliers.
http://www.chem-edata.com	Text and CAS number search capability. Fairly limited selection.
http://www.chemacx.com	A commercial product through CambridgeSoft, Available Chemicals Xchange features the complete catalogs of over 200 vendors.
http://www.ubichem.com	Ubichem is an independent British company which was founded in 1978 with the aim of supplying a wide range of fine chemicals and intermediates. Its literature review section at http://www.ubichem.com/lit.php has reports on a variety of chemical topics, including isoquinolines, palladium coupling, azides, and radiolabeling.

CHEMCATS is an online database accessible through Chemical Abstracts that contains over 13 million commercially available compounds, including pricing information when available from suppliers and for many also direct hyperlinks to suppliers' sites. CHEMCATS is routinely updated with new information provided by suppliers already in the database and with new suppliers and/or catalogs. This is another commercial product, but there is no up front fee: pricing is based on pay as you go and can be accessed through STN Easy (http://stneasy.cas.org/).

X. CHEMICAL SYNTHESIS

WebReactions (http://www.WebReactions.net) is a new, unique reaction search system offering direct retrieval of reaction precedents through the internet. The WebReactions system is easy to learn and use, the user merely draws the reactant and product using a Java-based chemical drawing program. It is virtually instantaneous in displaying matches, not just for the input reaction itself, but for as broad a range of analogs as desired.

The complete Organic Synthesis (OS) is now available free online at http://www.orgsyn.org/. Exact and substructure searches are supported following download of a free ChemDraw plugin as well as chemical name, formula, OS reference, keyword index searches. This site is available free of charge to all chemists and contains all of the ten Collective as well as Annual Volumes and Indices. Organic Syntheses (OS) is a compilation of 84 annual volumes containing selected and independently checked procedures and new reactions in the field of organic synthesis. Since the 1920s, volumes of OS consisting of synthetic procedures have been published annually. The first 6 Collective Volumes were published every 10 years, and the last 3 at 5-year intervals.

Two other sites related to chemical synthesis include http://orgchem.chem.uconn.edu/namereact/named.html which includes details of about 100 named reactions; and the reaction index at http://www.pmf.ukim.edu.mk/PMF/Chemistry/reactions/rindex.htm, which contains a very extensive list of named reactions in organic chemistry.

For biotechnological synthesis, there is a superb database containing information on microbial biocatalytic reactions and biodegradation pathways for primarily xenobiotic, chemical compounds. It is called the University of Minnesota Biocatalysis/Biodegradation Database (UM-BBD) at can be found at http://umbbd.ahc.umn.edu/search/index.html. The goal of the UM-BBD is to provide information on microbial enzyme-catalyzed reactions that are important for biotechnology.

There are a number of blogs oriented toward chemical synthesis (Table 13.5).

TABLE 13.5

Title	URL	Comments
Useful Chemistry	http://usefulchem.blogspot.com	A useful aggregative site that includes articles from a range of writers, specifically on various aspects of synthetic chemistry.
Carbon Tet	http://carbontet.blogspot.com/	Medicinal chemist with a focus on synthetic chemistry.
Chemical Musings	http://chemicalmusings.wordpress.com/	Thoughts on organic chemistry and my experiences as a new industrial chemist.
Heterocyclic chemistry	http://hetchem.blogspot.com/	Monthly articles on heteroscyclic syntheses.
Org. prep. daily	http://orgprepdaily.wordpress.com/	Procedures for various simple reactions.
Organic Chemistry Highlights	http://www.organic-chemistry.org	Stereoselective synthesis of natural products, new methods in synthetic organic chemistry, and computational organometallic chemistry in organic synthesis; 5–8 highlighted reactions per month, and short reviews of organic, bioorganic, organometallic and microwave chemistry, total synthesis of natural products and multi-component reactions.
She Blinded Me with Science	http://blind-science.blogspot.com/	Synthetic chemistry from a chemical biologist's point of view.
Totally Retrosynthetic	http://totallyretrosynthetic.blogspot.com/	Author has interest in synthesis of natural products.
Totally Synthetic	http://www.totallysynthetic.com/blog/index.php	Author is a synthetic chemist with Prof S Ley's group in Cambridge, England.

XI. CHEMICAL SOFTWARE PROGRAMS

A. Chemical drawing and viewing software

There are now a number of 2D and 3D molecule viewers available for free download, which work to make chemical structures visible on web pages. A summary of some of the available software is also held at http://www.indiana.edu/~cheminfo/mvts.html (Table 13.6).

B. Various chemoinformatics software

There are a few software programs capable of 3D structure, conformation generation, computer-aided drug design, and/or

TABLE 13.6

Viewer	Description	URL
Chime	The Chime plug-in displays 2D and 3D molecules directly within a web page and works with both Mozilla and Microsoft browsers. The molecules in the web page are "live," meaning they are not just pictures, but chemical structures that can be rotated, reformatted, and saved in various file formats for use in modeling or database software.	http://www.mdl.com/products/framework/chime/key_features.jsp
JMol	JMol, initiated by Dan Gezelter at Columbia University is JMol is a free, open source molecule viewer and editor. It is a collaboratively developed visualization and measurement tool for chemical scientists.	http://jmol.sourceforge.net/
ChemDraw plug-in	This is claimed as being more than a mere structure viewer or a slow Java applet, rather it runs as fast and is as familiar as the regular ChemDraw application. Available without charge, it enables searching of web databases by structure or substructure, and viewing of ChemDraw documents that others have placed on the web.	http://www.camsoft.com
JME molecular editor	JME is a free Java applet which allows generation and editing of molecules and reactions, and creation of molecule SMILES. The JME applet, written by Peter Ertl from Novartis, has become a standard for molecule structure input on the internet.	http://www.molinspiration.com/jme/index.html
MarvinSketch and MarvinView	A set of Java-based chemistry software including MarvinSketch, an applet for editing and visualizing molecules on a web page; MarvinView, an applet for viewing molecules in 2D or 3D on a web page; and MolConvertor, a command line program that converts between various file types.	http://www.chemaxon.com/product/live_examples.html
WebMolecules	Web visualization of molecules in 3D – in real-time – has now been achieved. Over 150,000 molecular models are available onsite, which may be searched by CAS number or exact formula. In addition partial formula searching is permitted to look into the Top 2,000 molecules, which includes molecules of commercial value, educational importance, and of topical interest. Thousands of common molecules are organized into over 30 categories.	http://www.webmolecules.com
Waltz and CSD	ChemViz (short for chemistry visualization) is a set of web-based applications created by NCSA designed to catalyze a better understanding of chemical processes through visualization. Two tools are currently supported: Waltz, which generates images and animations of desired molecules and ions; and CSD, which presents a 3D model of complex organic compounds.	http://chemviz.ncsa.uiuc.edu/
Depict	This service accepts a SMILES string as input and returns an HTML page with an embedded image. Unfortunately, there is no control on the output style and image size.	http://www.daylight.com/daycgi/depict
ACD/Structure Drawing Applet	A complete structure drawing, editing, and visualization tool written in pure Java that can be incorporated into HTML documents. The applet can be used for composing substructure queries to databases and visualizing results.	http://www.acdlabs.com/products/java/sda/
OpenBabel	A cross-platform program and library designed to interconvert between many file formats used in chemical drawing and molecular modeling.	http://sourceforge.net/mailarchive/forum.php?thread_id=8125020&forum_id=3042

TABLE 13.7

Program name	URL	Description
CORINA	http://www2.chemie.uni-erlangen.de/services/telespec/corina/	3D Structure Generator – 1,000 structures can be generated for free.
FANTOM (Fast Newton-Raphson Torsion Angle Minimizer)	http://bose.utmb.edu/fantom/fm_home.html	Calculation of low-energy conformations of polypeptides and proteins, compatible with distance and dihedral angle constraints obtained typically from nuclear magnetic resonance (NMR) experiments.
MMFF94	http://ccl.net/cca/data/MMFF94/	Set of validation molecules based on X-ray crystallographic data.
Moloc	http://www.moloc.ch	Molecular modeling package.
NEWLEAD	http://www.ccl.net/cca/software/SGI/newlead/README.shtml	Computer program for the automatic generation of candidate structures.
PADRE	ftp://ftp.CCL.net/pub/chemistry/software/UNIX/PADRE/	Analysis of the results of conformational searches, and measurement of similarity and differences between molecules.
Pgchem::tigress	http://pgfoundry.org/projects/pgchem/	Chemoinformatics extension to the PostgreSQL database management system, that enables PostgreSQL to handle molecules through SQL statements.
PyMOL	http://pymol.sourceforge.net/	Open source molecular visualization system.
RasMol	http://www.umass.edu/microbio/rasmol/index2.htm	Molecular visualization software.

molecular modeling are listed available on free licenses, at least for academic purposes (Table 13.7).

A more extensive list of software generally available for pharmaceutical and biotechnological R&D is available from NetSci, a public information exchange (http://www.netsci.org/Resources/Software/Cheminfo/index.html). This list includes chemical databases, reaction databases, QSAR, and other programs. The modeling section of NETSci is to be found at http://www.netsci.org/Resources/Software/Modeling/CADD, and includes both open license and commercial software.

Most of the references on this section of the NetSci site are to software programs which are not free, even to academic licensees. Their exclusion from explicit mention from this review is not intended to imply any value judgment on their worth. Interested readers are encouraged to make their own enquiries if they wish to review the available offerings.

C. Datasets for virtual screening

Information from published literature, particularly from chemical catalogs has been used to create virtual libraries for drug screening. (Incidentally, virtual screening has also been used in a co-ordinated fashion together with high-throughput screening, rather than in competition with it – see http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=12471601&dopt=Abstract.)

For example, ZINC is a free database of small molecules for docking that are commercially available (http://blaster.docking.org/zinc/). ZINC is a self-referential acronym for "ZINC is not commercial," and contains over 3.3 million compounds in ready-to-dock, 3D formats. The downloads are available in sdf, mol2, and SMILES formats. Subsets of the libraries are available, and can be browsed (http://blaster.docking.org/zinc/bysubset.shtml). There is a subset of "drug-like" molecules assembled by searching the database according to the Lipinski rules of 5 (logP < 5; mol wt < 500; number of H-bond donors < 5; number of H-bond acceptors < 10). It is also possible to create a bespoke subset by searching according to physical properties including structure, net charge, calculated logP, rotatable bonds, number of H-donors, number of H-acceptors, polar desolvation, apolar desolvation, and molecular weight (http://blaster.docking.org/zinc/choose.shtml).

Another resource is within the similarity searching section of the Chemoinformatics site run by Andreas Bender at the University of Cambridge (http://cheminformaticcs.org/simsearch) using the MOLPRINT 2D approach. Users can compare a test library and a reference molecule in hydrogen-depleted mol2 formats. (These can be generated using the CONVERT or OpenBabel programs referenced earlier.) Sample structures for the reference molecule are available for the 5-HT3 antagonist and angiotensin-converting enzyme (ACE) inhibitor pharmacophores at http://cheminformatics.org/simsearch/testcompounds.shtml.

Comparison of fingerprints is performed by using the Tanimoto coefficient, Tc, which is defined by the number of common features of the two structures (AND), divided by the number of features which are contained in at least one of the structures (OR).

Another place where one can submit SMILES or molfiles and have a similarity score computed is described at http://www.jchem.com/doc/user/Compr.html (library comparison based on similarity) and at http://www.jchem.com/doc/user/Jcsearch.html (for all kinds of searching including the Tanimoto-based similarity method). Both of these methods are included in the JChem software (http://www.chemaxon.com/products.html), which is free for academic use.

XII. ANALYSIS

Analytik (http://www.analytik.de/) is a very comprehensive German information site for analytical chemists. It relates discussions of analytical problems, contains a small but excellent link collection to chemical databases and literature (with an application database) and press releases from the German Chemical Society, etc.

Spectra Online (http://spectra.galactic.com/) provides about thousands of IR, MS, NMR, UV/VIS, and NIR spectra which can be consulted free on this site. Retrieval of information is possible by entering the compound name, CAS registry number, molecular formula, etc. Requires registration (free).

Macherey-Nagel (http://www.macherey-nagel.com) provides GC, HPLC, SFC, SPE, and TLC-methods with details and literature references listed on this site. A classification is made according to categories of products (e.g. HPLC analysis of amides, amines, etc.).

A new release of version 1.0 of an open access, open submission, open source NMR database NMRShiftDB has been announced at (http://nmrshiftdb.ice.mpg.de/). The software and database content can be downloaded via http://www.sourceforge.net/projects/nmrshiftdb. NMRShiftDB is a web database for organic structures and their NMR spectra. It allows for spectrum prediction (currently only for carbon) as well as for searching spectra, structures, and other properties. Currently, the database contains over 20,000 structures and over 23,000 measured spectra (as well as about 500 calculated spectra).

The Java applet that comes with NMRShiftDB includes an array of features for molecular display (such as ball-and-stick, wireframe, space-fill, etc.), translation into SMILES nomenclature, and possibility of structure editing. Searching of chemicals can be based on (amongst others) chemical name, keyword, CAS number, literature title/author, and chemical formula.

XIII. CHEMICAL PUBLICATIONS

A. Journals

Nearly all journals have a web presence, and an increasing majority have electronic versions of their publications (including archives) available through the web site. A convenient listing of them is available in the chemistry section of the WWW virtual library at http://www.liv.ac.uk/Chemistry/Links/journals.html. Salient journals related to medicinal chemistry include those shown in Table 13.8.

A piece of software to extract data from literature is the Experimental Data Checker and OSCAR toolkit available

TABLE 13.8

Publisher	Journal title	URL
American Chemical Society	Bioconjugate Chemistry	http://pubs.acs.org/journals/bcches
American Chemical Society	Journal of Natural Products	http://pubs.acs.org/journals/jnprdf
American Chemical Society	Journal of Pharmaceutical Sciences	http://pubs.acs.org/journals/jpmsae
American Chemical Society	Journal of Medicinal Chemistry	http://pubs.acs.org/journals/jmcmar/
American Chemical Society	Modern Drug Discovery	http://pubs.acs.org/journals/mdd/index.html
American Chemical Society	Organic Process Research & Development	http://pubs.acs.org/journals/oprdfk
Bentham Scientific Publishers	Current Medicinal Chemistry	http://www.bentham.org/cmc/
Bentham Scientific Publishers	Current Pharmaceutical Design	http://www.bentham.org/cpd/
Bentham Scientific Publishers	Current Drug Discovery Technologies	http://www.bentham.org/cddt/index.htm
Current Drugs	Current Opinion in Drug Discovery and Development	http://scientific.thomson.com/products/coddd/
Elsevier	Bioorganic & Medicinal Chemistry	http://www.elsevier.com/inca/publications/store/1/2/9/
Elsevier	Bioorganic & Medicinal Chemistry Letters	http://www.elsevier.com/inca/publications/store/9/7/2/
Elsevier	Drug Discovery Today	http://www.drugdiscoverytoday.com
Elsevier	European Journal of Medicinal Chemistry	http://www.elsevier.com/wps/find/journaldescription.cws_home/505813/description

from http://www.rsc.org/Publishing/ReSourCe/AuthorGuidelines/AuthoringTools/ExperimentalDataChecker/index.asp. Experimental data on new molecules in organic and inorganic chemistry is presented in a standard form which varies little from journal to journal. Typically, the appearance of the compound is described, followed by melting points (if applicable), R_f, infrared and NMR data, and mass spectral information. OSCAR will extract this information from either a paragraph of experimental data, or a full paper, and then run some checks to test the data for consistency. After the user pasting the experimental data into an HTML form, the program returns both the analytical information and a critical assessment of the same; it can also plot the 1H NMR spectrum from the analyzed information.

B. Open Access

The issue of Open Access is a huge one for scientific publishing. Advocates want to move from traditional subscription-based journals to a model that would make all research findings accessible to anyone with a computer. There are however a number of problems with the open-access approach.

1. *Availability of information*

A fundamental issue for open-access journals is the quality of the science. Resources for open-access publications include:

BioMed Central (http://www.biomedcentral.com) a for-profit open-access publisher, with a diverse group of peer-reviewed journals.

The Directory of Open Access Journals, DOAJ (http://www.doaj.org/), a clearinghouse for free, full text, quality-controlled scientific, and scholarly journals.

Public Library of Science (PLoS) (http://www.plos.org), a non-profit organization of scientists and physicians committed to making the world's scientific and medical literature a public resource. One of their major initiatives is PLoS Biology, (http://biology.plosjournals.org/perlserv/?request=index-html) a peer-reviewed, open-access journal published by PLoS.

SPARC, the Scholarly Publishing and Academic Resource Coalition, (http://www.arl.org/sparc/), an alliance of academic and research libraries and organizations working to correct market dysfunctions in the scholarly publishing system. SPARC is a partner of PLoS.

Various other sites have been set up to discuss the issues related to open access such as SciDev.Net (http://www.scidev.net) and in blog-form at http://www.earlham.edu/~peters/fos/fosblog.html.

2.. *Attitude of funders of science*

The Welcome Trust has announced it will require results from research funded by the Trust to be available in public repositories 6 months after publication, partly on the basis that an author-pays business model has the opportunity for a saving of 30% on publishing costs alone compared to reader-pays (http://science.slashdot.org/article.pl?sid=05/03/20/2043237). Similarly, NIH have published guidelines to ensure that publicly funded research is widely available (http://publicaccess.nih.gov/), and encourages investigators to make NIH-funded peer-reviewed manuscripts available to other researchers and the public through the NIH National Library of Medicine's (NLM) PubMed Central (PMC) (http://www.pubmedcentral.gov/) immediately after the final date of journal publication.

At present, all papers appearing in Royal Society (http://www.royalsoc.ac.uk) journals can be accessed free of charge 12 months after their publication, but the Royal Society has expressed concern that the proponents of open access should not aim unilaterally for an environment inimical to for-profit scientific publishing. The Society's publications include Biology Letters, Philosophical Transactions Parts A and B, and Proceedings of the Royal Society Parts A and B (http://www.royalsoc.ac.uk/page.asp?id=2462).

The Directory of Open Access Journals (http://www.doaj.org) includes a full list; in the area of chemistry, there are currently 40 entries (http://www.doaj.org/ljbs?cpid=61). In addition to journals, a significant proportion of open-access information is available through self-archiving in institutional archives.

3. *Impact factors?*

Currently, despite significant growth, only about 20% of the number of papers published annually are open access. Whether open access increases impact is still debatable (http://opcit.eprints.org/oacitation-biblio.html). Proponents argue that open access increases the number of citations; dissenters argue that this is so only for prestigious authors who publish in prestigious journals and whose article is already highly cited.

C. Theses

Many universities are installing searchable and accessible thesis archives and (at least theoretically) this is a welcome addition to the web-searchable pantheon of scientific literature. The practical difficulty associated with this task is the sheer diversity of information sources, which are not all archived in a central location. Lists such as these must be assembled by hand; not only are they vast but they are also constantly changing. A partial solution to this problem is addressed through the Networked Digital Library of Theses and Dissertations (NDLTD) (http://www.ndltd.org). Since its inception in 1996, over a hundred universities have joined the initiative, each of which has a process in place for archiving and distribution of electronic theses and dissertations (ETDs). The Union Catalog Project (http://rocky.dlib.vt.edu/%7Eetdunion/cgi-bin/OCLCUnion/UI/index.pl) is an attempt to make these individual collections appear as one seamless digital library of ETDs to students and researchers

seeking out theses and dissertations. ETDs are owned and maintained by the institutions at which they were produced or archived, while all the metadata (title, author, etc.) have been gathered into a central search engine.

MIT's (Massachusetts Institute of Technology) Libraries' Document Services department at http://dspace.mit.edu/handle/1721.1/7582, one of the foremost institutions in this effort, offer the full text of selected master's and doctoral theses from all MIT departments. These include theses that have been previously requested and scanned by Document Services as well as theses from the university's pilot project in electronically submitted theses. Users can search the database by keyword, perform an advanced search with separate fields, or browse by author or year. All theses can be viewed as low-resolution (100 dpi) greyscale inline .gif images. The theses of some of their Nobel Prize winning alumni are available at http://libraries.mit.edu/docs/nobeltheses.html.

XIV. PATENT INFORMATION

The major world patent databases are online and searchable, there is a plethora of tools available for the desk scientific researcher.

Esp@cenet at the European Patent Office (EPO) Databases (http://www.european-patent-office.org/index.htm) allows free online patent searching in over 30 million documents (in EPO member states and worldwide) by entering keywords, patent numbers, institute names, etc.

The US Patent and Trademark Office (USPTO) web patent database (http://www.uspto.gov/patft/) provides access to both the US Patent Bibliographic Database, which includes bibliographic data from 1976 to the present and the AIDS Patent Database, which includes the full text and images of AIDS related patents issued by the US, European, and Japanese Patent offices. There is a patent number search page as well as Boolean and Advanced search pages for text field searching. Both cited and citing patents are hyperlinked to each patent. There are hyperlinks between the classification numbers and their definitions and good help pages for each search type.

However, recent information suggests that, surprisingly, not all of the databases are complete. According to Univentio, a Dutch patent information company (http://www.univentio.com/), Espacenet (http://es.espacenet.com) is missing hundreds of thousands of patent documents from various European countries including the UK, France, and Germany. The problems may be caused by a variety of means such as errors in digitization, or because the original paper copies were not available. The loss of paper archives is a growing problem, with lack of shelf space compromising libraries' ability to house complete collections. This in turn represents a significant problem for electronic patent office searches, which are the default method for a patent examiner being able to identify prior art for a new application.

Some national offices such as the UK Patent Office (http://www.patent.gov.uk/) have plans to extend the electronic archive with scanned images of British patents; these will need to be processed for optical character recognition, in order to make the text within searchable: without such processing, the images are not identified by searching methods.

Although USPTO (http://www.uspto.gov) provide images of the actual hard copy, the user currently has to combine singly downloaded .tiff or .pdf files in order to generate a single-file document. This tedious process has been obviated by commercial patent engines such as MicroPatent (http://www.micropatent.com) and Delphion (http://www.delphion.com), but now there are free alternatives to address this issue. FreePatentsOnline (http://www.freepatentsonline.com) is, as the name suggests, a freely accessible database that contains all patents published by the USPTO since number 4,000,000. It is automatically updated weekly, is searchable and can retrieve images of the results from the patent text pages. The search methods are similar to those available at the USPTO site. Search terms can be entered in certain fields, such as Title, Abstract, Assignee (Owner), etc., to locate patents or published patent applications having the entered terms in the specified fields (in the specified sections of the patents or applications). Search strings can also be connected with Boolean terms such as AND, OR, and ANDNOT, and parentheses can be used to order the connected terms. The ends of search terms can be truncated and the "wildcard" symbol $ (note that a term may not be truncated to less than four characters). The most complete search method is based on searching the "specification" field. To identify US classes for particular fields of technology, users can access the Manual of US Patent Classification at http://www.uspto.gov/web/classification/.

Using the concept of the wiki, the WikiPatents Community contributes to the US patent system by reviewing issued patents and, soon, pending patent applications (http://www.wikipatents.com). The public can add prior art references for a given patent, vote on the relevancy of both original and user-added references, and make comments about how the prior art is related to a patent. It has an area for discussing prior art searching, patent litigation, law changes, and reviews on certain patents. There is an issue resulting from the prior art analysis that may be posted on this site with regard to US patent prosecution. Under US Law 37 C.F.R. 1.56, those materially involved in the preparation and prosecution of a patent application have a duty of candor with the US Patent and Trademark Office (USPTO; http://www.uspto.gov), which requires the submission of relevant references and other information to a patent Examiner during patent prosecution. Information posted on WikiPatents.com is potentially part of that requirement. The review sections also include assessments of potential commercial applicability and indeed of value of the application. This is extended into a section where patents can be offered for licensing or sale (http://www.wikipatents.com/marketplace.php). This is presently in an inchoate state, with featured examples mostly of designs rather than chemical

or biological patents, but the pace of change in community web sites (such as Wikipedia) can be alarmingly rapid.

A repository of general interest in patent literature and intellectual property-related news and decisions, is available through IP News Flash at http://www.ipnewsflash.com. It is updated hourly with up-to-date information on patents and other intellectual property-related matters. The site offers an e-mail news feed with content delivered daily or monthly, as well.

There are various other sites providing general information on patents, such as that from the USPTO about patents and patenting procedures (http://www.uspto.gov/web/offices/pac/doc/general/). Various other useful bits of information about patent terms and procedures in other countries are available from Derwent (http://www.derwent.com/). A comparison table of web patent databases from Duke University is presented by the university library to help users compare the various resources available and assess which is best for each individual's needs (http://www.lib.duke.edu/chem/patcomp.htm and at http://www.lib.duke.edu/reference/subjects/patents.htm).

A. Japanese patents

The Japanese Patent Office's (JPO) web site (http://www.jpo.go.jp/) now provides certain information free in English. It provides more information in Japanese including free legal status information from the JPO's intellectual property digital library (IPDL) pages. There are five methods of searching the IPDL patent database (http://www.ipdl.inpit.go.jp/homepg_e.ipdl). The form for retrieval of patent images based on patent number is somewhat difficult to navigate (http://www4.ipdl.inpit.go.jp/Tokujitu/tjsogodben.ipdl?N0000=115), but is backed up by a useful help area at http://www.ipdl.inpit.go.jp/HELP/tokujitu/db_en/help_index.html.

There are further resources available for English translations of Japanese patent documents. Paterra, Inc., (http://www.paterra.com) is pleased to present the InstantMT™ service for Japanese patents on the internet. The InstantMT™ service retrieves the requested patent by number and rapidly provides a translated version which is rendered for download in a two-column formatted Acrobat PDF file. The system covers all Japanese Kokai (A documents) published after January 1, 1993 and all granted Japanese patents (Toroku) published since May 27, 1996. New documents are entered into the system within 2 weeks of being published in Japan. In a related development, Protys (http://www.protys.info/) provides a full text English database of the latest Japanese patents in a specialty current awareness database.

Paterra have prepared a guide (http://cxp.paterra.com/FTerms/Guide.htm) to compare the Japanese F-term system with the International Patent Classification (IFC) (http://www.wipo.int/classifications/fulltext/new_ipc/) and File Index classification system. The guide is a browseable front-end to the concordance provided by the Japan Patent Office. Users can scroll through IPC/FI classes and view precisely how they are mapped onto the F-Term themes. Each entry is linked to the corresponding F-Term table on the JPO site so the user can move directly from an IPC/FI classification to the corresponding F-Term table.

XV. TOXICOLOGY

There is a good introduction to the subject, which should be interesting to students, called "Toxicology Tutor" at http://sis.nlm.nih.gov/toxtutor.htm.

There are a number of toxicology databases available on the internet, and recently there has been an amalgamation of the best in the form of TOXNET (http://toxnet.nlm.nih.gov/), a cluster of databases on toxicology, hazardous chemicals, and related areas. The web site provides access to an impressive array of files containing factual information related to the toxicity and other hazards of chemicals. Users can readily extract toxicology data and literature references, as well as toxic release information, on particular chemicals. Alternately, one can perform a search with subject terms to identify chemicals that cause certain effects. A variety of display and sorting options are available.

A summary of further resources in this area (including subsets of the TOXNET database array) is provided in Table 13.9.

PubMed now links to chemicals found in TOXNET's HSDB http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?HSDB through a LinkOut feature which appears when a user clicks the "Links" part of any PubMed reference, shown on the far right hand side of the screen. The links now appear as specific chemical names. LinkOut provides PubMed users with connections to full-text articles, consumer health information, and supplementary data, related to a PubMed citation.

Another useful resource related to toxicology is sample data for LD50 i.v. (mice) for a short list of compounds at http://toxicity.molecularsociety.org/LD50.htm. The source of these data is the Molecular Society, an "international forum for the collection and exchange of information for [those] who are actively involved in the multi-disciplinary field of Molecular Sciences."

NCMS has a rather interesting free and fairly extensive database of solvents that allows rather extensive input of physical property ranges and various user-specified limits such as "not a carcinogen" or "not listed on the Montreal protocol (ozone)" (http://solvdb.ncms.org/solvdb.htm).

Fee-based resources include the updated US EPA Toxic Substance Control Act (TSCA) Chemical Inventory of 62,000 chemicals, which is available cross-referenced with Superfund Amendments and Reauthorization Act (SARA) Title III RCRA reporting requirements on CD-ROM. It features SARA III fields integrated with TSCA information and Adobe™ Acrobat™ (PDF) format for instant

TABLE 13.9

Type of database	Database name	URL	Description
Toxicology Data Files	HSDB (Hazardous Substances Data Bank)	http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?HSDB	Factual data bank of over 4,500 potentially hazardous chemicals. In addition to toxicity data, the file covers emergency handling procedures, environmental fate, human exposure, detection methods, and regulatory requirements. The data are fully referenced and peer-reviewed by expert toxicologists and other scientists.
	International Toxicity Estimates for Risk (ITER) Database	http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?iter	A new database within the TOXNET site that contains human health risk values from major organizations worldwide for over 600 chemicals of environmental concern. It is a product of Toxicology Excellence for Risk Assessment (TERA), a non-profit group whose mission is to protect public health by developing and communicating risk assessment values, improving risk methods through research, and educating the public on risk assessment issues.
	IRIS (Integrated Risk Information System)	http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?IRIS.htm	Online database of the Environmental Protection Agency (EPA; http://www.epa.gov) containing carcinogenic and non-carcinogenic health risk information on over 500 chemicals. Data have been scientifically reviewed and represent EPA consensus.
	CCRIS (Chemical Carcinogenesis Research Information System)	http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?CCRIS	Sponsored by the National Cancer Institute (NCI; http://www.nci.nih.gov/), CCRIS contains scientifically evaluated data derived from carcinogenicity, mutagenicity, tumor promotion and tumor inhibition tests on about 8,000 chemicals.
	GENE-TOX (Genetic Toxicology)	http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?GENETOX	Another EPA database, contains genetic toxicology test results on over 3,000 chemicals. Selectively reviewed for each of the test systems under evaluation. The GENE-TOX data bank is the product of these data review activities.
	Columbia Environmental Research Center (CERC)	http://www.cerc.usgs.gov/data/acute/acute.html	Acute toxicity of over 400 chemicals and 60 aquatic animals. The results have been provided from aquatic acute toxicity tests conducted by the USGS CERC. The acute toxicity test provides a relative starting point for hazard assessment of contaminants and is required for federal chemical registration programs for fungicides, rodenticides and pesticides. Data is organized and searchable by combinations of compound and species data (e.g. LC-50 data for various chemicals and exposure times).
Toxicology Literature Files	TOXLINE	http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?TOXLINE	Bibliographic database covering the biochemical, pharmacological, physiological, and toxicological effects of drugs and other chemicals; over 2.5 million citations, almost all with abstracts and/or index terms and CAS registry numbers.
	DART/ETIC (Development and Reproductive Toxicology/Environmental Teratology Information Center)	http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?DARTETIC.htm	Bibliographic database covering literature on teratology and other aspects of developmental toxicology, contains over 90,000 references to teratology literature published since 1965.
	EMIC (Environmental Mutagen Information Center)	http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?EMIC	Bibliographic database containing over 100,000 references on chemical, biological, and physical agents that have been tested for genotoxicity, covers literature published since 1965.
Toxic Releases Files (TRI)	TRI (Toxic Chemical Release Inventory Files)	http://www.epa.gov/tri/	Contains information on the annual estimated releases of toxic chemicals to the environment for 1995–1997. It is based on data submitted to the EPA from industrial facilities throughout the US and includes the amounts of certain toxic chemicals released into the environment on over 650 chemicals and chemical categories. Pollution prevention data are also reported.
Carcinogenicity	Carcinogenic Potency Project	http://potency.berkeley.edu/listofpubs.topic.html	A useful resource related to carcinogenicity, includes a wide array of publications from the Carcinogenic Potency Project. The references include papers on methodological analysis of the relevance of carcinogenicity prediction from bioassays, species comparisons, target organs, mechanism of carcinogenesis, risk assessment techniques, possible cancer hazards of natural and synthetic chemicals, and causes and prevention of cancer.

search/retrieval. For details see http://www.env-sol.com/solutions/TSCASARA.HTML.

Finally, for prediction of toxicological parameters, the OSIRIS Property Explorer available at http://www.chemexper.com/tools/propertyExplorer/main.html (and listed above in the section "Prediction and calculation of molecular properties") has some interesting capabilities.

XVI. METASITES AND TECHNOLOGY SERVICE PROVIDER DATABASES

Metasites providing access to a range of resources devoted to chemistry that have not already been referred to are listed in Table 13.10.

TABLE 13.10

Title	URL	Comments
Chemistry Section of the WWW Virtual Library	http://www.liv.ac.uk/Chemistry/Links/links.html	Thorough, up-to-date and accurate listings of a large number of chemistry sites. The chemical database section at http://www.liv.ac.uk/Chemistry/Links/refdatabases.html gives details of a collection of about 80 chemical databases (among them: Analytical Abstracts, Beilstein, CCDC, CA Selects Plus, ChemFinder, DrugDB, FT-IR Library, STN, etc).
Organic Chemistry Resources Worldwide	http://www.organicworldwide.net	Created by Koen Van Aken, a Belgian chemist, a well organized and highly useful site for all engaged in synthetic organic chemistry research.
Caltech	http://library.caltech.edu/collections/chemistry.htm	Practically an indispensible point of call for databases and search engines for chemistry.
Liege, Belgium	http://www.ulg.ac.be/libnet/ud18.htm	An initiative from Simone Jérôme, chemistry librarian at the University of Liège, Belgium.
The Chemical Database Service (CDS)	http://cds.dl.ac.uk/	The CDS provides online access to numerous chemical databases, which are available free of charge to academics at UK universities. The chemistry links cover a large variety of topics (among them general information sites, reference databases, chemical sources, chemical web sites, UK universities, chemistry FTP sites). The solid-phase synthesis database is notable for its description of over 27,000 reactions described.
University of Cincinnati Online Database Collection	http://www.engrlib.uc.edu/selfhelp/alphlist.htm	Links to engineering, biology and chemistry databases, etc. are listed on this important site.
Chemiedatenbanken	http://www.chemie-datenbanken.de/	An excellent collection of German and international chemical databases e.g. free resources, general collections, and commercial database providers.
CHEMINFO	http://www.indiana.edu/~cheminfo	Metasite for chemical information resources on the internet and elsewhere, originating from the Indiana University chemical information courses. Usage has increased from nearly 100,000 successful requests in 2000 to over a quarter of a million requests in 2006. The main information page currently available at Indiana is Selected Internet Resources for Chemistry (SIRCh), which includes about 31 chemistry resource guides available on the internet (http://www.indiana.edu/~cheminfo/ca_gcisd.html). The ` also includes a link to the Chemical Acronyms database (http://www.oscar.chem.indiana.edu/cfdocs/libchem/acronyms/acronymsearch.html), which currently represents over 11,000 items linked to the full forms of the words.

Part III

Primary Exploration of Structure-Activity Relationships

Camille G. Wermuth

Section Editor

Chapter 14

Molecular Variations in Homologous Series: Vinylogues and Benzologues

Camille G. Wermuth

I. HOMOLOGOUS SERIES
A. Definition and classification
B. Shapes of the biological response curves
C. Results and interpretation
II. VINYLOGUES AND BENZOLOGUES
A. Applications of the vinylogy principle
B. Comments
REFERENCES

"Méthyle, éthyle, propyle, butyle … futile"

Old adage

I. HOMOLOGOUS SERIES

A. Definition and classification

The concept of a homologous series was introduced into organic chemistry by Gerhardt.[1] In medicinal chemistry the term has the same meaning, namely molecules differing from one another by only a methylene group.

The most frequently encountered homologous series in medicinal chemistry are monoalkylated derivatives, cyclo-polymethylenic compounds, straight chain difunctional systems, polymethylenic compounds and substituted cationic heads.

1. *Monoalkylated derivatives*

$$R—X \rightarrow R—CH_2—X \rightarrow R—CH_2—CH_2—X, \text{ etc...}$$

An example is provided by a series of neuraminidase inhibitors leading to the design of the avian flu agent tamiflu. These compounds result from a cyclohexene scaffold containing lipophilic side-chains.[2]

As shown in Figure 14.1, a 6.300-fold increase in potency is observed when the hydroxylic hydrogen is replaced by a diethyl-methyl side-chain. A comparable increase in potency is observed in a series of 1-methyl-1,2,3,4-tetrahydro-pyridyl-pyrazines described by Ward *et al.*[3] and exhibiting M_1 muscarinic agonistic properties. In changing from *O*-methyl to *O*-butyl, the affinity for the M_1 receptor varies from 850 to 17 nM. Another example is found in a series of 2-pyrone-derived elastase inhibitors.[4]

Neuraminidase inhibition	
R =	IC_{50} (nM)
H	6.300
CH_3—	3.700
CH_3—CH_2—	2.000
CH_3—CH_2—CH_2—	180
CH_3—CH_2—CH_2—CH_2—	300
$(CH_3)_2$—CH_2—CH_2—	200
CH_3—CH_2—$CH(CH_3)$—	10
$(CH_3$—$CH_2)_2$—CH—	1
$(CH_3$—CH_2—$CH_2)$CH—	16
Cyclopentyl	22
Cyclohexyl	60
Phenyl	530

FIGURE 14.1 Monoalkylated, cyclohexene-derived neuraminidase inhibitors.

Copyright © 2008, Elsevier Ltd
All rights reserved.

2. *Cyclopolymethylenic compounds*

Examples of such kind of structures with regularly increasing ring sizes are found for guanethidine (see Chapter 16), or for enalaprilat analogs (Figure 14.2).[5] In the latter example, a 4,000-fold increase in inhibition of angiotensin converting enzyme is obtained when changing from the five-membered ring ($n = 2$) to the eight-membered ring ($n = 5$).

Another example, published by scientists from Parke-Davis, relates to a series of "dipeptoid" analogs of cholecystokinin.[6] These compounds are α-methyl-tryptophan derivatives, N-substituted by carbamic esters of cyclanols with ring sizes increasing from cyclobutyl to cyclododecyl (Figure 14.3). Here again an optimal size was found (cyclononyl).

Size	IC_{50} (nM)
$n = 2$	19.000
$n = 3$	1.700
$n = 4$	19
$n = 5$	4.8
$n = 6$	8.1

FIGURE 14.2 Angiotensine-convertase inhibiting potency of enalaprilat analogs.[5]

Cycle	log *P*	IC_{50} (nM)
Cyclobutyl	3.88	12,100
Cyclopentyl	4.44	5,170
Cyclohexyl	5.00	520
Cycloheptyl	5.55	190
Cyclooctyl	6.11	125
Cyclononyl	6.67	85
Cyclodecyl	7.23	247
Cyclododecyl	8.34	1,437

FIGURE 14.3 Optimal ring size for a series of cyclanol carbamates.

3. *Open, difunctional, polymethylenic series*

$$X—(CH_2)_n—Y \rightarrow X—(CH_2)_{n+1}—Y$$

In the above general formula, X and Y can represent very diverse functional elements. The compounds can be symmetrical (X = Y; "dimers") or non-symmetrical (X ? Y); see Chapter 18. Usually, X and Y represent *polar* functions or *functionalized cyclic systems*.

When X and Y are polar functions, they are made essentially from functional groups such as alcohols, amines, acids, amides, amidines, or guanidines (Figure 14.4). A classical representative of a difunctionalized, symmetrical compound is decamethonium.

When X and Y are functionalized cyclic systems, they can be alicyclic or aromatic as well as homocyclic or heterocyclic (Figure 14.5). In any case, they bear some polar function or polar element. An example of this type of compounds is pentamidine.

Other examples are symmetrical bradykinin antagonists[7] and symmetrical lexitropsins (netropsine, distamycine), active against HIV-I viruses.[8] Non-symmetrical polymethylenic thromboxane synthetase inhibitors are described by Press *et al.*[9] The compounds contain a thiophene-2-carboxamide moiety, separated from an imidazole ring by 3–8 methylene units. Surprisingly, whereas most of the compounds show

FIGURE 14.4 Examples of functional groups encountered in open, difunctional, polymethylenic compounds and structure of decamethonium.

FIGURE 14.5 Examples of functionalized rings found in straight chain, difunctional, polymethylenic compounds. Structure of pentamidine.

similar thromboxane-synthetase inhibiting activities, only the two medium-sized ones ($n = 3$ and 4) showed hypotensive effects in spontaneously hypertensive rats (Figure 14.6).

In a series of benzimidazole-derived thromboxane A_2 receptor antagonists described by Nicolai *et al.*,[10] the crucial element is the distance between the carboxylic group and the benzimidazole ring. A 200-fold increase in affinity was observed when changing from propionic to a butyric side-chain (Figure 14.7).

4. *Substituted cationic heads*

With cationic head groups, homology achieves simultaneously a progressive increase in bulkiness and in lipophilicity. Figure 14.8 illustrates the influence of increasing bulkiness around the dopamine nitrogen in the antagonism of reserpine-induced catalepsy in mice.[11]

B. Shapes of the biological response curves

The most common curves are bell-shaped, the peak activity corresponding to a given value of the number n of carbon atoms (curve A, Figure 14.9). However, several other relationships were found among homologous series:

1. The activity can increase, without any particular rule, with the number of carbon atoms (curve B).
2. The biological activity can alternate with the number of carbon atoms, resulting in a zig-zag pattern (curve C).
3. In other series, the activity increases first with the number of carbon atoms and then reaches a plateau (curve D).
4. The activity can also decrease regularly, starting with the first member of the series (curve E). This was found for the toxicity of aliphatic nitriles or for the antiseptic properties of aliphatic aldehydes.
5. A last possibility resides in inversion of the pharmacological activity accompanying the increase in the number of carbon atoms (not shown on Figure 14.9, will be discussed below).

Size	IC_{50} (nM)
$n = 3$	600
$n = 4$	200
$n = 5$	200
$n = 6$	200
$n = 8$	70

FIGURE 14.6 Thromboxane synthetase inhibiting activity of a series of *N*-(imidazolyl-alkyl)-thiophene-5-carboxamides.[9]

Size	IC_{50}(nM)
$n = 0$	1,700
$n = 1$	7.8
$n = 2$	20

FIGURE 14.7 Affinity for the thromboxane A_2 receptor.[10]

FIGURE 14.8 Anticataleptic activity of substituted dopamines.[11]

FIGURE 14.9 Shapes of the biological response curves in homologous series.

C. Results and interpretation

1. *Curves with a maximum activity peak (bell-shaped curves) and curves with a continuous increase of activity*

In such series the continuous growth of an alkyl chain or of methylene units increases the hydrophobic part of the molecule. Various physicochemical parameters, such as solubility in water, partition coefficients, chromatographic R_f values, and critical micellar concentration, are precisely governed by the same fundamental property: the hydrophobic character.

a. Bell-shaped curves

Curves with an activity maximum are the most common ones, it is currently admitted that they reflect the existence of an optimal partition coefficient associated with the easiest crossing of the biological membranes. The relationship between the biological response and the partition coefficient is then illustrated by a parabolic curve. An example is found in structural analogs of PAF-acether.[12] In varying the length of the alkoxy chain from *n*-butyl to *n*-eicosanyl, the authors observed the peak activity for the *n*-hexadecyl chain, with a 1,200-fold interval between the most and the least active compounds (Figure 14.10). The drop in activity observed for the descending branch, is usually attributed to insufficient solubility in water (equal to incapability to cross the aqueous biophases), but can also be due to the formation of micelles. In this case, the concentration of the free drug, which represents the directly available form, lies under the critical threshold level. Bell-shaped curves are also seen when using isolated cells for which it can be demonstrated that the receptor is outside the membrane. In this case, the dominant factor is probably not the crossing of biological membranes. Changes in critical micellar concentration with increasing chain length could explain the effect in some cases, however, the curve is often too steep for this to be an acceptable explanation. Another possibility is that there is a lipophilic pocket of finite size. In many cases this pocket is not actually in the receptor protein. An argument in favor of this explanation are that the top of the bell is at C_{16} or C_{18} which fits with the length of the alkyl chains making up part of the bilayer, examples being PAF-acether analogs[12] and leucotriene D_4 agonists/antagonists. Another bit of evidence that supports this idea is the observation that the position of the peak of the curve can vary depending on which cell type is expressing the same receptor protein.

The study of the activities of *some* homologous compounds, can, through interpolation, identify which term is associated with the highest potency. The optimization method proposed by Bustard,[13] makes use of the Fibonacci numbers, and allows the identification of the most active compound (presumed to exist in a given interval) with the smallest possible number of syntheses (see also Refs. [14,15]).

FIGURE 14.10 Antiaggregant activity of structural analogs of PAF-acether.[12]

b. Apparently continuous increase

Actually, an apparently continuous increase of activity may correspond simply to the ascendant branch of the parabola (see the two curves in Table 14.1). The observed "pseudo-linear" curve usually occurs in an insufficiently explored series. A true linear correlation would imply the existence of compounds of infinite potency!

2. *Non-symmetrical curves with a maximum activity peak*

In some instances, curves with maximum activity peaks are not symmetrical and one side shows very sharp activity variations whereas the other one corresponds to a progressive variation. The shape of such a curve is represented on Figure 14.11.

For the $GABA_A$ antagonists represented on Figure 14.12, the peak activity corresponds to the branching of a *butyric* side-chain on the aminopyridazine system. The affinity diminishes drastically for shorter chains, but very progressively for longer chains.[18]

c. The particular case of polymethylenic bis-ammonium compounds

Compounds having the general formula $(CH_3)_3N^+(CH_2)_nN^+(CH_3)_3$ usually have high affinity for the cholinergic

receptors. When the values of n are intermediary ($n = 5$ or 6: penta- or hexamethonium) such compounds behave like cholinergic *agonist*s (toward the sympathic ganglions). For higher values ($n = 10$; decamethonium) the compounds become *antagonists* of acetylcholine (at the muscular end plate). In both cases, increasing acetylcholine levels displace them from their binding sites. When considering the neuro-muscular blockade, one observes again a curve with an asymmetric profile: sudden changes between $n = 6$ and 8, and then progressive diminution between $n = 9$ and 14.

Explanation: If we suppose that compounds of general structure $X(CH_2)_nY$ interact by means of their polar groups X and Y with complementary groups at the receptor, four

FIGURE 14.11 Non-symmetrical curve with a maximum activity peak.

TABLE 14.1 Local Anesthetic Activity[16] and Spasmolytic Activity[17] in Homologous Series

R	Duration of anesthesia in rabbit cornea (min)	R	Spasmolytic activity on guinea-pig isolated gut
Hydrogen	11	Methyl	8
Methyl	23	Ethyl	12
Ethyl	34	Propyl	24
Propyl	49	Butyl	98
Butyl	93	Pentyl	240
		Hexyl	410

interaction schemes can be foreseen, depending on the value of *n* (Figure 14.13: 1–5):

1. *n is small*: The molecule is too short and only one of its polar ends can establish an interaction with the complementary sites of the receptor (Figure 14.13: 1). The molecule will be inactive or poorly active. This is the case for the pyridazinyl-glycine of Figure 14.12.
2. *n possesses sufficient length*: A good interaction can be established with complementary sites of the receptor and trigger off the biological response (Figure 14.13: 2). This represents the optimal case.
3. *n is too great*: Two situations are foreseeable. If the molecule is *rigid* or if there is *steric hindrance*, the interaction is not possible for Y (Figure 14.13: 3). If the molecule is *flexible* and if the *steric tolerance is sufficient*, the fit can be entirely satisfactory (Figure 14.13: 4).
4. *n is very great*: In this case (Figure 14.13: 5), the fit with the receptor is again very good, but with a further located subsite Y″ instead of Y′, the substance can then behave as an *antagonist* (this was the above-discussed case of decamethonium).

FIGURE 14.12 Affinity of $GABA_A$ antagonists for the $GABA_A$ receptor site.[18]

3. *Serrated variations*

One sometimes observes alternating (serrated) variations of activity (zig-zag curves) according to whether the number of atoms of carbon is even or odd. Such an example is found for antimalarials derived from methoxy-6-amino-8-quinoline (Figure 14.14).

For these derivatives the antimalarial activity is greater if *n* represents an odd number (for studied values that vary from $n = 4$ to 10).[19] Another example is provided by leucotriene B_4 antagonists derived from hydroxyacetophenones[20] (Table 14.2).

FIGURE 14.13 Different modes of interaction of bifunctional molecules.

For both cases the findings are a reflection of the rotational energy curves for adjacent CH_2 groups. Similar observations were made in a series of 4,4′-dimethylamino-diphenoxyalkanes tested as potential schistosomicides.[21] For diamines where $n = 4$ until $n = 10$, the activity on the schistosomes varies in alternate manner (Figure 14.15).

FIGURE 14.14 Antimalarial activity in a homologous series of bifunctional methoxy-6-amino-8-quinolines. *Source*: After Magidson and Strukov[19].

TABLE 14.2 Zig-zag Variations of the Affinity of Hydroxyacetophenone Derivatives for the Human Peripheral Neutrophils. Inhibition of [^{3}H] LTB4 Binding at 0.1 mM.[20]

n = Length of the methylene chain	% Inhibition of [^{3}H] LTB_4 binding at 0.1 M
3	28
4	17
5	56
6	13
7	49

Alkyl-linked bis(amidinobenzimidazoles) with an even number of methylenes connecting the benzimidazole rings have a higher affinity for the minor groove of DNA than those with an odd number of methylenes (Figure 14.15).[22] Serrated variations of acetylcholinesterase inhibiting activity were also observed for donepezil analogs (Figure 14.16) (Box 14.1).[23]

Variations of the biological activity are not necessarily linked to the induction of effects at the level of a given receptor, but could have come from a pharmacokinetic factor (urinary or biliary excretion, plasma protein binding, differential metabolism). A case of differential metabolism is illustrated by the comparison of the toxicities of odd and even ω-fluoro acids.[24] The β-oxidation of odd chain length compounds leads to the extremely toxic fluoroacetic acid, while that of the acids with even numbers of carbon atoms generates β-fluoropropionic acid which is clearly less toxic (Table 14.3).

4. *Inversion of the activity*

It can happen that the lower members of a series possess one activity profile and that the higher terms possess a different activity, which contrasts with that of the lower members.

FIGURE 14.15 4,4′-Dimethylamino-diphenoxyalkanes[21] and alkyl-linked bis(amidinobenzimidazoles).[22]

n = length of the methylene chain	% Inhibition of acetylcholinesterase (nM)
1	5.7
2	30
3	1.5
4	35
5	14

FIGURE 14.16 Inhibition of acetylcholinesterase by donepezil and homologs.[23]

BOX 14.1 The Origin of the Zig-zag Variations

Zig-zag variations are well known in homologous series for physical properties such as melting points and solubilities. Thus, propane, with an odd number of carbon atoms, melts at −189.9°C whereas butane, with an even number, melts 51.6°C higher at −138.3°C. But odd-numbered pentane melts at 129.7°C, only 8.6°C higher than butane. Boese *et al.* studying X-ray structures of *n*-propane through *n*-nonane at −90°C say that the methyl groups on chains lying end-to-end are the culprits (Boese, R., Weiss, H.-C., Bläser, D. The melting point alternation in the short claim *n*-alkanes: single crystal X-ray analyses of propane at 30 K and of *n*-butane to *n*-nonane at 90 K. *Angew. Chem. Int. Ed.* 1999, **38**, 988–992).

In even-numbered chains the methyl groups dovetail nicely and stay out of one another's way. But in odd-numbered chains methyl groups on one end can only avoid one another by increasing the distance between chain ends. This less-than-tight packing in odd-numbered chains results in their anomalous melting points, the researchers suggest. In the examples above, the alkyl chain represents a spacer group between two binding groups. In some cases, it can be shown that the energy required to fold the molecule to obtain the required separation should change in a zig-zag manner with increasing chain length.

TABLE 14.3 Zig-zag Variations of the Toxicity of Aliphatic ω-Fluoro Derivatives (LD_{50} for Mice in mg/kg Intraperitoneally)[24]

	$F(CH_2)_nCOOH$		$F(CH_2)_nCO$		$F(CH_2)_nCOH$		$F(CH_2)_nCH_3$	
n	Odd	Even	Odd	Even	Odd	Even	Odd	Even
1	6.6		6		10			
2		60				46.5		
3	0.65		2		0.9			
4		>100		81		>100		
5	1.35		0.58		1.2		1.7	
6		40		>100		80		35
7	0.64		2		0.6		2.7	
8		>100		53		32		21.7
9	1.5		1.9		1		1.7	
10		57		>40		>100		15.5
11	1.25				1.5		2.5	

This phenomenon is particularly observed when the bulkiness of cationic heads is progressively increased. In *N-alkylated derivatives of norepinephrine*,[25] progressive alkylation reduces the hypertensive activity according to the sequence: NH2, NHMe, NHEt, NHNPro. Finally, the molecules become hypotensive for the values: NHIsoPro, NH*n*Bu and NHIsoBu (Table 14.4).

This anomaly is explained by the fact that norepinephrine can interact with two subclasses of receptors (α- and β-adrenergic receptors). The less hindered derivatives are able to bind to both α- and β-receptors, hindered ones solely to β-receptors. A similar inversion of properties is observed when the cholinergic agonist carbachol is modified by dibutylation at the carbamate function and exchange of one of its methyl groups for an ethyl group (Figure 14.17). The analog, dibutoline, is a powerful cholinolytic.

In morphine (agonist), the replacement of the *N*-methyl group by a more bulky radical such as *N*-allyl, *N*-cyclopropyl-methyl or *N*-cyclobutyl-methyl leads to powerful antagonists of the opiate receptors (see Section III in Chapter 20).

Introduction of bulkiness in a cationic head does not always cause a change from agonist to antagonist. Thus, the analog *N*-propyl-apomorphine is a more powerful dopaminergic agonist than the apomorphine itself. The creation of bulkiness is obviously not limited to cationic head groups and lipophilic groups can be attached to any other part of the molecule (Figure 14.18).[26]

5. *Conclusion*

Variations in homologous series are generally the search for the optimal lipophilicity. In the cyclopolymethylenic series

TABLE 14.4 Gradual Inversion of the Activity in a Homologous Series[25]

	Blood pressure of the cat	
R	Hypertensive	Hypotensive
Hydrogen	++	–
Methyl	++	–
Ethyl	+	+
Propyl	–	+
Isopropyl	–	++
Butyl	–	++
Isobutyl	–	++

FIGURE 14.17 Carbachol (left) and dibutoline (right).

FIGURE 14.18 Tolerance to bulkiness.[26]

conformational problems may be added. For difunctional polymethylenic derivatives, intercharge distances and possibly, elements of symmetry (see Chapter 18) can take over. Whereas the activity profile is generally preserved during homology changes, very large differences in potency can be found, that confound the old adage "methyl, ethyl, propyl, butyl … futile".

II. VINYLOGUES AND BENZOLOGUES

The vinylogy principle was first formulated by Claisen in 1926,[27] who observed for formylacetone acidic properties similar to that of acetic acid. The vinyl goup plays the role of an electron-conducting channel between the carbonyl and the hydroxyl group. The same effect explains the acidity of ascorbic acid (Figure 14.19).

Today the vinylogy principle is explained by the mesomeric effect and it applies to all conjugated systems: imine and ethynyl groups, phenyl rings, and aromatic heterocycles (Figure 14.20). For a review of the chemical aspects of the vinylogy principle see Krishnamurthy.[28]

A. Applications of the vinylogy principle

Although numerous applications of the vinylogy concept are found in the medicinal chemistry literature, only a very few of them are of practical interest, mainly because the preparation of vinylogues usually leads to compounds which are more sensitive to metabolic degradation and more toxic (reactivity of the conjugated double bond) than the parent drug, without being more active.

1. *Authentic vinylogues*

The vinylogues of phenylbutazone,[29] and of pethidine (C. G. Wermuth, unpublished results) have the same type of activity than the parent drug, but the duration of action, especially for the pethidine analog, is notably shorter than that of the initial molecule (Figure 14.21). This is probably due to an easier metabolic degradation of the styryl double bond.

In preparing the vinylogues of acetylcholine (Figure 14.21), Tenconi and Barzaghi.[30] succeeded in separating the nicotinic from the muscarinic activity (Table 14.5).

Tolcapone (Figure 14.22) was designed as an inhibitor of the enzyme catechol *O*-methyltransferase useful in the l-DOPA treatment of Parkinson's disease.[31] In avoiding the

FIGURE 14.19 Formylacetone (enolic form) and vitamin C are comparable in acidity to carboxylic acids.

FIGURE 14.20 Vinylogy and its extensions.

FIGURE 14.21 Vinylogues of phenylbutazone, pethidine and acetylcholine.

TABLE 14.5 Cholinergic Profile of the Vinylogue of Acetylcholine[30]

Compound	Nicotinic activity	Muscarinic activity	Sensitivity to ACh-esterase
Ach	+	+	Sensitive
Vinylogue	+	Insensitive	Insensitive

FIGURE 14.22 The vinylogy principle applied to the catechol *O*-methyltransferase inhibitor tolcapone.

methylation of l-DOPA as well as that of dopamine it prolonges the benefic activities of these molecules.

Catechol *O*-methyltransferase inhibition represents therefore a valuable adjuvant to the l-DOPA decarboxylase inhibition. Unfortunately, tolcapone exhibited severe liver damages and had to be removed from the market. The corresponding vinylogue entacapone is devoided of these side effects.[32]

Wright *et al.*[33] describe a series of hydroxamic acid vinylogues acting as dual inhibitors of 5-lipoxygenase (5-LO; $IC_{50} = 0.15 \times 10^{-6}$M) and of interleukin-1β (IL-1β) biosynthesis ($IC_{50} = 2.8 \times 10^{-6}$M) which might be useful as anti-inflammatory drugs (Figure 14.23).

For such compounds several possible isomeric and tautomeric forms can be considered. These include the (*E*) and (*Z*) geometrical isomers (a and b), as well as the tautomers nitrone (c), 5-hydroxy-isoxazolidine (d), and oxaziridine (e). Examination of the ^{1}H and ^{13}C NMR spectra of the

FIGURE 14.23 Hydroxamic acid vinylogues and their various isomeric forms.[33]

FIGURE 14.24 Cinnamic derivatives as vinylogues of benzaldehyde.[34]

vinylogues revealed that each of the tautomeric possibilities are present in solution in varying proportions. The relative proportion of each isomer was found to be dependent on the solvent, the pH, and its chemical structure.

In order to design compounds able to react covalently with the nucleophilic cysteine of human 3C rhinovirus protease, scientists from Agouron designed vinylogous derivatives of the prototype inhibitor 3-carbamoyl-benzaldehyde (Figure 14.24). In this particular case, Michael acceptor reactivity is eventually the wanted feature.[34]

FIGURE 14.25 Oxotremorine is an ethynologue of the acetylcholine pharmacophore.

2. Ethynologues

Some ethynologues of biologically active compounds were prepared by Dunoguès *et al.*, unfortunately they did not describe their biological activity: *aspirin ethynologue*,[35] *nicotinamide and isoniazide ethynologues*,[36] and *chalcones ethynologues*.[37] In some way the cholinergic antagonist oxotremorine can be considered as an ethynologue of the acetylcholine pharmacophore (Figure 14.25).

As a rule, simple azavinylogues are unstable compounds, due to the easy hydrolysis of the imino bond. However, the particular case of *O*-alkylated oximes (XCHNOY; with Y = R or Ar) can be interesting insofar that the oximic imino bond was shown to be biostable.[38] Preparing azavinylogues of β-blocking agents (Figure 14.26) led to some active compounds.[39,41] The proposal was made that the stable oxime CNOCH2 could mimic a portion of an aromatic ring, thus simulating an aryl or an aryloxymethylene group.[42]

Reduction of the imino bond results in a decrease but not a loss of activity and ether derivatives retain activity.[41] Tricyclic oxime β-blockers showed high selectivity for β_2 receptors.[39,40] Noxiptyline (Agedal®, Bayer) is the oximic equivalent of amitriptyline (*Review*: Hoffmeister;[38,43,44] *Cryst. Struct*: Bandoli[44]).

FIGURE 14.26 Oxime ethers as azavinylogues.

TABLE 14.6 Cyclovinylogues of Procainamide, Relative Activity with Regard to Procainamide[45]

Compound	Local anesthetic power	Antiarrhytmic activity
Procainamide	1	1
Ortho-cyclovinylogue	~0	0.17
Meta-cyclovinylogue	47	0
Para-cyclovinylogue	35	0

3. *Cyclovinylogues*

These vinylogues have the advantage of being more stable toward *in vivo* metabolism. In addition, they allow molecular variations with *ortho*, *meta*, and *para* positional isomers. Thus, for cyclovinylogues of procaïnamide the highest local anesthetic activity was found with the meta derivative, which showed also the best dissociation between local anesthetic and antiarrhytmic activity (Table 14.6).[45]

Similar results were observed with cyclovinylogues of *lidocaine*.[46] For other references, see Valenti.[47] Compound *TA-1801* [ethyl 2-(4-chlorophenyl)-5-(2-furyl)-4-oxazoleacetate],[48] can to some extent be considered as an arenologue of *clofibrate* (Figure 14.27).

FIGURE 14.27 TA-1801, an arenologue of clofibrate.[48]

FIGURE 14.28 The vinylogous relationship between the carbonyl and the amino group in pyrroline-3-ones gives them the reactivity of secondary amides.[49]

FIGURE 14.29 Linear and angular adenine benzologues.

Pyrroline-3-ones were used as peptidic bond surrogates by Hirschmann *et al.*[49] In such compounds (Figure 14.28), thanks to vinylogy, the carbonyl and the amino group show the same chemical reactivity to that of a secondary amide.

4. *Benzologues*

Linear and angular benzologues of guanine[50] and adenine,[51–53] were published without any indication of bio logical activity (Figure 14.29). A review article on chemistry and biochemistry of benzologues was published by Leonard

FIGURE 14.30 Linear benzologues derived from zaprinast.[57]

and Hiremath.[54] Linear and angular benzologues of xanthines showed submicromolar affinities for rat brain A_1 and A_2 adenosine receptors[55] and benzologues of quinolone antibacterials maintained high antimicrobial activity.[56]

A very convincing example of the usefulness of benzologues is provided by the synthesis of compound "A," a linear benzologue of the prototypical PDE5 inhibitor zaprinast (Figure 14.30) and its optimization to potent and selective PDE5 inhibitors such as "B".[57]

B. Comments

1. Due to important changes in the geometry, vinylogues often have unpredictable activity. For this reason, vinylogues play a minor role in medicinal chemistry. In addition, their metabolic vulnerability or their increased toxicity may represent a significant drawback.
2. However, the vinylogy principle is sometimes applied to the design of bioisosteres. Thus, the guanidinic group of the benzimidazole fenbendazole[58] can be compared to its vinylogue[59] in the corresponding imidazo[1,2-*a*]pyridine (Figure 14.31). Both compounds are anthelmintics of similar potency.
3. The vinylogy principle can account for unexpected chemical reactivity that is not always recognized at a first glance (Figure 14.32). So, for example, the basicity of the N1 nitrogen is strengthened in compound CGS 8216 thanks to the vinylogous influence of the quinoline nitrogen.

For a similar reason the carbonyl group of benzopiperidones or of 3-acyl-indoles behaves chemically more like an amidic carbonyl then like a ketonic one. In 2-methoxy-*para*-benzoquinone the reactivity of the methoxy group is that of a carboxylic ester, rendering it susceptible to attack by secondary amines.

FIGURE 14.31 Application of the vinylogy principle to the design of a fenbendazole bioisostere.[59]

FIGURE 14.32 Unexpected chemical reactivities attributable to vinylogy.

REFERENCES

1. Gerhardt, C. Principes de la classification sériaire. In *Traite de Chimie Organique*, Vol. 1. Firmin Didot Frères: Paris, 1853, pp. 121–142.
2. Kim, C. U., Lew, W., Williams, M. A., Wu, H., Zhang, L., Chen, X., Escarpe, P. A., Mendel, D. B., Laver, W. G., Stevens, R. C. Structure–activity relationship studies of novel carbocyclic influenza neuraminidase inhibitors. *J. Med. Chem.* 1998, *41*, 2451–2460.
3. Ward, J. S., Merritt, L., Klimkowski, V. J., Lamb, M. L., Mitch, C. H., Bymaster, F. P., Sawyer, B., Shannon, H. E., Olesen, P. H., Honoré, T., Sheardown, M. J., Sauerberg, P. Novel functional M_1 selective muscarinic agonists. 2. Synthesis and structure–activity relationships of 3-pyrazinyl-1,2,5,6-tetrahydro-1-methylpyridines. Construction of a molecular model for the M_1 pharmacophore. *J. Med. Chem.* 1992, *35*, 4011–4019.
4. Cook, L., Ternai, B., Ghosh, P. Inhibition of human sputum elastase by substituted 2-pyrones. *J. Med. Chem.* 1987, *30*, 1017–1023.
5. Thorsett, E. D. Conformationally restricted inhibitors of angiotensin converting enzyme. In *Actualités de Chimie Thérapeutique*

(Combet-Farnoux, C., Ed.), Vol. 13. Société de Chimie Thérapeutique: Chatenay-Malabry, 1986, pp. 257–268.
6. Eden, J. M., Higginbottom, M., Hill, D. R., Horwell, D. C., Hunter, J. C., Martin, K., Pritchard, M. C., Rahman, S. S., Richardson, R. S., Roiberts, E. Rationally designed "dipeptoid" analogues of cholecystokinin (CCK): N-terminal structure-affinity relationships of α-methyltryptophan derivatives. *Eur. J. Med. Chem.* 1993, *28*, 37–45.
7. Cheronis, J. C., Whalley, E. T., Nguyen, K. T., Eubanks, S. R., Allen, L. G., Duggan, M. J., Loy, S. D., Bonham, K. A., Blodgett, J. K. A new class of bradykinin antagonists: synthesis and *in vitro* activity of bis-succinimidoalkane peptide dimers. *J. Med. Chem.* 1992, *35*, 1563–1572.
8. Wang, W., Lown, J. W. Anti-HIV-I activity of linked lexitropsins. *J. Med. Chem.* 1992, *35*, 2890–2897.
9. Press, J. B., Wright, W. B., Jr., Chan, P. S., Haug, M. F., Marsico, J. W., Tomcufcik, A. S. Thromboxane synthetase inhibitors and antihypertensive agents. 3. *N*-[(1*H*-Imidazol-1-yl)alkyl]heteroaryl amides as potent enzyme inhibitors. *J. Med. Chem.* 1987, *30*, 1036–1040.
10. Nicolai, E., Goyard, J., Benchetrit, T., Teulon, J. M., Caussade, F., Virone, A., Delchambre, C., Cloarec, A. Synthesis and structure–activity relationships of novel benzimidazole and imidazo[4,5-*b*]pyridine acid derivatives as thromboxane A_2 receptor antagonists. *J. Med. Chem.* 1992, *36*, 1175–1187.
11. Ginos, J. Z., Stevens, J. M., Nichols, D. E. Structure-activity relationships of N-substituted dopamine and 2-amino-6,7-dihydroxy-1,2,3,4-tetrahydronaphtalene analogues: behavioral effects in lesioned and reserpinized mice. *J. Med. Chem.* 1979, *22*, 1323–1329.
12. Godfroid, J.-J., Broquet, C., Jouquey, S., Lebbar, M., Heymans, F., Redeuilh, C., Steiner, E., Michel, E., Coeffier, E., Fichelle, J., Worcel, M. Structure–activity relationship in PAF-acether. 3. Hydrophobic contribution to agonistic activity. *J. Med. Chem.* 1987, *30*, 792–797.
13. Bustard, T. M. Optimization of alkyl modifications by fibonacci search. *J. Med. Chem.* 1974, *17*, 777–778.
14. Santora, N. J., Auyang, K. Non-computer approach to structure–activity study. An expanded Fibonacci search applied to structurally diverse types of compounds. *J. Med. Chem.* 1975, *18*, 959–963.
15. Martin, Y. C. *Quantitative Drug Design, a Critical Introduction*. Marcel Dekker, Inc.: New York, 1987. pp. 257–261.
16. Koelzer, P. P., Wehr, K. H. Beziehungen zwischen chemischer Konstitution un pharmakologischer Wirkung bei mehreren Klassen neure Lokalanaesthetica. *Arzneimitt.-Forsch* 1958, *8*, 544–550.
17. Funcke, A. B. H., Ernsting, M. J. E., Rekker, R. F., Nauta, W. T. Untersuchungen über Spasmolytica. 1. Mandelsäureester. *Arzneimitt. -Forsch.* 1953, *3*, 503–506.
18. Wermuth, C. G., Bourguignon, J.-J., Schlewer, G., Gies, J. P., Schoenfelder, A., Melikian, A., Bouchet, M.-J., Chantreux, D., Molimard, J.-C., Heaulme, M., Chambon, J.-P., Bizière, K. Synthesis and structure–activity relationships of a series of aminopyridazine derivatives of γ-aminobutyric acid acting as selective $GABA_A$ antagonists. *J. Med. Chem.* 1987, *30*, 239–249.
19. Magidson, O. J., Strukow, I. T. Die Derivate des 8-Aminochinolins als Antimalariapräparate. Mitteilung II: Der Einfluß der Länge der Kette in Stellung 8. *Arch. Pharm.* 1933, *271*, 569–580.
20. Herron, D. K., Goodson, T., Bollinger, N. G., Swanson-Bean, D., Wright, I. G., Staten, G. S., Thompson, A. S., Froelich, L. L., Jackson, W. T. Leukotriene B4 receptor antagonists: the LY 255 283 series of hydroxyacetophenones. *J. Med. Chem.* 1992, *35*, 1818–1828.
21. Raison, C. G., Standen, O. D. The schistosomicidal activity of symmetrical diaminodiphenoxyalkanes. *Br. J. Pharmacol.* 1955, *10*, 191–199.
22. Fairley, T. A., Tidwell, R. R., Donkor, I., Naiman, N. A., Ohemeng, K. A., Lombardy, R. J., Bentley, J. A., Cory, M. Structure, DNA minor groove binding, and base pair specificity of alkyl- and aryl-linked bis(amidinobenzimidazoles) and bis(amidinoindoles). *J. Med. Chem.* 1993, *36*, 1746–1753.
23. Sugimoto, H., Iimura, Y., Yamanishi, Y., Yamatsu, K. Synthesis and structure–activity relationships of acetylcholinesterase inhibitors: 1-benzyl-4-[(5,6-dimethoxy-1-oxoindan-2-yl)methyl]piperidine hydrochloride and related compounds. *J. Med. Chem.* 1996, *38*, 4821–4829.
24. Pattison, F. L. M. *Toxic Aliphatic Fluorine Compounds*. Elsevier: Amsterdam, 1959.
25. Ariëns, E. J.*Molecular Pharmacology*, Vol. 1. Academic Press: New York, 1964.
26. Beyer, K. H., Baer, J. E. Physiological basis for the action of newer diuretic agents. *Pharmacol. Rev.* 1961, *Vol. 13*, pp. 517–562.
27. Claisen, L. Zu den *O*-Alkylderivaten des Benzoyl-acetons und den aus ihnen entstehenden Isooxazolen. *Ber. dtsch. chem. Ges.* 1926, *59*, 144–153.
28. Krishnamurthy, S. The principle of vinylogy. *J. Chem. Educ.* 1982, *59*, 543–547.
29. Yamamoto, H., Kaneko, S.-I. Synthesis of 1-phenyl-2-styryl-3,5-dioxopyrazolidines as antiinflammatory agents. *J. Med. Chem.* 1970, *13*, 292–295.
30. Tenconi, F., Barzaghi, F. Attivita nicotinica di vinil-analoghi di esteri della colina. *Boll. Chim. Pharmaceut.* 1964, *103*, 569–575.
31. Zürcher, G., Keller, H. H., Kettler, R., Borgulya, J., Bonetti, E. P., Eigenmann, R., Da Prada, M. Ro 40-7592, a novel, very potent, and orally active inhibitor of catechol-*O*-methyltransferase: a pharmacological study in rats. *Adv. Neurol.* 1990, *53*, 497–503.
32. Nissinen, E., Linden, I. B. Biochemical and pharmacological properties of a peripherally acing catechol-*O*-methyltransferase inhibitor: entacapone. *Naunyn Schmiedebergs Arch. Pharmacol.* 1992, *346*, 262–266.
33. Wright, S. W., Harris, R. R., Kerr, J. S., Green, A. M., Pinto, D. J., Bruin, E. M., Collins, R. J., Dorow, R. L., Mantegna, L. R., Sherk, S. R., Covington, M. B., Nurnberg, S. A., Welch, P. K., Nelson, M. J., Magolda, R. L. Synthesis, chemical, and biological properties of vinylogous hydroxamic acids: dual inhibitors of 5-lipoxygenase and IL-1 biosynthesis. *J. Med. Chem.* 1962, *35*, 4061–4068.
34. Reich, S. H., Johnson, T., Wallace, M. B., Kephart, S. E., Fuhrman, S. A., Worland, S. T., Mattews, D. A., Hendrickson, T. F., Chan, F., Meador, J., III, Ferre, R. A., Brown, E. L., DeLisle, D. M., Patick, A. K., Binford, S. L., Ford, C. E. Substituted benzamide inhibitors of human rhinovirus 3C protease: structure-based design, synthesis, and biological evaluation. *J. Med. Chem.* 2000, *43*, 1670–1683.
35. Babin, P., Bourgeois, P., Dunoguès, J. Synthèse de l'éthynologue de l'acide acétylsalicylique. *C. R. Acad. Sci. (Paris)* 1976, *283*, 149–152.
36. Babin, P., Cassagne, A., Dunoguès, J., Duboudin, F., Lapouyade, P. Ethynologues du nicotinamide et de l'isoniazide. *J. Heterocycl. Chem.* 1981, *18*, 519–523.
37. Babin, P., Lapouyade, P., Dunoguès, J. Synthesis of chalcone ethynologues with a pharmacological objective. *Canad. J. Chem.* 1982, *60*, 379–382.
38. Hoffmeister, F. Zur Frage pharmakologisch-klinischer Wirkungsbeziehungen bei Antidepressiva, dargestellt am Beispiel von Noxiptilin. *Arzneimitt.-Forsch.* 1969, *19*, 458–467.
39. Leclerc, G., Mann, A., Wermuth, C. G., Bieth, N., Schwartz, J. Synthesis and β-adrenergic blocking activity of a novel class of aromatic oxime ethers. *J. Med. Chem.* 1977, *20*, 1657–1662.
40. Imbs, J. L., Miesch, F., Schwartz, J., Velly, J., Leclerc, G., Mann, A., Wermuth, C. G. A potent new β_2-adrenoceptor blocking agent. *Br. J. Pharmacol.* 1977, 357–362.
41. Leclerc, G., Bieth, N., Schwartz, J. Synthesis and β-adrenergic blocking activity of new aliphatic oxime ethers. *J. Med. Chem.* 1980, *23*, 620–624.
42. Macchia, B., Balsamo, A., Lapucci, A., Martinelli, A., Macchia, F., Bresci, M. C., Fantoni, B., Martinotti, E. An interdisciplinary approach to the design of new structures active at the β-adrenergic receptor. Aliphatic oxime ether derivatives. *J. Med. Chem.* 1985, *28*, 153–160.
43. Aichinger, G., Behner, O., Hoffmeister, F., Schütz, S. Basische tricyclische Oximinoäther und ihre pharmakologischen Eigenschaften. *Arzneimitt.-Forsch.* 1969, *19*, 838–845.
44. Bandoli, G., Nicolini, M. Crystal structure of the antidepressant noxiptyline hydrochloride (5-dimethylaminoethyloximino-5*H*-dibenzo[*a*,*d*]-cyclohepta-1,4-diene hydrochloride). *J. Crystallogr. Spectrosc. Res.* 1983, *13*, 191–199.

45. Valenti, P., Mazzotti, M., Rampa, A., Magistretti, M. J. Cyclovinylogues of procainamide. *Arch. Pharm.* 1982, *315*, 1003–1007.
46. Valenti, P., Montanari, P., Da Re, P., Soldani, G., Bertelli, A. Synthesis and pharmacological properties of three lidocaine cyclovinylogues. *Arch. Pharm.* 1980, *313*, 280–284.
47. Valenti, P., Montanari, P., Fabbri, G., Giovannini, L., Giacomelli, A. Cyclo-vinylogues of some antimuscarinic drugs. *Arch. Pharm.* 1985, *318*, 222–224.
48. Mooriya, T., Seki, M., Takabe, S., Matsumoto, K., Takashima, K., Mori, T. Compound TA-1801 [ethyl 2-(4-chlorophenyl)-5-(2-furyl)-4-oxazoleacetate]. *J. Pharm. Sci.* 1987, *76*, S164.
49. Smith, A. B., III, Keenan, T. P., Holcomb, R. C., Sprengeler, P. A., Guzman, M. C., Wood, J. L., Caroll, P. J., Hirschmann, R. Design synthesis and crystal structure of a pyrrolinone-based peptidomimetic possessing the conformation of a β-strand: potential application to the design of novel inhibitors of proteolytic enzymes. *J. Am. Chem. Soc.* 1992, *114*, 10672–10674.
50. Cottis, S. G., Clarke, P. B., Tieckelmann, H. Pyrazolo[3,4-*b*]pyridines and pyrazolo[3′,4′:6,5]pyrido[2,3-*d*]pyrimidines. *J. Heterocycl. Chem.* 1965, *2*, 192–201.
51. Leonard, N. J., Morrice, A. G., Sprecker, M. A. Linear benzoadenine. A stretched-out analog of adenine. *J. Org. Chem.* 1975, *40*, 356–363.
52. Morrice, A. G., Sprecker, M. A., Leonard, N. J. The angular benzoadenines. 9-Aminoimidazo [4,5-*f*] quinazoline and 6-aminoimidazo [4,5-*h*] quinazoline. *J. Org. Chem.* 1975, *40*, 363–366.
53. Leonard, N. J., Sprecker, M. A., Morrice, A. G. Defined dimensional changes in enzyme substrates and cofactors. Synthesis of lin-benzoadenosine and enzymatic evaluation of derivatives of the benzopurines. *J. Am. Chem. Soc.* 1976, *98*, 3987–3994.
54. Leonard, N. J., Hiremath, S. P. Dimensional probes of binding and activity. *Tetrahedron* 1986, *42*, 1917–1961.
55. Schneller, S. W., Ibay, A. C., Christ, W. J., Bruns, R. F. Linear and proximal benzo-separated alkylated xanthines as adenosine-receptor antagonists. *J. Med. Chem.* 1989, *32*, 2247–2254.
56. Jordis, U., Sauter, F., Rudolf, M., Cai, G. Synthesen neuer Chinolon-Chemotherapeutika 1: Pyridochinoline und Pyridophenanthroline als “lin-benzo-Nalidixinsäure”-Derivate. *Monatsh. Chem.* 1988, *119*, 761–780.
57. Rotella, D. P., Sun, Z., Zhu, Y., Krupinski, J., Pongrac, R., Seliger, L., Normandin, D., Macor, J. E. N-3-substituted imidazoquinazolinones: potent and selective PDE5 inhibitors as potential agents for treatment of erectile dysfunction. *J. Med. Chem.* 2000, *43*, 1257–1263.
58. Averkin, E. A., Beard, C. C., Dvorak, C. A., Edwards, J. A., Fried, J. H., Kilian, J. G., Schiltz, R. A., Kistner, T. P., Drudge, J. H., Lyons, E. T., Sharp, M. L., Corwin, R. M. Methyl 5(6)-phenylsulfinyl-2-benzimidazolecarbamate, a new, potent anthelmintic. *J. Med. Chem.* 1972, *15*, 1164–1166.
59. Bochis, R. J., Dybas, R. A., Eskola, P., Kulsa, P., Linn, B. O., Lusi, A., Meitzner, E. P., Milkowski, J., Mrozik, H., Olen, L. E., Peterson, L. H., Tolman, R. L., Wagner, A. F., Waksmunski, F. S. Methyl 6-(phenylsulfinyl) imidazo [1,2-*a*] pyridine-2-carbamate, a potent, new anthelmintic. *J. Med. Chem.* 1978, *21*, 235–237.

Chapter 15

Molecular Variations Based on Isosteric Replacements

Paola Ciapetti and Bruno Giethlen

I. INTRODUCTION
II. HISTORY: DEVELOPMENT OF THE ISOSTERISM CONCEPT
A. The molecular number
B. The isosterism concept
C. The notion of pseudoatoms and Grimm's hydride displacement law
D. Erlenmeyer's expansion of the isosterism concept
E. Isoserism criteria: Present conceptions
F. The bioisosterism concept: Friedman's and Thornber's definitions
III. CURRENTLY ENCOUNTERED ISOSTERIC AND BIOISOSTERIC MODIFICATIONS
A. Replacement of univalent atoms or groups
B. Interchange of divalent atoms and groups
C. Interchange of trivalent atoms and groups
D. Ring equivalents
E. Groups with similar polar effects: functional equivalents
F. Reversal of functional groups
IV. SCAFFOLD HOPPING
A. Successful examples of serendipitous scaffold hopping
B. Scaffold hopping and virtual screening
V. ANALYSIS OF THE MODIFICATIONS RESULTING FROM ISOSTERISM
A. Structural parameters
B. Electronic parameters
C. Solubility parameters
D. Anomalies in isosterism
VI. MINOR METALLOIDS-TOXIC ISOSTERS
A. Carbon–silicon bioisosterism
B. Carbon–boron isosterism
C. Bioisosteries involving selenium
REFERENCES

Things are not always what they seem; the first appearance deceives many; the intelligence of a few perceives what has been carefully hidden …

Anonymous

I. INTRODUCTION

In a biologically active molecule the replacement of an atom or a group of atoms by another one presenting the same physiochemical properties is based on the concept of *isosterism*. The notion of isosterism was introduced in 1919 by Langmuir[1] who was mainly focused in the similarities of electronic and steric arrangement of atoms, groups, radicals, and molecules.

The concept of isosters was then broadened by Grimm in 1925 with the statement of Hydride Displacement Law, and, further on, Erlenmeyer extended Grimm's classification defining isosters as atoms, ions, and molecules in which the peripheral layers of electrons can be considered identical. The extensive application of isosterism to modify a part of a biologically active molecule to get another one of similar activity, has given rise to the term of *bioisosterism* or non-classical isosterism. As initially defined by Friedman,[2] bioisosters include all atoms and molecules which fit the broadest definition for isosters and that elicit the similar biological activity. In medicinal chemistry, the concept of bioisosterism is a research tool of the utmost importance widely used in analogs design.

Copyright © 2008, Elsevier Ltd
All rights reserved.

FIGURE 15.1 Isosteric replacements of —NH— with —O— and —S— in clozapine analogs.

FIGURE 15.2 Example of bioisosteric replacement in the benzodiazepine series.

The term "analog" is derived from the Greek word αηαλσγια (analogia) and has been used to point out structural and functional similarity. Applied to drugs this means that an analog of an existing drug shares chemical and therapeutic similarity with the parent compound. This definition implies that three categories of drug analogs can be listed: (a) those presenting chemical and therapeutic similarity; (b) those presenting only similar chemical features; and (c) those eliciting the same pharmacological effect but displaying a completely different chemical scaffold. Isosters have been classified as either classical or non-classical according to the degree of electronic and steric alikeness. Bioisosterism is used interchangeably with the term non-classical isosters. In Figure 15.1 it is reported clozapine and two analogs as an example of classic isosteric replacement of NH with O and S.

An illustration of bioisosteric change is given in Figure 15.2 where the amide moiety of diazepam has been replaced by a triazole ring to give alprazolam.

Taken in its broadest meaning, bioisosterism include the replacement of the initial molecular scaffold by a different one, keeping the same biological activity. This approach, called scaffold hopping, is well illustrated by molecules like diazepam, zolpidem, zaleplon, and zopiclone which exert the same biological response acting as full agonists of GABA-A (γ-aminobutyric acid) receptor at the benzodiazepine site though being structurally different (see Figure 15.3).

Isosterism, bioisosterism, and scaffold hopping will be discussed in more details on the following paragraphs.

II. HISTORY: DEVELOPMENT OF THE ISOSTERISM CONCEPT

The development of the concept of isosterism takes its roots in the attempts to extend to the whole molecules the knowledge acquired for elements, namely that two elements possessing an identical peripheral electronic distribution do also possess similar chemical properties.

A. The molecular number

Allen, in 1918[3] defined the *molecular number* of a compound in a similar way of the atomic number:

$$N = aN_1 + bN_2 + cN_3 + \cdots + zN_i$$

where N = molecular number

$N_1, N_2, N_3, \ldots N_i$ = respective atomic numbers of each element of the molecule.

$a, b, \ldots z$ = number of atoms of each element present in the molecule.

Example: Comparison of the ammonium and the sodium cations. The atomic number of nitrogen is 7 and that of hydrogen is 1. Thus the molecular number of the ammonium

FIGURE 15.3 Scaffold hopping can be considered as the broadest example of bioisosterism: structurally different molecules elicit the same biological activity.

cation can be calculated and compared to that of the sodium ion:

	Atomic number		Molecular number
NH_4^+	$7 + (4 \times 1)$	=	11
Na^+	11	=	11

Possessing the same molecular number, the ammonium cation should resemble the sodium cation. This is roughly true. More generally, two compounds with identical molecular numbers present at least some similar physical properties (e.g. specific heat).

B. The isosterism concept

Independently from Allen, Langmuir in 1919[4] defined the concept of isosterism:

> *"Comolecules are thus isosteric if they contain the same number and arrangement of electrons. The comolecules of isosteres must, therefore, contain the same number of atoms. The essential differences between isosteres are confined to the charges on the nuclei of the constituent atoms."*

On the basis of these similarities, Langmuir identified a list of 21 groups of isosteres.[5] Some of these are listed in Table 15.1. He further deduced from the octet theory that the number and arrangement of electrons in these molecules are the same. Thus, isosteres were initially defined as those compounds or groups of atoms that have the same number and arrangement of electrons. Then, he defined other relationships in a similar manner. Argon was viewed as an isostere of K^+ ion and methane as an isostere of NH_4^+ ion. He deduced that K^+ ions and NH_4^+ ions must be similar because argon and methane are very similar in physical properties. The biological similarity of molecules such as

TABLE 15.1 Groups of Isosteres as Identified by Langmuir

Groups	Isosteres
1	H^-, He, Li^+
2	O^{2-}, F^-, Ne, Na^+, Mg^{2+}, Al^{3+}
3	S^{2-}, Cl^-, Ar, K^+, Ca^{2+}
↓	↓
8	N_2, CO, CN^-
9	CH_4, NH_4^+
10	CO_2, N_2O, N^{3+}, CNO^-
↓	↓
21	SeO_4^{2-}, AsO_4^{3-}

CO_2 and N_2O was later coincidentally acknowledged as both compounds were capable of acting as reversible anesthetics to the slime mold *Physarum plycephalum*.[6]

The first example clearly demonstrates that isosterism does not inevitably imply "isoelectric" structures (having the same total electric charge), but it becomes evident that isoelectronic isosteres show the closest analogies:

C═O and N═N; CO_2 and NO_2; N═N═N^- and N═C═O^-

In the field of organic chemistry, Langmuir predicted the analogy between diazomethane and ketene, which was only discovered later.

FIGURE 15.4 The notion of pseudoatoms.

TABLE 15.2 Hydride Displacement Law: In Each Vertical Column the Atom is Followed by its Pseudoatoms

Number of electrons					
6	7	8	9	10	11
C^{4-}	N^{3-}	—O—	—F	Ne	Na^+
	CH^{3-}	—NH—	—OH	FH	
		$—CH_2—$	$—NH_2$	OH_2	
			$—CH_3$	NH_3	OH_3^+
				CH_4	NH_4^+

C. The notion of pseudoatoms and Grimm's hydride displacement law

Later on, in 1925, Grimm[7] formulated the "hydride displacement law" according to which the addition of hydrogen to an atom confers on an aggregate the properties of the atom of next highest atomic number. An isoelectronic relationship[8] exists among such aggregates which were named *pseudoatoms*. Thus, when a proton is "added" to the O^{2-} ion in the nuclear sense, an isotope of fluorine is obtained Figure 15.4).

When the same proton is introduced at the peripheral electronic level, a "pseudo-F," in other words an OH^-, is created. In this context, the H^+ ion having penetrated into the electronic shell of the oxygen is assumed to be masked by the greater atom and to exert only negligible effect toward the outside. The fluoride anion F^- and the hydroxyl anion OH^- show therefore some analogies. The generalization of the pseudoatom concept represents the so-called "hydride displacement law" proposed in a tabular form by Grimm[8] in each vertical column, the original atom is followed by its isosteric pseudoatoms (Table 15.2).

D. Erlenmeyer's expansion of the isosterism concept

Starting in 1932, Erlenmeyer[9] published a series of detailed studies about the isosterism concept, and particularly, about its first applications to biological problems. Erlenmeyer proposed his own definition of isosteres as

> *elements, molecules or ions in which the peripheral layers of electrons may be considered identical.*

Erlenmeyer also proposed three expansions of the isosterism concept:

1. To the whole group of elements present in a given column of the periodic table. Thus silicon becomes isosteric to carbon, sulfur to oxygen, etc.
2. To the pseudoatoms, with the aim of including groups which at a first glance seem totally different, but which, in practice, possess rather similar properties. This is the case for the pseudohalogens (e.g. Cl≅CN≅SCN, etc.)
3. To the ring equivalents: the equivalence between —CH═CH— and —S— explaining the well-known analogy between benzene and thiophene (Table 15.3).

TABLE 15.3 The Sulphur Atom is Approximately Equivalent to an Ethylenic Group (Size, Mass, Capacity to provide an Aromatic Lone Pair)

Compound	E°C	Isostere	E°C
Benzene	80°	Thiophene	84°
Methylbenzene	110°	2-Methyl-thiophene	113°
Chlorobenzene	132°	2-Chloro-thiophene	130°
Acetylbenzene	200°	2-Acetyl-thiophene	214°

E. Isoserism criteria: Present conceptions

The main criterion for isosterism is that two isosteric molecules must present similar, if not identical, volumes, and shapes. Ideally, isosteric compounds should be isomorphic and able to cocrystallize. Among the other physical properties that isosteric compounds usually share one can cite: boiling point, density, viscosity, and thermal conductivity. However, certain properties must be different: dipolar moments, polarity, polarization, size, and shape (e.g. in comparing F^- and OH^-, the size and the shape of H cannot be totally neglected). After all, the external orbital may be hybridized differently.

In conclusion, it became evident to the physicists that the concept of isosterism, developed before quantum-mechanical theories, could not provide at the molecular level the same results as those that the periodic classification

FIGURE 15.5 5,5′-Disubstituted oxazolidine-diones **1** and hydantoins **2** show similar antiepileptic profiles.

had provided for the elements, namely a correlation between electronic structure and physical and chemical properties. In the field of medicinal chemistry the isosterism concept, taken in its broadest sense, has proved to be a research tool of the utmost importance. The main reason for this is because isosteres are often much more alike in their biological than in their physical and chemical properties. An illustrative example is found in the comparison of oxazolidine-diones (**1**) and hydantoins (**2**) which possess different chemical reactivities but present a similar antiepileptic profile (Figure 15.5).

F. The bioisosterism concept: Friedman's and Thornber's definitions

Recognizing the usefulness of the isosterism concept in the design of biologically active molecules, Friedman[2] proposed to call *bioisosteres* compounds

> *"which fit the broadest definition of isosteres and have the same type of biological activity."*

This definition received rapid acceptance and is now commonly used. Moreover, Friedman considers that isosteres that exhibit opposite properties (antagonists) have also to be considered as bioisosteres, since usually they interact with the same recognition site. This is the case for para-aminobenzoic acid and para-aminobenzene-sulfonamide[10,11] and also for glutamic acid and its phosphonic analogs.[12,13]

The use of the word isosterism has been largely taken beyond its original meaning when employed in medicinal chemistry and Thornber[14] proposes a loose and flexible definition of the term bioisostere:

> *"Bioisosteres are groups or molecules which have chemical and physical similarities producing broadly similar biological effects."*

III. CURRENTLY ENCOUNTERED ISOSTERIC AND BIOISOSTERIC MODIFICATIONS

In 1970, Burger[15] classified and subdivided bioisosteres into two broad categories: classic and non-classic. Grimm's Hydride Displacement Law and Erlenmeyer's definition of isosteres outline a series of replacement which have been referred to as classical bioisosteres. Classical bioisosteres[5,16] have been traditionally divided into several distinct categories: (a) monovalent atoms or groups; (b) divalent atoms or groups; (c) trivalent atoms or groups; (d) tetravalent atoms, and (e) ring equivalents (Table 15.4).

TABLE 15.4 Classic Bioisostere Atoms and Groups

Monovalent	Divalent	Trivalent	Tetravalent
—OH, —NH_2, —CH_3, —OR	—CH_2—	=CH—	=C=
—F, —Cl, —Br, —I, —SH, —PH_2	—O—	=N—	=Si=
—Si_3, —SR	—S—	=P—	=N^+=
	—Se—	=As—	=P^+=
	—Te—	=Sb—	=As^+=
			=Sb^+=

Some non-classical isosteres are reported in Table 15.5 and from a brief glance it can be noticed that they do not obey the steric and electronic definition of classical isosteres. A second notable characteristic of non-classical bioisosteres is that they do not have the same number of atoms as the substituent or moiety for which they are used as a replacement.

As the distinction between isosteres and bioisosteres is rather of academic interest, it is preferred, in this chapter to treat both categories together. Consequently, for example, divalent series such as O=, HN=, and H_2C= can be discussed together with S=. However, the correct nomenclature will be used as much as possible, keeping in mind nevertheless that "isosteric replacement" embraces both true isosteres and bioisosteres.

A. Replacement of univalent atoms or groups

Halogens (particularly chlorine) can be replaced by other electron-attracting functions such as trifluoromethyl or cyano groups. In the antibiotic chloramphenicol, both the chlorine atoms of the dichloroacetic moiety and of the *para*-nitro-phenyl group yielded productive isosteric replacements (Table 15.6). Many other examples of univalent atoms or groups replacements are found in the chapter dealing with substituent effects (Chapter 20) and with quantitative structure–activity relationships (Chapter 23).

B. Interchange of divalent atoms and groups

A first series of frequently interchanged divalent atoms or groups is represented by O, S, NH, and CH_2 and many

TABLE 15.5 Non-Classical Isosteres

—CO—	—COOH	—SO_2NH_2	—H	—CONH—	—COOR	—$CONH_2$
—CO_2—	—SO_3H	—PO(OH)NH_2	—F	—NHCO—	—ROCO—	—$CSNH_2$
—SO_2—	—tetrazole					
—SO_2NR—	—SO_2NHR —SO_2NH_2		—OH —CH_2OH		—catechol	
—CON—	—3-hydroxyisoxazole				—benzimidazole	
—CH(CN)—	—2-hydroxychromones		—$NHCONH_2$			C_4H_4S
R—S—R			—NH—CS—NH_2			—C_5H_4N
(R—O—R)	═N—					—C_6H_5
R—N(CN)—	C(CN)═R′		—NH—C(═$CHNO_2$)—NH_2 —NH—C(═CHCN)—NH_2			
—halide						—C_4H_4NH
	—CF_3					
	—CN					
	—$N(CN)_2$					
	—$C(CN)_3$					

TABLE 15.6 Isosteric Replacements in the Amphenicol Family

Compound	X	Y
Chloramphenicol	—NO_2	—CH—Cl_2
Thiamphenicol	CH_3—SO_2—	—CH—Cl_2
Cetophenicol	CH_3—CO—	—CH—Cl_2
Azidamphenicol	—NO_2	—CH_2—N_3

TABLE 15.7 Meperidine Analogs[17]

Meperidine

X	Analgesic potency (Meperidine = 1)
O	12
NH	80
CH_2	20
S	1,5

interesting examples are found in the literature. In a study on meperidine analogs (Table 15.7) potent analgesic compounds were found for X = O, NH, and CH_2.[17] Surprisingly, the sulfur analog showed only moderate activity. As an *in vivo* test was used to assess the activity, the weaker effect may be attributable to a faster metabolism (sulfoxide or sulfone formation?).

Similar changes can be applied to cyclic series, for example, to a series such as piperidine–morpholine–thiomorpholine–piperazine or in introducing oxygen or sulfur atoms into cyclic ketoprofen analogs.[18]

Nice isosteric variations were observed in a series of thermolysin inhibitors.[19] For these isosteres the replacement of the phosphonamide (X = NH) function by a phosphonate (X = O) or a phosphinate (X = CH_2) function demonstrated clearly that the maximal activity was associated with the phosphonamide which is able to establish a hydrogen bond with alanine 113 (Figure 15.6).

In order to have a more stable analog of the acetylcholinesterase inhibitor alkaloid physostigmine, Chen *et al.*[20] prepared some 8-carbaisosteres of physostigmine (Table 15.8). The authors envisioned that replacing the *N*-methyl group at N_8 of the physostigmine nucleus by a methylene group would increase its chemical and metabolic stability, thanks to the change of the less stable aminal group to a more stable amino group.

The carbaisosteres are as potent as or even more potent than the corresponding physostigmines. In addition, the (−)-enantiomers which possess the same absolute configuration at C_{3a} and C_{8a} as that of physostigmine are generally 6 to 12 times more potent in inhibiting acetylcholinesterase than the corresponding (+)-enantiomers.

Other interesting bioisosteric replacements of oxygen atoms were devised during the search of non-hydrolyzable phosphotyrosyl (pTyr) mimetics. The phosphoryl ester oxygen of pTyr has been replaced either by a methylene (Pmp)[21] or by a difluoromethylene (F_2Pmp).[22]

FIGURE 15.6 Isostery in thermolysin inhibitors.[19]

C. Interchange of trivalent atoms and groups

The substitution of —CH═ by —N═ in aromatic rings has been one of the most successful applications of classical isosterism (see Section III.D.). Interchange of trivalent atoms are found also in non aromatic rings. For example the 4-dimethylamino-antipyrine and its carba-isostere are about equally active as antipyretics[23](Figure 15.7).

FIGURE 15.7 Interchange of trivalent atoms and groups.

TABLE 15.8 Physostigmines (a) and Carbaisosteres (b)[20]

Compound	R1	R2	IC_{50} (nM)	LD_{50} (mg/kg)
(−)-Physostigmine	CH_3	CH_3	128	0.88
(−)-Heptyl physostigmine	n-C_7H_{13}	CH_3	110	24
(±)-Carba-isostere 1	n-C_7H_{13}	CH_3	114	21
(−)-Carba-isostere 2a	n-C_7H_{13}	C_2H_5	36	6
(+)-Carba-isostere 2b	n-C_7H_{13}	C_2H_5	211	18

Similar interchanges are found in the proceeding from desipramine to nortriptyline and protriptyline (Figure 15.7) or among the antihistaminics, when comparing tripelennamine with chlorpheniramine (Figure 15.7).

D. Ring equivalents

The importance of chemical rings in drug discovery is never overstated. Looking at the best selling drugs in 2006 reported in Table 15.9, it can be noticed that 8 out of 10 molecules are small molecules and that all of them have a chemical ring in their structure. This trend is reflected also by the other marketed drugs and by the ensemble of bioactive molecules. In bioactive molecules, rings have a multiple role. Rings are responsible of the basic shape of the molecule rendering it rigid or flexible and are also responsible for a suitable spatial orientation of the pharmacophoric groups. In other molecules, chemical rings are directly responsible of the biological activity as they interact directly with the receptors either *via* heteroatom forming hydrogen bonding either *via* hydrophobic interactions.

The ADME (absorption, distribution, metabolism, excretion) profile as well as the toxicity of biologically active molecules can rely on the nature of a given ring: such

TABLE 15.9 Best Selling Drugs in 2006

Structure	Leading brands	Indication	Company	2006 sales (US$BN)
	Lipitor (atorvastatin)	High cholesterol	Pfizer	13.6
	Nexium (esomeprazole)	Gastroesophageal reflux	Astra-Zeneca	6.7
	Seretide/Advair (fluticasone + salmeterol)	Asthma	GlaxoSmithKlin	6.3
	Plavix (clopidogrel)	Coronary artery disease	Sanofi-Aventis Brustol-Myers Squibb	5.8

(Continued)

TABLE 15.9 (Continued)

Structure	Leading brands	Indication	Company	2006 sales (US$BN)
	Norvasc (amlodipine)	Hypertension	Pfizer	5
165-Amino-acid-glycosylated protein hormone which contains 5 N-linked oligosaccharide chains	Aranesp (darbepoetin alfa)	Anemia	Amgen	5
	Zyprexa (olanzapine)	Schizophrenia	Eli Lilly	4,7
	Ripserdal (risperidone)	Antipsychotic	Janssen Pharmaceutica	4.6
Recombinant human soluble tumor necrosis factor-alpha (TNF-α) receptor protein	Enbrel (etanercept)	Autoimmune diseases	Amgen-Wyeth	4.5
	Effexor (venlafaxine)	Depression	Wyeth	4

as the hydrophobicity, the polarity, the electronic properties, etc. Medicinal chemists use to change one ring with a bioisosteric one in order to improve the ADME-tox profile of a given molecule keeping the desired activity and selectivity.

In the conception of "me too" drugs as well as in getting a stronger patent position, bioisosteric replacement of ring systems has played a master role in drug design. The substitution of —CH═ by —N═ or —CH═CH— by —S— in aromatic rings has been one of the most successful applications of classical isosterism. Early examples are found in the sulfonamide antibacterial with the development of sulphapyridine, sulphapyrimidine, sulphathiazole, etc. (Figure 15.8).

Other examples are found in the neuroleptic or antidepressant tricyclics, in the benzodiazepine tranquilizers and antiepileptics, and in the development of semi-synthetic penicillins and cephalosporins with broader spectra of activity and greater stability toward β-lactamases. In some instances. an aromatic ring can be replaced by an ethynyl group. In some instances an aromatic ring can a replaced by an ethyln group. Such a replacement was reported by Wallace et al.[24] in a series of inhibitors of Endothelin-converting Enzyme (ECE-1) deriving from the biphenyl compound CGS 26 303 and its analogs (Figure 15.8).[25]

In all these cases no essential activity difference is found between the original drug and its isostere. However,

FIGURE 15.8 Classical ring equivalents.

FIGURE 15.9 The thiophene isostere of zonisamide is practically inactive as an anticonvulsant.

it can happen that the procedure fails. Binder *et al.*,[26] for example, reported that thieno[2,3-*d*]isoxazole-3-methanesulfonamide, the thiophene analog of the anticonvulsant drug zonisamide (Figure 15.9),[27] was practically inactive against pentetrazole- or electroshock-induced convulsions in mice, even at high doses.

Because of the paramount role of ring systems in the drug discovery process a lot of attention has been given to the equivalence between rings. Indeed the efforts to replace rings or other functional groups in an active molecule form a large part of medicinal chemistry practice devoted also to extend the space of bioactive molecules in the chemistry universe. A typical lead optimization program often involves a trial and error process of replacing chemical functionality including rings and in most cases this intellectual exercise is based on a chemist's knowledge and recollection of his or her past experience. As a source of inspiration in the following paragraphs are listed some examples of bioisosters of the most encountered rings in biologically active molecules.

FIGURE 15.10 Ligands for central cholinergic receptors with different non-classical bioisosteres of the pyridine ring.

1. *Bioisosteres of pyridine*

One of the most used cycles in medicinal chemistry is the pyridine ring. The bioisosteres of this heterocycle are well known as ligands for the central nicotinic cholinergic receptors. The pyridine ring of nicotine can be replaced by different other rings like methyl-isoxazole or methyl-isothiazole[28–30] (Figure 15.10a). A novel series of nicotinic agonists was described by Olesen *et al.*[30] In their paper, the bioisosteric replacement of the isoxazole ring in the (3-methyl-5-isoxazoly)methylene-azacyclic compound (**3**, Figure 15.10b) with pyridine, pyrazine, oxadiazole, or an acyl group resulted in ligands with moderate to high affinity

for the central nicotinic cholinergic receptors (IC_{50} = 2.0 to IC_{50} > 1,000 nM) (Figure 15.10b).

Two other publications on the bioisosteric replacement of pyridine show similar results. The first paper reports the study of bioisosteric potential of diazines in the field of combined antithrombic thromboxane A_2 synthetase inhibitors and receptor antagonists.[31,32] On the basis of the structure–activity relationships (SAR) observed in this study, it turned out that only the 2-pyrazinyl, 4-pyridazinyl and 5-pyrimidinyl systems are appropriate bioisosteric moieties for the 3-pyridyl system in the dual active platelet antiaggregatory compound Ridogrel (Figure 15.11). Gohlke *et al.*[31] also observed the bioisosteric potential of diazines in the SAR of the compound DUB-165. The replacement of the 3-pyridyl group of DUB-165 by a 4-pyridazinyl, 5-pyrimidinyl, or 2-pyrazinyl moiety, resulted in ligands retaining affinity for nAChRs subtypes, thus demonstrating that the three isomeric diazines are appropriate bioisosteres of the 3-pyridyl moiety (Figure 15.11).

FIGURE 15.11 Bioisosteric moieties for the 3-pyridyl ring.

N-(6-Chloronaphthalen-2-)sulfonylpiperazine derivatives **4** and **5** (Figure 15.12) are potent factor Xa inhibitors. Haginoya *et al.*[33] proposed to replace the pyridine-phenyl or the pyridine-piperidine residue by a fused-bicyclic ring which contains an aliphatic amine and a pyridine to yield the compound **6** that has an interesting factor Xa inhibitor activity. The bioisosteric replacement of the pyridine moiety of the 6-methyl-5,6,7,8-tetrahydro-[1,6]naphthyridine by phenyl, thiophene, or thiazole analogs yielded analogs with similar or better antifactor Xa activity, but also to conserve a moderate bioavailability.

A nice example of pharmacophore equivalent of the 2,3-diaminopyridine with cyclopropylamino acid amide has been reported recently.[34] A series of 2,3-diaminopyridines represented by **7** acting as bradykinin (BK) B_1 antagonists, was thoroughly investigated with the aim to replace the metabolically labile diaminopyridine ring with a more bioavailable moiety. An accurate design strategy led researchers

FIGURE 15.12 Bioisosteric replacements of the pyridine ring in a series of factor Xa inhibitors.

at Merck to replace the 2,3-diamino pyridine with a series of alkylamino acid amides. Among this series compound **8**, though less active than **7** as BK B1 antagonist, showed a better pharmacokinetic profile with an improved bioavailability, an increased half-life, and a decreased clearance compared to **7** (Figure 15.13).

To demonstrate the generality of this bioisosteric equivalence, the same group applied the same substitution to **9**, a factor Xa inhibitor that yielded **10**. Compound **10** reflects the same loss in potency that was observed for **8** towards **7**, but reoptimization of the *N*-acyl group afforded quickly a significant gain of potency.

2. *Imidazo*[1,2-*a*]*pyridine bioisosteres*

Abe *et al.* [35] reported that the imidazol [1,2-a]pyridine moiety of the basic framework of a class of the non-peptide bradykinin B_2 receptor antagonists (**11**, Figure 15.14) could be successfully replaced by several heterocyclic bioisosteres. Among those, the 1-methyl-2-methoxy-1*H*-benzimidazole, 2-methylquinoxaline and 2-methylquinoline derivatives showed potent B2 binding affinities against both human and guinea pig B2 receptors (Figure 15.14).

3. *Pyridazine bioisosteres*

Since the antidepressant minaprine has been launched, several pyridazine analogs were proposed and synthesized. Most of them showed different activities on the central nervous system. A series of these compounds (Figure 15.15) are acetylcholinesterase inhibitors with variable bioisosteres of the central pyridazine.[36,37] The replacement by pyridine, 1,2,4-thiadiazole and triazines yields compounds with weaker but still acceptable activity.

Examples from literature show that there are several non-classical bioisosteres for the pyridine system or the pyridazine system. Which is the rationale behind this in order to choose the most appropriate analog to start with? An indication could be given by the comparison of the boiling points

FIGURE 15.13 Cyclopropylamino acid amide as pharmacophoric replacement for 2,3-diaminopyridine. Application to the design of novel bradykinin B_1 receptor antagonist and factor Xa inhibitor.

FIGURE 15.14 Novel class of orally active non-peptide bradykinin B_2 receptor antagonists showing valuable examples of imidazo[1,2-a]pyridine analogs.

FIGURE 15.15 *In vitro* inhibition of acetylcholinesterase in rat striatum homogenates: pyridine, thiadiazole, and triazine replacement of the pyridazine ring.

of these heterocycles assuming that the most similar the boiling point is, the most appropriate the bioisostere is (Figure 15.16).[38] For example, while searching for a bioisostere of the pyridine ring, looking at the boiling points it appears that the best candidates are the pyrimidine, the pyrazine, and the 1,2,4-thiadiazole systems (Figure 15.16). The same could be extended to the pyridazine ring where this comparison permits the selection of the 1,2,4-triazine or the 1,3,4-thiadiazole rings (Figure 15.16). In fact, those findings confirm the examples found in the literature and reported in the previous part of this chapter about isosteres or bioisosteres of the pyridine and of the pyridazine ring.

A possible interpretation of these results can be the fact that in the heterocyclic series, the boiling point is correlated to the dipolar moment of the molecule and that, for two heterocyclic rings having the same aromatic geometry, the similarity of the dipolar moments may represent the dominant feature.

FIGURE 15.16 Structures and boiling points of pyridine and pyridazine isosteres.

4. *Bioisosteres of other heterocycles*

Chemical rings are fundamental for drug molecules as a typical drug molecule consists of a combination of chemical rings, chains, and functional groups. The high proportion of chemical rings in drug molecules talks in favor of the large effort to synthesize and find new chemical ring equivalents, whether it is to design a new platform for a better patent position, replace a metabolically unstable moiety, look for a more favorable receptor interaction or have a better pharmacodynamic profile.

There are many examples of ring equivalents and so far it seems impossible to rationalize a method that could give a reliable indication which heterocycle could be the best substitute of another one. The best approach is to look at the examples reported in the literature and learn from those.

Selective cyclooxygenase-2 inhibitors (COX-2 inhibitors) give a nice example of bioisosters of heterocycles. The comparison of the most potent selective COX-2 inhibitors (Figure 15.17) suggests that isoxazoles, pyridines, and pyrazoles are good bioisosteres of each other as well as nitrophenol, and indanones.[39]

FIGURE 15.17 Examples of selective cyclooxygenase-2 inhibitors showing different ring equivalents.

Another nice example of heterocycle bioisosteres is given by the oxazolidinone antibacterials. Since the discovery of linezolid as potent antibacterial against Gram positive organisms,[40] many researchers have taken inspiration from the oxazolidinone scaffold in order to create new and potentially improved antibacterial agents. The structure of Lienzolid (Figure 15.18) is divided in three parts: the A ring corresponding to the oxazolidinone moiety, the B ring coresponding to the Phenyl ring and the C ring corresponding to the morpholine. Some research groups have worked around the modification of the B and C rings.[41,43]

The work of Snyder *et al.*, together with other examples found in the literature, shows that several level of activity can be identified amongst the different isosteres. It has to be kept in mind that the replacement of a ring by its isosteres does not always lead to an iso-active or more active compound (Figure 15.18). But in general, it is possible to obtain at least one or more equivalent systems.[42]

Another particularly interesting example of ring bioisosterism is found in the development of the antiulcer H2-receptor histamine antagonists in which the initial imidazole ring was changed to various other "equivalents" such as a furan, a thiazole, and finally a phenyl ring (Figure 15.19). A detailed and very interesting account of the discovery and the development of these compounds is found in Ganellin and Roberts' book.[44]

Better bioisosteric design possibilities are provided by quantum-chemical calculations. Mallamo *et al.*[45] made use of electrostatic potential surface maps complementarity in defining sulfonyl heterocycles bioisosteric to the steroidal antiandrogenic drug zanoterone (Figure 15.20). Striking differences in the electrostatic potential surfaces accounted for the observed variability in the furan (active) and the thiophene (inactive) analogs of zanoterone. Good androgen receptor affinity was then anticipated and effectively found for the oxazole and the thiazole analogs of zanoterone.

The apparent failure of the isosterism concept for the inactive thiophene, inversed furan and pyrimidine is thus interpretable on a rational basis. In Table 15.10 are listed some more "exotic" examples of bioisosteric replacements of cyclic systems.

E. Groups with similar polar effects: functional equivalents

1. *Carboxylic acid bioisosteres*

Medicinal chemists have frequently to face the problem of developing surrogates for carboxylic acid groups. The availability of bioisosteric replacement for the carboxylic acid group has been critical to the development of novel medicinal agents especially in the area of neurochemistry. Less polar and more hydrophobic groups were needed to replace the carboxylic group of GABA in order to cross the blood–brain barrier, but keeping some of the physiochemical properties of the parent functional group like pK_a, partition coefficient, etc.

Thus, the two GABAergic agonists, isoguvacine and THIP (Figure 15.21), have similar pharmacological properties

FIGURE 15.18 Oxazolidinone bioisosteres synthesized as antibacterial agents.

FIGURE 15.19 Antiulcer H_2-receptor histamine antagonists: evolution of structures in the course of the time. Note the progressive use of a furan, a thiazole, and finally a phenyl ring in place of the original imidazole ring. On the same series it is noteworthy the bioisosterism of the urea moiety with different surrogates: thiourea, *N*-cyanoguanidine, *N*-nitro-ethene diamine, and *N*-aminosulfonil guanidine.

FIGURE 15.20 Zanoterone isosteres.[45]

to GABA itself. The key parameters in these compounds are the acidic ($pK_a < 4$) and the basic (protonated nitrogen) functions with a <5.1 Å intercharge distance.

For the carboxylic function of active compounds, three different classes of bioisosters can be distinguished: (a) the direct derivatives (Table 15.11); (b) the planar acidic heterocycles (Table 15.12); and (c) the non-planar sulfur- or phosphorus-derived acidic functions (Table 15.5).

a. Direct derivatives

Direct derivatives of carboxylic acids are different functional groups which maintain an acidic proton and a hydrogen bond acceptor (the carbonyl) in order to have similar specific interactions with the receptor. Among the most common direct derivatives can be listed: the hydroxamic acids (R—CO—NH—OH), the acylcyanamides (R—CO—NH—CN) and the acylsulfonamides (R—CO—NH—SO_2—R′) (Table 15.11).

Hydroxamic acids: The hydroxamic acids have a high chelating power and they have found a wide application as inhibitors of enzymes having a metal ion like Zn^{2+} in the active site. Examples note of worth are found in the inhibitors of matrix metalloproteinases (MMP),[65–67,76] tumor necrosis factor α converting enzyme (TNFα),[69] and histone deacetylase (HDAC)[59,60] where the hydroxamic acid derivatives have given potent and bioavailable compounds despite this group is particularly prone to hydrolysis, reduction, and glucoronidation.[68] Another interesting example of exploitation of the hydroxamic acid is given by the antiinflammatory hydroxamates bufexamac,[77] ibuproxam,[78] and oxametacin[79] (Figure 15.22) where the hydroxamates can act either as a prodrug-like ibuproxam being metabolized to ibuprofen (CONHOH → COOH) in man,[80] either as a true bioisostere like oxametacin being metabolically stable in man.[81,82]

Acylsulfonamides: Acylsulfonamides are also quite used in drug discovery and have been employed in several therapeutic areas such as β_3 adrenergic receptor agonist,[74]

TABLE 15.10 Other Ring Bioisosteres

Original ring	Bioisostere	Activity	Reference
Indole	Indazole	5-HT_3 antagonists	Fludzinski *et al.*[46]
Indole	Furo[3,2-b] pyridine	5-HT_{1F} agonist	Mathes *et al.*[47]
Indole	6H-thieno[2,3-b] pyrrole	Hallucinogen serotonin agonist	Blair *et al.*[48]
Indole	Napthalene	Agonist at melatonin receptors and antagonist at $5HT_{2c}$	Yous *et al.*,[49] Depreux *et al.*[50]
3,4-Dialkoxy-phenyl	Indole	Phospho diesterase inhibitors	Blaskó *et al.*[51]
		GABA uptake inhibitors	Kardos *et al.*[52]
Quinoline-2-carboxylate	Indole2-carboxylate	Glycine antagonists	Salituro *et al.*[53]
o-Nitro-phenyl	Furoxane	Calcium antagonist	Calvino *et al.*[54]
Spiro-hydantoin	Spiro hydroxyacetic acid unit	Aldose reductase inhibitor	Lipinski *et al.*[55]
Thiazol-2-ylamine	Pyrazolo[1,5-a] pyridine	Dopamine D_3 receptor agonist	Löber *et al.*[56]

chemokine receptor 1 (CXCR1) inhibitors,[83] hepatitis C virus (HCV) inhibitors,[75] angiotensin receptor agonists,[84] antitumoral compounds with antiproliferatives properties,[85,86] and prostaglandin EP4 receptor antagonists.[87]] Acylcyanamides are mainly of academic interest.

b. Planar acidic heterocycles and aryl derivatives

In Table 15.12 are reported the planar acidic heterocycles and aryl derivatives that are commonly used as carboxylic acid bioisosters. The most used is the tetrazole ring which has found wide application in different therapeutic fields spanning from type 2 diabetes[98,99] to hepatitis C virus (HCV),[75] malaria,[100] Alzheimer's disease (AD),[101] anxiety treatment,[102] and pain management.[103] The tetrazole has been employed to fix different issues, for example to improve bioavailability,[100] to enhance the blood–brain barrier penetration,[102,103] to increase potency,[99,104] to get a better chemical stability,[101] to bring some selectivity (the GABA tetrazole analog inhibits GABA-transaminase, but not succinic semialdehyde dehydrogenase)[105] or to be used as a prodrug.[106] However, in some instances, tetrazole analogs are poorly active.[107]

FIGURE 15.21 An example of bioisosterism, or non-classical isosterism: GABA, isoguvacine, and THIP are all agonists for the GABA-A receptor. The 3-hydroxy-isoxazole ring has a comparable acidity to that of a carboxylic acid function.[57]

Hydroxy-isoxazoles and other cognate heterocyclic phenols encompassing an acidity range from 3.0 to 7.1 were incorporated in GABA agonists, antagonists, and uptake inhibitors.[108,109] The experience gained with 3-hydroxy-isoxazoles in the GABA field was also transferable to glutamate receptor ligands and led to selective antagonists for glutamic acid receptor subtypes.[108]

Thiazolidinediones are another class of heterocyclic carboxylic acid surrogates commonly used for peroxisome proliferator-activated receptors (PPAR) agonists,[93,94] as potent antihyperglycemic and lipid activity modulators. Other interesting, but less studied heterocyclic surrogates are: 3,5-dioxo-1,2,4-oxadiazolidine,[110] 3-hydroxy-1,2,5-thiadiazoles,[91] 1,2,4-oxadiazole-5(4*H*)-ones,[95] 1,2,4-thiadiazole-5(4*H*)-ones[95], 3,5-difluoro-4-hydroxyphenyl,[96,97] and 3-hydroxy-γ-pyrones.[92,111]

c. Non-planar sulfur- or phosphorous-derived acidic functions

The most extensive use of phosphonates (Table 15.13) occurred in the design of amino acid neurotransmitter antagonists such as glutamate[114] and GABA-B antagonists.[112]

Amongst a set of cholecystokinin (CCK) antagonists derived from the non-peptide CCK-B selective antagonist CI-988, Drysdale *et al.*[115] prepared a series of carboxylate surrogates spanning a pK_a range from <1 (sulfonic acid) to >9.5 (thio-1,2,4-triazole). The affinity and the selectivity of the compounds were rationalized by considering pK_a values,

TABLE 15.11 Carboxylic Acid Isosteres: Direct Derivatives

	Hydroxamic acids	High chelating power Histone deacetylase inhibitors Matrix metalloproteinases inhibitors Tumor necrosis factor α converting enzyme	Almquist *et al.*[58] Massa *et al.*,[59] Lu *et al.*[60] Remiszewski *et al.*,[61] Plumb *et al.*,[62] Kelly *et al.*,[63] Buggy *et al.*[64] Hanessian *et al.*[65] Aranapakam *et al.*[66,67] Noe *et al.*[68] Duan *et al.*[69]
	Acyl-cyanamides	Mainly academic interest	von Kohler *et al.*[70] Kwon *et al.*[71]
	Acyl-sulfonamides	Glycine, GABA, and β-alanine analogs Antiatherosclerotics pK_a # 4,5 β_3 Adrenergic receptor agonist hepatitis C virus	Drummond and Johnson[72] Albright *et al.*[73] Uehling *et al.*[74] Johansson *et al.*[75]

TABLE 15.12 Carboxylic Acid Isosteres: Planar Acidic Heterocycles and Aryl Derivatives

	Tetrazoles	Very popular, great number of publications. pK_a = 6.6 to 7.2	Bovy *et al.*[88] Marshall *et al.*[89]
	Mercaptoazoles + Sulfinylazoles + Sulfonylazoles	Phosphonate isosteres pK_a mercapto: 8.2–11.5 pK_a sulfinyl: 5.2–9.8 pK_a sulfonyl: 4.8–8.7	Chen *et al.*[20]
 X = O or S	Isoxazoles Isothiazoles	GABA and glutamic acid analogs	Krogsgaard-Larsen *et al.*[57] Krogsgaard-Larsen[90]
	Hydroxy-thiadiazole	Isoxazole isostere pK_a # 5	Lunn *et al.*[91]
	Hydroxy-chromones	Kojic acid derivatives (as GABA agonists)	Atkinson *et al.*[92]
	Thiazolidinediones	Dual PPAR α/γ agonists	Hulin *et al.*,[93] Henke[94]
	1,2,4-Oxadiazole-5(4H)-ones	Antimycobacterial	Gezginci *et al.*[95]
	1,2,4-Thiadiazole-5(4H)-ones	Antimycobacterial	Gezginci *et al.*[95]
	3,5-Difluoro-4-hydroxyphenyl	Aldose reductase inhibitor GABA analog	Nicolaou *et al.*[96] Qiu *et al.*[97]

charge distribution, and geometry of the respective acid mimics (Table 15.14).

In order to choose the best carboxylic acid bioisoster pK_a and log *P* are important parameters to look at before proceeding with one surrogate or another. An interesting study comparing pK_a and log *P* values of some aryl phosphonic acids, aryl tetrazoles, and aryl sulfonamides has been published by Franz.[116] The values of pK_a and log *P* are very important parameters in drug design. Several published examples show that the interchange of carboxylic acid for tetrazole and sulfonamide often results in useful drugs. This study indicates that the log *P* will be lowered by about 1 log unit substituting a phosphonic acid by a carboxylic acid group. There is an increase in acidity which may be a limiting factor in the absence of active transport. Replacing a carboxylic acid group by a phosphonic acid group gives a much more acidic compound; replacing it with a sulfonamide group results in a much less acidic compound; and with a tetrazole replacement, acidity is essentially unchanged (Figure 15.23). If one wants to lower the log *P*, and if increased acidity of the compound was not a limiting factor, substitution of a phosphonic acid group for a carboxylic acid would be a viable approach (Figure 15.24).

FIGURE 15.22 Hydroxamate isosteres of antiinflammatory drugs.

Diamino-cyclobutene-diones as α-amino carboxylic acid surrogates: Diamino-cyclobutene-dione was proposed by Kinney *et al.*[117] as an original surrogate of the α-amino carboxylic acid function (Figure 15.25).

Carboxylic functions as phosphonates surrogates: pTyr mimetics serve as important components of many competitive protein-tyrosine kinases inhibitors. To date, the most potent of these inhibitors have relied on phosphonate-based structures to replace the 4-phosphoryl group of the parent pTyr residue (Figure 15.26). Interestingly, it was found that carboxy-based pTyr analogs can be utilized to introduce the anionic oxygen functionality of the parent phosphate. Particularly, when *p*-(2-malonyl)phenylalanine (Pmf) was incorporated as pTyr replacement in the high-affinity Grb2 SH2 domain binding sequence, potencies approaching one of phosphonate mimetics were obtained.[118]

The above example is an elegant illustration of the possibility to mimick the pyramidal structure of phosphates or phosphonates by means of two planar carboxylic groups, the three-dimensionality originating from the malonic methylenic carbon atom.

TABLE 15.13 Carboxylic Acid Isosteres: Non Planar Sulfur- or Phosporus-Derived Acidic Functions

	X = H X = OH X = NH_2 X = $CH(OR)_2$	Phosphinates Phosphonates Phosphonamides	Many examples in glutamate and in $GABA_B$ antagonist series	Froestl *et al.*[112]
		Sulfonates	Sulphonic analogs of GABA and glutamic acid	Rosowski *et al.*[113]
		Sulfonamides	Weak acids, used rather as equivalents of phenolic hydroxyls in the design of catecholamine analogs	von Kohler *et al.*[70]

TABLE 15.14 Exploration of the Carboxyl Isosterism Possibilities in a Series of CCK Antagonists.

R	IC_{50} (nM) CCK-B	IC_{50} (nM) CCK-A	A/B ratio	pK_a
R—CH_2—COOH	1.7	4500	2500	5.6
Charge distributed monoanionic acid mimics				
	6.0	970	160	5.4
	2.6	1700	650	6.5
	2.4	620	260	4.3
	2.5	680	270	>9.5
	16	850	53	>9.5
	4.3	660	150	7.7
	1.7	940	550	7.0
	6.3	1300	200	5.2

(Continued)

TABLE 15.14 (Continued).

Charge distributed monoanionic acid mimics				
	18	600	33	>8.2
	14	1300	93	>9.5
Point charge monoanionic acid mimics				
	70	300	4.3	>9.5
	77	680	9	7.9
	110	790	7	>9.5
	80	510	6.4	>9.5
	21	1500	71	>9.5
Tetrahedral acid mimics				
$P(O)(OH)_2$	27	5200	190	3.4; 7.7
CH_2—$P(O)(OH)_2$	23	2700	120	3.4; 7.8
P(O)(OH)(OEt)	12	480	40	6.5
P(O)(OH)Me	12	1700	140	3.8
CH_2—P(O)(OH)Me	23	4400	190	3.7
CH_2—SO_3Na	1.3	1010	780	–

2. *Carboxylic esters bioisosteres*

Carboxylic esters are nice tools for SAR studies investigation but they are not valuable group to be left in drug molecules because they lack the proper of *in vivo* stability. As matter of fact they are hydrolized quite quickly by the esterases which are quite ubiquitous and are present in the blood, the liver, the kidneys, and other organs. To overcome this problem the carboxylic ester in the most promising compounds has been replaced by bioisosteric groups leading to molecules with improved pharmacokinetic profile.

One of the first examples of ester bioisoster is the amide bond as it is illustrated by the local anesthetic procaine where the readily hydrolyzable ester group was replaced with an amide to give procainamide as more stable analog (Figure 15.27).

FIGURE 15.23 pK_1 values for aromatic compounds (redrawn after Franz[116]).

FIGURE 15.24 Log P values for aromatic compounds (redrawn after Franz[116]).

FIGURE 15.25 3,4-Diamino-3-cyclobutene-1,2-dione as surrogate of the α-amino carboxylic acid function.

Similarly, the lactone ring of the muscarinic agonist pilocarpine was changed into various, still active isosteres such as the corresponding thiolactone, lactam, lactol, and thiolactol.[119] A series of aspirin isosteres has been prepared by replacing the carboxylic ether oxygen successively by a nitrogen, sulfur, or carbon isosteric equivalent.[120] None of the isosteric compounds showed any activity. This result is readily understood since the particular role of aspirin as an acylating agent of the enzyme cyclooxygenase has been demonstrated.[121]

Pilocarpine is widely used as a topical miotic for controlling the elevated intraocular pressure associated with glaucoma. Beside its low lipophilicity, which stimulated the search for prodrugs,[122] pilocarpine has a short duration of action, its lactonic ring being rapidly opened to yield pilocarpic acid. In pilocarpine, by substituting the lactonic ester function by its carbamate equivalent, a much more stable analog, which is as effective as pilocarpine, was obtained.[123]

In addition to these classical changes more exotic groups and radicals have been proposed and tested as valuable carboxylic ester surrogates like five-membered heterocycles and alkyl oximes. Studies about the design and the synthesis

FIGURE 15.26 Malonates as surrogates of phosphonates.

FIGURE 15.27 Amide as an ester surrogate of the local anesthetic procaine.

of more metabolically stable muscarinic ligands have shown that it is possible to replace the methyl ester of arecoline **12** and of the quinuclidine **13** with a series of five-membered rings like oxadiazoles, thiadiazoles, triazoles, and tetrazoles to give potent muscarinic agonist (Figure 15.28).[124–130]

Note of worth is also the successful replacement of the methyl ester with oxime ethers and *N*-methoxy imidoyl nitrile in the design of muscarinic agonists (Figure 15.28).[131,132]

Several 1,2,4-oxadiazoles and the other five-membered heterocycles are employed also for other therapeutic indications like monoamine transporter and opioid receptors,[133,134] 5-HT agonists,[135] bradykinin B_1 receptor antagonists,[136] and vascular endothelial growth factor receptor (VEGFR-2) inhibitors.[137]

Recently it has been published a nice example of the use of the methyl ester bioisosters in the design of a series of metabolically stable HIV inhibitors (Figure 15.29).[138,139] Starting from an alkenyldiarylmethane containing three methyl esters, the systematic substitution of on ester at the time with the appropriate biomimetic group let to a molecule with an enhanced metabolic stability in rat plasma ($t_{1/2}$ = 61 hour) along with the ability to inhibit HIV-1 reverse transcriptase (Figure 15.29).

A confirmation of the alkyl oxime ether as metabolically stable isoster of an ester has been reported by Watson *et al.* in the synthesis of a capsid binder active against Human Rhinovirus[140] (Figure 15.30).

The change in (−)-cocaine of the carbomethoxy substituent into carbethoxyisoxazole doubles the potency in [^{3}H] mazindol binding and [^{3}H] dopamine (DA) uptake (Figure 15.31). Astonishingly the replacement of the carbomethoxy group by a chlorovinyl moiety produces a comparable gain in potency thus arguing against the involvement of the carbomethoxy group in H-bonding.[141]

Another rather unusual example of ester isosterism is the replacement of the ether oxygen by a fluoro-nitrogen (Figure 15.32a) as mentioned by Lipinski.[142] Other

FIGURE 15.28 Muscarinic ligands derived from the ester bioisoster replacement of arecoline **12** and quinuclidine methyl ester **13**.

FIGURE 15.29 Design of new metabolically stable HIV inhibitor.

FIGURE 15.30 Ethyl oxime ether as ester replacement gave an improved bioavailability.

uncommon examples are found in the replacement of the ester function of acetylcholine by exo–endo amidinic functions of 3-aminopyridazines in muscarinic agonists (Figure 15.32b)[143] and of the carbomethoxy group of α-yohimbine (rauwolscine) by a *N*-methylsulfonamide function (Figure 15.32c).[144]

3. *Carboxamide bioisosteres*

Biologically active molecules containing amide bonds suffer usually of pharmacokinetic liability. In order to increase their stability, bioisosteric transformation of the carboxamide have been performed and yielded a lot of successful examples especially in the area of petidomimetic. The isosteric replacements for peptidic bonds have been summarized by Spatola[145] and by Fauchère.[146] The most used and well-established modifications are: *N*-methylation, configuration change (D-configuration at Cα), formation of a retroamide or an α-azapeptide, use of aminoisobutyric or dehydroamino acids, replacement of the amidic bond by an ester [depsipeptide], ketomethylene, hydroxyethylene or thioamide functional group, carba replacement of the amidic carbonyl, and use of an olefinic double bond (Figure 15.33).

In other cases carboxamides are converted to sulfonamides as illustrated by the synthesis of the hypoglycemic sulfonyl isostere of glybenclamide.[147]

FIGURE 15.31 Replacement in (−)-cocaine of the carbomethoxy group by a carbethoxyisoxazole and a chlorovinyl moiety.

FIGURE 15.32 (a) Replacement of ester ether oxygen by a fluoro-nitrogen; (b) Exo–endo amidine in place of a carboxylic ester functionality; (c) *N*-methylsulfonamide analog of α-yohimbine (rauwolscine).

More unusual isosteric replacements for the peptidic bond were proposed and are reported in Figure 15.34. Among them, hydroxyethylureas were used in the design of a novel class of potent HIV-1 protease inhibitors, diacylcyclopropanes in the design of novel renin inhibitors, and pyrroline-3-ones for various proteolytic enzyme inhibitors.[148,149] Vinyl fluorides can probably be considered as representing the closest possible bioisosteres of the peptide bond. The synthetic methods available allow, by an appropriate selection of the precursors, the preparation of analogs of dipeptidic combinations of amino acids bearing no other functionalities in their side chains, for example Gly, Ala, Val, Phe, and Pro.[150] Vinyl fluorides have been used in the design of bioisosteres of peptide bonds as in the case of the analgesic dipeptide 2,6,-dimethyl-L-tyrosyl-D-alanine-phenylpropionamide.[151]

Other structural modifications of the amide bond proposed to create more chemically stable and orally available molecules include heterocyclic rings such as 1,2,4-oxadiazoles and 1,3,4-oxadiazoles (Table 15.15). 1,2,4-Oxadiazoles have been used as bioisosters of the carboxamide moiety in SH2 inhibitors of tyrosine kinase ZAP-70,[152] 5-HT_{1D} receptor agonists,[135] histamine H_3 receptor antagonists,[153] and 5-HT_3 receptor antagonists.[154] An example worthy of note is given by the β_3 adrenergic receptor agonists where the replacement of the amide bond with 1,2,4-oxadiazoles lead to compounds with improved oral bioavailability retaining the β_3AR agonist activity.[155–159] 1,3,4-Oxadiazoles have also been used as amide bond surrogates in several therapeutic areas like benzodiazepine receptor agonists,[160] muscarinic receptor agonists,[161] NK1 receptor antagonists[162] and they have also been used as Phe-Gly peptidomimetics.[163]

Other heterocyclic replacements of the amide bond include oxazole rings,[164,165] cyclic amidines,[166,167] pyrrole rings,[168] and phenylimidazole[169] (Table 15.15).

In the case of aryl amides, heterocyclisation was explored in order to overcome the pharmacokinetic liability. The search for transient receptor potential vanilloid 1 (TRPV1) antagonists as potential analgesic has led to the discovery of

FIGURE 15.33 Well-established isosteric replacements for peptidic bonds.

FIGURE 15.34 Unusual isosteric replacements for the peptidic bond.

a series of biaryl amides such as those reported in (Figure 15.35a) which suffer of pharmacokinetic liability.[170] In order to overcome this problem heterocyclization was explored and has yielded benzimidazoles and indazolones as promising carboxamides bioisosters[170] (Figure 15.35a). Other examples of benzimidazoles as aryl amides isosters has also been reported as melanin-concentrating hormone receptor 1 (MCH R1) antagonists.[171,172]

TABLE 15.15 Heterocyclic Surrogates of the Amide Bond

	1,2,4-Oxadiazoles	β_3 adrenergic receptor agonist	Naylor *et al.*,[157,158] Feng *et al.*,[156] Biftu *et al.*,[155] Parmee *et al.*[159]
		Fatty acid oxidation inhibitors	Elzein *et al.*,[164] Koltun *et al.*[165]
	1,3,4-Oxadiazoles	Phe-Gly peptidomimetics NK1 receptor antagonists	Borg *et al.*[163] Ladduwahetty *et al.*[162]
	Oxazoles	Fatty acid oxidation inhibitors	Elzein *et al.*,[164] Koltun *et al.*[165]
	Cyclic amidines	Dopamine D4 receptor agonists	Einsiedel *et al.*[166,167]
	Pyrroles	D3 antagonists	Einsiedel *et al.*[168]
	Phenylimidazoles	D4 receptor ligand	Thurkauf *et al.*[169]

FIGURE 15.35 Benzimidazoles, indazolones, and indoles as carboxamides bioisosters.

The indole ring has also been reported as valuable carboxamide bioisoster (Figure 15.35b) of the non-steroidal antiandrogen bicalutamide.[173]

4. *Urea and thiourea bioisosteres*

Chemical modifications of the thiourea group of the metiamide led to the discovery of *N*-cyanoguanidine, *N*-nitro-ethene diamine, and *N*-aminosulfonil guanidine as successful bioisosters of the urea group[44] (Figure 15.19). This successful example of bioisosterism has been applied to the sulfonyl-urea function of torsemide,[174] a diuretic loop (Figure 15.36). This investigation led to the design and synthesis of the corresponding thiourea, cyanoguanidine, and diaminonitroethylene group. The structure, the electronic features (charges, molecular electrostatic potential, molecular orbitals) and the lipophilic properties of these three analogs were determined. In conclusion, cyanoguanidine and diaminonitroethylene turned out to be valuable isosteres of the (thio)urea function as they share common geometric and electronic properties. In addition, they cover a range of lipophilicity that makes them suitable for compound pharmacomodulation as illustrated by the diuretic and antiepileptic properties of analogs of torsemide.

In a study undertaken on antagonists of the dihydropyridine neuropeptide Y_1 receptor, common bioisosteres such as thiourea and cyanoguanidine were examinated as urea replacement (Figure 15.37) and also more uncommon derivatives like squaric acid and thiadiazole oxide.[175,176] Both cyanoguanidine and thiourea derivatives demonstrated

FIGURE 15.36 Structure of torsemide and its urea bioisosteres.

FIGURE 15.37 Binding affinities for various urea bioisosteres acting as antagonists at the NPY Y_1 receptor.

FIGURE 15.38 Less common urea equivalents.

FIGURE 15.39 Bioisosteric replacements of the phenol function in the design of *N*-methyl-D-aspartate (NMDA) receptor antagonists.

potent binding affinity at the Y_1 receptor. The two heterocycle replacements, squaric acid, and thiadiazole oxide, demonstrated good binding affinity although both were 10 times less active compared to parent urea.

Among more exotic surrogates, the 3,4-diamino thiadiazole dioxide moiety was proposed as a weakly acidic urea equivalent (Figure 15.38).[177] The similar thiatriazole dioxide is found in the H_2 antagonist tuvatidine (HUK 978). Other bioisosteres are exo–endo amidinic heterocyles bearing an electron-attracting function in the α position[178,179] (Figure 15.38).

5. *Bioisosteres of the phenol function*

The optimal bioisosters of the phenolic function should have approximately the same size of the hydroxyl itself and should have approximately the same acidity range (weak acid) and be able to form hydrogen bonds.

The most popular surrogates for phenolic functions are NH groups rendered acidic through the presence of an electron-attracting group. In Figure 15.39 it is reported an application of this bioisostery in the design of *N*-methyl-D-aspartate (NMDA) receptor antagonists.[180] In this case, the replacement of the phenol by heterocyclic NH-containing rings was performed in order to slow metabolism and hence to improve oral bioavailability. Indeed the potent and NR1A/2B-receptor selective benzimidazolone analog was obtained demonstrating oral activity in a rodent model of Parkinson's disease at 10 and 30 mg/kg.

Using this analogy, Wu *et al.*[181] prepared a series of phenolic bioisosteres of benzazepine D_1/D_5 antagonists (Figure 15.40).

FIGURE 15.40 Phenol bioisosteres of benzazepine D_1/D_5 antagonists.

Compared to the reference compound SCH 38393 they exhibit similar or more potent activities, high selectivity over D_2–D_4 receptors and improved *in vivo* pharmacokinetics. The optimization of the hydrogen bond donating capacity of various heterocycles allowed the identification of several potent D_1/D_5 antagonists with high selectivity D_2–D_4, α_{2a} adrenergic receptors, and the 5-HT transporter keeping at the same time excellent pharmacokinetic profiles.

Other popular phenol bioisoteric substituents[182,183] include methanesulfamide (CH_3SO_2NH—), hydroxymethyl ($HOCH_2$—) or hydroxyisopropyl ($HOC(CH_3)_2$—), various amide groups (—NHCHO, —$NHCOCH_3$, —$NHCOC_6H_5$, methanesulfamidomethyl ($CH_3SO_2NHCH_2$—), dimethylaminosulfonamide ($(CH_3)_2NSO_2NH$—), and others groups with an ionizeable proton next to or near an aromatic ring.

Hacksell *et al.*[184] described the compound (±)-3-(1-propyl-3-piperidinyl)phenol [(±)-3-PPP] as highly selective for presynaptic brain DA receptors. However, the clinical potential of these molecules (Figure 15.41) as antischizophrenic or antiparkinsonian agents, may be limited by their relatively low oral bioavailability and their short duration of action.[185] Thus, a heterocyclic analog of 3-PPP might retain some of its pharmacological activity while having improved oral bioavailability and duration of action.[186] In the case of talipexole (B-HT 920) (Figure 15.41), a DA agonist which has been reported to be selective for DA autoreceptor, 2-aminothiazolyl moiety was used as a replacement for the phenol ring.

This successful example led successively to the discovery of the PD 118440, a dopamine-autoreceptor agonist,[186] and pramipexole, a DA agonist with preference for the DA D_3 (over DA D_2) receptor.[187–189]

6. *Catechol bioisosteres*

All the previous examples can be applied to the catechol family. Many of these bioisosteres share with the catechols the ability to chelate metal atoms and to form hydrogen-bonded second rings. In benzimidazole this H-bond ring is mimicked by way of a covalent ring structure (Figure 15.42).

A recent case described the synthesis of more stable bioisosteres of inhibitors of the insulin-like growth factor-1 receptor kinase (Figure 15.43). Based on the structure of AG 538[190] which contains two catechol rings and is sensitive to oxidation in cell, a new series of kinase inhibitors were developed. The catechol moiety was replaced by a benzoxazolone ring resistant to oxidation yielding two compounds GB19 and AGL2263 which maintain same potency as AG 538.

Until now, 2-aminothiazole derivatives or other aromatic heterocyclic systems have been used as a bioisosteric

FIGURE 15.41 2-Aminothiazole bioisostery applied to the dopamine agonist 3-PPP.

FIGURE 15.42 The bioisostery catechol/benzimidazole.

surrogate for the catechol nucleus (*cf.* studies around the pramipexole).[187] Hübner *et al.*[191–193] evaluated non-aromatic catechol bioisosteres (Figure 15.44). They were able to demonstrate that the π-electronic system of the non-aromatic endiyne FAUC 88 and 73 can be used as an efficient bioisostere for the catechol fragment of dopamine. Combined with conformational rigidization, this approach led to dopamine-receptor agonists with high affinity for the subtypes of D_2-family and especially for D_3.

7. *Sulfonamides bioisosters*

The introduction of the 5-HT_{1D} receptor agonist sumatriptan for the acute treatment of migraine has marked an intense research effort to discover more potent and selective 5-HT_{1D} receptor agonists with improved pharmacokinetic profiles.[194] The H-bond acceptor ability of the 5-position is crucial for 5-HT_{1D} receptor affinity and selectivity as it was revealed by reference compounds (5-HT and sumatriptan). The search for other analogs with better bioavailability led to the replacement of the methyl sulfonamide with different heterocycles which are able to be H-bond acceptors. The amino-oxadiazole analog, L-695,894 (Figure 15.45), has a better oral bioavailability but lack of the good selectivity in fact it has a significant affinity for 5-HT_{2A} and 5-HT_{2C} receptors.[135] Alternative five-membered heteroaromatic rings were explored to find analogs with log Ds<-0.5 in order to minimize the central nervous system penetration.[194] Simple unsubstituted imidazoles, triazoles and tetrazoles (Table 15.16) have high affinity and good selectivity for the 5-HT_{1D} receptor. The 1,2,4-triazole 1, MK-462, was shown to have the optimal pharmacokinetic profile with rapid oral absorption and high bioavailability.

Using a different approach, Glen *et al.*[195] built a pharmacophore based on the pharmacological activities of a series of novel C- and *N*-linked hydantoin analogs. Their model was used as a framework for the design of a diverse series of analogs with good affinity and selectivity for 5-HT_{1D}. Suitable with the constraints required for good oral absorption, a potent selective 5-HT_{1D} agonist, S-5-methyl-2-oxazolidinone analog, has been described.

The dipolar azido group is a bioisostere[196] of the SO_2NH_2 and SO_2Me hydrogen bonding pharmacophores present in many selective COX-2 inhibitors and the azido analogs **14** and **15** are useful biochemical agents for photoaffinity labeling of the COX-2 enzyme (Figure 15.46).

F. Reversal of functional groups

The reversal of the peptidic functional groups is often used in peptide chemistry. The obtained *retropeptides* are generally more resistant to enzymatic attacks (Figure 15.33).[146,197] For thiorphan and *retro*-thiorphan an identical binding mode to the zinc protease thermolysin was demonstrated.[198] Similar inhibition values for thermolysin and neutral endopeptidase were observed, whereas, for another zinc protease, angiotensin-converting enzyme (ACE), noticeable differences for inhibition were found (Figure 15.47).

Compounds	Name	IC_{50} Values [μM] IGF-1R	IR	SRC	PKB
	AG 538	0.06	0.12	2	76
	GB 19	0.37	0.6	1.5	21
	AGL 2263	0.43	0.4	2.2	55

FIGURE 15.43 Inhibition of IGF-1R and other kinases by AG 538 bioisosteres and analogs.

FIGURE 15.44 Aromatic and non-aromatic dopamine isosteres.

FIGURE 15.45 Replacement of the phenolic function of serotonine by various functions able to serve as H-bond acceptors.

TABLE 15.16 Displacement of [3H]-5-HT to 5-HT1D Recognition Sites in Pig Caudate Membranes by N-Linked Imidazoles, Triazoles and Tetrazoles and Standard 5-HT1D Agonists

Compound	Het	pIC_{50}	log D[a]
5-HT		8.0	
Sumatriptan		7.7	−1.17
L-695,894		7.6	−0.67
1, MK-462		7.3	−0.74
2		7.5	−0.53
3		7.2	−0.74
4		7.3	−0.70
5		6.6	−0.20
6		7.4	−0.64
7		7.4	−0.34

[a] *log* P *measured at pH* 7.4.

FIGURE 15.46 Selective COX-2 inhibitors and their azido bioisosteres.

FIGURE 15.47 Inhibition values of thiorphan and *retro*-thiorphan for three zinc proteases.

But the strategy of functional inversion can also be applied to non-peptidic compounds. A historical example is the change from orthoform to neo-orthoform (orthocaine; Figure 15.48). The unwanted side effects, often encountered with aromatic para-amino substituted compounds ("para effects," essentially of allergic origin) are abolished in the meta amino isomer whereas the local anesthetic activity is maintained. Similarly the "meta" isomer of benoxinate has a local anesthetic activity identical to that benoxinate itself.[199]

The β-blocking agent practolol which was one of the first cardioselective β-blocker, was rapidly replaced by its isomeric analog atenolol which presents less side effects (Figure 15.49).

The inversion of the ester function of *meperidine* leads to 1-methyl-4-phenyl-4-propionoxy piperidine (Figure 15.50) which is five times more potent as an analgesic drug than meperidine and represents the model compound of the series of inverted esters.[17]

The change from indomethacin to clometacin, although representing a clean example of functional group reversal, causes more profound alterations than that shown in the

FIGURE 15.48 Positional isomery in local anesthetics.[199]

FIGURE 15.49 Reversal of functional goups in practolol yields atenolol.

FIGURE 15.50 Meperidine and the corresponding inverted ester.[17]

previous examples (Figure 15.51). At a first glance, this change can even seem too drastic; however, in returning the molecule of clometacin by 180°, the resemblance with the parent molecule becomes evident.

Indomethacin is mainly used as a non-steroidal antiinflammatory agent and occasionally as an analgesic; clometacin on the other hand is usually recommended as an analgesic and shows weak antiinflammatory properties. Applied to serotonin, a similar reversal of a functional moiety yielded $5HT_{2C}$ receptor selective agonists[200] (Figure 15.52).

IV. SCAFFOLD HOPPING

Scaffold hopping is a useful technique used by medicinal chemists to discover structurally novel compounds starting from known active molecules by changing the molecular backbone of the template. Medicinal chemists have quite often the need of "jumping" in a different chemical space: new chemical entities offer the choice of easier chemical accessibility and so possibility of speeding up the lead optimization process. Switching to a different molecular backbone can help to avoid some undesirable ADME-tox properties. Scaffold hopping come to a rescue also to move from complex natural molecules to easier synthesisable small molecules. Moreover "jumping" in a new chemical series isofunctional with a given one is fundamental when new IP position is looked for.

A. Successful examples of serendipitous scaffold hopping

The concept of scaffold hopping relies upon the assumption that structurally and chemically distinct templates interact with the same receptor inducing equivalent biological activity. This seems to contradict the leading principle of the lead optimization process and drug design that is based on the concept that chemical similarity is reflected by similar biological activity. On the other hand this concept is not always true. In fact not always structural and chemical analogy predicts similar biological affinity as it is shown in Figure 15.53. In fact promethazine, imipramine, and chlorpromazine that are strictly similar from a structure point of view, interact with different receptors and present complete different therapeutic profiles acting respectively as H1 antagonist, uptake inhibitor and dopamine antagonist.

Indeed scaffold hopping can be interpreted as the broadest expansion of the bioisosterism concept. The idea of scaffold hopping is indeed not new, it is rather a new term to designate known experimental evidences. A retrospective analysis of the marketed drugs and/or of the ligands acting

FIGURE 15.51 Functional inversion applied to indomethacin.

FIGURE 15.52 Serotonin analogs resulting from a functional reversal.[200]

FIGURE 15.53 Structure and chemical similarity are not always synonymous of similar biologically active compounds.

on the same receptors brings up some interesting examples of structurally different molecules eliciting the same biological activity. A nice example is given by the non-steroidal antiinflammatory drugs (NSAIDs) COX-2 inhibitors as illustrated

FIGURE 15.54 An example of scaffold hopping: four structurally different molecules have similar therapeutic properties by interacting with the same receptor.

in Figure 15.54 where there are reported four know traditional NSAIDs: indomethacin, nimesulide, diclofenac, and celecoxib.

Actually the GABA-A ligands at the benzodiazepine site give a nice picture of matching and mismatching of scaffold hopping and bioisosterism (Figure 15.55). Looking at diazepam, zolpidem, zopiclone, and zaleplon, a structural analogy is barely found. Despite this they are all GABA-A ligands acting as full agonists. The same can be said about the inverse agonists DMCM, **16**, **17**, FG 8094, and **18**: chemically different compounds have comparable biological activity. On the other hand similar compounds like FG 8094, Bretazenil, and Ro 15-1788, do not give the same biological response.

FIGURE 15.55 GABA-A ligands binding at benzodiazepine site.

B. Scaffold hopping and virtual screening

The examples cited in the above paragraph illustrated that similar biological activity can be obtained with structurally different molecules which seems to contradict the bioisosterism principle discussed in this chapter. The serendipitous examples of scaffold hopping described above have been generated either by high-throughput screening (HTS) either by selective optimization of side activity, the SOSA approach.[201] In other studies the "hopping" in a different molecular scaffold has relied upon the intuition and/or the experience of the medicinal chemist working on the project. Most of the virtual screening studies aiming to hop in isofunctional molecules are based on the bioisosterism concept that structural likeness predicts analogous biological activity. At a first glance it can be concluded that a rational approach to scaffold hopping is not possible. Indeed molecular modeling and chemoinformatics experts have set up a lot of softwares and methods to address the paradox of scaffold hopping and bioisosterism and find a logical process.[202,203]

Among the virtual screening techniques reported so far four types of computational approaches for the scaffold hopping can be distinguished: (a) shape matching, (b) pharmacophore searching, (c) fragment replacement, and (d) similarity searching. Most of the programs offer combinations of several approaches and they are used, according to the case scenario of a given project, either as on a ligand-based virtual screening base either on a structure (receptor)-based screening. Other studies have reported a combination of the two methods to increase the performance of the virtual screening. In a case where more active ligands are known and no information are known about the receptor structure, it will be more useful to proceed through shape matching combined with a pharmacophore searching. Instead when the receptor structure is known, a virtual screening based on molecular docking is a powerful and successful technique. A detailed description of the computational approaches of scaffold hopping goes beyond the scope of this chapter and only few illustrative examples will be reported. In Figure 15.56 it is reported an example of scaffold hopping

FIGURE 15.56 Design of a novel CCK_2 antagonist via a scaffold hopping approach from a known series of indoles.

FIGURE 15.57 Discovery of non-peptide inhibitor of β-secretase by high-throughput docking.

related to the design of a novel cholecystokinin-2 (CCK_2) antagonist.[204] Starting from a known series of indoles (**19**) with high *in vitro* affinity and selectivity but poor bioavailability, a CCK_2 pharmacophore receptor has been identified and used to design new scaffolds to overcome the problems associated with the high level of biliary elimination of the compound **19**. Two news series of pyrroles and imidazoles (**20**) were identified and, although they are not as potent as the original indole derivative *in vitro*, they have greater potential as drug candidates because of their reduced tendency to biliary excretion.

Scaffold hopping is an extremely useful technique to switch from peptide ligands to small molecule ligands. A fragment-based docking procedure followed by substructure search were used to identify a set of small molecules as β-secretase inhibitors.[205] In Figure 15.57 is reported one of a ligand identified with the virtual screening that may serve as starting point for further optimization for Alzheimer disease.

Most of the studies based on virtual screening use a hierarchical approach combining several conceptually different techniques.[206] The most common combination used is a 3D pharmacophore modeling followed by a docking study. In Figure 15.58 is reported an illustrative example of this combined use: a set of 10 known inhibitors of Human Rhinovirus Coat protein was used to design a 3D pharmacophore that was subsequently used for a docking study of the Maybridge DB.[206]

A set of 10 compounds was retrieved according to the score fit of the molecules. The 10 molecules were tested *in vitro* and all of them displayed antirhinoviral activity in the micromolar range. Particularly promising results were achieved with the compound reported in Figure 15.58 which shows an EC_{50} of 4.3 μM.

V. ANALYSIS OF THE MODIFICATIONS RESULTING FROM ISOSTERISM

It is rare that the replacement of a part of a molecule by an isosteric or bioisosteric group leads to a *strictly identical* active principle. In practice that is even not sought, and one prefers that the new compound produces a change as compared to the parent molecule. In general the isosteric replacement, even though it represents a subtle structural change, results in a modified profile: some properties of the parent molecule will remain unaltered, others will be changed. Bioisosterism will be productive if it increases the potency, the selectivity and the bioavailability or decreases the toxicity and undesirable effects of the compound. In proceeding to isosteric modifications one will focus *predominantly* on a given parameter: structural, electronic,

FIGURE 15.58 A set of 10 known HRV coat protein inhibitors was used for 3D pharmacophore modeling followed by docking study afforded new candidate structure suitable for an optimization process.

hydrophilic, but it is all but impossible not to alter simultaneously several parameters simultaneously.

A. Structural parameters

These will be important when the portion of the molecule involved in the isosteric change serves to maintain other functions in a particular geometry. That is the case for tricyclic psychotropic drugs (Figure 15.59).

In the two antidepressants (imipramine and maprotiline), the bioisosterism is geometrical insofar that the dihedral angle α formed by the two benzo rings is comparable: $\alpha = 65°$ for the dibenzazepine and $\alpha = 55°$ for the dibenzocycloheptadiene.[207] This same angle is only 25° for the neuroleptic phenothiazines and the thioxanthenes. In these examples the part of the molecule modified by isosterism is not involved in the interaction with the receptor. It serves only to position correctly the other elements of the molecule.

The structure of various bioisosteric retinoic acid receptor agonists highlights the dominantly geometric parameter of this bioisostery (Figure 15.60).[208]

FIGURE 15.59 The tricyclic antidepressants (imipramine and maprotiline) are characterized by an dihedral angle of 55° to 65° between the two benzo rings, this angle is only 25° for the tricyclic neuroleptics (chlorpromazine, chlorprothixene).[207]

B. Electronic parameters

Electronic parameters govern the nature and the quality of ligand–receptor or ligand–enzyme interactions. The relevant parameters will be inductive or mesomeric effects, polarizability, pK_a, capacity to form hydrogen bonds, etc. Despite their very different substituents in the *meta* position, the two epinephrine analogs (Figure 15.61) exert comparable biological effects: they are both β-adrenergic agonists. In fact the key parameter resides in the very close pK_a values.[197]

C. Solubility parameters

When the functional group involved in the isosteric change plays a role in the absorption, the distribution or the excretion of the active molecule, the hydrophilic–lipophilic parameters become important.

Imagine in an active molecule the replacement of $-CF_3$ ($\pi = +0.88$) by $-CN$ ($\pi = -0.57$) (Figure 15.62). The electron-attracting effect of the two groups will be comparable, but the molecule with the cyano function will be

FIGURE 15.60 Bioisostery in its broadest sense.

FIGURE 15.61 An example of bioisosterism, or non-classical isosterism, the methylsulfonamide substituent has comparable acidity to the phenolic hydroxyl group.[197]

clearly more hydrophilic. This loss in lipophilicity can then be corrected in attaching elsewhere on the molecule a propyl, isopropyl, or cyclopropyl group.

D. Anomalies in isosterism

In this paragraph two applications of the bioisosterism concept that imply unusual behaviors of commonly encountered atoms or groups are discussed.

1. *Fluorine–hydrogen isosterism*

There is a anomaly residing in the fact that fluorine does not resemble other halogens, notably chlorine, and that, on the other hand, it often mimics an atom of hydrogen.[209]

1. *Steric aspects*: The fluorine atom is considerably smaller than the rest of the halogen atoms. Seen from the steric point of view it resembles more hydrogen than chlorine (Table 15.17). Effectively fluoro-derivatives differ from the other halogenated derivatives because fluorine forms with carbon particularly stable bonds and, in contrast to other halogens, is only rarely ionized or displaced. Because it is both chemically inert and of small size organic fluorine is often compared to hydrogen.

   One understands especially the incorporation by living organisms of fluoroacetic acid in place of acetic[210] acid or of 5-fluoro-nicotinic acid and 5-fluoro-uracil as antimetabolites. This "fraudulent" incorporation leads to lethal syntheses.[211] This is generally not the case with the corresponding chlorinated, brominated, or iodinated analogs.
2. *Electronic aspects*: Fluorine is the most electronegative of the halogens (Table 15.17) and forms particularly stable bonds with carbon atoms. This chemical inertia explains why fluoro-derivatives are more resistant to metabolic degradation (Figure 15.63). Thus for the β-haloalkylamines (nitrogen mustards), the alkylating activity is lost when chlorine or bromine are replaced by fluorine or by hydrogen.[212]

   The isosterism between H and F will therefore often serve to give analogs that are more resistant to metabolic degradation (obstructive halogenation: flunarizine and in flufenisal; Figure 15.63). Similarly the CF_3 group is biostable whereas CH_3 is easily oxidized.[209]
3. *Absence of d orbitals*: An other difference between fluorine and the other halogens comes from the absence of a d orbital for fluorine, and thus its incapacity to participate in resonance effects with a donor of π electrons (Figure 15.64).

   This explains among why *para*-fluorophenol is slightly less acidic than phenol while for other *para*-halogenated phenols the acidity changes in parallel with the atomic number (Table 15.18[209]).
4. *Case study*: A good example of continuous variation of activity in halogenated compounds is provided by a series of antihistaminic drugs related to tripelennamine (Figure 15.65, X = H).

Apparently we are dealing here with a classical isosteric series: F, Cl, Br, I, but sensitive to steric hindrance in the para position. Probably what happens *in vivo*, is para-hydroxylation of the benzene ring. The best candidate becomes then the para-fluoro compound, since it is not bulkier than the unsubstituted compound while being biostable.

FIGURE 15.62 The loss in lipophilicity resulting from the bioisosteric exchange of a CF_3 for a CN has to be compensated by the equivalent of a three carbon residue.

TABLE 15.17 Fluorine–Hydrogen Isosterism. Observe the Comparable Sizes of the Two Atoms, whereas Chlorine is Close to the Methyl and Trifluoromethyl

Parameter	H	F	Cl	CH_3	CF_3
Atomic radius	0.29	0.64	0.99	–	–
Van der Waals radius	1.2	1.35	1.80	<2	<2
Molecular refractivity	1.03	0.92	6.03	5.65	5.02
Electronic effect (*para* σ)[a]	0.00	0.06	0.23	−0.17	0.54
Resonance effect (R)[a]	0.00	−0.34	−0.15	−0.13	0.19
Electronic effect (σ^*)[b]	–	3.08	2.68	0.00	2.85

[a]*For aromatic systems.*
[b]*for aliphatic systems.*

FIGURE 15.63 In flunarizine and in flufenisal fluorine atoms in para position prevent metabolic hydroxylation.

FIGURE 15.64 The resonance between the OH lone pair and the X group is not possible if X = F.

2. *Exchange of ether oxygen and methylene group*

Ether oxygen atoms and methylene groups possess a similar tetrahedral structure and should normally be isosteric. In fact the isosterism between O and CH_2 yields very often anomalous results and brought Friedman [214] to the interesting observation *"that the omission of the ether oxygen changes biological activity much less in some cases than the replacement by the isosteric methylene group"* (Figure 15.66). In the meperidine series, for example, the change from the *N*-phenoxypropyl derivative to the isosteric phenoxybutyl decreases the analgesic potency by

a factor of ten, whereas the omission of the ether oxygen yields a slightly more potent compound.[17] A list of seven other examples is given by Schatz in the second edition of Burger's Medicinal Chemistry.[215]

The explanation for this anomalous behavior may be that the omission of the ether oxygen yields a closer compound in terms of lipophilicity than its replacement by a methylene. An example which can be compared to Friedman's paradox is found in the resemblance of the phenylethyl type β-blockers (e.g. dichloroisoprenaline, sotalol) with the phenoxypropanol type (e.g. practolol, acebutolol).

TABLE 15.18 Dissociation Constants of *p*-Halogenated Phenols[209]

Compound	Dissociation constant $K_a \times 10^{-10}$
Phenol	0.32
p-Fluorophenol	0.26
p-Chlorophenol	1.32
p-Bromophenol	1.55
p-Iodophenol	2.19

FIGURE 15.65 Variation of activity in a series of antihistaminic compounds as a function of the halogated para-substituent.[213]

VI. MINOR METALLOIDS-TOXIC ISOSTERS

In this section we describe some "exotic" applications of the bioisostery concept implying the utilization of unusual elements such as silicon, boron, selenium, arsenic, and antimony. The use of those elements as bioisosteres of carbon in existing drugs is a different approach enabling the introduction of a new drug-like chemical space into the drug discovery and development process.

A. Carbon–silicon bioisosterism

Silicon is directly below carbon in the periodic table so, according to the Erlenmeyer's expansion of the isosterism concept, carbon and silicon can be considered as true isosters. Sila-substitution (C/Si exchange) of biologically active substances is an approach to search for new drug-like candidates with improved pharmacological properties and stronger IP position. The application of this isosterism remains however limited. For reviews on the subject, see Fessenden and Fessenden,[216] Tacke and Zilch,[217,218] Ricci *et al.*,[219] and Showell and Mills.[220]

Silicon is more electropositive than carbon (and even more if compared to oxygen and nitrogen) and the covalent silicon–carbon bonds in the sp^3 hybridization state, are

FIGURE 15.66 Friedman's ether oxygen-methylene group paradox.[2]

FIGURE 15.67 Organosilicon active substances.

20% longer than the corresponding carbon–carbon bond. This difference is reflected on the properties of the sila-isosters compared to the corresponding carbon analogs. For example acidity, the hydrogen bond strength of the sila-nol is more favorable as a donor than that of carbinol. In this respect the silanol isosters may be beneficial in those pharmacophores where the H-bonding has a predominant role. From a lipophilicity point of view the silicon-containing analogs are more lipophilic than their carbon analogs and when an increase in lipophilicity is seeked this stratagem could be used. The C = O double bond of a ketone is quite stable, whereas the formation of a Si = O double bond is disfavored over its hydrate form, the silicon diol. The chemistry of silicon relies mainly on the chemistry of single bonds which have led to its appropriate use as a tetrahedral bioisostere of carbon. In Figure 15.67 are reported some of examples of the sila-substitution of existing drugs. Among these, m-trimethylsilyl-phenyl *N*-methylcarbamate and m-trimethylsilyl-α-trifluoroacetophenone (zifrosilone) are acetylcholinesterase inhibitors[221–223], sila-meprobamate is a CNS depressant[224] sila-pridinol is an anticholinergic[225], flusilazole is a fungicide for agricultural use[226], and (+)-RP 71,602 is a potent and selective 5-HT_{2A} antagonist.[227]

A silicon-containing hypocholesterolemic squalene epoxidase inhibitor,[228,229] the silicon analog of the α_2-adrenergic antagonist atipamezole,[230], a highly potent, stable and CNS-penetrating silatecan[231] some ACE inhibitors[232], a HIV protease inhibitors[233], the (*R*)-sila-analog of the antidepressant venlafaxine[234], and a trimethylsilylpyrazole as novel inhibitor of p38 MAP kinase[235] (mitogen-activating protein) were published (Figure 15.68). As for the preceding molecules, the silicon atom of these compounds is quaternary and thus expected to be less sensitive to metabolic degradation (Figure 15.68).

Compared to their carbon bioisosteres, silicon-containing molecules are more sensitive to hydrolysis and to nucleophilic attack in general, even when located in the center of a quaternary structure, the risk exists that the silicon atom will be attacked. Thus, 1-chloro-1-sila-bicyclo-(2,2,1)-heptane can still be hydrolyzed by an attack on the vacant d orbital;[236] this attack is *lateral* and therefore possible even in the cases where the corresponding carbon derivative would have been inert toward a S_N2 reaction (Figure 15.69).

This sensitivity toward lateral attacks explains the four times shorter duration of action of sila-meprobamate compared to its carbon isostere on a model of tranquillizing activity in mice (rotarod test, potentiation of hexobarbital-induced sleep, and intraperitoneal injection).[224] On the other hand, when given orally, sila-meprobamate is practically inactive. One of the first metabolites formed has been characterized as being a di-siloxane[216] (Figure 15.67). For the two phenyl-trimethylsilyl-derived AChE inhibitors, the rather positively charged trimethyl-silyl group mimics the trimethyl-ammonium function present in acetylcholine. For these compound metabolic oxidation does not take place on the silicon, but on one of methyl groups ($Si\text{-}CH_3 \rightarrow Si\text{—}CH_2\text{—}OH$).[221]

B. Carbon–boron isosterism

Boron is essential for plant growth and development. In medicinal chemistry its main use is related to that of coupling reagent. The most important employment of boron as drug relies upon the treatment of certain tumors by Boron Neutron

FIGURE 15.68 Structures of silicon-containing compounds which are biologically active.

FIGURE 15.69 Due to the presence of a vacant d orbital, a lateral attack can substitute for dorsal attacks in organo-silicon derivatives.[236]

Capture Therapy (BNCT).[237–239] BNCT is a binary treatment modality that can selectively irradiate tumor tissue. The delivery to the tumor cells of a drug containing the ^{10}B isotope followed by a low energy irradiation (neutrons) causes the ^{10}B to split, releasing an alpha particle a lithium nucleus. These products of the $^{10}B(n,\alpha)^{7}Li$ reaction are very damaging to cells but have a combined path length in tissue of approximately 14 μm, or roughly the diameter of two cells. Thus, most of the ionizing energy impared to tissue is localized to 10B-loaded cells. The problem here is to insure a sufficient concentration of the product in the tumor being treated.

Few medicines based on boron are known, in general boric acid or a boronic acid serve to esterify an α-diol or an ortho-diphenol. This is the case for the emetic antimony borotartrates of the ancient pharmacopoieas, for the injectable catecholamine solutions, for tolboxane[240], that is close to meprobamate and that was commercially available as a tranquillizer some decades ago, or also for the phenylboronic esters of chloramphenicol.[241] Boromycine was the first natural product containing boron isolated.[242] It is a complex between boric acid and a polyhydroxylated tetradentate macrocycle.[243] Another natural product is aplasmomycin with antibiotic properties.[244]

Some of the boron-containing molecules biologically active are reported in Figure 15.70. Some boronic analogs of amino acids were prepared as chymotrypsine and elastase

FIGURE 15.70 Boron-containing molecules with a biological activity.

FIGURE 15.71 Ebselen and its main metabolites.[256]

inhibitors,[245] and more recently as antineoplastic agent (Velcade, a proteosome inhibitor reported in Figure 15.70).[246] Carboxyboranes complexed with a tertiary amine ($R_3N.H_2B$-COOH) are considered as boron amino acid isosteres due to the isoelectronic features. Compounds-containing carboxy boranes have shown anticancer, hypolidemic, and antifungal activity.[247] Diazaborines are active against malaria,[248] and oxazaborolidines possess antibacterial activity (Figure 15.70).[249] Boronic chalcones are reported to be antitumor agents (Figure 15.70).[250]

Organoboron derivatives, even more than organosilicon compounds, are sensitive to hydrolytic degradation that always leads to the final formation of boric acid. But boric acid has teratogenic properties in chickens. It produces the same malformations as those produced by a riboflavine (vitamin B_2) deficiency and the administration of riboflavine prevents these toxic effects.[251,252] The mechanism by which boric acid produces a deficiency in riboflavine is not known. In man the chronic utilization of boron derivatives results in cases of borism (dry skin, cutaneous eruptions, and gastric troubles).[253]

C. Bioisosteries involving selenium

Selenium can be considered the best isoster of sulfur as it is just below it in the periodic table. These two atoms have very similar physical properties: the radius of selenium is only 12.5% bigger than that of sulfur, and their electronegativity is rather similar.

Selenium and its derivatives are highly toxic and with the exception of ^{75}Se derivatives which serve diagnostic purposes (e.g. ^{75}*Se-selenomethionine*, used as a radioactive imaging agent in pancreatic scanning), there is no chemically defined seleno-organic drug on the market. Klayman reviewed a large number of selenium derivatives as chemotherapeutic agents in 1973.[254] Selenium bioisosteres of sulfur compounds are mainly used as research tools (e.g. bis [2-chloroethyl] selenide as selenium bioisostere of the classical sulfur mustards[255]). Selenocysteine is present in the catalytic site of mammalian glutathione-peroxidase and this explains the importance of selenium as an essential trace element.

The only selenium-containing drug candidate is *ebselen* (Figure 15.71) which owes its antioxidant and antiinflammatory properties to its interference with the selenoenzyme glutathione-peroxidase.[257] Because of its strongly bound selenium moiety only metabolites of low toxicity are formed.[256]

REFERENCES

1. Langmuir, I. The structure of atoms and the octet theory of valence. *Proc. Natl. Acad. Sci. USA* 1919, *5*(7), 252–259.
2. Friedman, H. L. Influence of isosteric replacements upon biological activity. *NASNRS* 1951, *206*, 295–358.
3. H.S. Allen, *Molecular Frequency and Molecular Number*. London, 1918; Vol. parts I–III.
4. Langmuir, I. Isomorphism, isosterism and covalence. *J. Amer. Chem. Soc.* 1919, *41*, 1543–1559.
5. Patani, G. A., LaVoie, E. J. Bioisosterism: a rational approach in drug design. *Chem. Rev.* 1996, *96*(8), 3147–3176.
6. Seifriz, W. Pathogenicity and isosterism. *Science* 1948, *107*, 15–16.
7. Grimm, H. G. Structure and size of the non-metallic hybrids. *Z. Electrochem.* 1925, *31*, 474–480.
8. Grimm, H. G. On the systematic arrangement of chemical compounds from the perspective of research on atomic composition; and on some challenges in experimental chemistry. *Naturwissenschaften* 1928, *17*, 557–564.
9. Erlenmeyer, H., Leo, M. On pseudoatoms. *Helv. Chim. Acta.* 1932, *15*, 1171–1186.
10. Gelmboldt, V. O., Ennan, A. A., Ganin, E. V., Simonov, Y. A., Fonari, M. S., Botoshansky, M. M. Synthesis and structure of fluorosilicic acid compounds with 4-aminobenzoic acid and with 4-aminobenzenesulfonamide: the role of H-bonding in crystal structure formation. *J. Fluorine Chem.* 2004, *125*, 1951–1957.
11. McLeod, J. W., Mayr-Harting, A., Walker, N. Observations on the bactericidal and bacteriostatic actions of *p*-aminobenzenesulfonamide and *p*-hydroxylaminobenzenesulfonamide, with special reference to their suppression by *p*-aminobenzoic acid. *Br. J. Exp. Pathol.* 1944, *25*, 27–37.
12. Brauner-Osborne, H., Krogsgaard-Larsen, P. Pharmacology of (S)-homoquisqualic acid and (S)-2-amino-5-phosphonopentanoic acid [(S)-AP5] at cloned metabotropic glutamate receptors. *Br. J. Pharmacol.* 1998, *123*(2), 269–274.
13. Malachowski, W. P., Coward, J. K. The chemistry of phosphapeptides: investigations on the synthesis of phosphonamide, phosphonate, and phsophinateanalogues of glutamyl-γ-glutamate. *J. Org. Chem.* 1994, *59*(25), 7625–7634.
14. Thornber, C. W. Isosterism and molecular modification in drug design. *Chem. Soc. Rev.* 1957, *8*, 563–580.
15. Burger, A., (Ed.) In *Medicinal Chemistry*, 3rd ed. Wiley-Interscience ED.: New York, 1970, pp. 64–80.
16. Lima, L. M., Barreiro, E. J. Bioisosterism: a useful strategy for molecular modification and drug design. *Curr. Med. Chem.* 2005, *12*(1), 23–49.
17. Janssen, P. A. J., Van der Eycken, C. A. M. The chemical anatomy of potent morphine-like analgesics (p. 42). In *Drugs Affecting the Central Nervous system* (Burger, A., Ed.), Marcel Dekker, Inc.: New York, 1968, pp. 25–60.
18. Boyle, E. A., Mangan, F. R., Markwell, R. E., Smith, S. A., Thomson, M. J., Ward, R. W., Wyman, P. A. 7-Aroyl-2, 3-dihydrobenzo[*b*]furan-3-carboxylic acids and 7-benzoyl-2, 3-dihydrobenzo[*b*]thiophene-3-carboxylic acids as analgesic agents. *J. Med. Chem.* 1986, *29*, 894–898.
19. Morgan, B. P., Scholtz, J. M., Ballinger, M. D., Zipkin, I. D., Bartlett, P. A. Differential binding energy: a detailed evaluation of the influence of hydrogen bonding and hydrophobic groups on the inhibition of thermolysin by phosphorous-containing inhibitors. *J. Am. Chem. Soc.* 1991, *113*, 297–307.
20. Chen, Y. L., Nielsen, J., Hedberg, K. D., Dunaiskis, A., Jones, S., Russo, L., Johnson, J., Ives, J., Liston, D. Syntheses, resolution, and structure-activity relationships of potent acetylcholinesterase inhibitors: 898-carbaphysostigmine analogues. *J. Med. Chem.* 1992, *35*, 1429–1434.
21. Marseigne, I., Roques, B. P. Synthesis of new amino-acids mimicking sulfated and phosphorylated tyrosine residues. *J. Org. Chem.* 1988, *53*, 3621–3624.
22. Burke, T. R., Jr, Smyth, M., Nomizu, M., Otaka, A., Roller, P. P. Preparation of fluoro- and hydroxy-4-phosphonomethyl-D,L-phenylalanine suitably protected for solid phase synthesis of peptides containing hydrolytically stable analogues of *O*-phosphotyrosine. *J. Org. Chem.* 1993, *58*, 1336–1340.
23. Erlenmeyer, H., Willi, E. Zusammenhänge zwischen Konstitution und Wirkung bei Pyrazolonderivaten. *Helv. Chim. Acta.* 1935, *18*, 740–743.
24. Wallace, E. M., Moliterni, J. A., Moskal, M. A., Neubert, A. D., Marcopoulos, N., Stamford, L. B., Trapani, A. J., Savage, P., Chou, M., Jeng, A. Y. Design and synthesis of potent selective inhibitors of endothelin-converting enzyme. *J. Med. Chem.* 1998, *41*, 1513–1523.
25. De Lombaert, S., Stamford, L. B., Blanchard, L., Tan, J., Hoyer, D., Diefenbacher, C. G., Wei, D., Wallace, E. M., Moskal, M. A., Savage, P., Jeng, A. Y. Potent non-peptidic dual inhibitors of endothelin-converting enzyme and neutral endopeptidase 24.11. *Bioorg. Med. Chem. Lett.* 1997, *7*, 1059–1064.
26. Binder, D., Noe, C. R., Holzer, W., Baumann, K. Thiophen als Strukturelement Physiologisch Aktiver Substanzen, 16. Thienoisoxazole Durch Substitution am Oximstickstoff. *Arch. Pharm.* 1987, *320*, 837–843.
27. Uno, H., Kurokawa, M., Masuda, Y., Nishimura, H. Studies on 3-substituted 1,2-benzisoxazole derivatives. 6. Syntheses of 3-(sulfamoylmethyl)-1,2-benzisoxazole derivatives and their anticonvulsant activities. *J. Med. Chem.* 1979, *22*, 180–183.
28. Garvey, D. S., Wasicak, J. T., Decker, M. W., Brioni, J. D., Buckley, M. J., Sullivan, J. P., Carrera, G. M., Holladay, M. W., Arneric, S. P., Williams, M. Novel isoxazoles which interact with brain cholinergic channel receptors have intrinsic cognitive enhancing and anxiolytic activities. *J. Med. Chem.* 1994, *37*(8), 1055–1059.
29. Garvey, D. S., Wasicak, J. T., Elliott, R. L., Lebold, S. A., Hettinger, A. M., Carrera, G. M., Lin, N. H., He, Y., Holladay, M. W., Anderson, D. J. *et al*. Ligands for brain cholinergic channel receptors: synthesis and in vitro characterization of novel isoxazoles and isothiazoles as bioisosteric replacements for the pyridine ring in nicotine. *J. Med. Chem.* 1994, *37*(26), 4455–4463.
30. Olesen, P. H., Tonder, J. E., Hansen, J. B., Hansen, H. C., Rimvall, K. Bioisosteric replacement strategy for the synthesis of 1-azacyclic compounds with high affinity for the central nicotinic cholinergic receptors. *Bioorg. Med. Chem.* 2000, *8*(6), 1443–1450.
31. Gohlke, H., Gundisch, D., Schwarz, S., Seitz, G., Tilotta, M. C., Wegge, T. Synthesis and nicotinic binding studies on enantiopure diazine analogues of the novel (2-chloro-5-pyridyl)-9-azabicyclo[4.2.1]non-2-ene UB-165. *J. Med. Chem.* 2002, *45*(5), 1064–1072.
32. Heinisch, G., Holzer, W., Kunz, F., Langer, T., Lukavsky, P., Pechlaner, C., Weissenberger, H. On the bioisosteric potential of diazines: diazine analogues of the combined thromboxane A2 receptor antagonist and synthetase inhibitor ridogrel. *J. Med. Chem.* 1996, *39*(20), 4058–4064.
33. Haginoya, N., Kobayashi, S., Komoriya, S., Yoshino, T., Suzuki, M., Shimada, T., Watanabe, K., Hirokawa, Y., Furugori, T., Nagahara, T. Synthesis and conformational analysis of a non-amidine factor Xa inhibitor that incorporates 5-methyl-4,5,6,7-tetrahydrothiazolo[5,4-c]pyridine as S4 binding element. *J. Med. Chem.* 2004, *47*(21), 5167–5182.
34. Wood, M. R., Schirripa, K. M., Kim, J. J., Wan, B. L., Murphy, K. L., Ransom, R. W., Chang, R. S., Tang, C., Prueksaritanont, T., Detwiler,

T. J., Hettrick, L. A., Landis, E. R., Leonard, Y. M., Krueger, J. A., Lewis, S. D., Pettibone, D. J., Freidinger, R. M., Bock, M. G. Cyclopropylamino acid amide as a pharmacophoric replacement for 2,3-diaminopyridine. Application to the design of novel bradykinin B1 receptor antagonists. *J. Med. Chem.* 2006, *49*(4), 1231–1234.

35. Abe, Y., Kayakiri, H., Satoh, S., Inoue, T., Sawada, Y., Inamura, N., Asano, M., Aramori, I., Hatori, C., Sawai, H., Oku, T., Tanaka, H. A novel class of orally active non-peptide bradykinin B2 receptor antagonists. 3. Discovering bioisosteres of the imidazo[1,2-a] pyridine moiety. *J. Med. Chem.* 1998, *41*(21), 4062–4079.
36. Contreras, J. M., Parrot, I., Sippl, W., Rival, Y. M., Wermuth, C. G. Design, synthesis, and structure-activity relationships of a series of 3-[2-(1-benzylpiperidin-4-yl)ethylamino]pyridazine derivatives as acetylcholinesterase inhibitors. *J. Med. Chem.* 2001, *44*(17), 2707–2718.
37. Contreras, J. M., Rival, Y. M., Chayer, S., Bourguignon, J. J., Wermuth, C. G. Aminopyridazines as acetylcholinesterase inhibitors. *J. Med. Chem.* 1999, *42*(4), 730–741.
38. Wermuth, C. G. *The Practice of Medicinal Chemistry*, 2nd ed. Academic Press: London, 2003. p. 768.
39. Biava, M., Porretta, G. C., Cappelli, A., Vomero, S., Manetti, F., Botta, M., Sautebin, L., Rossi, A., Makovec, F., Anzini, M. 1,5-Diarylpyrrole-3-acetic acids and esters as novel classes of potent and highly selective cyclooxygenase-2 inhibitors. *J. Med. Chem.* 2005, *48*(9), 3428–3432.
40. Barbachyn, M. R., Ford, C. W. Oxazolidinone structure-activity relationships leading to linezolid. *Angew. Chem. Int. Ed. Engl.* 2003, *42*(18), 2010–2023.
41. Renslo, A. R., Luehr, G. W., Gordeev, M. F. Recent developments in the identification of novel oxazolidinone antibacterial agents. *Bioorg. Med. Chem.* 2006, *14*(12), 4227–4240.
42. Nilius, A. M. Have the oxazolidinones lived up to their billing? Future perspectives for this antibacterial class. *Curr. Opin. Investig. Drugs* 2003, *4*(2), 149–155.
43. Snyder, L. B., Meng, Z., Mate, R., D'Andrea, S. V., Marinier, A., Quesnelle, C. A., Gill, P., DenBleyker, K. L., Fung-Tomc, J. C., Frosco, M., Martel, A., Barrett, J. F., Bronson, J. J. Discovery of isoxazolinone antibacterial agents. Nitrogen as a replacement for the stereogenic center found in oxazolidinone antibacterials. *Bioorg. Med. Chem. Lett.* 2004, *14*(18), 4735–4739.
44. Ganellin, C. R. Discovery of cimetidine, ranitidine and other H_2-receptor histamine antagonists. In *Medicinal Chemistry: The Role of Organic Chemistry in Drug Research*. (Ganellin, C. R., Roberts, S. M., Eds.), Academic Press: London, 1993, pp. 227–255.
45. Mallamo, J. P., Pilling, G. M., Wetzel, J. R., Kowalczik, P. J., Bell, M. R., Kullnig, R. K., Batzold, F. H., Juniewiecz, P. E., Winnecker, R. C., Luss, H. R. Antiandrogenic steroidal sulfonyl heterocycles. Utility of electrostatic complementarity in defining bioisosteric sulfonyl heterocycles. *J. Med. Chem.* 1992, *35*, 1663–1670.
46. Fludzinski, P., Evrard, D. A., Bloomquist, W. E., Lacefield, W. B. Indazoles as indole bioisosteres: synthesis and evaluation of the tropanyl ester and amide of indazole-3-carboxylate as antagonists to the serotonin $5HT_3$ receptor. *J. Med. Chem.* 1987, *30*, 1535–1537.
47. Mathes, B. M., Hudziak, K. J., Schaus, J. M., Xu, Y. C., Nelson, D. L., Wainscott, D. B., Nutter, S. E., Gough, W. H., Branchek, T. A., Zgombick, J. M., Filla, S. A. Substituted furo[3,2-b]pyridines: novel bioisosteres of 5-HT 1F receptor agonists. *Bioorg. Med. Chem. Lett.* 2004, *14*(1), 167–170.
48. Blair, J. B., Marona-Lewicka, D., Kanthasamy, A., Lucaites, V. L., Nelson, D. L., Nichols, D. E. Thieno[3,2-b]- and thieno[2,3-b]pyrrole bioisosteric analogues of the hallucinogen and serotonin agonist *N*, *N*-dimethyltryptamine. *J. Med. Chem.* 1999, *42*(6), 1106–1111.
49. Yous, S., Andrieux, J., Howell, H. E., Morgan, P. J., Renard, P., Pfeiffer, B., Lesieur, D., Guardiola-Lemaitre, B. Novel naphthalenic ligands with high affinity for the melatonin receptor. *J. Med. Chem.* 1992, *35*(8), 1484–1486.
50. Depreux, P., Lesieur, D., Mansour, H. A., Morgan, P., Howell, H. E., Renard, P., Caignard, D. H., Pfeiffer, B., Delagrange, P., Guardiola, B. *et al*. Synthesis and structure-activity relationships of novel naphthalenic and bioisosteric related amidic derivatives as melatonin receptor ligands. *J. Med. Chem.* 1994, *37*(20), 3231–3239.
51. Blaskó, G., Major, E., Blaskó, G., Rózsa, I., Szántay, C. Pyrimido[1,6-*a*]pyrido[3,4-*b*] indoles as new platelet inhibiting agents. *Eur. J. Med. Chem.* 1986, *21*, 91–95.
52. Kardos, J., Blaskó, G., Simonyi, M., Szántay, C. Octahydroindolo[2,3-a]quinolizin-2-one, a novel structure for γ-aminobutyric acid (GABA) uptake inhibition. *Eur. J. Med. Chem.* 1985, *21*, 151–154.
53. Salituro, F. G., Harrison, B. L., Baron, B. M., Nyce, P. L., Stewart, K. T., Kehne, J. H., White, H. S., McDonald, I. 3-(2-Carboxyindol-3-yl)propionic acid-based antagonists of the *N*-methyl-D-aspartic receptor associated glycine binding site. *J. Med. Chem.* 1992, *35*, 1791–1799.
54. Calvino, R., Stilo, A. Di., Fruttero, R., Gasco, A. M., Sorba, G., Gasco, A. Pharmacochemistry of the furoxan ring: recent developments. *Il Farmaco* 1993, *48*, 321–334.
55. Lipinski, C. A., Aldinger, C. E., Beyer, T. A., Bordner, J., Burdi, D. F., Bussolotti, D. L., Inskeep, P. B., Siegel, T. W. Hydantoin isosteres. *In vivo* active spiro hydroxy acetic aldose reductase inhibitors. *J. Med. Chem.* 1992, *35*, 2169–2177.
56. Lober, S., Hubner, H., Gmeiner, P. Fused azaindole derivatives: molecular design, synthesis and in vitro pharmacology leading to the preferential dopamine D3 receptor agonist FAUC 725. *Bioorg. Med. Chem. Lett.* 2002, *12*(17), 2377–2380.
57. Krogsgaard-Larsen, P., Hjeds, H., Falch, E., Jørgensen, F. S., Nielsen, L. Recent advances in GABA agonists, antagonists and uptake inhibitors: structure-activity relationships and therapeutic potential. In *Advances in Drug Research* (Testa, B., Ed.), Vol. 17. Academic Press: London, 1988, pp. 381–456.
58. Almquist, R. G., Chao, W. R., Jennings-White, C. Synthesis and biological activity of carboxylic acid replacement analogues of the potent angiotensin converting enzyme inhibitor 5(S)-benzamido-4-oxo-6-phenylhexanoyl-L-proline. *J. Med. Chem.* 1985, *28*, 1067–1071.
59. Massa, S., Mai, A., Sbardella, G., Esposito, M., Ragno, R., Loidl, P., Brosch, G. 3-(4-Aroyl-1H-pyrrol-2-yl)-*N*-hydroxy-2-propenamides, a new class of synthetic histone deacetylase inhibitors. *J. Med. Chem.* 2001, *44*(13), 2069–2072.
60. Lu, Q., Yang, Y. T., Chen, C. S., Davis, M., Byrd, J. C., Etherton, M. R., Umar, A., Chen, C. S. Zn2+-chelating motif-tethered short-chain fatty acids as a novel class of histone deacetylase inhibitors. *J. Med. Chem.* 2004, *47*(2), 467–474.
61. Remiszewski, S. W., Sambucetti, L. C., Bair, K. W., Bontempo, J., Cesarz, D., Chandramouli, N., Chen, R., Cheung, M., Cornell-Kennon, S., Dean, K., Diamantidis, G., France, D., Green, M. A., Howell, K. L., Kashi, R., Kwon, P., Lassota, P., Martin, M. S., Mou, Y., Perez, L. B., Sharma, S., Smith, T., Sorensen, E., Taplin, F., Trogani, N., Versace, R., Walker, H., Weltchek-Engler, S., Wood, A., Wu, A., Atadja, P. *N*-hydroxy-3-phenyl-2-propenamides as novel inhibitors of human histone deacetylase with *in vivo* antitumor activity: discovery of (2E)-*N*-hydroxy-3-[4-[[(2-hydroxyethyl)[2-(1H-indol-3-yl)ethyl]amino]methyl]phenyl]-2-propenamide (NVP-LAQ824). *J. Med. Chem.* 2003, *46*(21), 4609–4624.
62. Plumb, J. A., Finn, P. W., Williams, R. J., Bandara, M. J., Romero, M. R., Watkins, C. J., La Thangue, N. B., Brown, R. Pharmacodynamic response and inhibition of growth of human tumor xenografts by the novel histone deacetylase inhibitor PXD101. *Mol. Cancer Ther.* 2003, *2*(8), 721–728.
63. Kelly, W. K., O'Connor, O. A., Krug, L. M., Chiao, J. H., Heaney, M., Curley, T., MacGregore-Cortelli, B., Tong, W., Secrist, J. P., Schwartz, L., Richardson, S., Chu, E., Olgac, S., Marks, P. A., Scher, H., Richon, V. M. Phase I study of an oral histone deacetylase inhibitor,

suberoylanilide hydroxamic acid, in patients with advanced cancer. *J. Clin. Oncol.* 2005, *2*(17), 3923–3931.

64. Buggy, J. J., Cao, Z. A., Bass, K. E., Verner, E., Balasubramanian, S., Liu, L., Schultz, B. E., Young, P. R., Dalrymple, S. A. CRA-024781: a novel synthetic inhibitor of histone deacetylase enzymes with antitumor activity *in vitro* and *in vivo*. *Mol. Cancer Ther.* 2006, *5*(5), 1309–1317.
65. Hanessian, S., Moitessier, N., Gauchet, C., Viau, M. *N*-aryl sulfonyl homocysteine hydroxamate inhibitors of matrix metalloproteinases: further probing of the S(1), S(1)′, and S(2)′ pockets. *J. Med. Chem.* 2001, *44*(19), 3066–3073.
66. Aranapakam, V., Davis, J. M., Grosu, G. T., Baker, J., Ellingboe, J., Zask, A., Levin, J. I., Sandanayaka, V. P., Du, M., Skotnicki, J. S., DiJoseph, J. F., Sung, A., Sharr, M. A., Killar, L. M., Walter, T., Jin, G., Cowling, R., Tillett, J., Zhao, W., McDevitt, J., Xu, Z. B. Synthesis and structure-activity relationship of *N*-substituted 4-aryl-sulfonylpiperidine-4-hydroxamic acids as novel, orally active matrix metalloproteinase inhibitors for the treatment of osteoarthritis. *J. Med. Chem.* 2003, *46*(12), 2376–2396.
67. Aranapakam, V., Grosu, G. T., Davis, J. M., Hu, B., Ellingboe, J., Baker, J. L., Skotnicki, J. S., Zask, A., DiJoseph, J. F., Sung, A., Sharr, M. A., Killar, L. M., Walter, T., Jin, G., Cowling, R. Synthesis and structure-activity relationship of alpha-sulfonylhydroxamic acids as novel, orally active matrix metalloproteinase inhibitors for the treatment of osteoarthritis. *J. Med. Chem.* 2003, *46*(12), 2361–2375.
68. Noe, M. C., Natarajan, V., Snow, S. L., Mitchell, P. G., Lopresti-Morrow, L., Reeves, L. M., Yocum, S. A., Carty, T. J., Barberia, J. A., Sweeney, F. J., Liras, J. L., Vaughn, M., Hardink, J. R., Hawkins, J. M., Tokar, C. Discovery of 3,3-dimethyl-5-hydroxypipecolic hydroxamate-based inhibitors of aggrecanase and MMP-13. *Bioorg. Med. Chem. Lett.* 2005, *15*(11), 2808–2811.
69. Duan, J. J., Chen, L., Wasserman, Z. R., Lu, Z., Liu, R. Q., Covington, M. B., Qian, M., Hardman, K. D., Magolda, R. L., Newton, R. C., Christ, D. D., Wexler, R. R., Decicco, C. P. Discovery of gamma-lactam hydroxamic acids as selective inhibitors of tumor necrosis factor alpha converting enzyme: design, synthesis, and structure-activity relationships. *J. Med. Chem.* 2002, *45*(23), 4954–4957.
70. Kohler, H.v., Eichler, B., Salewski, R. Untersuchungen zum sauerstoffanalogen Charakter der $C(CN)_2$- und NCN- Gruppen. *Z. Anorg. Chem.* 1970, *379*, 183–192.
71. Kwon, C.-H., Nagasawa, H. T., DeMaster, E. G., Shirota, F. N. Acyl, *N*-protected α-aminoacyl, and peptidyl derivatives as prodrug forms of the alcohol deterrent agent cyanamide. *J. Med. Chem.* 1986, *29*, 1922–1929.
72. Drummond, J. T., Johnson, G. Convenient procedure for the preparation of alkyl end aryl substituted *N*-(aminoalkylacyl)sulfonamides. *Tetrahedron Let* 1988, *29*, 1653–1656.
73. Albright, J. D., DeVries, V. G., Du, M. D., Largis, E. E., Miner, T. G., Reich, M. F., Shepherd, R. G. Potential antiatherosclerotic agents. 3. Substituted benzoic and non benzoic analogues of cetabon. *J. Med. Chem.* 1983, *26*, 1393–1411.
74. Uehling, D. E., Donaldson, K. H., Deaton, D. N., Hyman, C. E., Sugg, E. E., Barrett, D. G., Hughes, R. G., Reitter, B., Adkison, K. K., Lancaster, M. E., Lee, F., Hart, R., Paulik, M. A., Sherman, B. W., True, T., Cowan, C. Synthesis and evaluation of potent and selective beta(3) adrenergic receptor agonists containing acylsulfonamide, sulfonylsulfonamide, and sulfonylurea carboxylic acid isosteres. *J. Med. Chem.* 2002, *45*(3), 567–583.
75. Johansson, A., Poliakov, A., Akerblom, E., Wiklund, K., Lindeberg, G., Winiwarter, S., Danielson, U. H., Samuelsson, B., Hallberg, A. Acyl sulfonamides as potent protease inhibitors of the hepatitis C virus full-length NS3 (protease-helicase/NTPase): a comparative study of different C-terminals. *Bioorg. Med. Chem.* 2003, *11*(12), 2551–2568.
76. Noe, M. C., Natarajan, V., Snow, S. L., Wolf-Gouveia, L. A., Mitchell, P. G., Lopresti-Morrow, L., Reeves, L. M., Yocum, S. A., Otterness, I., Bliven, M. A., Carty, T. J., Barberia, J. T., Sweeney, F. J., Liras, J. L., Vaughn, M. Discovery of 3-OH-3-methylpipecolic hydroxamates: potent orally active inhibitors of aggrecanase and MMP-13. *Bioorg. Med. Chem. Lett.* 2005, *15*(14), 3385–3388.
77. Buu-Hoï, N. P., Lambelin, G., Lepoivre, C., Gillet, C., Gautier, M., Thiriaux, J. Un nouvel agent antiinflammatoire de structure non stéroïdique: l'acide *p*-butoxyphénylacéthydroxamique. *C. R. Acad. Sci.(Paris).* 1965, *261*, 2259–2262.
78. Orzalesi, G., Selleri, R., Pharmaceutical 2-(4-isobutylphenyl) propionohydroxamic acid. *Ger. Offen. 2,400,531 (24 Jul. 1974, to Societa Italo-Britannica L. Manetti & H.Roberts e C.). 1974, C.A. 1974; 81; 120272i*.
79. De Martiis, F., Franzone, J. S., Tamietto, T. Sintesi e Proprieta Antiflogistiche di Alcuni Acidici Indolil-Acetoidrossammici. *Bol. Chim. Farm.* 1975, *114*, 309–318.
80. Orzalesi, G., Mari, F., Bertol, E., Selleri, R., Pisaturo, G. Anti-inflammatory agents: determination of ibuproxam and its metabolite in humans. *Arzneim.-Forsch.* 1980, *30*, 1607–1609.
81. Demay, F., De Sy, J. A new non-steroidal anti-inflammatory drug (NSAID) in current rheumatologic practice (oxamethacin). *Curr. Ther. Res.* 1982, *31*, 113–118.
82. Vergin, H., Ferber, H., Brunner, F., Kukovetz, W. R. Pharmakokinetik und Biotransformation von Oxametacin bei gesunden Probanden. *Arzneim.-Forsch.* 1981, *31*, 513–518.
83. Allegretti, M., Bertini, R., Cesta, M. C., Bizzarri, C., Di Bitondo, R., Di Cioccio, V., Galliera, E., Berdini, V., Topai, A., Zampella, G., Russo, V., Di Bello, N., Nano, G., Nicolini, L., Locati, M., Fantucci, P., Florio, S., Colotta, F. 2-Arylpropionic CXC chemokine receptor 1 (CXCR1) ligands as novel noncompetitive CXCL8 inhibitors. *J. Med. Chem.* 2005, *48*(13), 4312–4331.
84. Wan, Y., Wallinder, C., Johansson, B., Holm, M., Mahalingam, A. K., Wu, X., Botros, M., Karlen, A., Pettersson, A., Nyberg, F., Fandriks, L., Hallberg, A., Alterman, M. First reported nonpeptide AT1 receptor agonist (L-162,313) acts as an AT2 receptor agonist *in vivo*. *J. Med. Chem.* 2004, *47*(6), 1536–1546.
85. Lobb, K. L., Hipskind, P. A., Aikins, J. A., Alvarez, E., Cheung, Y. Y., Considine, E. L., De Dios, A., Durst, G. L., Ferritto, R., Grossman, C. S., Giera, D. D., Hollister, B. A., Huang, Z., Iversen, P. W., Law, K. L., Li, T., Lin, H. S., Lopez, B., Lopez, J. E., Cabrejas, L. M., McCann, D. J., Molero, V., Reilly, J. E., Richett, M. E., Shih, C., Teicher, B., Wikel, J. H., White, W. T., Mader, M. M. Acyl sulfonamide anti-proliferatives: benzene substituent structure-activity relationships for a novel class of antitumor agents. *J. Med. Chem.* 2004, *47*(22), 5367–5380.
86. Mader, M. M., Shih, C., Considine, E., Dios, A. D., Grossman, C. S., Hipskind, P. A., Lin, H. S., Lobb, K. L., Lopez, B., Lopez, J. E., Cabrejas, L. M., Richett, M. E., White, W. T., Cheung, Y. Y., Huang, Z., Reilly, J. E., Dinn, S. R. Acyl sulfonamide anti-proliferatives. Part 2: activity of heterocyclic sulfonamide derivatives. *Bioorg. Med. Chem. Lett.* 2005, *15*(3), 617–620.
87. Hattori, K., Tanaka, A., Fujii, N., Takasugi, H., Tenda, Y., Tomita, M., Nakazato, S., Nakano, K., Kato, Y., Kono, Y., Murai, H., Sakane, K. Discovery of diphenyloxazole and Ndelta-Z-ornithine derivatives as highly potent and selective human prostaglandin EP(4) receptor antagonists. *J. Med. Chem.* 2005, *48*(9), 3103–3106.
88. Bovy, P. R., Reitz, D. B., Collins, J. T., Chamberlain, T. S., Olins, G. M., Corpus, V. M., Mc Mahon, E. G., Palomo, M. A., Koepke, J. P., McGraw, D. E., Gaw, G. J. Nonpeptide angiotensin II antagonists: *N*-phenyl-1-*H*-pyrrole derivatives are angiotensin II receptor antagonists. *J. Med. Chem.* 1993, *36*, 101–110.
89. Marshall, W. S., Goodson, T., Cullinan, G. J., Swanson-Bean, D., Haisch, K. D., Rinkema, L. E., Fleisch, J. H. Leukotriene receptor antagonists. 1. Synthesis and structure-activity relationships of alkoxyacetophenone derivatives. *J. Med. Chem.* 1987, *30*, 682–689.
90. Krogsgaard-Larsen, P. In (Hansch, C., Sammes, P. G., Taylor, J. B., Emmet, J. C., Eds) *Comprehensive Medicinal Chemistry*, Vol. 3. Pergamon Press: Oxford, 1990, pp. 493–537.

91. Lunn, W. H. W., Schoepp, D. D., Lodge, D., True, R. A., Millar, J. D. Poster N° P-041.A. LY262466, DL-2-amino-3-(4-hydroxy-1,2,5-thiazol-3-yl) propanoic acid hydrochloride, a novel and selective agonist at the AMPA excitatory amino acid receptor, *XIIth International Symposium on Medicinal Chemistry*, Basel, Switzerland, September 13–17, 1992, pp. 113–117.
92. Atkinson, J. G., Girard, Y., Rokach, J., Rooney, C. S., McFarlane, C. S., Rackham, A., Share, N. N. Kojic amine-A novel γ-aminobutyric acid analogue. *J. Med. Chem.* 1979, *22*, 90–106.
93. Hulin, B., McCarthy, P. A., Gibbs, E. M. The glitazone family of antidiabetic agents. *Curr. Pharm. Des.* 1996, *2*, 85–102.
94. Henke, B. R. Peroxisome proliferator-activated receptor alpha/gamma dual agonists for the treatment of type 2 diabetes. *J. Med. Chem.* 2004, *47*(17), 4118–4127.
95. Gezginci, M. H., Martin, A. R., Franzblau, S. G. Antimycobacterial activity of substituted isosteres of pyridine- and pyrazinecarboxylic acids. 2. *J. Med. Chem.* 2001, *44*(10), 1560–1563.
96. Nicolaou, I., Zika, C., Demopoulos, V. J. [1-(3,5-Difluoro-4-hydroxyphenyl)-1H-pyrrol-3-yl]phenylmethanone as a bioisostere of a carboxylic acid aldose reductase inhibitor. *J. Med. Chem.* 2004, *47*(10), 2706–2709.
97. Qiu, J., Stevenson, S. H., O'Beirne, M. J., Silverman, R. B. 2,6-Difluorophenol as a bioisostere of a carboxylic acid: bioisosteric analogues of gamma-aminobutyric acid. *J. Med. Chem.* 1999, *42*(2), 329–332.
98. Liljebris, C., Larsen, S. D., Ogg, D., Palazuk, B. J., Bleasdale, J. E. Investigation of potential bioisosteric replacements for the carboxyl groups of peptidomimetic inhibitors of protein tyrosine phosphatase 1B: identification of a tetrazole-containing inhibitor with cellular activity. *J. Med. Chem.* 2002, *45*(9), 1785–1798.
99. Momose, Y., Maekawa, T., Odaka, H., Ikeda, H., Sohda, T. Novel 5-substituted-1H-tetrazole derivatives as potent glucose and lipid lowering agents. *Chem. Pharm. Bull. (Tokyo)* 2002, *50*(1), 100–111.
100. Biot, C., Bauer, H., Schirmer, R. H., Davioud-Charvet, E. 5-Substituted tetrazoles as bioisosteres of carboxylic acids. Bioisosterism and mechanistic studies on glutathione reductase inhibitors as antimalarials. *J. Med. Chem.* 2004, *47*(24), 5972–5983.
101. Kimura, T., Shuto, D., Hamada, Y., Igawa, N., Kasai, S., Liu, P., Hidaka, K., Hamada, T., Hayashi, Y., Kiso, Y. Design and synthesis of highly active Alzheimer's beta-secretase (BACE1) inhibitors, KMI-420 and KMI-429, with enhanced chemical stability. *Bioorg. Med. Chem. Lett.* 2005, *15*(1), 211–215.
102. Roppe, J., Smith, N. D., Huang, D., Tehrani, L., Wang, B., Anderson, J., Brodkin, J., Chung, J., Jiang, X., King, C., Munoz, B., Varney, M. A., Prasit, P., Cosford, N. D. Discovery of novel heteroarylazoles that are metabotropic glutamate subtype 5 receptor antagonists with anxiolytic activity. *J. Med. Chem.* 2004, *47*(19), 4645–4648.
103. Kozikowski, A. P., Zhang, J., Nan, F., Petukhov, P. A., Grajkowska, E., Wroblewski, J. T., Yamamoto, T., Bzdega, T., Wroblewska, B., Neale, J. H. Synthesis of urea-based inhibitors as active site probes of glutamate carboxypeptidase II: efficacy as analgesic agents. *J. Med. Chem.* 2004, *47*(7), 1729–1738.
104. Valgeirsson, J., Nielsen, E. O., Peters, D., Mathiesen, C., Kristensen, A. S., Madsen, U. Bioisosteric modifications of 2-arylureidobenzoic acids: selective noncompetitive antagonists for the homomeric kainate receptor subtype GluR5. *J. Med. Chem.* 2004, *47*(27), 6948–6957.
105. Kraus, J. L. Isosterism and molecular modification in drug design: tetrazole analogue of GABA: effects on enzymes of the gamma-aminobutyrate system. *Pharmacol. Res. Commun.* 1983, *15*, 183–189.
106. Hallinan, E. A., Tsymbalov, S., Dorn, C. R., Pitzele, B. S., Hansen, D. W., Jr, Moore, W. M., Jerome, G. M., Connor, J. R., Branson, L. F., Widomski, D. L., Zhang, Y., Currie, M. G., Manning, P. T. Synthesis and biological characterization of L-N(6)-(1-iminoethyl)lysine 5-tetrazole-amide, a prodrug of a selective iNOS inhibitor. *J. Med. Chem.* 2002, *45*(8), 1686–1689.
107. Schlewer, G., Wermuth, C. G., Chambon, J.-P. Analogues tétrazoliques d'agents GABA-mimétiques. *Eur. J. Med. Chem.* 1984, *19*, 181–186.
108. Krogsgaard-Larsen, P., Ferkany, J. W., Nielsen, E. O., Madsen, U., Ebert, B., Johansen, J. S., Diemer, S. H., Bruhn, T., Beattie, D. T., Curtis, D. R. Novel class of amino acid antagonists at non-*N*-methyl-D-aspartic acid excitatory amino acid receptors. Synthesis, *in vitro* and *in vivo* pharmacology, and neuroprotection. *J. Med. Chem.* 1991, *34*, 123–130.
109. Krogsgaard-Larsen, P., Rodolskov-Christiansen, T. GABA agonists. Synthesis and structure-activity studies on analogues of isoguvacine and THIP. *Eur. J. Med. Chem.* 1979, *14*, 157–164.
110. Kraus, J. L. Isosterism and molecular modification in drug design: new n-dipropylacetate analogs as inhibitors of succinic semi aldehyde dehydrogenase. *Pharmacol. Res. Commun.* 1983, *15*, 119–129.
111. Lichtenthaler, F. W., Heidel, P. Intermediates in the formation of γ-pyrones from hexose derivatives: a simple synthesis of kojic acid and hydroxymaltol. *Angew. Chem. Int. Ed.* 1969, *8*, 978–979.
112. Froestl, W., Furet, P., Hall, R. G., Mickel, S. J., Strub, D., Sprecher, G.v., Baumann, P. A., Bernasconi, R., Brugger, F., Felner, A., Gentsch, C., Hauser, K., Jaeckel, J., Karlsson, G., Klebs, K., Maître, L., Marescaux, C., Moser, P., Pozza, M. F., Rihs, G., Schmutz, M., Steinmann, M. W., Riezen, H.v., Vassout, A., Mondadori, C., Olpe, H.-R., Waldmeier, P. C., Bittiger, H. $GABA_B$ antagonists: novel CNS-active compounds. In *Perspectives in Medicinal Chemistry* (Testa, B., Kyburz, E., Fuhrer, W., Giger, R., Eds), VHC: Weinheim, 1993, pp. 259–272.
113. Rosowski, A., Forsch, R. A., Freisheim, J. H., Moran, R. G., Wick, M. Methotrexate analogues. 19. Replacement of the glutamate side chain in classical antifolates by L-homocysteic acid and L-cysteic acid: effect on enzyme inhibition and antitumor activity. *J. Med. Chem.* 1984, *27*, 600–604.
114. Watkins, J. C., Krogsgaard-Larsen, P., Honoré, T. Structure-activity relationships in the development of excitatory amino acid receptor agonists and competitive antagonists. *Trends Pharm. Sci.* 1990, *11*, 25–33.
115. Drysdale, M. J., Pritchard, M. C., Horwell, D. C. Rationally designed "dipeptoid" analogues of CCK. Acid mimics of the potent and selective non peptide CCK-B receptor antagonist CI-988. *J. Med. Chem.* 1992, *35*, 2573–2581.
116. Franz, R. D. Comparisons of *p*Ka and log *P* values of some carboxylic and phosphonic acids: synthesis and measurement. *AAPS PharmSci.* 2001, *3*(2), E10.
117. Kinney, W. A., Lee, N. E., Garrison, D. T., Podlesny, E. J., Jr, Simmonds, J. T., Bramlet, D., Notvest, R. R., Kowal, D. M., Tasse, R. P. Bioisosteric replacement of the α-amino carboxylic functionality in 2-amino-5-phosphonopentanoic acid yields unique 3,4-diamino-3-cyclobutene-1,2-dione containing NMDA antagonists. *J. Med. Chem.* 1992, *35*, 4720–4726.
118. Gao, Y., Luo, J., Yao, Z.-J., Guo, R., Zou, H., Kelley, J., Voigt, J. H., Yang, D., Burke, T. R., Jr Inhibition of Grb2 SH2 domain binding by non-phosphate-containing ligands. 2. 4-(2-Malonyl)phenylalanine as a potent phosphotyrosyl mimetic. *J. Med. Chem.* 2000, *43*, 911–920.
119. Shapiro, G., Floersheim, P., Boelsterli, J., Amstutz, R., Bolliger, G., Gammenthaler, H., Gmelin, G., Supavilai, P., Walkinshaw, M. Muscarinic activity of the thiolactone, lactam, lactol, and thiolactol analogues of pilocarpine and a hypothetical model for the binding of agonists to the m_1 receptor. *J. Med. Chem.* 1992, *35*, 15–27.
120. Thompkins, L., Lee, K. H. Comparison of analgesic effects of isosteric variations of salicylic acid and aspirin (acetylsalicylic acid). *J. Pharm. Sci.* 1975, *64*, 760–763.
121. Roth, G. J., Stanford, N., Majerus, P. W. Acetylation of prostaglandine synthase by aspirin. *Proc. Nat. Acad. Sci. USA* 1975, *72*, 3073–3076.

122. Bundgaard, H. *Design of Prodrugs*. Elsevier: Amsterdam, 1985. p. 55–61.
123. Sauerberg, P., Chen, J., WoldeMussie, E., Rapoport, H. Cyclic carbamate analogues of pilocarpine. *J. Med. Chem.* 1989, *32*, 1322–1326.
124. Sauerberg, P., Kindtler, J. W., Nielsen, L., Sheardown, M. J., Honoré, T. Muscarinic cholinergic agonists and antagonists of the 3-(3-alkyl-1,2,4-oxadiazol-5-yl)1,2,5,6-tetrahydropyridine type synthesis and structure-activity relationships. *J. Med. Chem.* 1991, *34*, 687–692.
125. Sauerberg, P., Olesen, P. H., Nielsen, S., Treppendahl, S., Sheardown, M. J., Honoré, T., Mitch, C. H., Ward, J. S., Pike, A. J., Bymaster, F. P., Sawyer, B. D., Shannon, H. E. Novel functional M_1selective muscarinic agonists. Synthesis and structure-activity relationships of 3-(1,2,5-thiadiazolyl)-1,2,5,6-tetrahydro-1-methylpyridines. *J. Med. Chem.* 1992, *35*, 2263–2274.
126. Sauerberg, P., Larsen, J. J., Falch, E., Krogsgaard-Larsen, P. A novel class of conformationally restricted heterocyclic muscarinic agonists. *J. Med. Chem.* 1986, *29*(6), 1004–1009.
127. Saunders, J., Cassidy, M., Freedman, S. B., Harley, E. A., Iversen, L. L., Kneen, C., MacLeod, A. M., Merchant, K., Snow, R. J., Baker, R. Novel quinuclidine-based ligands for the muscarinic cholinergic receptor. *J. Med. Chem.* 1990, *33*, 1128–1138.
128. Street, L. J., Baker, R., Book, T., Reeve, A. J., Saunders, J., Willson, T., Marwood, R. S., Patel, S., Freedman, S. B. Synthesis and muscarinic activity of quinuclidinyl- and (1-azanorbornyl)pyrazine derivatives. *J. Med. Chem.* 1992, *35*, 295–305.
129. Wadsworth, H. J., Jenkins, S. M., Orlek, B. S., Cassidy, F., Clark, M. S. G., Brown, F., Riley, G. J., Graves, D., Hawkins, J., Naylor, C. Synthesis and muscarinic activities of quinuclidin-3-yltriazole and -tetrazole derivatives. *J. Med. Chem.* 1992, *35*, 1280–1290.
130. Jenkins, S. M., Wadsworth, H. J., Bromidge, S., Orlek, B. S., Wyman, P. A., Riley, G. J., Hawkins, J. Substituent variation in azabicyclic triazole- and tetrazole-based muscarinic receptor ligands. *J. Med. Chem.* 1992, *35*(13), 2392–2406.
131. Bromidge, S. M., Brown, F., Cassidy, F., Clark, M. S., Dabbs, S., Hadley, M. S., Hawkins, J., Loudon, J. M., Naylor, C. B., Orlek, B. S., Riley, G. J. Design of [R-(Z)]-(+)-alpha-(methoxyimino)-1-azabicyclo[2.2.2]octane-3-acetonitri le (SB 202026), a functionally selective azabicyclic muscarinic M1 agonist incorporating the *N*-methoxy imidoyl nitrile group as a novel ester bioisostere. *J. Med. Chem.* 1997, *40*(26), 4265–4280.
132. Toja, E., Bonetti, C., Butti, A., Hunt, P., Fortin, M., Barzaghi, F., Formento, M. L., Maggioni, A., Nencioni, A., Galliani, G. 1-Alkyl-1,2,5,6-tetrahydropyridine-3-carboxaldehyde-*O*-alkyl-oximes: a new class of potent orally available active muscarinic agaonists related to arecoline. *Eur. J. Med. Chem.* 1991, *22*, 853–868.
133. Carroll, F. I. 2002 Medicinal Chemistry Division Award address: monoamine transporters and opioid receptors. Targets for addiction therapy. *J. Med. Chem.* 2003, *46*(10), 1775–1794.
134. Petukhov, P. A., Zhang, M., Johnson, K. J., Tella, S. R., Kozikowski, A. P. Sar studies of piperidine-based analogues of cocaine. Part 3: oxadiazoles,. *Bioorg. Med. Chem. Lett.* 2001, *11*(16), 2079–2083.
135. Street, L. J., Baker, R., Castro, J. L., Chambers, M. S., Guiblin, A. R., Hobbs, S. C., Matassa, V. G., Reeve, A. J., Beer, M. S., Middlemiss, D. N. *et al* Synthesis and serotonergic activity of 5-(oxadiazolyl)tryptamines: potent agonists for 5-HT1D receptors. *J. Med. Chem.* 1993, *36*(11), 1529–1538.
136. Kuduk, S. D., Ng, C., Feng, D. M., Wai, J. M., Chang, R. S., Harrell, C. M., Murphy, K. L., Ransom, R. W., Reiss, D., Ivarsson, M., Mason, G., Boyce, S., Tang, C., Prueksaritanont, T., Freidinger, R. M., Pettibone, D. J., Bock, M. G. 2,3-Diaminopyridine bradykinin B1 receptor antagonists. *J. Med. Chem.* 2004, *47*(26), 6439–6442.
137. Borzilleri, R. M., Zheng, X., Qian, L., Ellis, C., Cai, Z. W., Wautlet, B. S., Mortillo, S., Jeyaseelan, R., Sr., Kukral, D. W., Fura, A., Kamath, A., Vyas, V., Tokarski, J. S., Barrish, J. C., Hunt, J. T., Lombardo, L. J., Fargnoli, J., Bhide, R. S. Design, synthesis, and evaluation of orally active 4-(2,4-difluoro-5-(methoxycarbamoyl)phe nylamino)pyrrolo[2,1-f][1,2,4]triaz ines as dual vascular endothelial growth factor receptor-2 and fibroblast growth factor receptor-1 inhibitors. *J. Med. Chem.* 2005, *48*(12), 3991–4008.
138. Sakamoto, T., Cullen, M. D., Hartman, T. L., Watson, K. M., Buckheit, R. W., Pannecouque, C., De Clercq, E., Cushman, M. Synthesis and anti-HIV activity of new metabolically stable alkenyldiarylmethane non-nucleoside reverse transcriptase inhibitors incorporating *N*-methoxy imidoyl halide and 1,2,4-oxadiazole systems. *J. Med. Chem.* 2007, *50*(14), 3314–3321.
139. Deng, B. L., Hartman, T. L., Buckheit, R. W., Jr, Pannecouque, C., De Clercq, E., Cushman, M. Replacement of the metabolically labile methyl esters in the alkenyldiarylmethane series of non-nucleoside reverse transcriptase inhibitors with isoxazolone, isoxazole, oxazolone, or cyano substituents. *J. Med. Chem.* 2006, *49*(17), 5316–5323.
140. Watson, K. G., Brown, R. N., Cameron, R., Chalmers, D. K., Hamilton, S., Jin, B., Krippner, G. Y., Luttick, A., McConnell, D. B., Reece, P. A., Ryan, J., Stanislawski, P. C., Tucker, S. P., Wu, W. Y., Barnard, D. L., Sidwell, R. W. An orally bioavailable oxime ether capsid binder with potent activity against human rhinovirus. *J. Med. Chem.* 2003, *46*(15), 3181–3184.
141. Kozikowski, A. P., Roberti, M., Xiang, L., Bergmann, J. S., Callahan, P. M., Cunningham, K. A., Johnson, K. M. Structure-activity relationship studies of cocaine: replacement of the C-2 ester group by vinyl argues against H-bonding and provides an esterase-resistant, high-affinity cocaine analogue. *J. Med. Chem.* 1992, *35*, 4764–4766.
142. Lipinski, C. A. Bioisosterism in drug design. In *Annual Report in Medicinal Chemistry* (Bailey, D. M., Ed.), Vol. 21. Academic Press: San Diego, 1986, pp. 283–291.
143. Wermuth, C. G. Aminopyridazines – an alternate route to potent muscarinic agonists with no cholinergic syndrome. *Il Farmaco* 1993, *48*, 253–274.
144. Huff, J. R., Anderson, P. S., Baldwin, J. J., Clineschmidt, B. V., Guare, J. P., Lotti, V. J., Pettibone, D. J., Randall, W. C., Vacca, J. P. *N*-(1,3,4,6,7,12b-hexahydro-2*H*-benzo[b]furo[2,3-a]quinolizin-2-yl)-*N*-methyl-2-hydroxyethane-sulfonamide: a potent and selective α_2-adrenoreceptor antagonist. *J. Med. Chem.* 1985, *28*, 1756–1759.
145. Spatola, A. F. Peptide backbone modifications: structure-activity analysis of peptides containing amide bond surrogates. In *Chemistry and Biochemistry of Amino Acids, Peptides and Proteins* (Weinstein, B., Ed.), Vol. 7. Marcel Dekker: New York, 1983, pp. 267–357.
146. Fauchère, J.-L. Elements for the rational design of peptide drugs. In *Advances in Drug Research* (Testa, B., Ed.), Vol. 15. Academic Press: London, 1986, pp. 29–69.
147. Fournier, J.-P., Moreau, R. C., Narcisse, G., Choay, P. Synthèse et propriétés pharmacologiques de sulfonylurées isostères du glibenclamide. *Eur. J. Med. Chem.* 1982, *17*, 81–84.
148. Smith, A. B., III, Holcomb, R. C., Guzman, M. C., Keenan, T. P., Sprengeler, P. A., Hirschmann, R. An Effective Synthesis of Scalemic 3,5,5-Trisubstituted Pyrrolin-4-ones. *Tetrahedron Lett.* 1993, *34*, 63–66.
149. Smith, A. B., III, Keenan, T. P., Holcomb, R. C., Sprengeler, P. A., Guzman, M. C., Wood, J. L., Caroll, P. J., Hirschmann, R. Design, synthesis, and crystal structure of a pyrrolinone-based peptidomimetic possessing the conformation of a β-strand: potential application to the design of novel inhibitors of proteolytic enzymes. *J. Amer. Chem. Soc.* 1992, *114*, 10672–10674.
150. Allmendinger, T., Felder, E., Hungerbuehler, E. Fluoroolefin dipeptide isosteres. In *Selective Fluorination in Organic and Bioorganic Chemistry* (Weldi, J. T., Ed.). American Chemical Society: Washington, 1991, pp. 186–195. Vol. ACS Symp. 456
151. Chandrakumar, N. S., Yonan, P. K., Stapelfeld, A., Svage, M., Rorbacher, E., Contreras, P. C., Hammond, D. Preparation and opioid

activity of analogues of the analgesic dipeptide 2,6-dimethyl-l-tyrosyl-*N*-(3-phenylpropyl)-d-alanylamide. *J. Med. Chem.* 1992, *35*, 223–233.

152. Vu, C. B., Corpuz, E. G., Merry, T. J., Pradeepan, S. G., Bartlett, C., Bohacek, R. S., Botfield, M. C., Eyermann, C. J., Lynch, B. A., MacNeil, I. A., Ram, M. K., van Schravendijk, M. R., Violette, S., Sawyer, T. K. Discovery of potent and selective SH2 inhibitors of the tyrosine kinase ZAP-70. *J. Med. Chem.* 1992, *42*(20), 4088–4098.
153. Clitherow, J. W., Beswick, P., Irving, W. J., Scopes, D. I. C., Barnes, J. C., Clapham, J., Brown, J. D., Evans, D. J., Hayes, A. G. Novel 1,2,4-oxadiazoles as potent and selective histamine H_3 receptor antagonists. *Bioorg. Med. Chem. Lett.* 1996, *6*, 833–838.
154. Swain, C. J., Baker, R., Kneen, C., Moseley, J., Saunders, J., Seward, E. M., Stevenson, G., Beer, M., Stanton, J., Watling, K. Novel 5-HT3 antagonists. Indole oxadiazoles. *J. Med. Chem.* 1991, *34*(1), 140–151.
155. Biftu, T., Feng, D. D., Liang, G. B., Kuo, H., Qian, X., Naylor, E. M., Colandrea, V. J., Candelore, M. R., Cascieri, M. A., Colwell, L. F., Jr, Forrest, M. J., Hom, G. J., MacIntyre, D. E., Stearns, R. A., Strader, C. D., Wyvratt, M. J., Fisher, M. H., Weber, A. E. Synthesis and SAR of benzyl and phenoxymethylene oxadiazole benzenesulfonamides as selective beta3 adrenergic receptor agonist antiobesity agents. *Bioorg. Med. Chem. Lett.* 2000, *10*(13), 1431–1434.
156. Feng, D. D., Biftu, T., Candelore, M. R., Cascieri, M. A., Colwell, L. F., Jr, Deng, L., Feeney, W. P., Forrest, M. J., Hom, G. J., MacIntyre, D. E., Miller, R. R., Stearns, R. A., Strader, C. D., Tota, L., Wyvratt, M. J., Fisher, M. H., Weber, A. E. Discovery of an orally bioavailable alkyl oxadiazole beta3 adrenergic receptor agonist. *Bioorg. Med. Chem. Lett.* 2000, *10*(13), 1427–1429.
157. Naylor, E. M., Colandrea, V. J., Candelore, M. R., Cascieri, M. A., Colwell, L. F., Jr, Deng, L., Feeney, W. P., Forrest, M. J., Hom, G. J., MacIntyre, D. E., Strader, C. D., Tota, L., Wang, P. R., Wyvratt, M. J., Fisher, M. H., Weber, A. E. 3-Pyridylethanolamines: potent and selective human beta 3 adrenergic receptor agonists. *Bioorg. Med. Chem. Lett.* 1998, *8*(21), 3087–3092.
158. Naylor, E. M., Parmee, E. R., Colandrea, V. J., Perkins, L., Brockunier, L., Candelore, M. R., Cascieri, M. A., Colwell, L. F., Jr, Deng, L., Feeney, W. P., Forrest, M. J., Hom, G. J., MacIntyre, D. E., Strader, C. D., Tota, L., Wang, P. R., Wyvratt, M. J., Fisher, M. H., Weber, A. E. Human beta3 adrenergic receptor agonists containing imidazolidinone and imidazolone benzenesulfonamides. *Bioorg. Med. Chem. Lett.* 1999, *9*(5), 755–758.
159. Parmee, E. R., Naylor, E. M., Perkins, L., Colandrea, V. J., Ok, H. O., Candelore, M. R., Cascieri, M. A., Deng, L., Feeney, W. P., Forrest, M. J., Hom, G. J., MacIntyre, D. E., Miller, R. R., Stearns, R. A., Strader, C. D., Tota, L., Wyvratt, M. J., Fisher, M. H., Weber, A. E. Human beta3 adrenergic receptor agonists containing cyclic ureidobenzenesulfonamides. *Bioorg. Med. Chem. Lett.* 1999, *9*(5), 749–754.
160. Tully, W. R., Gardner, C. R., Gillespie, R. J., Westwood, R. 2-(Oxadiazolyl)- and 2-(thiazolyl)imidazo[1,2-a]pyrimidines as agonists and inverse agonists at benzodiazepine receptors. *J. Med. Chem.* 1991, *34*(7), 2060–2067.
161. Orlek, B. S., Blaney, F. E., Brown, F., Clark, M. S., Hadley, M. S., Hatcher, J., Riley, G. J., Rosenberg, H. E., Wadsworth, H. J., Wyman, P. Comparison of azabicyclic esters and oxadiazoles as ligands for the muscarinic receptor. *J. Med. Chem.* 1991, *34*(9), 2726–2735.
162. Ladduwahetty, T., Baker, R., Cascieri, M. A., Chambers, M. S., Haworth, K., Keown, L. E., MacIntyre, D. E., Metzger, J. M., Owen, S., Rycroft, W., Sadowski, S., Seward, E. M., Shepheard, S. L., Swain, C. J., Tattersall, F. D., Watt, A. P., Williamson, D. W., Hargreaves, R. J. *N*-heteroaryl-2-phenyl-3-(benzyloxy)piperidines: a novel class of potent orally active human NK1 antagonists. *J. Med. Chem.* 1996, *39*(15), 2907–2914.
163. Borg, S., Vollinga, R. C., Labarre, M., Payza, K., Terenius, L., Luthman, K. Design, synthesis, and evaluation of Phe-Gly mimetics: heterocyclic building blocks for pseudopeptides. *J. Med. Chem.* 1999, *42*(21), 4331–4342.
164. Elzein, E., Ibrahim, P., Koltun, D. O., Rehder, K., Shenk, K. D., Marquart, T. A., Jiang, B., Li, X., Natero, R., Li, Y., Nguyen, M., Kerwar, S., Chu, N., Soohoo, D., Hao, J., Maydanik, V. Y., Lustig, D. A., Zeng, D., Leung, K., Zablocki, J. A. CVT-4325: a potent fatty acid oxidation inhibitor with favorable oral bioavailability. *Bioorg. Med. Chem. Lett.* 2004, *14*(24), 6017–6021.
165. Koltun, D. O., Marquart, T. A., Shenk, K. D., Elzein, E., Li, Y., Nguyen, M., Kerwar, S., Zeng, D., Chu, N., Soohoo, D., Hao, J., Maydanik, V. Y., Lustig, D. A., Ng, K. J., Fraser, H., Zablocki, J. A. New fatty acid oxidation inhibitors with increased potency lacking adverse metabolic and electrophysiological properties. *Bioorg. Med. Chem. Lett.* 2004, *14*(2), 549–552.
166. Einsiedel, J., Hubner, H., Gmeiner, P. Benzamide bioisosteres incorporating dihydroheteroazole substructures: EPC synthesis and SAR leading to a selective dopamine D4 receptor partial agonist (FAUC 179). *Bioorg. Med. Chem. Lett.* 2001, *11*(18), 2533–2536.
167. Einsiedel, J., Hubner, H., Gmeiner, P. Cyclic amidines as benzamide bioisosteres: EPC synthesis and SAR studies leading to the selective dopamine D4 receptor agonist FAUC 312. *Bioorg. Med. Chem. Lett.* 2003, *13*(5), 851–854.
168. Einsiedel, J., Thomas, C., Hubner, H., Gmeiner, P. Phenyloxazoles and phenylthiazoles as benzamide bioisosteres: synthesis and dopamine receptor binding profiles. *Bioorg. Med. Chem. Lett.* 2000, *10*(17), 2041–2044.
169. Thurkauf, A., Yuan, J., Chen, X., He, X. S., Wasley, J. W., Hutchison, A., Woodruff, K. H., Meade, R., Hoffman, D. C., Donovan, H., Jones-Hertzog, D. K. 2-Phenyl-4(5)-[[4-(pyrimidin-2-yl)piperazin-1-yl]methyl]imidazole. A highly selective antagonist at cloned human D4 receptors. *J. Med. Chem.* 1997, *40*(1), 1–3.
170. Fletcher, S. R., McIver, E., Lewis, S., Burkamp, F., Leech, C., Mason, G., Boyce, S., Morrison, D., Richards, G., Sutton, K., Jones, A. B. The search for novel TRPV1-antagonists: from carboxamides to benzimidazoles and indazolones. *Bioorg. Med. Chem. Lett.* 2006, *16*(11), 2872–2876.
171. Wu, W. L., Burnett, D. A., Caplen, M. A., Domalski, M. S., Bennett, C., Greenlee, W. J., Hawes, B. E., O'Neill, K., Weig, B., Weston, D., Spar, B., Kowalski, T. Design and synthesis of orally efficacious benzimidazoles as melanin-concentrating hormone receptor 1 antagonists. *Bioorg. Med. Chem. Lett.* 2006, *16*(14), 3674–3678.
172. Wu, W. L., Burnett, D. A., Spring, R., Qiang, L., Sasikumar, T. K., Domalski, M. S., Greenlee, W. J., O'Neill, K., Hawes, B. E. Synthesis and structure-activity relationships of piperidine-based melanin-concentrating hormone receptor 1 antagonists. *Bioorg. Med. Chem. Lett.* 2006, *16*(14), 3668–3673.
173. Lanter, J. C., Fiordeliso, J. J., Allan, G. F., Musto, A., Hahn do, W., Sui, Z. A bioisosteric approach to the discovery of indole carbinol androgen receptor ligands. *Bioorg. Med. Chem. Lett.* 2006, *16*(21), 5646–5649.
174. Wouters, J., Michaux, C., Durant, F., Dogne, J. M., Delarge, J., Masereel, B. Isosterism among analogues of torasemide: conformational, electronic and lipophilic properties. *Eur. J. Med. Chem.* 2000, *35*(10), 923–929.
175. Luo, G., Mattson, G. K., Bruce, M. A., Wong, H., Murphy, B. J., Longhi, D., Antal-Zimanyi, I., Poindexter, G. S. Isosteric *N*-arylpiperazine replacements in a series of dihydropyridine NPY1 receptor antagonists. *Bioorg. Med. Chem. Lett.* 2004, *14*(24), 5975–5978.
176. Poindexter, G. S., Bruce, M. A., Breitenbucher, J. G., Higgins, M. A., Sit, S. Y., Romine, J. L., Martin, S. W., Ward, S. A., McGovern, R. T., Clarke, W., Russell, J., Antal-Zimanyi, I. Dihydropyridine neuropeptide Y Y1 receptor antagonists 2. Bioisosteric urea replacements. *Bioorg. Med. Chem.* 2004, *12*(2), 507–521.
177. Lumma, W. C., Jr, Anderson, P. S., Baldwin, J. J., Bolhofer, W. A., Habecker, C. N., Hirshfield, J. M., Pietruszkewicz, A. M., Randall,

W. C., Torchiana, M. L., Britcher, S. F., Clineschmidt, B. V., Denny, G. H., Hirschmann, R., Hoffman, J. M., Phillips, B. T., Streeter, K. B. Inhibitors of gastric acid secretion: 3,4-diamino-1,2,5-thiadiazole 1-oxides and 1,1-dioxides as urea equivalents in a series of histamine H_2-receptor antagonists. *J. Med. Chem.* 1982, *25*, 207–210.

178. Young, R. C., Ganellin, C. R., Graham, M. J., Grant, E. H. The dipole moments of 1,3-dimethylthiourea, 1,3-dimethyl-2-cyanoguanidine and 1,1-bis-methylamino-2-nitroethene in aqueous solution. *Tetrahedron* 1982, *38*, 1493–1497.
179. Young, R. C., Ganellin, C. R., Graham, M. J., Roantree, M. J., Grant, E. H. The dielectric properties of seven polar amidine-containing compounds of biological interest. *Tetrahedron Lett.* 1985, *26*, 1897–1900.
180. Wright, J. L., Gregory, T. F., Kesten, S. R., Boxer, P. A., Serpa, K. A., Meltzer, L. T., Wise, L. D. Subtype-selective *N*-methyl-d-aspartate receptor antagonists: synthesis and biological evaluation of 1-(heteroarylalkynyl)-4-benzylpiperidines. *J. Med. Chem.* 2000, *43*, 3408–3419.
181. Wu, W. L., Burnett, D. A., Spring, R., Greenlee, W. J., Smith, M., Favreau, L., Fawzi, A., Zhang, H., Lachowicz, J. E. Dopamine D1/D5 receptor antagonists with improved pharmacokinetics: design, synthesis, and biological evaluation of phenol bioisosteric analogues of benzazepine D1/D5 antagonists. *J. Med. Chem.* 2005, *48*(3), 680–693.
182. Burger, A. Isosterism and bioisosterism in drug design. *Progr. Drug Res.* 1991, *37*, 287–371.
183. Uloth, R. H., Kirk, J. R., Gould, W. A., Larsen, A. A. Sulfonanilides. I. Monoalkyl- and arylsulfonamidophenethanolamines. *J. Med. Chem.* 1966, *9*(1), 88–97.
184. Hacksell, U., Arvidsson, L. E., Svensson, U., Nilsson, J. L., Sanchez, D., Wikstrom, H., Lindberg, P., Hjorth, S., Carlsson, A. 3-Phenylpiperidines. Central dopamine-autoreceptor stimulating activity. *J. Med. Chem.* 1981, *24*(12), 1475–1482.
185. Bhaird, N. N., Fowler, C. J., Thorberg, O., Tipton, K. F. Involvement of catechol-*O*-methyl transferase in the metabolism of the putative dopamine autoreceptor agonist 3-PPP(3-(3-hydroxyphenyl)-*N*-*n*-propylpiperidine). *Biochem. Pharmacol.* 1985, *34*(19), 3599–3601.
186. Jaen, J. C., Wise, L. D., Caprathe, B. W., Tecle, H., Bergmeier, S., Humblet, C. C., Heffner, T. G., Meltzer, L. T., Pugsley, T. A. 4-(1,2,5,6-Tetrahydro-1-alkyl-3-pyridinyl)-2-thiazolamines: a novel class of compounds with central dopamine agonist properties. *J. Med. Chem.* 1990, *33*(1), 311–317.
187. Mierau, J., Schneider, F. J., Ensinger, H. A., Chio, C. L., Lajiness, M. E., Huff, R. M. Pramipexole binding and activation of cloned and expressed dopamine D2, D3 and D4 receptors. *Eur. J. Pharmacol.* 1995, *290*(1), 29–36.
188. Schneider, C. S., Mierau, J. Dopamine autoreceptor agonists: resolution and pharmacological activity of 2,6-diaminotetrahydrobenzothiazole and an aminothiazole analogue of apomorphine. *J. Med. Chem.* 1987, *30*(3), 494–498.
189. van Vliet, L. A., Rodenhuis, N., Wikstrom, H., Pugsley, T. A., Serpa, K. A., Meltzer, L. T., Heffner, T. G., Wise, L. D., Lajiness, M. E., Huff, R. M., Svensson, K., Haenen, G. R., Bast, A. Thiazoloindans and thiazolobenzopyrans: a novel class of orally active central dopamine (partial) agonists. *J. Med. Chem.* 2000, *43*(19), 3549–3557.
190. Blum, G., Gazit, A., Levitzki, A. Development of new insulin-like growth factor-1 receptor kinase inhibitors using catechol mimics. *J. Biol. Chem.* 2003, *278*(42), 40442–40454.
191. Hubner, H., Haubmann, C., Utz, W., Gmeiner, P. Conjugated enynes as nonaromatic catechol bioisosteres: synthesis binding experiments, and computational studies of novel dopamine receptor agonists recognizing preferentially the D subtype. *J. Med. Chem.* 2000, *43*(4), 756–762.
192. Lenz, C., Boeckler, F., Hubner, H., Gmeiner, P. Analogues of FAUC 73 revealing new insights into the structural requirements of nonaromatic dopamine D3 receptor agonists. *Bioorg. Med. Chem.* 2004, *12*(1), 113–117.
193. Lenz, C., Haubmann, C., Hubner, H., Boeckler, F., Gmeiner, P. Fancy bioisosteres: synthesis and dopaminergic properties of the endiyne FAUC 88 as a novel non-aromatic D3 agonist. *Bioorg. Med. Chem.* 2005, *13*(1), 185–191.
194. Street, L. J., Baker, R., Davey, W. B., Guiblin, A. R., Jelley, R. A., Reeve, A. J., Routledge, H., Sternfeld, F., Watt, A. P., Beer, M. S. *et al* Synthesis and serotonergic activity of *N*,*N*-dimethyl-2-[5-(1,2,4-triazol-1-ylmethyl)-1H-indol-3-yl]ethylamine and analogues: potent agonists for 5-HT1D receptors. *J. Med. Chem.* 1995, *38*(10), 1799–1810.
195. Glen, R. C., Martin, G. R., Hill, A. P., Hyde, R. M., Woollard, P. M., Salmon, J. A., Buckingham, J., Robertson, A. D. Computer-aided design and synthesis of 5-substituted tryptamines and their pharmacology at the 5-HT1D receptor: discovery of compounds with potential anti-migraine properties. *J. Med. Chem.* 1995, *38*(18), 3566–3580.
196. Habeeb, A. G., Praveen Rao, P. N., Knaus, E. E. Design and synthesis of celecoxib and rofecoxib analogues as selective cyclooxygenase-2 (COX-2) inhibitors: replacement of sulfonamide and methylsulfonyl pharmacophores by an azido bioisostere. *J. Med. Chem.* 2001, *44*(18), 3039–3042.
197. Larson, A. A., Lish, P. M. A new bio-isostere: alkylsulphonamidophenethanolamines. *Nature (London)* 1964, *203*, 1283–1285.
198. Roderick, S. L., Fournié-Zaluski, M. C., Roques, B. P., Matthews, B. W. Thiorphan and *retro*-thiorphan display equivalent interactions when bound to crystalline thermolysin. *Biochemistry* 1989, *28*, 1493–1497.
199. Büchi, J., Stünzi, E., Flury, M., Hirt, R., Labhart, P., Ragaz, L. Über lokalanästhetisch wirksame basische ester und amide verschiedener alkoxy-amino-benzoesäuren. *Helv. Chim. Acta.* 1951, *34*, 1002–1013.
200. Bös, M., Jenck, F., Martin, J. R., Moreau, J.-L., Sleight, A. J., Wichmann, J., Widmer, U. Novel agonists of 5HT2C receptors. Synthesis and biological evaluation of substituted 2-(indol-1-yl)-1-methylethylamines and 2-(indeno[1,2-*b*]pyrrol-1-yl)-1-methylethylamines. Improved therapeutics for obsessive compulsive disorder. *J. Med. Chem.* 1997, *40*, 2762–2769.
201. Wermuth, C. G. Selective optimization of side activities: another way for drug discovery. *J. Med. Chem.* 2004, *47*(6), 1303–1314.
202. Böhm, H.-J., Flohr, A., Stahl, M. Scaffold hopping. *Drug Disc. Today Technol.* 2004, *1*(3), 217–224.
203. Schneider, G., Schneider, P., Renner, S. Scaffold_Hopping: how far can you jump?. *QSAR Comb. Sci.* 2006, *25*(12), 1162–1171.
204. Low, C. M., Buck, I. M., Cooke, T., Cushnir, J. R., Kalindjian, S. B., Kotecha, A., Pether, M. J., Shankley, N. P., Vinter, J. G., Wright, L. Scaffold hopping with molecular field points: identification of a cholecystokinin-2 (CCK2) receptor pharmacophore and its use in the design of a prototypical series of pyrrole- and imidazole-based CCK2 antagonists. *J. Med. Chem.* 2005, *48*(22), 6790–6802.
205. Huang, D., Luthi, U., Kolb, P., Edler, K., Cecchini, M., Audetat, S., Barberis, A., Caflisch, A. Discovery of cell-permeable non-peptide inhibitors of beta-secretase by high-throughput docking and continuum electrostatics calculation. *J. Med. Chem.* 2005, *48*(16), 5108–5111.
206. Steindl, T. M., Crump, C. E., Hayden, F. G., Langer, T. Pharmacophore modeling, docking, and principal component analysis based clustering: combined computer-assisted approaches to identify new inhibitors of the human rhinovirus coat protein. *J. Med. Chem.* 2005, *48*(20), 6250–6260.
207. Wilhelm, M. The chemistry of polycyclic psycho-active drugs: serendipity or systematic investigation?. *Pharm. J.* 1975, *214*, 414–416.
208. Yoshimura, H., Kikuchi, K., Hibi, S., Tagami, K., Satoh, T., Yamauchi, T., Ishibahi, A., Tai, K., Hida, T., Tokuhara, N., Nagai, M. Discovery of novel and potent retinoic acid receptor alpha agonists: syntheses and evaluation of benzofuranyl-pyrrole and benzothiophenyl-pyrrole. *J. Med. Chem.* 2000, *43*, 2929–2937.

209. Chenoweth, M. B., McCarthy, L. P. On the mechanism of the pharmacophoric effect of halogenation. *Pharmacol. Rev.* 1963, *15*, 673–707.
210. Goldman, P. The carbon–florine bond in compounds of biological interest. *Science* 1969, *164*, 1123–1130.
211. Peters, R. A. *Biochemical Lesions and Lethal Synthesis*. Pergamon Press: Oxford, 1963.
212. Chapman, N. B., James, J. W., Graham, J. D. P., Lewis, G. P. Chemical reactivity and pharmacological activity among 2-haloethylamine derivatives with a naphtylmethyl group. *Chem. Ind. (London)* 1952, 805–807.
213. Vaughan, J. R. J., Anderson, G. W., Clapp, R. C., Clark, J. H., English, J. P., Howard, K. L., Marson, H. W., Sutherland, L. H., Denton, J. J. Antihistamine agents. IV. Halogenated *N*,*N*-dimethyl-*N′*-benzyl-*N*-(2-pyridyl)-ethylenediamines. *J. Org. Chem.* 1949, *14*, 228–234.
214. Friedman, H. L. Antihistamine, In *Influence of Isosteric Replacements upon Biological Activity*, Symposium on Chemical–Biological Correlation, Washington, DC, 1951; Sciences, N. A. o., Ed. National Research Council Publication: Washington, DC, 1951; p. 295.
215. Schatz, V. B. Isosterism and bioisosterism. In *Medicinal Chemistry* (Burger, A., Ed.). Interscience Publishers, Inc.: New York, 1963, pp. 72–88.
216. Fessenden, R. J., Fessenden, J. S. The biological properties of silicon compounds. In *Advances in Drug Research* (Harper, N. J., Simmonds, A. B., Eds), Vol. 4. Academic Press: London, 1967, pp. 95–132.
217. Tacke, R., Zilch, H. Sila-substitution-a useful strategy for drug design?. *Endeavour, New Series* 1986, *10*, 191–197.
218. Tacke, R., Zilch, H. Drug-design by sila-substitution and microbial transformations of organosilicon compounds: some recent results. *L'Actualité Chimique* 1986, 75–82.
219. Ricci, A., Seconi, G., Taddei, M. Bioroganosilicon chemistry: trends and perspectives. *Chimica Oggi-Chemistry Today* 1989, *7*, 15–21.
220. Showell, G. A., Mills, J. S. Chemistry challenges in lead optimization: silicon isosteres in drug discovery. *Drug Disc. Today* 2003, *8*(12), 551–556.
221. Anonymous. Zifrosilone. *Drugs Fut.* 1994, *19*, 854–855.
222. Hornsperger, J.-M., Collard, J.-N., Heydt, J. G., Giacobini, E., Funes, S., Dow, J., Schirlin, D. Trimethylsilylated trifluoromethyl ketones, a novel class of acetylcholinesterase inhibitors: biochemical and pharmacological profile of MDL 73,745. *Biochem. Soc. Trans.* 1994, *22*, 758–763.
223. Metcalf, R. L., Fukuto, T. R. Silicon-containing carbamate insecticides. *J. Econ. Entomol.* 1965, *58*, 1151.
224. Fessenden, R. J., Coon, M. D. Silicon-substituted medicinal agents. Silacarbamates related to meprobamate. *J. Med. Chem.* 1965, *8*, 604–608.
225. Tacke, R. Sila-Pharmaka, XIX. Sila-Pridinol und Pridinol: Darstellung und Eigenschaften sowie Strukturen im kristallinen und gelösten Zustand. *Chem. Ber.* 1980, *113*, 1962–1980.
226. Moberg, W. K. Fungicidal 1,2,3-triazole derivatives. *US Patent 4,510,136 (Apr. 9, 1985 to E.I. Du Pont de Nemours & Co) 1985*.
227. Damour, D., Barreau, M., Dutruc-Rosset, G., Doble, A., Piot, O., Mignani, S. 1,1-Diphenyl-3-dialkylamino-1-silacyclopentane derivatives: a new class of potent and selective 5-HT_{2A} antagonists. *Bioorg. Med. Chem. Lett.* 1994, *4*, 415–420.
228. Gotteland, J.-P., Brunel, I., Gendre, F., Désiré, J., Delhon, A., Junquéro, D., Oms, P., Halazy, S. (Aryloxy)methylsilane derivatives as new cholesterol biosynthesis inhibitors: synthesis and hypocholesterolemic activity of a new class of squalene epoxidase inhibitors. *J. Med. Chem.* 1995, *38*, 3207–3216.
229. Gotteland, J.-P., Delhon, A., Junquéro, D., Oms, P., Halazy, S. Design and synthesis of new hypocholesteremic organosilanes with antioxidant properties. *Bioorg. Med. Chem. Lett.* 1996, *6*, 533–538.
230. Heinonen, P., Sipilä, H., Neuvonen, K., Lönnberg, H., Cockroft, V. B., Wurster, S., Virtanen, R., Savola, M. K. T., Salonen, J. S., Savola, J. M. Synthesis and pharmacological properties of 4(5)-(2-ethyl-2,3-dihydro-2-silainden-2-yl)imidazole, a silicon analogue of atipamezole. *Eur. J. Med. Chem.* 1996, *31*, 725–729.
231. Bom, D., Curran, D. P., Kruszewski, S., Zimmer, S. G., Strode, J. T., Kohlhagen, G., Du, W., Chavan, A. J., Fraley, K. A., Bingcang, A. L., Latus, L. J., Pommier, Y., Burke, T. G. The novel silatecan 7-*tert*-butyldimethylsilyl-10-hydroxycamptothecin displays high lipophilicity, improved human blood stability, and potent anticancer activity. *J. Med. Chem.* 2000, *43*, 3970–3980.
232. Mutahi, M., Nittoli, T., Guo, L., Sieburth, S. M. Silicon-based metalloprotease inhibitors: synthesis and evaluation of silanol and silanediol peptide analogues as inhibitors of angiotensin-converting enzyme. *J. Am. Chem. Soc.* 2002, *124*(25), 7363–7375.
233. Chen, C. A., Sieburth, S. M., Glekas, A., Hewitt, G. W., Trainor, G. L., Erickson-Viitanen, S., Garber, S. S., Cordova, B., Jeffry, S., Klabe, R. M. Drug design with a new transition state analog of the hydrated carbonyl: silicon-based inhibitors of the HIV protease. *Chem. Biol.* 2001, *8*(12), 1161–1166.
234. Showell, G. A., Barnes, M. J., Daiss, J. O., Mills, J. S., Montana, J. G., Tacke, R., Warneck, J. B. (R)-sila-venlafaxine: a selective noradrenaline reuptake inhibitor for the treatment of emesis. *Bioorg. Med. Chem. Lett.* 2006, *16*(9), 2555–2558.
235. Barnes, M. J., Conroy, R., Miller, D. J., Mills, J. S., Montana, J. G., Pooni, P. K., Showell, G. A., Walsh, L. M., Warneck, J. B. Trimethylsilylpyrazoles as novel inhibitors of p38 MAP kinase: a new use of silicon bioisosteres in medicinal chemistry. *Bioorg. Med. Chem. Lett.* 2007, *17*(2), 354–357.
236. Sommer, L. H., Bennet, O. F., Campbell, P. G., Weyenberg, D. R. Stereochemistry of hydride ion displacement from silicon. Enhanced rates at bridgehead and 4-ring silicon atoms. *J. Amer. Chem. Soc.* 1957, *79*, 3295–3296.
237. Alam, F., Soloway, A. H., Bapat, B. V., Barth, R. F., Adams, D. M. Boron compounds for neutron capture therapy. *Basic Life Sci.* 1989, *50*, 107–111.
238. Gabel, D. Tumor-seeking for boron neutron capture therapy: synthesis and biodistribution. *Basic Life Sci.* 1989, *50*, 233–241.
239. Kahl, S. B., Joel, D. D., Finkel, G. C., Micca, P. L., Nawrocky, M. M., Coderre, J. A. A carboranyl porphyrin for boron neutron capture therapy of brain tumors. *Basic Life Sci.* 1989, *50*, 193–203.
240. Caujolle, F., Chan, Pham.-Huu. Structure Chimique et Activité Spasmolytique des Organoboriques. *Arch. Int. Pharmacodyn. Ther.* 1968, *172*, 467–474.
241. Mubarak, S. I. M., Stanford, J. B., Sugden, J. K. Some aspects of the antimicrobial and chemical properties of phenyl boronate esters of chloramphenicol. *Drug Dev. Ind. Pharm.* 1984, *10*, 1131–1160.
242. Hutter, R., Keller-Schierlein, W., Knusel, F., Prelog, V., Rodgers, G. C., Jr, Suter, P., Vogel, G., Voser, W., Zahner, H. The metabolic products of microorganisms. Boromycin. *Helv. Chim. Acta.* 1967, *50*(6), 1533–1539.
243. Dünitz, J. D., Hawley, D. M., Miklos, D., White, D. N. J., Berlin, Y., Marusik, R., Prelog, V. Structure of boromycin. *Helv. Chim. Acta.* 1971, *54*, 1709–1713.
244. Okami, Y., Okazaki, T., Kitahara, T., Umezawa, H. Studies on marine microorganisms. V. A new antibiotic, aplasmomycin, produced by a streptomycete isolated from shallow sea mud. *J. Antibiot. (Tokyo)* 1976, *29*(10), 1019–1025.
245. Kinder, D. H., Katzenellenbogen, J. A. Acylamino boronic acids and difluoroborane analogues of amino acids: potent inhibitors of chymotrypsine and elastase. *J. Med. Chem.* 1985, *28*, 1917–1925.
246. Adams, J., Behnke, M., Chen, S., Cruickshank, A. A., Dick, L. R., Grenier, L., Klunder, J. M., Ma, Y. T., Plamondon, L., Stein, R. L. Potent selective inhibitors of the proteasome: dipeptidyl boronic acids. *Bioorg. Med. Chem. Lett.* 1988, *8*(4), 333–338.

247. Dembitsky, V. M., Srebnik, M. Synthesis and biological activity of -aminoboronic acids, amine-carboxyboranes and their derivatives. *Tetrahedron* 2003, *59*(5), 579–593.
248. Surolia, N., RamachandraRao, S. P., Surolia, A. Paradigm shifts in malaria parasite biochemistry and anti-malarial chemotherapy. *Bioessays* 2002, *24*(2), 192–196.
249. Jabbour, A., Steinberg, D., Dembitsky, V. M., Moussaieff, A., Zaks, B., Srebnik, M. Synthesis and evaluation of oxazaborolidines for antibacterial activity against streptococcus mutans. *J. Med. Chem.* 2004, *47*(10), 2409–2410.
250. Kumar, S. K., Hager, E., Pettit, C., Gurulingappa, H., Davidson, N. E., Khan, S. R. Design, synthesis, and evaluation of novel boronic-chalcone derivatives as antitumor agents. *J. Med. Chem* 2003, *46*(14), 2813–2815.
251. Landauer, W. On the chemical production of developmental abnormalities and of phenocopies in chicken embryos. *J. Cell. Comp. Physiol.* 1954, *43*(1), 261–305.
252. Landauer, W., Clark, E. M. On the role of riboflavin in the teratogenic activity of boric acid. *J. Expt. Zool.* 1964, *156*, 307–312.
253. Browning, E. *Toxicity of Industrial Metals*, 2nd ed. Appleton-Century-Crofts: New York, 1969. p. 90–97.
254. Klayman, D. L., Günther, W. H. H. *Organic Selenium Compounds: Their Chemistry and Biology*. Wiley-Interscience: New York, 1973.
255. Kang, S.-I., Spears, C. P. Linear free energy relationships and cytotoxicities of para-substituted 2-haloethyl aryl selenides and bis (-chloroethyl) selenides. *J. Med. Chem.* 1987, *30*, 597–602.
256. Fischer, H., Terlinden, R., Löhr, J. P., Römer, A. A novel biologically active selenoorganic compound. VIII. Biotransformation of ebselen. *Xenobiotica* 1988, *18*, 1347–1359.
257. Parnham, M. J., Graf, E. Seleno-organic compounds and the therapy of hydroperoxide-linked pathological conditions. *Biochem. Pharmacol.* 1987, *36*, 3095–3102.

Chapter 16

Ring Transformations

Christophe Morice and Camille G. Wermuth

I. INTRODUCTION
II. ANALOGICAL APPROACHES
A. Analogy by ring opening: open-chain analogs
B. Analogy by ring closure
C. Other analogies
III. DISJUNCTIVE APPROACHES
A. Cocaïne-derived local anesthetics
B. Morphinic analgesics
C. Dopamine autoreceptor agonists
D. CCK antagonists
IV. CONJUNCTIVE APPROACHES
A. Dopaminergic antagonists
B. Glutamate NMDA and AMPA receptor antagonists
C. Norfloxacin analogs
D. Melatonin analogs
V. CONCLUSION
REFERENCES

"Rien ne se perd, rien ne se crée, tout se transforme."

Antoine Lavoisier (1743–1794)

I. INTRODUCTION

When active molecules contain cyclic systems, these can be opened, expanded, contracted and modified in many other ways, or even abolished. Conversely, non-cyclic molecules can be cyclized, attached to, or included in, ring systems.

In the daily practice of medicinal chemistry three kinds of approaches are currently used. The first approach does not affect the global complexity of the cyclic system and yields generally close analogs (or "me-too" compounds) of the original active principle. For this we propose the term *analogical approach*. It consists of ring-chain transformations, ring contractions or expansions and various other ring transformations. The second strategy, called the *disjunctive approach*[1] aims at the progressive simplification of the original active principle (which is often a natural compound). The objective is to extract information about the minimal structure that is required for activity. Finally the *conjunctive approach*[1] is based on the creation or addition of supplementary rings. The objective is to constrain an originally flexible compound and to impose precise conformations and configurations. The preparation of such molecules is of prime importance in the exploration of ligand–receptor interaction and for molecular modeling studies.

II. ANALOGICAL APPROACHES

A. Analogy by ring opening: open-chain analogs

Open ring analogs of cyclic active principles (open drugs, open-chain analogs) can be designed and synthesized. The usefulness of such compounds is rather questionable, and

Copyright © 2008, Elsevier Ltd
All rights reserved.

FIGURE 16.1 Cycloguanyl is the active metabolite of proguanyl.[3]

it appears that most of them were prepared for me-too purpose. Actually two possibilities can be foreseen:

1. The open analog is again cyclized after oxidation or dehydration by a metabolic enzyme. We deal here with potential rings, which represent nothing more than metabolic precursors of the active species.
2. The open analog does not cyclize *in vivo* but can present some conformational analogy with the ring-containing active principle. These kinds of analogs are known as pseudocycles.

1. *Potential rings*: in vivo *return to the cyclic derivative*

Compounds generating the active form after *in vivo* cyclization are in fact prodrugs and will be discussed in a more detailed manner in Chapter 38.

a. Proguanil

A historical example is the antimalarial drug proguanyl[2] (Figure 16.1). It was observed that this compound is inactive *in in vitro* cultures of *Plasmodium gallinaceum* but that the serum of animals treated with proguanyl is active in these cultures.

It was concluded that the actual active principle was a metabolite[4] which was subsequently identified as cycloguanyl.[3] In tropical medicine proguanyl is preferred to cycloguanyl, the latter compound being too rapidly eliminated by the kidney.

b. Potassium canrenoate

This is a water-soluble prodrug and can be administered parenterally (Figure 16.2). It has no intrinsic activity, but it can exert its diuretic activity (as an aldosterone antagonist) because of its interconversion with canrenone. Canrenone is itself the major metabolite of spironolactone.[5]

2. *Irreversible open compounds*: *pseudocycles*

The open analog is assumed to present a similar conformation to that of the cyclic one. Before adopting the term

FIGURE 16.2 *In vivo* the inactive potassium canrenoate cyclizes to canrenone.[5]

FIGURE 16.3 Open analogs of estradiol.

pseudocycles it must be ascertained by nuclear magnetic resonance (NMR) or X-ray crystallography that they really mimic the ring-closed analog.

a. Diethylstilbestrol

It is currently accepted that the theory of pseudocycles was elaborated to account for the estrogenic activity of compounds such as bisdehydrodoisynolic acid, allenestrol and diethylstilbestrol (Figure 16.3).

The similarity with the natural hormone estradiol is striking. Nevertheless it is highly probable that for receptor binding, the general shape of the molecules, and the distances separating the functional groups are more important than their degree of cyclization.[6]

b. Non-phenolic estradiol analogs

Aromatic open ring analogs are much less described in the literature. The difficulty to perform a pseudocycle that mimics the aromaticity and the planarity of the original ring

FIGURE 16.4 Aldoximes: as bioisosteres of the phenol ring.

FIGURE 16.5 Open analogs of clonidine (the hypotensive activity is expressed as variation of the arterial blood pressure 30 min after i.v. injection in pentobarbital anesthetized rats).[9]

limits its use. For example, salicylaldoxime **1**[7] and anthranylaldoxime **2**[8] generate a pseudocycle that reproduces the naphthol structure **3**[7] (Figure 16.4). The hydrogen bond between the oxime group and the OH (in **1**) or the NH (in **2**) groups are essential to create a stable pseudo-ring that insures a high-affinity interaction with the estrogen receptors via the hydroxyl group. These pseudocycles represent bioisosteres of the phenolic A group of estradiol.

c. Clonidine

Open analogs of the centrally acting hypotensive agent clonidine (Figure 16.5) have a similar activity profile but with a 30 to 100-fold loss in potency.[9] Surprisingly seco-clonidine that is the closest analog of clonidine, was found to be less active than the corresponding monomethyl derivative.

d. Cromakalim

A more recent example is the open-chain analog of cromakalim that was prepared as a more flexible pyrrolidone replacement (Figure 16.6). It retains about a third of the potency of cromakalim.[10]

FIGURE 16.6 (3*S*,4*R*) cromakalim and its open-chain analog.[10]

B. Analogy by ring closure

Cyclizing open structures or creating an additional ring system in a given structure represents one of the useful methods in the search for biologically active conformations. The end result is a more constraint molecule, with an imposed conformation. This strategy is related to the *conjunctive approach* developed further on in this chapter.

The inconvenience is that additional isomeric centers may be introduced and that the selected cyclization mode might not lead to the active conformation adopted by the open-chain drug. A particularly convincing example is given by the ring-closed analog of the thrombin inhibitor NAPAP which is 100 times more potent than the corresponding open-chain drug[11] (Figure 16.7).

1. *Inhibition of gastric* H^+/K^+ ATP*ase by substituted Imidazo*[1,2-*a*]*pyridines*

Apparently the substitution in the *para* position of the pyridine ring is detrimental (R = H → R = Me, Figure 16.8). However, the ring closure achieving a conformational restriction yields a highly potent compound.[13]

FIGURE 16.7 The ring-closed analog of the thrombin inhibitor NAPAP is 100 times more potent than the corresponding open-chain drug.[12]

FIGURE 16.8 Ring closure achieves a conformational restriction and yields a highly potent compound.

2. *Mevinolin and compactin*

An example of reversible ring closure is found with mevinolin and compactin that are both potent inhibitors of hydroxy-methyl-glutaryl-coenzyme A reductase (HMG-CoA reductase), the rate-determining enzyme in the *de novo* biosynthesis of cholesterol.

In vivo these ring-closed derivatives (Figure 16.9) are hydrolyzed to the open-chain 3,5-dihydroxyvaleric acid, form that mimics the structure of the proposed intermediate in the reduction of HMG-CoA by HMG-CoA reductase.[14]

3. *Arylpropionic analgesic and anti-inflammatory drugs*

The potent analgesic benzoyl-indane carboxylic acid TAI-901[15] is the cyclized analog of the well-known anti-inflammatory analgesic agent ketoprofene (Figure 16.10). The corresponding heterocyclic analogs were also prepared.[16] The compounds show potent analgesic activities with low gastric irritation.

4. *The sulpiride side-chain*

Among the numerous benzamide drugs developed by the Delagrange scientists some have diethylaminoethyl side-chains (e.g. tiapride) whereas others, such as sulpiride, have *N*-ethyl-pyrrolidinyl-methyl side-chains that can be considered as the corresponding ring-closed analogs (Figure 16.11).

Both compounds are dopaminergic antagonists with neurotropic and antiemetic activity. Note that the ring closure creates an asymmetric center; the commercial form is the racemate, the slightly more active isomer being the *S*-(−)-sulpiride.[17] An additional constraining factor results from the establishment of a hydrogen bond between the amidic N-H hydrogen and the methoxy oxygen.[18] A conformationally restricted remoxipride analog in which the intramolecular hydrogen bond is replaced by a covalent bond (Figure 16.11) is equipotent in D2 receptor preparations.[19]

5. *Cyclized dopamine: the* ADTN's

The 2-amino-5,6-dihydroxy- and the 2-amino-6,7-dihydroxy-1,2,3,4-tetrahydro-naphtalenes are cyclized analogs of dopamine, corresponding to the α- and the β-rotamer, respectively (Figure 16.12).

As the cyclization generates a chiral center, four different ADTN's are possible, showing differential affinities for the dopamine receptors.[20] An extensive study of the aminotetralins and analogs containing additional rings (octahydrobenzo[g]quinolines) and compounds resulting from ring enlargements was published by Seiler *et al*.[21]

6. GABA*ergic agonists*

The transition from γ-aminobutyric-acid (GABA) to *trans*-4-amino-crotonic acid, followed by cyclization into isoguvacine, and finally into THIP (Figure 16.13), achieves simultaneously the rigidification of the flexible GABA molecule and the production of THIP, a metabolically stable and still potent GABA agonist.[22]

7. *Ring-closed analog of nicotine*

Alzheimer's disease has received the most attention as a therapeutic target for nicotinic drugs, as nicotinic receptor binding was found to be significantly reduced in distinct brain regions of Alzheimer's patients.[23] Because the

FIGURE 16.9 The chemically stable lactones mevinolin and compactin represent the ring-closed forms of the *in vivo* active parent substituted 3,5-dihydroxyvaleric acid.

FIGURE 16.10 Ring-closed ketoprofene analogs.

FIGURE 16.11 The typical sulpiride side-chain results formally from a ring closure of the diethylaminoethyl side-chain of earlier prepared derivatives such as tiapride. Observe the intramolecular hydrogen bond that creates a pseudocycle[15] and which can be mimicked by a covalent and constraint analog.[16]

FIGURE 16.12 Ring-closed analogs of the two rotamers of dopamine.

FIGURE 16.13 GABAergic agonists.[22]

identification of various adverse physiological effects[23,24] of the alkaloid (*S*)-nicotine (**4**, Figure 16.14), synthetic modifications of its structure have been performed in order to improve potency and selectivity while reducing the toxicity.

Several groups[25,26] pursued the concept of conformational restriction in nicotine with the objective of forcing both pharmacophore nitrogen atoms into well-defined angles ("up" and "down" conformations), which would help in elucidating the "active" conformation of nicotine on one hand and obtaining a tool to study receptor subtype specificity on the other hand. These endeavors led to, among others, the class of hexahydropyrroloisoquinolines (**5**, Figure 16.14).[25] The correct stereochemistry began to play an increasingly decisive role: while the relatively low enantioselectivity of nicotine (the affinity of (*S*)-nicotine is 10–100 times higher than that of (*R*)-nicotine)[27] has been an intriguing phenomenon for many years, the premise that conformational restraint of nicotine should enhance enantioselectivity has been well established.[28]

The (−)-enantiomer of **5** shows low relative affinity (K_i = 605 nM) for the nAChR [3*H*]nicotine binding site, whereas the (+)-enantiomer fails to displace the radioligand even at 10 μM concentrations.[26]

Spiro-annulated compounds **6a**,**b** (Figure 16.14) exhibited much higher binding affinities.[29] *nor*-Derivative (+)-**6a** bound at K_i = 53.1 nM with a 10-fold higher affinity than its enantiomer (K_i = 533 nM). Introduction of the *N*-methyl substituent resulted in a significant improvement,

FIGURE 16.14 Natural (*S*)-(–)-nicotine and its bridged [25] and spiro[29] analogs.

probably due to the gain of fortified receptor interaction: (+)-**6b** (K_i = 4.79 nM) appeared to be the most interesting ligand of this set as it not only bounded in the low nanomolar range but also exhibited a 30-fold higher affinity than its enantiomer (K_i = 148 nM).[29]

8. *Cyclic analogs of β-blockers*

Conventional β-blockers possess a number of pharmacological properties, for example, β-blocking, quinidine-like, local anesthetic and hypotensive. With the hope of achieving some specificity, Basil *et al.*[30] considered the possibility of synthesizing ring-closed analogs (Figure 16.15). One of the compounds prepared, 3,4-dihydro-3-hydroxy-6-methyl-1,5-benzoxazocine was a potent β-blocker. This activity is unlikely to be due to hydrolysis to the open-chain derivative since the corresponding primary amine, formed by hydrolysis of the benzoxazocine ring has less than 0.25 the activity of the latter. On the other hand, it is difficult to reconcile the benzoxazocine configuration with the structural requirements associated with the occupation of β-receptors.

Later studies by Evans *et al.* also envisaged the synthesis of cyclized analogs of the phenylpropanolamine type of β-blockers.[31,32] The authors hoped that by restricting the conformation (by cyclizing the carbon atom bearing the terminal amino group to the aromatic ring, see Figure 16.15), β-blocking activity would be lost but antihypertensive activity might be retained. This turned out to be true in animal tests and in double-blind clinical studies, so the potassium channel activator cromakalim was developed.[32]

9. *Cyclized diphenhydramine*

Nefopam (Figure 16.16) is the representative of a new class of centrally acting skeletal muscle relaxants, also possessing a benzoxazocine structure.[33] Formally nefopam is a cyclized analog of orphenadrine and diphenhydramine.

In contrast to the parent molecules, nefopam has no antihistaminic activity, keeping only the muscle relaxant effects. Clinically nefopam is used as muscle relaxant, but also, and this was originally not anticipated, it is an antidepressant and an analgesic. These clinical indications

FIGURE 16.15 Cyclized analogs of β-blocking phenylpropanolamines.[31,32]

FIGURE 16.16 Nefopam is a cyclized analog of orphenadrine and diphenhydramine.[33]

can be explained by the interference of nefopam with the serotonergic transmission. More precisely, studies of the separated stereoisomers of nefopam explain its serotonin uptake properties and suggest that descending serotonergic pathways are involved in its antinociceptive activity.[34]

10. *Rimonabant analogs*

Several groups have recently reported constrained *me-too* analogs of rimonabant (compound **7**, Figure 16.17), the most advanced antagonist of cannabinoid type 1 receptor (CB-1) that was recently approved in the European Union for the treatment of obesity.[35] Compounds that incorporate conformational constraints into the tetrasubstituted pyrazole structure have the potential for increased binding activity and subtype selectivity (toward CB-2 receptor), resulting from preferential bioactive conformation (Figure 16.17). Tricycles analogs have been described including the central pyrazole ring and the 5-aryl group (compounds **8a–c**).[36] The orientation of the 5-aryl group is imposed by the size of the central cycle. This leads to dramatic differences of affinities for the mouse CB-1 receptor, from micromolar to sub-picomolar values. Surprisingly, the selectivity for mCB-1 versus mCB-2 varies from 60,000:1 to 1:6,000.

Compounds having a constraint between the pyrazole core C-4 position and the carboxamide moiety have also been reported (Figure 16.18, compound **9**). These pyrazolopirimidinone derivatives are slightly less active in the human CB-1 binding assay than rimonabant (K_i = 12 versus 2.1 nM)).[37] This could be explained by the splitting of carboxamide group in a non-favorable orientation.

To constrain this carboxamide hydrogen donor/acceptor group to exhibit the preferential bioactive conformation, a novel series of analogs have been prepared.[38] In this case, an ethylene bridge is incorporated between the pyrazole N-2 position and the carboxamide nitrogen (Figure 16.18, compound **10**). The potency of compound **10** is similar to that observed for compound **7** in the same *in vitro* human CB-1 assay (K_i = 2.2 versus 1.1 nM).

The slight structural variation of constrained analogs compared to rimonabant is shown in Figure 16.19. The superposition of low-energy conformations[39] of **7** with the tricyclic analogs (**8c**, **9** and **10**) are in accordance with the results obtained in affinity binding assays, regarding the position of the three substitutions around the central pyrazole core.

11. *Ring variations around phenylbutazone*

The anti-inflammatory drug phenylbutazone (Compound **11**, Figure 16.20) led to many me-too copies, such as the ring-opened analog bumadizon **12** (Ca^{++} salt = Eumotol®[40]) or the ring-closed analog apazone **13** (Prolixan®).[41] The cinnoline derivative **14**[42] results again from a ring closure and served as model for the design of its open counterpart, the styrylbutazone **15**.[43] The quinolinyl-3,5-dioxopyrazolidine **16**[44] represents another interesting ring variation with a 7-chloro-quinoline moiety in the butazone portion.

C. Other analogies

Applied to ring systems, the following molecular modifications seem to be conducted mainly with the objective of bypass patent protections and to allow the synthesis of me-too products.

FIGURE 16.17 Cannabinoid mouse receptor (mCB-1 and mCB-2) affinity and selectivity for rimonabant **7** and compounds **8a–c**.[36]

1. *Ring enlargement and ring contraction*

Ring enlargement and ring contraction can be considered as homologous variations in the cyclic series (mentioned in Chapter 14).

a. Barbiturics and opioids

In Figure 16.21, two additional examples taken from the barbituric and from the opiate series, respectively, are shown. In the case of the change of the barbiturics to hydantoins the contraction is accompanied by the loss of a carbonyl group (Figure 16.22). Nevertheless potent antiepileptics are found in both series.

b. Inogatran and melagatran

The classical motif of thrombin inhibitors is the D-Phe-Pro-Arg sequence mimicking thrombin's natural substrate, fibrinogen. In development candidates such as inogatran[45] and melagatran[46] (Figure 16.23), the proline unit was replaced by its ring-expanded and its ring-contracted equivalent, respectively.

c. AMPA receptor antagonists

Several 2,3-benzodiazepin-4-ones such as compound GYKI 53 655 (Figure 16.24) are AMPA receptor antagonists and possess noteworthy anticonvulsant and neuroprotective properties. The corresponding ring-contracted analogs, 6,7-methylene-dioxydihydrophtalazines (compound SYM 2207) and 6,7-methylene-dioxyphtalazin-1(2*H*)-ones, possess a similar activity profile (compound **17**).[47,48]

d. Oxotremorine

Oxotremorine and its ring-opened analog Oxo-2 (Figure 16.25) are partial muscarinic agonists producing large guanine nucleotide shifts in the heart (32 and 23, respectively), suggesting strong M2 agonist-like effects.[49]

The corresponding piperidinic analog Oxo-Pip[50] having a predicted antagonist [^{3}H]QNB/[^{3}H] CD ratio of 2.2, produced only a weak shift (5.0) in the concentration-response curve with the addition of the stable guanine nucleotide analog.[49] Thus, the change from a pyrrolidine to a piperidine ring is able to change a partial agonist into an antagonist.

FIGURE 16.18 Cannabinoid human receptor (hCB-1) affinity for rimonabant **7**, compounds **9** and **10**.[37,38]

(A) (B) (C)

FIGURE 16.19 Overlay of the low-energy conformation of rimonabant 7, successfully with compound 8c (A), compound 10 (B) and compound 9 (C).[39]

The replacement of the core cyclopentane ring of the prostaglandin FP agonist cloprostenol with a cyclohexane ring (compound **18**, Figure 16.26) yielded a clearly less active agent (EC_{50} = 319 nM instead of 1 nM).[51]

2. *Reorganization of ring systems*

The four molecular variations described below represent some more "exotic" approaches to the design and manipulation of the original ring systems. They may bring useful alternatives allowing escape from overcrowded avenues of research.

a. Transforming simple rings into spiro derivatives or into bi- or tricyclic systems

A first example is in the guanethidine analogs.[52] As the original guanethidine patents covered ring sizes varying from five-membered to ten-membered rings a possible way to get

FIGURE 16.20 Successive ring openings and closures in phenylbutazone-derived anti-inflammatory drugs.

FIGURE 16.21 Six-membered rings exchanged for seven-membered rings.

FIGURE 16.22 (a) Barbiturates (left) and hydantoins (right).

round them was the design of isolipophilic *spiro* systems (Dausse compounds **a** and **b**,[53,54] Figure 16.27). Another possibility, originating from Takeda scientists, implies the use of an azetidine surrogate for the ethylene-diamine chain.[55] Finally polycyclic systems can replace the octahydroazocine ring, as illustrated by the bicyclic compounds **c** and **d** from Dausse[56] or by the tricyclic compound from Lumière Laboratories.[57,58] Many other imaginative solutions were proposed; they are well reviewed by Mull and Maxwell.[52] More recently similar variations were applied to

FIGURE 16.23 Structures of the thrombin inhibitors inogatran and melagatran.[45,46]

FIGURE 16.24 The heptacyclic AMPA antagonist GYKI 53 655 and ring-contracted analogs.[47,48]

FIGURE 16.25 Oxotremorine yields ring-opened and ring-extended analogs.[49]

FIGURE 16.26 Cloprostenol and ring-extended analog.[51]

the design of an impressive number of analogs of the anticonvulsant drug gabapentin (Figure 16.28).[59,60]

b. Splitting benzo compounds ("Benzo Cracking")

Dissociation of a fused ring system (Figure 16.29), particularly by splitting a benzo compound, can sometimes improve its solubility and nevertheless only alter slightly its pharmacokinetic profile and its long-term toxicity.

c. Restructuring ring systems

Among the above-mentioned molecular variations on ring systems, some can be used *simultaneously*. Thus, the splitting of the benzimidazole heterocycle in the anthelmintic thiabendazole, and the concomitant association of the two five-membered rings, yield tetramizole (Figure 16.30). One of the two enantiomers of tetramizole, the L-(–)-form, or levamizole, is also a potent anthelmintic.

The change from the D_1-selective dopaminergic agonist **DPTI** (3-(3,4-dihydroxyphenyl)-1,2,3,4-tetrahydroisoquinoline) to the equally D_1-selective compound **SKF 38 393** is a combination of benzo cracking, a new benzo fusion and ring enlargement (Figure 16.30). As a result, the compounds still resemble each other and are both recognized by the dopamine D_1-receptor.[21]

An interesting example of a restructured ring system was designed by Meyer *et al.*[61] for the design of 5-methoxy-hexahydro-1*H*-benz[f]isoindole as a surrogate of the very frequently used *ortho*-methoxy-*N*-arylpiperazine. It led

FIGURE 16.27 Alternative possibilities in the design of guanethidine analogs.

FIGURE 16.28 Alternative possibilities in the design of gabapentin analogs.

both to an enhancement of affinity and improved selectivity for α_{1A} receptor antagonism (Figure 16.31).

d. Ring dissociation

The natural compound khellin generated two families of cardioactive drugs: on one side the benzopyrones, illustrated by the 3-methyl-chromone[62] and chromonar (carbochromen), on the other side the benzofurans, illustrated by amiodarone (Figure 16.32). Both families possess anti-arrhythmic and anti-anginal properties.

III. DISJUNCTIVE APPROACHES

Starting from a polycyclic structure (which is often from natural origin), the chemist proceeds to progressive pruning of the molecule ("molecular strip tease"). Sometimes very simple reasoning guides the medicinal chemist and the final

FIGURE 16.29 Splitting of fused rings often yields drugs with similar activity, sometimes with improved solubility and/or less toxicity.

FIGURE 16.30 Restructured ring systems.

compound has only a remote resemblance to the model compound. Such an exercise led to the transformation of the natural compound asperlicin to a totally synthetic simplified benzodiazepine derivative (see below Figure 16.36).[63]

A. Cocaïne-derived local anesthetics

Figure 16.33 illustrates how simplified synthetic copies of cocaine were designed. The change from cocaine to procaine retains the local anesthetic effects without keeping the narcotic properties.

B. Morphinic analgesics

Probably more than a thousand of simplified analogs of the alkaloid *morphine* were investigated.[64] Many of them were just inactive, but it was soon recognized that the phenylpiperidine unit was crucial for the central analgesic properties (Figure 16.34). In contrast to what was observed for cocaine, no clear discrimination between the analgesic and the narcotic properties could be achieved.

C. Dopamine autoreceptor agonists

The discovery of 3-(3-hydroxyphenyl)-*N*-*n*-propyl piperidine ((±)-3-PPP), a centrally acting dopamine receptor agonist with selectivity for dopaminergic autoreceptors,[65] offers a potential alternative to neuroleptics in the treatment of schizophrenia. The structure of (±)-3-PPP (Figure 16.35) can be considered as a product resulting from a disjunctive approach applied to pergolide.[66,67] Surprisingly, an increase in the pergolide-like character of 3-PPP, through incorporation of a methylmercaptomethyl group (compound **19**), did not improve the potency.[68]

D. CCK antagonists

After the discovery of the potent CCK antagonistic activity of the natural compound asperlicin,[69] the scientists from the Merck group prepared first some simple semi-synthetic derivatives.[70] Then, recognizing in asperlicin the elements of a benzodiazepinone and a tetrahydroindole, they followed the hunch that these elements alone may confer some CCK antagonistic activity.[63] This reasoning proved to be valid (Figure 16.36).

FIGURE 16.31 Hexahydr obenz[f]isoindole as a surrogate of the very frequently used *ortho*-methoxy-*N*-arylpiperazine.

FIGURE 16.32 Cardioactive drugs obtained by dissociation of the khellin molecule into benzopyrones and benzofurans.

FIGURE 16.33 Progressive simplification of the cocaine molecule.

IV. CONJUNCTIVE APPROACHES

As already mentioned at the beginning of this chapter, the purpose of the conjunctive method lies in the design of compounds structurally more complex than the lead compound. In practice this is generally achieved in creating and/or adding supplementary ring systems to constrain the molecule and to impose specific conformations.

A. Dopaminergic antagonists

Starting from the flexible haloperidol molecule, Humber *et al.*[71] designed the rigid (+)-butaclamol that contains three clearly defined stereocenters (Figure 16.37). The same stereochemical requirements as in (+)-butaclamol are found in compound Ro 14-8625 prepared by Imhof *et al.*[72]

FIGURE 16.34 Progressive simplification of the morphine molecule.

FIGURE 16.35 3-PPP is a result of the disjunctive approach applied to pergolide.[65,68]

FIGURE 16.36 Productive disjunction of the asperlicin molecule.[63]

FIGURE 16.37 The conjunctive method applied to the design of haloperidol-derived dopaminergic antagonists.

FIGURE 16.38 The conjunctive method applied to the design of NMDA and AMPA receptor antagonists.

FIGURE 16.39 Norfloxacin and its tetracyclic analog.[78]

B. Glutamate NMDA and AMPA receptor antagonists

The progressive change of glutamic acid to D-AP5,[73] to the piperidine analog CGS 19755[74] and finally to the tetrahydroisoquinoline PD 134705,[75] led to NMDA receptor antagonists. Similarly rigidification into the perhydroquinolines **20**, **21** and **22**[76] (Figure 16.38) illustrates another application of the conjunctive approach that led to potent AMPA antagonists. In addition to the elements enhancing structural rigidity, these latter compounds contain three new chiral centers.

FIGURE 16.40 Cilansetron, an annelated analog of ondansetron.[79]

C. Norfloxacin analogs

Since the development in 1980 of *norfloxacin*[77] as a useful antibacterial agent, a large number of analogs have been synthesized. Among them, the conjunctive approach led to highly potent tetracyclic analogs (Figure 16.39).[78]

In a comparable way, a number of annelated analogs of the 5-HT_3 antagonist ondansetron were investigated (Figure 16.40). Among them, cilansetron ($n = 1$) was found to be about 10 times more potent without loss in selectivity.[79]

FIGURE 16.41 Ring-fused melatonin analogs.[80]

FIGURE 16.42 ACE inhibitors derived from captopril.

D. Melatonin analogs

Tetracyclic analogs in which the melatonin indole ring is fused to a dihydroindole, a tetrahydroisoquinoline and a benzazepine system, respectively (compounds **23**, **24** and **25**; Figure 16.41) represent ligands for the melatonin MT2 receptor as potent as melatonin but selective with regard to the MT1 receptor. Interestingly, the passage from the six-membered (compound **24**) to the seven-membered fused ring system (compound **25**) produced a switch from agonist to antagonist activity.[80]

The well-known angiotensin converting enzyme (ACE) inhibitor captopril has a relatively simple structure and therefore represents an excellent starting material for the conjunctive approaches (Figure 16.42).

Modeling studies based on a template structure constructed from the superposition of the energy-minimized benzo-fused ACE inhibitors shown in Figure 16.42 (compounds **27** and **29**) suggested the synthesis of the 13-membered heterocyclic lactam analog (compound **28**).[81]

V. CONCLUSION

Molecular variations involving the study of homologous series or the application of the vinylogy concept induce relatively minor changes of the pharmacological profile and rather result in optimizing the potency. Modifying ring systems – ring-chain transformation, ring contractions and expansions, reorganization of cyclic systems – represents a highly productive approach in the design of new drug analogs and in the exploration of the drug–receptor interactions.

REFERENCES

1. Schueller, F. W. *Chemobiodynamics and Drug Design*. McGraw-Hill: New York, 1960.
2. Curd, F. H. S., Davey, D. G., Rose, F. L. Studies on synthetic antimalarial drugs-X. Some biguanide derivatives as new types of antimalarial substances with both therapeutic and causal prophylactic activity. *Ann. Trop. Med.* 1945, *39*, 208–214.
3. Crowther, A. F., Levi, A. A. Proguanyl – the isolation of a metabolite with high antimalarial activity. *Br. J. Pharmacol.* 1953, *8*, 93–101.
4. Hawking, F., Perry, W. L. M. Activation of paludrine. *Br. J. Pharmacol.* 1948, *3*, 320–331.
5. Weiner, I. M. Drugs affecting renal function and electrolyte metabolism. In *Goodman and Gilman's the Pharmacological Basis of Therapeutics* (Goodman-Gilman, A., Rall, T. W., Nies, A. S., Taylor, P., Eds), 8th Edition. Pergamon Press: New York, 1990, pp. 708–731.
6. Buzetta, B., Hospital, M. Relations structure-activité. In *Pharmacologie Moléculaire* (Cohen, Y., Ed.). Masson & Cie: Paris, 1978, pp. 27–36.
7. Minutolo, F., Bertini, S., Papi, C., Carlson, K. E., Katzenellenbogen, J. A., Macchia, M. Salicylaldoxime moiety as a phenolic "A-Ring" substitute in estrogen receptor ligands. *J. Med. Chem.* 2001, *44*(24), 4288–4291.
8. Minutolo, F., Antonello, M., Bertini, S., Ortore, G., Placanica, G., Rapposelli, S., Sheng, S., Carlson, K. E., Katzenellenbogen, B. S., Katzenellenbogen, J. A., Macchia, M. Novel estrogen receptor ligands based on an anthranylaldoxime structure: role of the phenol-type pseudocycle in the binding process. *J. Med. Chem.* 2003, *46*(19), 4032–4042.
9. Rouot, B., Leclerc, G., Wermuth, C. G., Miesch, F., Schwartz, J. Synthèse et essais pharmacologiques d'arylguanidines, analogues ouverts de la clonidine. *Eur. J. Med. Chem.* 1978, *13*, 337–342.
10. Ashwood, V. A., Cassidy, F., Coldwell, M. C., Evans, J. M., Hamilton, T. C., Howlett, D. R., Smith, D. M., Stemp, G. Synthesis and antihypertensive activity of 4-(substituted-carbonylamino)-2*H*-1-benzopyrans. *J. Med. Chem.* 1990, *33*, 2667–2672.
11. Mack, H., Pfeiffer, T., Hornberger, W., Bohm, H. J., Hoffken, H. W. Design, synthesis and biological activity of novel rigid amidinophenylalanine derivatives as inhibitors of thrombin. *J. Enzyme Inhib.* 1995, *9*(1), 73–86.
12. Kubinyi, H. Molecular similarity. 2. The structural basis of drug design. *Pharm Unserer Zeit* 1998, *27*(4), 158–172.
13. Kaminski, J. J., Wallmark, B., Briving, C., Andersson, B. M. Antiulcer agents. 5. Inhibition of gastric H+/K(+)-ATPase by substituted imidazo[1,2-a]pyridines and related analogues and its implication in modeling the high affinity potassium ion binding site of the gastric proton pump enzyme. *J. Med. Chem.* 1991, *34*(2), 533–541.
14. Nakamura, C. E., Abeles, R. H. Mode of interaction of β-hydroxy-β-methylglutaryl coenzyme A reductase with strong binding inhibitors: compactin and related compounds. *Biochemistry* 1985, *24*, 1364–1376.
15. Kawai, K., Tamura, S., Morimoto, S., Ishii, H., Kuzuna, S. Pharmacology of 4-benzoyl-1-indancarboxylic acid (TAI-901) and 4-(4-methylbenzoyl)-1-indancarboxylic acid (TAI-908). *Arzneim. Forsch.* 1982, *32*, 113–117.
16. Boyle, E. A., Mangan, F. R., Markwell, R. E., Smith, S. A., Thomson, M. J., Ward, R. W., Wyman, P. A. 7-Aroyl-2,3-dihydrobenzo[*b*]furan-3-carboxylic acids and 7-benzoyl-2,3-dihydrobenzo[*b*]thiophene-3-carboxylic acids as analgesic agents. *J. Med. Chem.* 1986, *29*, 894–898.
17. Jenner, P., Clow, A., Reavill, C., Theodoru, A., Marsden, C. D. Stereoselective actions of substituted benzamide drugs on cerebral dopamine mechanisms. *J. Pharm. Pharmacol.* 1980, *32*, 39–44.
18. Waterbeemd van de, H., Testa, B. Theoretical conformational studies of some dopamine antagonistic benzamide drugs: 3-pyrrolidyl- and 4-piperidyl derivatives. *J. Med. Chem.* 1983, *26*, 203–207.
19. Norman, M. H., Kelley, J. L., Hollingsworth, E. B. Conformationally restricted analogues of remoxipride as potential antipsychotic agents. *J. Med. Chem.* 1993, *36*, 3417–3423.
20. Cannon, J. G., Lee, T., Goldman, H. D., Costall, B., Naylor, R. J. Central dopamine agonist properties of some 2-aminotetralin derivatives after peripheral and intracerebral administration. *J. Med. Chem.* 1977, *20*, 1111–1116.
21. Seiler, M. P., Bölsterli, J. J., Floersheim, P., Hagenbach, A., Markstein, R., Pfäffli, P., Widmer, A., Wüthrich, H. Recognition at dopamine receptor subtypes. In *Perspectives in Medicinal Chemistry* (Testa, B., Kyburz, E., Fuhrer, W., Giger, R., Eds), Verlag Helvetica Chemica Acta and VCH: Basel and Weinheim, 1993, pp. 221–237.
22. Krogsgaard-Larsen, P., Hjeds, H., Falch, E., Jørgensen, F. S., Nielsen, L. Recent advances in GABA agonists, antagonists and uptake inhibitors: structure–activity relationships and therapeutic potential. In *Adv. Drug Res.* (Testa, B., Ed.), Vol. 17. Academic Press: London, 1988, pp. 381–456.
23. Levin, E. D., Simon, B. B. Nicotinic acetylcholine involvement in cognitive function in animals. *Psychopharmacology (Berl)* 1998, *138*(3–4), 217–230.
24. Williams, M., Arneric, S. P. Beyond the tobacco debate: dissecting out of the therapeutic potential of nicotine. *Expert Opin. Invest. Drugs* 1996, 1035–1045.
25. Glassco, W., Suchocki, J., George, C., Martin, B. R., May, E. L. Synthesis, optical resolution, absolute configuration, and preliminary pharmacology of (+)- and (−)-*cis*-2,3,3a,4,5,9b-hexahydro-1-methyl-1*H*-pyrrolo-[3,2-*h*]isoquinoline, a structural analog of nicotine. *J. Med. Chem.* 1993, *36*, 3381–3385.
26. Glennon, R. A., Dukat, M. Nicotine receptor ligands. *Med. Chem. Res.* 1996, 465–486.

27. Holladay, M. W., Lebold, S. A., Lin, N. H. Structure–activity relationships of nicotinic acetylcholine receptor agonists as potential treatments for dementia. *Drug. Dev. Res.* 1995(35), 191–213.
28. Damaj, M. I., Glassco, W., Marks, M. J., Slobe, B., James, J. R., May, E. L., Rosecrans, J. A., Collins, A. C., Martin, B. R. Pharmacological investigation of (+)- and (−)-*cis*-2,3,3a,4,5,9b-hexahydro-1-methyl-1*H*-pyrrolo-[3,2-*h*]isoquinoline, a bridged-nicotine analog. *J. Pharmacol. Exp. Ther.* 1997, *282*(3), 1425–1434.
29. Ullrich, T., Krich, S., Binder, D., Mereiter, K., Anderson, D. J., Meyer, M. D., Pyerin, M. Conformationally constrained nicotines: polycyclic, bridged, and spiro-annulated analogues as novel ligands for the nicotinic acetylcholine receptor. *J. Med. Chem.* 2002, *45*(18), 4047–4054.
30. Basil, B., Coffee, E. C. J., Gell, D. L., Maxwell, D. R., Sheffield, D. J., Wooldridge, K. R. H. A new class of sympathic β-receptor blocking agents. 3,4-Dihydro-3-hydroxy-1,5-benzoxazocines. *J. Med. Chem.* 1970, *13*, 403–406.
31. Evans, J. M., Fake, C. S., Hamilton, T. C., Poyser, R. H., Watts, E. A. Synthesis and antihypertensive activity of substituted *trans*-4-amino-3,4-dihydro-2,2-dimethyl-2*H*-1-benzopyran-3-ols. *J. Med. Chem.* 1983, *26*, 1582–1589.
32. Stemp, G., Evans, J. M. Discovery and development of cromokalim and related potassium channel activators. In *Med. Chem.* (Ganellin, C. R., Roberts, S. M., Eds). Academic Press: London, 1993, pp. 141–162.
33. Heel, R. C., Brogden, R. N., Pakes, G. E., Speight, T. M., Avery, G. S. Nefopam: a review of its pharmacological properties and therapeutic efficacy. *Drugs* 1980, *19*, 249–267.
34. Glaser, R., Donnel, D. Stereoisomer differentiation for the analgesic drug nefopam hydrochloride using modeling studies of serotonin uptake area. *J. Pharm. Sci.* 1989, *78*, 87–90.
35. Antel, J., Gregory, P. C., Nordheim, U. CB1 cannabinoid receptor antagonists for treatment of obesity and prevention of comorbid metabolic disorders. *J. Med. Chem.* 2006, *49*(14), 4008–4016.
36. Murineddu, G., Ruiu, S., Loriga, G., Manca, I., Lazzari, P., Reali, R., Pani, L., Toma, L., Pinna, G. A. Tricyclic pyrazoles. 3. Synthesis, biological evaluation, and molecular modeling of analogues of the cannabinoid antagonist 8-chloro-1-(2′,4′-dichlorophenyl)-*N*-piperidin-1-yl-1,4,5,6-tetrahydrobenzo [6,7]cyclohepta[1,2-c]pyrazole-3-carboxamide. *J. Med. Chem.* 2005, *48*(23), 7351–7362.
37. Carpino, P. A., Griffith, D. A., Sakya, S., Dow, R. L., Black, S. C., Hadcock, J. R., Iredale, P. A., Scott, D. O., Fichtner, M. W., Rose, C. R., Day, R., Dibrino, J., Butler, M., Debartolo, D. B., Dutcher, D., Gautreau, D., Lizano, J. S., O'Connor, R. E., Sands, M. A., Kelly-Sullivan, D., Ward, K. M. New bicyclic cannabinoid receptor-1 (CB1-R) antagonists. *Bioorg. Med. Chem. Lett.* 2006, *16*(3), 731–736.
38. Smith, R. A., Fathi, Z., Brown, S. E., Choi, S., Fan, J., Jenkins, S., Kluender, H. C., Konkar, A., Lavoie, R., Mays, R., Natoli, J., O'Connor, S. J., Ortiz, A. A., Podlogar, B., Taing, C., Tomlinson, S., Tritto, T., Zhang, Z. Constrained analogs of CB-1 antagonists: 1,5,6,7-Tetrahydro-4*H*-pyrrolo[3,2-c]pyridine-4-one derivatives. *Bioorg. Med. Chem. Lett.* 2007, *17*(3), 673–678.
39. 3D structures were generated and minimized using DS viewerPro 6.0, © 2005, by Accelrys Software Inc.
40. Pfister, R., Sallmann, A., Hammerschmidt, W., Substituted malonic acid hydrazides. *US Patent3,455,999* (July 15, 1969; Geigy Chemical Corporation).
41. Mixich, G. Zum chemischen Verhalten des Antiphlogisticums "Azapropazon" (Mi 85) = 3-dimethylamino-7-methyl-1,2-(*n*-propylmalonyl)-1,2-dihydro-1,2,4-benzotriazin. *Helv. Chim. Acta.* 1968, *51*, 532–538.
42. Jahn, U., Wagner-Jauregg, T. Vergleich von Zwei neuen Klassen Antiphlogisticher Substanzen im Collier-Test. *Arzneim.-Forsch.* 1968, *18*, 120–121.
43. Yamamoto, H., Kaneko, S.-I. Synthesis of 1-phenyl-2-styryl-3,5-dioxopyrazolidines as antiinflammatory agents. *J. Med. Chem.* 1970, *13*, 292–295.
44. Wermuth, C. G., Choay, J., Nouveaux dérivés de la pyrazolidine, leur procédé de fabrication et médicaments contenant ces nouveaux dérivés. *Fr. Patent 2 244 513* (July 12, 1973, Choay S.A.), 1973.
45. Teger-Nilsson, A. C., Bylund, R., Gustafsson, D., Gyzander, E., Eriksson, U. *In vitro* effects of inogatran, a selective low molecular weight thrombine inhibitor. *Thromb. Res.* 1997, *85*, 133.
46. Gustafsson, D., Antonsson, T., Bylund, R., Eriksson, U., Gyzander, E., Nilsson, I., Elg, M., Mattson, C., Deinyum, J., Pehrsson, S., Karlsson, A., Soerensen, H. Effects of melagatran, a low-molecular-weight thrombin inhibitor, on thrombin and fibrinolytic enzymes. *Thromb. Haemostasis* 1998, *79*, 110.
47. Pelletier, J. C., Hesson, D., Jones, K. A., Costa, A. M. Substituted 1,2-dihydrophtalazines: potent, selective, and noncompetitive inhibitors of the AMPA receptor. *J. Med. Chem.* 1996, *39*, 343–346.
48. Grasso, S., De Sarro, G., De Sarro, A., Micale, N., Zappala, M., Puja, G., Baraldi, M., De Micheli, C. Synthesis and anticonvulsant activity of novel and potent 6,7-methylenedioxyphtalazin-1 (2*H*)-ones. *J. Med. Chem.* 2000, *43*, 2851–2859.
49. Trybulski, E. J., Zhang, J., Kramss, R. H., Mangano, R. M. The synthesis and biochemical pharmacology of enantiomerically pure methylated oxotremorine derivatives. *J. Med. Chem.* 1993, *36*, 3533–3541.
50. Ringdahl, B., Jenden, D. J. Pharmacological properties of oxotremorine and its analogues. *Life Sci.* 1983, *32*, 2401–2413.
51. Klimko, P. G., Davis, T. L., Griffin, B. W., Sharif, N. A. Synthesis and biological activity of a novel 11a-homo (cyclohexyl) prostaglandin. *J. Med. Chem.* 2000, *43*, 3400–3407.
52. Mull, R. P., Maxwell, R. A. Guanethidine and related adrenergic neuronal blocking agents. In *Antihypertensive Agents* (Schlittler, E., Ed.), Academic Press: New York, 1967, pp. 115–149.
53. Giudicelli, R., Najer, H., Lefèvre, F. Comparaison des durées d'action du *N*-β-guanidino-éthyl-aza-6 spiro[2,5] octane (LD 3598) et de la guanéthidine. *Compt. Rend. Acad. Sci.* 1965, *260*, 726–729.
54. Najer, H., Giudicelli, R., Sette, J. Guanidines douées d'action antihypertensive, 4e mémoire: *N*-β-guanidinoéthyl azaspiro alcanes (1ère partie). *Bull. Soc. Chim. Fr.* 1964, 2572–2581.
55. Toda, N., Usui, H., Shimamoto, K. Modification by AZ-55, guanethidine and bretylium of responses of atria and aortic strips to transmural stimulation. *Jpn. J. Pharmacol.* 1972, *22*, 125–135.
56. Najer, H., Giudicelli, R., Sette, J. Guanidines douées d'activité antihypertensive, 3e mémoire: *N*-β-guanidinoéthyl azabicyclo alcanes. *Bull. Soc. Chim. Fr.* 1962, 1593–1597.
57. Anonymous, New basic derivatives of 6,9-endoxo-3-azabicyclo-[4.3.0]-nonane. *Br. Patent 973,533* (February 21, 1962, Laboratoire Lumière, S.A.), 1964.
58. Anonymous, New basic derivatives of 6,9-endomethylene-3-azabicyclo-[4.3.0]-nonane. *Br. Patent 972,088* (February 19, 1962, Laboratoire Lumière, S.A.), 1964.
59. Bryans J. S., Morrell, A. I., Novel stereoselective processes for preparation of gabapentin analogues W09914184 (Warner-Lambert Company (USA)), 1999.
60. Horwell, D. C., Bryans, J. S., Kneen, C. O., Morrell, A. I., Ratcliffe, G. S. Preparation of novel bridge cyclic amino acids as pharmaceutical agents W09733859 (Warner-Lambert Company (USA)), 1997.
61. Meyer, M. D., Altenbach, R. J., Basha, F. Z., Carroll, W. A., Condon, S., Elmore, S. W., Kerwin, J. F., Jr., Sippy, K. B., Tietje, K., Wendt, M. D., Hancock, A. A., Brune, M. E., Buckner, S. A., Drizin, I. Structure–activity studies for a novel series of tricyclic substituted hexahydrobenz[e]isoindole alpha(1A) adrenoceptor antagonists as potential agents for the symptomatic treatment of benign prostatic hyperplasia (BPH). *J. Med. Chem.* 2000, *43*(8), 1586–1603.
62. Jongebreur, G. Relation between the chemical constitution and the pharmacological action, especially on the coronary vessels of the heart, of some synthesized pyrones and khellin. *Arch. Intern. Pharmacodynamie.* 1952, *90*, 384–411.
63. Evans, B. E., Rittle, K. E., Bock, M. G., DiPardo, R. M., Freidinger, R. M., Whitter, W. L., Lundell, G. F., Veber, D. F., Anderson, P. S.,

Chang, R. S. L., Lotti, V. J., Cerino, D. J., Chen, T. B., Kling, P. J., Kunkel, K. A., Springer, J. P., Hirshfield, J. Methods for drug discovery: development of potent, selective, orally effective cholecystokin antagonists. *J. Med. Chem.* 1988, *31*, 2235–2246.

64. Janssen, P. A. J., Van der Eycken, C. A. M. The chemical anatomy of potent morphine-like analgesics. In *Drugs Affecting the Central Nervous system* (Burger, A., Ed.), Marcel Dekker, Inc.: New York, 1968, pp. 25–60.
65. Hjorth, S., Carlsson, A., Wikström, H., Lindberg, P., Sanchez, D., Hacksell, U., Arvidsson, L.-E., Svensson, U., Nilsson, J. L. G. 3-PPP, a new centrally acting DA-receptor agonist with selectivity for autoreceptors. *Life Sci.* 1981, *28*, 1225–1238.
66. Bach, N. J., Kornfeld, E. C., Jones, N. D., Chaney, M. O., Dorman, D. E., Paschal, J. W., Clemens, J. A., Smalstig, E. B. Bicyclic and tricyclic ergoline partial structures. Rigid 3-(2-aminoethyl)pyrazoles as dopamine agonists. *J. Med. Chem.* 1980, *23*, 481–491.
67. Fuller, R. W., Clemens, J. A., Kornfeld, E. C., Snoddy, H. D., Smalstig, E. B., Bach, N. J. Effects of (8β)-8-[(methylthio)methyl]-6-propylergoline on dopaminergic function and brain dopamine turnover in rats. *Life Sci.* 1979, *24*, 375–382.
68. Kelly, T. R., Howard, H. R., Koe, K., Sarges, R. Synthesis and dopamine autoreceptor activity of a 5-(methylmercapto)methyl-substituted derivative of (±)-3-PPP (3-(3-hydroxyphenyl)-1-*n*-propylpiperidine). *J. Med. Chem.* 1985, *28*, 1368–1371.
69. Chang, R. S. L., Lotti, V. Y. Biochemical and pharmacological characterization of an extremely potent and selective nonpeptide CCK antagonist. *Proc. Natl. Acad. Sci. USA* 1986, *83*, 4923–4926.
70. Bock, M. G., DiPardo, R. M., Rittle, K. E., Evans, B. E., Freidinger, R. M., Veber, D. F., Chang, R. S. L., Chen, T.-B., Keegan, M. E., Lotti, V. J. Cholecystokinin antagonists. Synthesis of asperlicin analogues with improved potency and water solubility. *J. Med. Chem.* 1986, *29*, 1941–1945.
71. Bruderlein, F. T., Humber, L. G., Voith, K. Neuroleptic agents of the benzocycloheptapyridoisoquinoline series. 1. Syntheses and stereochemical and structural requirements for activity of butaclamol and related compounds. *J. Med. Chem.* 1975, *18*, 185–191.
72. Imhof, R., Kyburz, E., Daly, J. J. Design, synthesis, and X-ray data of novel potential antipsychotic agents. Substituted 7-phenylquinolizidines: stereospecific, neuroleptic, and antinociceptive properties. *J. Med. Chem.* 1984, *27*, 165–175.
73. Evans, R. H., Francis, A. A., Jones, A. W., Smith, D. A. S., Watkins, J. C. The effects of a series of ω-phosphonic α-carboxylic amino acids on electrically evoked and excitant amino acid-induced responses in isolated spinal cord preparations. *Br. J. Pharmacol.* 1982, *75*, 65–75.
74. Hutchinson, A. J., Williams, M., Angst, C., de Jesus, R., Blanchard, L., Jackson, R. H., Wilusz, E. J., Murphy, D. E., Bernard, P. S., Schneider, J., Campbell, T., Guida, W., Sills, M. A. 4-(Phosphonoalkyl)- and 4-(phosphonoalkenyl)-2-piperidinecarboxylic acids: synthesis, activity at *N*-methyl-d-aspartic acid receptors, and anticonvulsant activity. *J. Med. Chem.* 1989, *32*, 2171–2178.
75. Humblet, C., Johnson, G., Malone, T., Ortwine, D. F. In Design, synthesis and molecular modeling of phosphonoalkyl-substituted tetrahydroisoquinolines as competitive NMDA antagonists. *Proceedings of the XIth International Symposium on Medicinal Chemistry*, Jerusalem, Israël, 1990, p. 26.
76. Ornstein, P. L., Arnold, M. B., Augenstein, N. K., Lodge, D., Leander, J. D., Schoepp, D. D. (3SR,4aRS,6RS,8aRS)-6-[2-(1*H*-tetrazol-5-yl)ethyl]decahydroisoquinoline-3-carboxylic acid: a structurally novel, systematically active, competitive AMPA receptor antagonist. *J. Med. Chem.* 1993, *36*, 2046–2048.
77. Koga, H., Itoh, A., Murayama, S., Suzue, S., Irikura, T. Structure–activity relationships of antibacterial 6,7- and 7,8-disubstituted 1-alkyl-1,4-dihydro-4-oxoquinoline-3-carboxylic acids. *J. Med. Chem.* 1980, *23*, 1358–1363.
78. Jinbo, Y., Taguchi, M., Inoue, Y., Kondo, H., Miyasaka, T., Tsujishita, H., Sakamoto, F., Tsukamoto, G. Synthesis and antibacterial activity of a new series of tetracyclic pyridone carboxylic acids. 2. *J. Med. Chem.* 1993, *36*, 3148–3153.
79. Wijngaarde van, I., Hamminga, D., Hes, R.v., Standaar, P. J., Tipker, J., Tulp, M. T. M., Mol, F., Olivier, B., Jonge de, A. Development of high-affinity 5-HT3 receptor antagonists. Structure-affinity relationships of novel 1,7-annelated indole derivatives. 1. *J. Med. Chem.* 1993, *36*, 3693–3699.
80. Faust, R., Garrat, P. J., Yeh, L.-K., Tsotinis, A., Panoussopoulou, M., Calogeropoulou, T., Teh, M.-T., Sugden, D. Mapping the melatonin receptor. 6. Melatoonin agonists and antagonists derived from 6*H*-isoindolo[2,1-a]indoles, 5,6-dihydroindolo[2,1-a]isoquinolines, and 6,7-dihydro-5*H*-benzo[c]azepino[2,1-*a*]indoles. *J. Med. Chem.* 2000, *43*, 1050–1061.
81. Stanton, J. L., Sperbeck, D. M., Trapani, A. J., Cote, D., Sakane, Y., Berry, C. J., Ghai, R. D. Heterocyclic lactam derivatives as dual angiotensin converting enzyme and neutral endopeptidase 24.11 inhibitors. *J. Med. Chem.* 1993, *36*, 3829–3833.

Chapter 17

Conformational Restriction and/or Steric Hindrance in Medicinal Chemistry

André Mann

I. INTRODUCTION
A. Theoretical points
B. On constrained analogs
C. On conformational analysis
D. Steric effects
E. Rigid compounds and bioavailability
II. CASE STUDIES
A. Bradykinin
B. Allylic constraints for inducing conformational rigidity
C. Diversity-Oriented Synthesis
D. Epibatidine bioactive conformation
E. Ligands for the Hepatitis C virus
F. Nociceptin
G. Opioid receptors ligands
H. Peptidomimetics for SH2 domains
III. SUMMARY AND OUTLOOK
REFERENCES

If a change occurs in one of the conditions of a system in equilibrium, the system will adjust itself so as to annul, as far as possible, the effect of that change.

Henri-Louis Le Chatelier (1850–1936)

I. INTRODUCTION

For the practice of medicinal chemistry the understanding of the thermodynamics of the events that are governing molecular recognition is essential. The binding (non-covalent interactions) between a drug and a protein is mediated by several interactions: ion–ion, dipole–dipole and lipophilic interactions, hydrogen bonding and shape complementarity. The strength of the binding equilibrium is obtained by thermodynamic data that are measuring the relative contributions of each of those interactions. If two molecules bind, they must display greater affinity for each other than they do for the solvent (in many cases water). The probability that a ligand is catched by a receptor is depending on the existence of intermolecular forces: *the promotion of binding is favored by favorable interactions of polar functional groups and the formation of hydrophobic contacts*; *the adverse factors for binding are mainly related to the restriction of translational and vibrational energies upon complex formation*.[1–4] Indeed, the consequences of the binding of a ligand to a receptor are a loss of conformational energy which is gained by the surroundings (entropy). The net result for the receptor–ligand system is an energy penalty, detrimental for the strength of the binding. In order to overcome such entropy penalties and to try to improve the potency of a ligand, a suitable strategy is to restrict its *conformational flexibility*. Therefore once a lead compound has been identified for a targeted biological receptor, the optimization toward potency and/or selectivity is usually the next step. Among several strategies, depending on the chemical structure of the identified lead, the introduction of conformational restriction and/or steric hindrance are popular tactics in medicinal chemistry. This approach can tentatively be conceptualized as following: it is the optimization of the *free energy* gained during the association of a ligand with a receptor by the relative spatial disposition of functional groups in introducing: *bulk, unsaturation or cyclization*. The expected benefits will be: receptor selectivity, increase potency, pharmacophore identification, and metabolic stabilization. Additionally, a chemical *bonus* can be expected in producing original compounds with innovative chemistry. Finally, there are also some risks: any structural change in

Copyright © 2008, Elsevier Ltd
All rights reserved.

a compound alters more than one property with unknown consequences. Indeed, this approach is valuable in cases where the lead compounds are exhibiting flexibility and where the low energy conformations are not representative for a good fit for the receptor. In this chapter some theoretical points related to thermodynamics will be addressed, and then several examples will be presented where the introduction of rigidity has produce a positive or a negative impact in structure-based drug design.

A. Theoretical points

Molecular recognition is the start for all biological processes. Indeed, many of the important issues in medicinal chemistry hinge on understanding the bimolecular non-covalent interactions between a biomacromolecule (receptor) and a small ligand (drug). Under equilibration conditions, binding affinity can be expressed in terms of difference in free energy (ΔG) of free and bound states, or in terms of the equilibrium constant (K), experimentally accessible by a direct measure of the energy of the interaction with the Gibbs-van't Hoff relation (equation 17.1, or more conveniently equation 17.1a, with conversion to decimal log):

$$\Delta G^\circ = RT \ln K \quad \text{with} \quad \Delta G^\circ = \Delta H^\circ - T\Delta S^\circ \qquad R = 8.13 \text{ J/mol/K} \tag{17.1}$$

$$\Delta G^\circ = -5.7^* \log K \quad {}^*\text{in [kJ/mol] at } 30^\circ\text{C} \tag{17.1a}$$

The more negative ΔG°, the larger is the affinity of the ligand for the bioreceptor (the binding constant is high). As a reminder if a non-covalent association is expected ΔG° must be negative and the Gibbs equation is demanding a negative value for ΔH° and a positive for $T\Delta S^\circ$. If $\Delta G^\circ = -30$ kJ/mol the association constant is in the micromolar range. By using the Gibbs equation (equation 17.1), we can monitor the effects of changes in bonding within the system (ΔH°) and changes in the entropy of the system ($T\Delta S^\circ$). The values of the standard function ΔH° can be considered a quantitative indicator of the changes in intermolecular bond energies (hydrogen bonding and van der Waals interactions) occurring during the binding, while the standard entropy ΔS° is a good indicator of the reorganization undergone by the water molecules during the same process. As a general consideration if ΔG° is negative, it means that a chemical process (a reaction, for the forming of covalent bonds, or an equilibrium, for the forming of an association) are favored. If for the formation of a *covalent bond* the contribution to ΔG° is usually dominated by ΔH° (the bond is formed in an *irreversible* process), in the case of a *non-covalent interaction* or equilibrium the part taken by ΔH° and $T\Delta S^\circ$ in the energy balance are often comparable. These two considerations are the result of the respective bond strengths from 100 to 500 kJ/mol for the covalent bond (the enthalpy term is largely prevailing: $\Delta H^\circ >> T\Delta S^\circ$) and, from 1 to 60 kJ/mol for the irreversible binding (the enthalpy and entropy are of similar importance: $\Delta H^\circ \cong T\Delta S^\circ$)[5–8] (Box 17.1).

BOX 17.1 Non-covalent Association: Balance Between Energy Benefit and Penalty

Before the non-covalent association, the ligand and receptor surrounded by water molecules have conformational flexibility. But once fixed to the receptor the flexibility of the ligand and the receptor (in part) are severely reduced (Scheme 1). The energy balance of the complex changes in the following way: the polar interactions (**A … X**, or **B … Y**), are contributing for a favorable ΔG (the energy term $\Delta H < 0$), in contrast the reduced mobility of the ligand in the complex will result in an entropy penalty, (energy will be lost $\Delta S < 0$), an unfavorable contribution, but the release of fixed water will favor the entropy (ΔS of the system is increasing). The water molecules attached on the ligand or on the receptor are more ordered that in the water bulk, liberation during the binding will be entropy driven and favorable to the binding process. Therefore, a binding equilibrium is set if an overall negative change in free energy is accounted ($\Delta G < 0$, equation 17.1)[9–10] (Figure 17.1).

FIGURE 17.1 Non-covalent association between a ligand with two polar atoms (**A** and **B**) linked by three methylene units with a receptor containing complementary polarizable atoms (**X** and **Y**). *Source*: Adapted from Ref. [9].

The values of elemental binding energies for non-covalent association are small and comparable to the thermal energy that opposes them in the binding experiment in aqueous media. Binding constants have been estimated by using an extension of the Gibbs equation. Böhm was using a set of interactions arising from experimentally known binding constants resulting from the association of small molecular weight ligands with proteins.[11] Five types of ΔG contributions were identified (see Table17.1) and their values were extracted from a regression analysis. These values are generally used for a semi-quantitative evaluation of binding constants and are useful in the design of ligands (Table 17.1).

1. Entropy–enthalpy compensation

A key point is that a decrease in motion implies a decrease in entropy, since it results in fewer accessible arrangements. Consider the formation of a specific non-covalent bond (e.g. $\mathbf{L}^{\cdots}\mathbf{R}$ for the transformation $\mathbf{L} + \mathbf{R} \rightarrow \mathbf{L}^{\cdots}\mathbf{R}$). An increase in its strength (which corresponds to an increasing negative contribution to $\Delta H°$, favorable to the binding process) will be accompanied by an increasing restriction in the relative motion of **L** and **R** in $\mathbf{L}^{\cdots}\mathbf{R}$ (which corresponds to a negative contribution to $\Delta S°$, unfavorable to the binding process). *This opposing interplay between enthalpy and entropy is known as enthalpy/entropy compensation and is a fundamental property of non-covalent interactions*. It arises because bonding opposes motion and, also reciprocally, motion opposes bonding. The two effects can be traded off against each other because the strength of non-covalent bonds is, at room temperature, comparable to the thermal energies that oppose them. The above considerations have experimental basis. The thermodynamic parameters ($\Delta G°$, $\Delta H°$, $\Delta S°$) for 136 ligands, binding to 10 biological receptors (mainly membrane receptors) have been analyzed.[12] It appeared that $\Delta H°$ and $\Delta S°$ values correlate according to a regression equation (see equation 17.2). The regression seems remarkable in view of its high correlation and because of the large number of measurements ($n = 186$). The observed *enthalpy–entropy compensation* seems to be indicative of common factors that control the binding processes, irrespective of the nature of ligands and of their macromolecular targets:

$$\Delta H° = -39 \text{ kJ/mol} + 287\Delta S° \quad \text{in (kJ/mol)} \qquad (17.2)$$
$$(n = 186,\ r = 0.981,\ s = 2.06)$$

This correlation is reflecting another concept: any tightening of the intermolecular bonds (*the enthalpic factor*) is compensated by a loss of degrees of freedom (*the entropic factor*). The binding association can be enthalpy- or entropy-driven. The origin of the $\Delta H/\Delta S$ compensation is probably related to an intrinsic property of the hydrogen bond which is determining the association of the participants (water, drug, and binding site) in the drug-receptor binding equilibrium.[12] As a conclusive statement the thermodynamic laws teach us that three factors are involved for improving binding affinity: (i) improving ligand protein interactions over those with the solvent in order to obtain a favorable (negative) enthalpy change; (ii) making the ligand more hydrophobic in order to make the solvation entropy large and positive and (iii) pre-shaping the ligand to the geometry of the binding site in order to minimize the loss of conformational entropy upon binding.[13]

TABLE 17.1 Average Values for Elemental ΔG Contributions

Physical process	ΔG values kJ/mol[a]
Energy cost of bimolecular association	+5.4
Energy cost of restriction of an internal rotor	+2[b]
Benefit of the hydrophobic effect (per Å of buried hydrocarbon)	−0.17 Å^2
Benefit of making a neutral hydrogen bond of ideal geometry	−4.7
Benefit of making an ionic hydrogen bond of ideal geometry	−8.3

Source: From Böhm[11] and Williams[9].
[a]Elemental contribution to the global $\Delta G°$ value.
[b]Per rotor.

2. Enthalpy- versus entropy-driven binding

The following example – a case where thermodynamics evidenced a different binding behavior of agonist and antagonists – is illustrative for the enthalpy/entropy compensation in the binding of a series of agonists and antagonists of the β-adrenergic receptor (Table 17.1).[14,15] β-Adrenergic agonists and/or partial agonists are found to bind with large negative enthalpies ($\Delta H°$ from −79 to −56 kJ/mol), indicating strong electrostatic interactions between the bound conformation on the receptor, associated with a large loss of motional entropy (a tight complex is formed: $\Delta S°$ from −45.7 to −16.6 J/mol/K). Antagonists, on the other hand, do not fulfill the requirements of a good complementarity as evidenced by the criterion that their enthalpies of binding are small ($\Delta H°$ from −21 to +14 kJ/mol), the loss of motional entropy is therefore less than that of the agonists (a loose complex is formed $\Delta S°$ from +17 to +52.5 J/mol/K). But the agonists and antagonists have similar affinity constants, supporting the enthalpy/entropy compensation (Table 17.2).

In order to get a better understanding of the enthalpy–entropy compensation, the data from table 17.2 are best

TABLE 17.2 Thermodynamic Parameters of Ligands Binding to the β-adrenergic Receptor of Turkey Erythrocytes and the Equilibrium Constants K_D at 37°

	ΔG° (kJ/mol)	ΔH° (kJ/mol)	ΔS° (J/mol/K)	K_D/mol (at 310°K, 37°C)
Agonists				
(−) isoproterenol	−39.3	−56	−54	$2.5 \cdot 10^{-7}$
(−)Norephedrine	−33	−79	−45.7	$2.5 \cdot 10^{-6}$
Partial agonists				
Soteronol	−34.2	−32.6	+5	$1.6 \cdot 10^{-6}$
Fenoterol	−32.2	−25.1	+23	$4 \cdot 10^{-6}$
Terbutaline	−25.1	−17.2	+27.5	$4 \cdot 10^{-5}$
Antagonists				
(−)Propranolol	−52.2	−15.9	+116.6	$1.5 \cdot 10^{-9}$
IPS-339	−51	+0.8	+169	$2.5 \cdot 10^{-9}$
Pindolol	−49	−20.9	+91.1	$4 \cdot 10^{-9}$
Atenolol	−31	−14.2	+54	$7 \cdot 10^{-6}$
Practolol	−31	+16.3	+153	$7 \cdot 10^{-6}$
Sotalol	−34.5	−8.7	+81.5	$1.2 \cdot 10^{-6}$

Source: Adapted from Ref. [15].

FIGURE 17.2 Plots for ΔH versus $T\Delta S$ for the binding of agonists (●), antagonists (○) and partial agonists (⬬) to the β-adrenergic receptors. The data are extracted from Table 17.1. *Source*: Adapted with modification from Ref. [9].

represented with a plot ΔH versus $T\Delta S$. Indeed, in Figure 17.2 clusters are clearly differentiating the agonist, from the antagonist or partial agonists. An interesting explanation based on the dimerization of the membrane-bound receptors (G-protein-coupled receptors) has been given by Williams[9] about this observation: the agonists (●) (activating the receptor after binding) induce receptor aggregation which is beneficial in bonding (negative contribution to ΔH) but adverse in entropy (negative contribution to $T\Delta S$). The antagonist (○) (without activation after binding) should therefore be (in comparison to agonist binding) relatively adverse in bounding and favorable in entropy. The consequences are seen in Figure 17.1, the agonists (●) are binding to the β-receptors by an enthalpy-driven process, whereas the antagonists (○) are binding with entropy and/or enthalpy–entropy driven processes.

In contrast, if agonists induce dissociation of receptor multimers (e.g. the adenosine A1 or A2 receptor), antagonists may induce the formation of oligomers. Such antagonist binding should therefore be (in comparison to agonist binding) relatively favorable in overall bonding and unfavorable in entropy in opposition to the agonists with adverse bonding and favorable entropy. A plot of the thermodynamic data of the binding of agonists (◉) or antagonists (○) to adenosine A1 receptors is depicted in Figure 17.3. The antagonists and the agonists are clearly in an adverse situation the former are enthalpy-driven and the second entropy-driven.[9,16] Many other receptors systems have been studied from a thermodynamic point of view, including G-protein-coupled receptors[17] (Figure 17.3).

Finally, an important point to stress out is that hydrophobicity is a major source of binding in drug-receptor interactions. Indeed, from a study over 415 oral drugs it appeared that on an average drugs contain only one to two hydrogen donors and three to four acceptors, whereas the average number of hydrophobic atoms in a drug is 16.[10] The contributions between polar and hydrophobic interactions in molecular recognition may be considered similarly as the balance between enthalpy and entropy.

B. On constrained analogs

Let us consider the interaction of a biological receptor with a small ligand. Before the interaction the two entities

have translational and rotational flexibility, contributive to their respective entropy. If now the association occurs, degrees of motion, as well as internal rotations about single bonds are lost, the consequence is an entropy cost about −58 kJ/mol at room temperature for a small ligand less than 1,000 Da.[4,7,8] As a general statement for the binding of a ligand to a receptor: there is an *entropy cost* associated with any bimolecular interaction that is a consequence of the lost of degrees of motion when two molecules are rigidly constrained within a complex. The torsional entropy (the free rotation of a bond) is related to the rotatable bonds in the ligand and the freezing of one of them has a cost of 2–3 kJ/mol at room temperature.[4,7,9] Considering equation (17.1), the entropy penalty will render ΔS negative reducing the binding energy. As a consequence for building an association constant of a reasonable value ($\Delta G < 0$), the above adverse free energy costs in bringing about conformational order, must be offset by favorable intermolecular interactions such as hydrogen bonds, van der Waals packing, hydrophobic and Coulombian interactions. One consequence of this qualitative thermodynamic analysis is that a flexible ligand has to compensate by the enthalpic, the entropic parameter. Therefore binding optimization can be accomplished either by making ΔH more negative, or making ΔS more positive, or an appropriate combination of both. Theoretically, the highest value for a given binding constant is accessible to a flexible ligand if the receptor recognized the ligand in its low energy conformation and with an optimum orientation of the functional groups. Therefore, all gained informations toward the knowledge of the active conformation of the ligand will serve the chemical construction of a better lead. There are some illustrative examples obtained by the incorporation of constrained amino acids into bioactive compounds.[18] But as yet the access to the thermodynamic parameters of ligand/receptor interactions are still in its infancy, the structure-based drug design relays in the general cases on semi-empirical strategies using structure–activity relationships for the optimization steps. The restriction of flexibility is one of them. Search for rigid analogs, introduction of conformational constraints, steric hindrance, or finding locked conformations are the terms used in medicinal chemistry to account for that approach. Finally, the energy to introduce by chemical synthesis is expected to be restored by the increase of the binding value (Box 17.2).

FIGURE 17.3 Plots of ΔH versus ΔS for the binding of agonists (◉) and antagonists (○) to the adenosine A1 receptor. *Source*: Adapted with modification from Refs. [9,16].

C. On conformational analysis

To find the energetically favorable conformations for a flexible ligand, there are many possible calculation methods. These available methods calculate the energy for *in vacuo* isolated ligands, that did not account for the real binding conditions.[19]

It is almost obvious that during the binding process, flexible ligands are deformed when binding to receptors. This deformation is a general phenomenon and the reason

BOX 17.2 Phenylalanine: A Case Study for Conformational Rigidification

Phenylalanine, a lipophilic amino acid has been a popular substrata for evaluating the concept of rigidification. Indeed, the phenylethyl side chain may adopt a large variety of conformations, which maintain the two polar functions under various orientations. Non-natural phenylalanines have been synthesized by the introduction of bulky substituents, the construction of cycles with a carbon skeleton or by the incorporation of a functional group. Some rigidified phenylalanines are depicted in Figure 17.4.

FIGURE 17.4 Several modes of rigidification for phenylalanine.

for it, is presumed to lie in the ligand's search for hydrogen bonds on the protein to replace the solute-solvent hydrogen bonds which are lost as the molecule enters the binding site or in the unmasking of hydrophobic pockets not accessible by water. A study has been performed on 33 ligands whose single X-ray structures, as well as their conformations obtained by cocrystallization with their receptor have been recorded.[20] From that study it appears that the degree of conformational change depend somewhat on the number of rotors (rotatable bonds) in the ligand: for ligands with five or more rotatable bounds, the crystal structure does not represent the protein-bound conformation in shape. Therefore the solid state structure of flexible ligand remains only an existing conformation, but with limited use. Of course, in the more favorable cases, when a crystal structure of the complex-ligand-receptor can be obtained the design of improved ligand is facilitate because the main interaction can be quantified and then optimized. But at present time for membrane-located receptor (RCGP's) the crystal complexes are still awaiting.

D. On the nature of steric effects

Steric effects may arise in a number of ways.[21] Primary steric effects result from repulsions between valence electrons or non-bonded atoms. Such repulsions can only result in an increase in the energy of a group of atoms. In a chemical reaction the overall steric effect may be either favorable or unfavorable. For example, if sterics effects in the reactant are larger than that in the product (or transition state) then the reaction is favored (steric augmentation); if the reverse is the case, the reaction is disfavored (steric diminution). The same arguments can be put forward in biological systems for the formation of receptor–ligand complex. If one compare the binding of a ligand to a biological receptor with or without a subsequent chemical reaction, let say an enzyme or a hormonal receptor, there are some obvious difference. The enzymatic reactions involve only structural groups in proximity to those atoms which are actually participating in bond making or breaking. Therefore, the enzyme tolerates structurally different ligands provided that the spot of the ligands, where the reaction should take place is accessible. Even in some case, if the direct affinity of a ligand for an enzyme is low, the subsequent chemical transformation can take place. In the formation of a ligand–receptor complex any group of atoms that is in van der Waals contact with the receptor or the biopolymer can be or is involved in the binding event. If the receptor lies in a pocket that can adjust any bioactive substance no matter what its size or shape, then no steric effect will be observed. If however, the parent biopolymer has limited conformational flexibility, and, as it likely, this flexibility is not the same in all directions, then steric effect will be observed. Furthermore, the steric effect will be conformationally dependent, and it is probable that the minimal steric interaction principle will be observed: this principle states that a substituent whose steric effect is conformationally dependant will prefer that conformation which minimizes steric repulsions and will give rise to the smallest steric effect. Finally, there are secondary steric effects on receptor binding that are produced by a substituent: (i) lowering the accessibility to an important group due to steric hindrance; (ii) changing the concentration of a conformer due to steric effects; (iii) shielding the active site from attack by a bulky reagent; (iv) variation in the electronical resonance of a π-bonded substituent by out of plane repulsion.

E. Rigid compounds and bioavailability

If the management of potency and selectivity of a set of drug candidates is often an easy task, for the prediction of bioavailability no real rationale is at hand. Therefore, any effort in this direction is of great importance in the drug discovery process. An intriguing correlation between the bioavailability of a compound and the number of its rotatable bounds was found by Veber using an empirical approach based on a set of 1,100 drug candidates (this study was conducted in rat).[22] Indeed, 10 or fewer rotatable bonds together with a polar surface of 140 Å (or 12 or fewer H-bond donors and/or acceptors) irrespective of the molecular weight will give a high probability of good oral bioavailability in the rat! This finding is of great importance because until now the limit of the molecular weight for a drug candidate is set to be of 500 Da by the Lipinsky rule.[23] Veber *et al.* suggest that by freezing some of the rotatable bounds, the molecular weight (although the upper limit is not yet predictable) is no more an essential parameter to be considered. Of course, further data have to be piled up in order to validate the above observations; however, it seems that the introduction of conformational constraints in a drug candidate has to be considered also in the case when pharmacokinetics is an issue to be solved.

II. CASE STUDIES

A. Bradykinin

Bradykinin-related peptides are potent vasoactive molecules implicated in inflammation and pain that mediate their acute actions via two receptors one inductible (B1) and the other expressed (B2). It appears that a B1-selective antagonist would be useful as anti-inflammatory agents. The majority of non-peptidic B1 antagonists have a basic amine and a lipophilic sulfonamide as exemplified by compound **1** (SR240612, $K_i = 0.48$ nM),[24] the distance between the basic nitrogen and the sulfonamide has been recognized as crucial for binding selectivity to the B1 subtype over the B2. Lead

identification efforts sought to optimize the spatial placement between the basic nitrogen (piperidine ring) and the naphtylsulfonamide of the beta-phenylalanine. Compound **2** (R,S), a simplified surrogate of **1** was identified as a selective B1antagonist (K_i = 382 nM).[25] A second lead compound **3** (K_i = 520 nM) was the one methylene unit homolog of **2**.[25] Subsequent studies showed that the *R* enantiomer of **2** contained [**2**(R), K_i = 132 nM] the biological activity. Making use of the conformational restriction strategy the benzylic position on compound **2**(S) was fixed to the arene ring via a six-membered carbocycle. The resulting racemic tetraline **4** was 75-fold more potent (K_i = 6.4 nM) than the acyclic **2**, when a similar cyclization was applied to phenethylamine **3** (to afford bicycle **6**) a modest increase in potency was observed (IC_{50} = 570 nM for **4** and IC_{50} = 380 for **6**). Epimer **7**, from **4** has the optimal configuration for optimal potency [K_i = 0.24 nM]. The chromane analog **8** with similar configuration was also very potent [K_i = 0.77 nM]. The basic side chain in **4** and **6** is located in a different space region, probably the explication for the 50-fold increase in binding. To understand the reasons for the 550-fold increase in K_i (**2** to **7**), a computational study was performed to provide a quantitative measurement for the entropy component responsible for the potency of **7** or **8**. The *Shannon entropy*[26] was applied as this method has been used to characterize protein conformations[27] (Figure 17.5).

The conformational entropy of compound **7** was calculated by converting their respective energy profiles into probability functions. First the global change in $d\Delta G°$ (difference in affinity) for tetralin **7** relative to **4** is significant: 15.5 kJ/mol. From the calculations with the Shannon paradigm the respective enthalpy $d\Delta H°$ and $Td\Delta S°$ entropy differences are 11 kJ/mol and 4.5 kJ/mol from the flexible compound **4** to its rigidified analog **7**. Clearly the entropy decrease justify only 30% of the affinity gain and probably in the rigidified **7** favorable enthalpy interactions (van der Waals interactions) are accounting for the remaining 70% binding energy increase. The addition of the two contributions are accounting for the 550-fold affinity enhancement. This unexpected repercussion of the rigidification approach is in line with the cooperativity effect described by Williams arising from the enthalpy–entropy compensation.[28] Finally, compound **7** or **8**, potent selective antagonists, disclosed a credible B1 pharmacophore.

B. Allylic constraints for inducing conformational rigidity

For the design of constraint analogs, it is essential that the analogs should be as similar as possible to the parent compound in size, shape and molecular weight in order to introduce as less modification as possible to the parent structure. As a small rigid structure, the cyclopropyl ring is likely to be effective in restricting the conformation of a molecule without changing the chemical and physical properties of the lead compound. Adjacent substituents on a cyclopropane ring exert significant mutual steric repulsion, because they are fixed in eclipsed orientations. But a simple alkene can replace the cyclopropyl ring by the virtue of 1,3-allylic strain. The two examples below will highlight the concepts.

FIGURE 17.5 Design of new Bradykinin ligands by the conformational locking approach. The sulfonamide and the basic side chain (piperidine) are the major anchor functions.[25]

1. *Histamine* H_3 *ligands*

Homeostatic processes related to the neurotransmitter histamine are mediated by four identified subtypes H_{1-4}. Much attention has been focused on the H_3 receptor that is mainly present in the central nervous system. Histamine is a neurotransmitter with a flexible backbone, thus the conformational restriction may improve the specific binding to one of the receptor subtypes. The cyclopropane unit as exposed above is an ideal element to be incorporated into a flexible side chain for rigidification purposes without altering to much the remaining structure elements. This strategy has been called stereochemical diversity-oriented restriction because once the cyclopropyl unit has been incorporated, the remaining side chain can be oriented in *cis* or *trans* directions via the stereoisomerism of the cyclopropane. Using this tactic several agonists and antagonist for the H_3 or H_4 receptors in nanomolar range were disclosed such as compounds **1**, AEIC or GT-2331.[29–31] The stereochemical diversity-oriented approach can be a powerful strategy in medicinal chemistry studies, when the structure of the target protein is not available or if few structural informations of selective ligands are known (Figure 17.6).

FIGURE 17.6 Constrained histamine analogs for subtype targeting. The cyclopropyl ring orients the appended side chains in various space direction depending if the substituent is located *cis* or *trans*.

2. *Milnacipran analogs*

Milnacipran (**2**), a clinically effective antidepressant, and (1S, 2R)-1-Phenyl-2[(S)-1-aminoproppy1]-N, N-diethylcyclopropranecarboxamide (PPDC) (**3**), are exampling the allylic strain. Compound **4,** a congested analog of **2** was prepared in an attempt to amplify the *N*-methyl-d-aspartate (NMDA) component of **3**. The conformation of the aminomethyl moiety of PPDC (**3**) which is essential for the biological activity, proved to be restricted to conformer **B** by X-ray crystallography. The conformers **A** and **B** are less populated because of severe steric hindrances (Figure 17.7).

But the synthesis of **2** is rather troublesome because of presence of the contiguous three asymmetric carbons. Therefore, it was speculated that the conformational restriction might possibly occur if the cyclopropane is replaced by an ethylene group, since a cyclopropyl and an ethylenyl functions have similar structural features.[32] The adjacent *cis*-substituents on alkenyl bond mutually exert quite significant steric repulsion because of their eclipsed conformation, known as "1,3 allylic strain." Based on that concept two analogs **3** and **4** were synthesized and both were potent ligand to the MK 801/NMDA binding site in respect to milnacipran. The allylic strain in **4** was clearly observed in its X-ray structure and in the solid state, **4** is restricted to the conformation corresponding to the conformer **B′**. Therefore the use of allylic strain, a determining factor in organic synthesis for regio and/or stereocontrol can be a useful way to introduce discrete conformational constraint with significant biological effects.

FIGURE 17.7 Cyclopropyl constraints can be replaced by allylic strains. The less crowded conformers **B** and **B′** were observed in the X-ray structures of **3** and **4**.

C. Diversity-oriented synthesis

Inspired by the structure and bioactivity of natural compounds, the concept of Diversity-oriented synthesis (DOS) has emerged as a powerful tool for the discovery of new lead compounds as well as new biological functions.[33,34] DOS is attractive for medicinal chemists because in three- to five-step a huge amount of compounds can be obtained, even better a broad chemical space is populated with small molecules having skeletal diversity. Indeed, collections of products with distinct molecular skeletons are particularly effective at achieving a diverse display of chemical functionality in the 3D space. The guiding hypothesis of DOS is that synthetic compounds that embody features characteristic of bioactive natural products may prove equally suited to provide the enthalpic and/or entropic driving forces for protein binding: (i) rigidity from covalent or hydrogen bonding to reduce conformational flexibility; (ii) stereochemistry to fit into the chiral active sites of proteins and multiple hydrophilic and hydrophilic groups. The skeletal diversity can be generated by two strategies: the reagent-based approach, a *common substrate* is transformed by *different reagents* to a collection of products with distinct skeletons; the substrate-based approach, a *common reagent* transform a *collection of substrates* with identical skeletons but with different appendages that pre-encode skeletal informations. The examples below illustrated the two concepts developed by Schreiber's group. In the reagent-based approach the following sequence was developed. Starting from triene **1**, a two-fold partner in Diels–Alder reaction, the dienophiles **R1**-**R4** afforded the diverse rigidified scaffolds **2–5**, amenable to further transformations (Figure 17.8).

In the substrate-based approach, compounds **6–8** are transformed with a mild oxidant to **9–11** that exhibit indeed different molecular skeletons. The appendages (called σ elements) are responsible for the obtained diversity. In this approach, the maintenance of structural similarities until late in the synthesis is facilitating the chemical process. In the drug discovery process, applying the DOS principles that are considering the conformational constraints as a part of the diversity design, will probably facilitate the lead generation step and the emergence of new molecular entities (Figure 17.9).

D. Epibatidine bioactive conformation

Epibatidine (EPI) **1**[35] isolated from the skin of an ecuadorian frog has an improbable story from its discovery to its biological activity. EPI is a non-opioid analgesic agent with a potency 200-fold greater than that of morphine in mice via the neuronal nicotinic acetylcholine receptor (nAChR). Both natural (−) and unnatural (+) EPI possesses high and similar affinity for the nAChR. However, the respective *N*-Me analogs have different behavior. The nAChR receptor does not distinguish between (−) *N*-Me EPI and (−) EPI, the two compounds are equipotent; but the nAChR is 15-fold less sensitive to the (+) *N*-Me EPI than to (+) EPI. In regard to the paradoxical structure–activity and the adverse side effects of EPI, studies were performed to

FIGURE 17.8 The reactions of different reagents **R1**–**R4** to a common substrate produce different skeletons.[34]

FIGURE 17.9 Similar substrates with different appendages (σ-elements) **6–8** produces a diversity of skeletons **9–11**with a unique set of reagent.[34]

determine the pharmacophore of EPI in order to discover better candidates. Molecular modeling studies have suggested that EPI adopts two local energy minimum conformations **1A** and **1B** with similar energies, by rotation around the bond between the two putative pharmacophore elements, the rigid 7-azabicyclo[2.2.1]heptane system and the 2-chloropyridin. Conformer **1A** has the two nitrogen atoms on the same side, conformer **1B** has the two nitrogen atoms on opposite sides with inter-nitrogen distance of *ca.* 4.5 and 5.5 Å, respectively. Albeit the two conformers have similar energies, the difference in the N–N distance is supporting two different pharmacophoric models. Therefore, compounds have been designed and synthesized as rotationally locked analogs in which the important pharmacophoric elements are held rigidly with the shorter and longer N-N distances corresponding to those in **1A** and **1B**. Two sets of compounds syn-**2′** and anti-**2′**, spirodihydrofuranopyridine and syn-**2′** and anti-**2′**, methanopyrrolonaphthyridine have been synthezised as constrained EPI analogs (Figure 17.10).

The former were obtained by a methylen-oxo bridge between the two heterocycles and the second by linking the basic nitrogen to the pyridine ring. If the two compounds syn-**2′**[36] and anti-**2′**[37] were devoid of affinity for the nAChR, syn-**2** with the shorter N-N distance retain some affinity although modest in respect to EPI (K_i = 136 nM). The study points out that conformer **1A** of EPI constitute the best starting point for the design of further analogs of EPI.

E. Ligands for the hepatitis C virus

The hepatitis C virus (HCV) is at the origin of severe liver diseases. The current available treatment is based on a cytokine with low efficacy and many side effects. The NS5B protein plays a central role in the replication of HCV and is an attractive target for drug development. Recently, a compound with a benzimidazole structure (**1**) has been identified as inhibitor of the NS5B protein with an IC_{50} value of 3.2 μM. Chemical efforts have been engaged toward the finding of more potent analogs. The breakthrough came when a series of constraint analogs were synthesized.[38,39] The different steps in the elaboration of those compounds will be discussed below. When a cyclohexyl replaces a cyclopentyl rest in **1** and an extra aryl group is attached on the benzyl appendage the affinity was increase of one order of magnitude in compounds **2–7**. The affinity of compound **4** (IC_{50} = 0.10 μM), exhibiting a fluorine atom at C-2 in the ortho-position next to the imidazole ring, suggests that the presence of the fluorine in compound **4** does introduce a conformational restriction from the biphenyl subunit, beneficial for the activity. Then the synthesis of more restricted analogs focused at C-2 was engaged. Before the imidazole in compound **8** was converted to pyrrole ring yielding **9** with an enhanced affinity. These in between manipulation improve not only the biological activity but also simplify the chemical exploration. The former hypothesis was confirmed in comparing the activities of compounds **9** (IC_{50} = 0.12 μM) and **10** (IC_{50} = 0.026 μM) where the introduction of an extra ring: bridging the 6,5-bicyclic core and the C-2 ring led to tetracyclic scaffold (**10**). Then the effects of the size of this ring and C or O in position X (five- to eight-membered rings, **11–15**) were examined. Regardless of the atom type, the optimal ring size was found to be a seven-membered ring. Oxygen as atom X (**14**) seems to be better than a CH_2 (**12**). Next, the ring size was fixed as seven-membered and other heteroatoms were examinated. Compound **16** with a nitrogen (X = N) was equipotent to **14**. Interestingly, a benzyloxy residue was removed without damage from **10** to **11–16**. The value of the dihedral angle for the compounds **11–16** were then determined from the most stable conformations (Figure 17.11).

FIGURE 17.10 Epibatidine and analogs. The two conformers **1A** and **1B** were locked into two different scaffolds. Only syn-**2** conserved some affinity in binding experiments.

	X	IC50 μM
3	H	0.30
4	F	0.10
5	Cl	0.35
6	OMe	1.1
7	CF_3	0.93

	X	n	IC50 μM	ϕ(°)
11	CH_2	0	0.21	0
12	CH_2	2	0.077	53
13	O	1	0.14	24
14	O	2	0.026	43
15	O	3	0.047	61
16	NH	2	0.021	46

FIGURE 17.11 With a dihedral angle at 47°, compound **17** was the best ligand in the series. The relative orientation of the indolyl part and the aryl ring are crucial for binding interactions.

The biological potency roughly correlates with the measured dihedral angle, suggesting that the measured angle reflects its value in the bioactive conformation, the value is approximately 46° (**14** and **16**). The final confirmation came when **17** a close analog of compound **14** was cocrystallized with the NS5B polymerase domain of HCV, the observed dihedral angle in the complex structure was 47° in good agreement with the value predicted by modeling.

FIGURE 17.12 Rigidifying of nociceptine ligands J-1 and J-2 after computational determination of a low energy scaffold SR14136 is equipotent to J-2, a pharmacophore can be adjusted. *Source*: Adapted with modifications from Ref. [42]

The discovery of the potent HCV inhibitors **14** and **16** resulted mainly from conformational rigidification, a gain of entropy is probably responsible for the affinity increase (cpd **8** to cpd **14**) of about a factor 50 (8 kJ/mol). The identification of compound **17** follows a classical step when an ortho-effect (binding increase with an ortho-substituent) is identified; the design is conducted so as to find the optimum dihedral angle.

F. Nociceptin

Nociceptin is the endogenous ligand for the opioid receptor-like called NOP receptor. NOP is a G-protein-coupled receptor but it does not bind opiates with high affinity.[40] Several studies suggest that NOP agonists may be useful for treatment of anxiety, and antagonists may have analgesic activity. Since nociceptine is a 17-amino acid peptide, the discovery of low molecular weight ligands is highly desirable. From a series of benzimidazole-based ligands reported as potent selective NOP antagonists, a rational ligand design was undertaken from two structurally similar NOP ligands, J-1 (IC_{50} = 200 nM) and J-2 (IC_{50} = 5.9 nM) but with different binding affinities for the NOP receptor.[41,42] It is likely that the steric interactions of the two added substituents (Me and Cl in J-2 in respect of J-1) restricts the rotation of the *N*-benzyl substituent, resulting in increased affinity and selectivity. Computer structural analysis revealed that J-1 has two low energy conformers J-1A and J-1B with only 0.4 kcal/mol of energy difference whereas J-2A is a more favorable conformer, compared with J-2B, whose energy is 1.5 kcal/mol higher. Therefore the potent affinity of J-2 may largely be due to contribution of the global low-energy conformer J-2A. Therefore SR14136 was designed with the piperidine *N*-substituent is locked in a defined space with respect to the piperidine ring. The phenyl ring in SR14136 has an equatorial location and the low energy conformers J-1A and J-2A. The *N*-α-methylbenzyl substituent was locked into a quinolizidine ring system such that the phenyl group of the *N*-substituent now occupies a spatial position resembling that in J-1A and J-2B. SR14136 is a full agonist for the NOP receptor with an IC_{50} value of 5.7 nmol comparable to J-2. Finally, the conformational rigidification did not improve the affinity but provides a better understanding of the NOP receptor binding pockets. In the present case starting from a compound with high affinity, the rigidification strategy shows limitations, but in compensation a new scaffold is disclosed and a pharmacophore of the series can be disclosed (Figure 17.12).

G. Opioid receptors ligands

In 1978, Zimmermann disclosed opioid antagonists in a series of 4-(3-hydroxyphenyl) piperidines.[43] These compounds were structurally unique since prior to their discovery, opioid antagonists were generally analogs of morphine. The maximum potency (nanomolar range affinity for μ, κ and δ opioid receptors) was achieved with compound **3** (K_i = 1.8 nM on the μ subtype) when the *N*-substituent incorporated a lipophilic entity (phenyl ring) separated from the piperidine nitrogen by two methylene units. The fact that the whole phenethyl side chain of **3** can freely rotate relative to the piperidine ring, it was not possible to pinpoint the exact location of this binding site relative to the rest of the molecule. A chemical effort was accomplished for the synthesis of more rigid analogs of **3** by linking the piperidine 2- or 6-position to the benzylic position of the *N*-phenethyl moiety by an ethylene chain linker. This modification yielded a novel series of fused bicyclic structures, the octahydroquinolizidine derivatives **4–11**[44] (Figure 17.13).

Two of the eight constrained analogs compounds **6** and **9** showed the highest affinities. Interestingly, **6** is a potent antagonist (K_i = 0.6 nM) and **9** was an agonist (K_i = 0.9 nM) against the μ-opioid receptor. A conformational search on flexible compound **3** and the best rigidified

FIGURE 17.13 The rigidification of compound **3** produces a couple of diastereomers and enantiomers that were separated. The study provide decisive information for the location of the 3-hydroxyphenyl ring relative to the opiate receptors, equatorial orientation corresponds to the pharmacophore for agonists and equatorial orientation to the pharmacophore for antagonists.

analogs **6** and **9** gave the following results. Compound **3** in its overall low energy conformation has the hydroxyphenyl moiety in equatorial orientation, indeed if the axial orientation is imposed the energy penalty is 12 kJ/mol. For compound **6** the penalty is 36 kJ/mol for the same conformational analysis. Therefore, the low-energy conformation of **3** could be regarded as the bioactive conformation. The high affinity of the rigid octahydroquinolizine **6** supports this assertion. Compounds **3** and **6** superimpose perfectly in their low-energy conformers and have an equatorial orientation for the hydrophenyl residue. In contrast, compound **9** achieves the lowest energy conformer with the hydrophenyl in axial position, in equatorial a penalty credit of 16 kJ/mol falls on it. Finally, the conformational restriction tactic applied to phenylpiperidine **3** culminated in the proposal that the differences of functional activity (agonist versus antagonist) relied on different spatial locations of the hydrophenyl moiety (axial versus equatorial).

H. Peptidomimetics for SH2 domains

1. *Rigidification by metathesis*

Restriction of conformational flexibility is an important consideration for peptidomimetic design. Global constraint may be obtained by backbone cyclization either in a head-to-tail fashion or through amino acid residue side chains. A modification of this latter approach can be achieved using the Ru-catalyzed ring-closing metathesis. However, one limitation of the use of this method is that the resulting macrocycle frequently lacks side-chain functionality at the sites of ring closure, originally present in the open-chain parent compounds. If the side chains are incorporated into ring-closing segments this type of loss can be avoided as demonstrated in the synthesis of a Grb2 SH2 domain inhibitor (kinase traffic). Rigid peptide **2**,[45] with a *trans* double bond was a perfect mimic of **1**[46] lacking only one amino functionality. The use of a metathesis reaction

FIGURE 17.14 Compound **2**, obtained by metathesis, is an excellent mimic of **1**.

FIGURE 17.15 A pseudo peptide is obtained if a cyclopropyl unit is embedded in a peptide, the amide bound is lost but the side chains maintain the spatial disposition of the side chain as in the native peptide.

between two terminal alkenes was the key step for the successful synthesis of **2**. But in the absence of biological data for **2**, the pertinence of the rigidification was not settled[45] (Figure 17.14).

2. *Cyclopropyl peptides*

The thermodynamic and structural consequences of the introduction of selected conformational constraints into biological relevant peptides have been illustrated by a body of work by the Martin's group. A novel class of peptide mimic was invented wherein a cyclopropane ring served as rigid replacement for the Cα and Cβ carbon atoms and the NH group of an amino acid residue **1**. The *trans* relationship of the substituents at C(1) and C(3) in **2** was predicted to locally stabilize a β-strand. The R^1 group at C(2) in **2** is oriented so to it occupies the same region of space relative to the backbone as in **1** (Figure 17.15).

Although the cyclization of **1** to **2** removes an amide N–H and potential hydrogen bonding capability from the resultant pseudopeptide, the favorable energetic advantage that arises from restricting at least two rotors should approximately offset the energetic cost of losing an intermolecular hydrogen bond (Figure 17.16).

With an appropriate chemical tactic, compound **4** was prepared as a constraint analog of the potent renin inhibitor **3**, and it was discovered that **4** was equipotent in the subnanomolar range. This finding was supporting that in the 3D orientation the functional groups in **3** and **4** are matching reasonably. But as pointed out by the authors no bonus was registered after the pre-organization introduced by the cyclopropyl ring. Another biological system was investigated: the interactions with SH2 domains. Structural studies of complexes of Src SH2 domains with phosphopeptides related to pTyr-Glu-Glu-Ile **5** (pYEEI) reveal that these antagonists bind to the Src2 SH2 domain in extended conformations in which two binding pockets accommodate the pTyr residue and the hydrophobic residue (e.g. Ile….) at the pY + 3 position. On the other hand, the phosphotyrosine peptide having sequences related to pTyr-Val-Asn **6** (pYVN) binds in turned conformations to the Grb2 SH2 domain. Despite the different binding modes of phosphotyrosine-derived peptides to the Src and Grb2 SH2 domains the pY moieties invariably bind with a high degree of similarity. Therefore compounds **7** and **9** could serve as good pre-organized analogs of peptides **5** and **6**. Furthermore, compounds **8** and **10** would serve then as appropriate flexible controls. The thermodynamic parameters (ΔG, ΔH, ΔS) for forming complexes of **7** (rigid) and **8** (flexible) with the Src SH2 domain were determined by isothermal titration microcalorimetry (ITC). Consistent with the expectations, an entropic advantage of 36 J/mol/K arose from pre-organizing (**7** versus **8**). However, the favorable entropy of binding affinities for the constrained ligand was offset by an enthalpic penalty that resulted in approximately equal binding affinities for **7** and **8**. In this case, the entropy advantage was balanced by an enthalpy penalty. The same ITC study was performed with **9** and **10** on the Grb2 SH2 domain. Both compounds have low affinities in the micromolar range. The constrained ligand **9** was two-fold better than its flexible counterpart **10**, but it is the enthalpy gain which increased the affinity and not

FIGURE 17.16 Pseudo peptide **4** has a comparable affinity as the renin ligand **3**.

FIGURE 17.17 Compounds **8** and **10** with their constrained analogs **7** and **9**, potent ligands for SH2 domains.

the entropy. A complete unexpected result, regarding the conventional wisdom about the energetic effect associated with pre-organization. To explain these results, the flexibility of the binding pocket is advocated. The number of closer polar contacts is increasing from **9** to **10** and the enthalpy component is increasing in the free energy whereas the entropy of the complexes is decreasing owing the adaptability of the protein to the ligand. Finally, from the above results, it must be concluded that the measure of the thermodynamic parameters (not always possible) are solely able to answer the question if a protein–ligand interaction is enthalpy- or entropy-driven. It can happen that pre-organization (rigidification) can be entropically unfavorable[47–50] (Figure 17.17).

III. SUMMARY AND OUTLOOK

In biological processes, the non-covalent bonding is orchestrated by complex thermodynamic rules. If a lead compound has been identified, one strategy is to introduce structural elements that reduce the conformational flexibility (if any) for pre-organizing the ligand in an ideal conformation best recognized by the bioreceptor. As a consequence the energy penalty (entropic factor) associated with the binding can be diminished in respect to the flexible parent compound. This energy gain is then possibly recovered by increasing the value of the affinity constant. This popular tactic in medicinal chemistry based on thermodynamic considerations remains in some extent empirical, but has met successes, mainly for the elaboration of working or preliminary pharmacophores. Furthermore, selected structural modifications of the ligand frames may improve both potency and selectivity. But the rigidification of flexible ligands is challenging, sophisticated chemistry is needed to achieve the syntheses of constrained compounds with complex architectures. Finally, it is expected in a near future that the extensive use of ITC will allow, next to binding constant, the measurements of the thermodynamic elemental parameters ($\Delta H°$, $\Delta S°$…) for a better design of constrained ligands.

REFERENCES

1. Jencks, W. On the attribution and additivity of binding energies. *Proc. Natl. Acad. Sci. USA* 1981, *78*, 4046–4050.
2. Raffa, R. B., Porreca, F. Thermodynamic analysis of the drug-receptor interaction. *Life Sci.* 1989, *44*, 245–258.
3. Andrews, P. R., Craik, D. J., Martin, J. L. Functional group contributions to drug-receptor interactions. *J. Med. Chem.* 1984, *27*, 1657–1684.
4. Page, M. I. Entropie, Bindungsenergie und enzymatische Katalyse. *Angew. Chem.* 1977, *89*, 456–467.
5. Bardi, S. J., Luque, I., Freire, E. Structure-based thermodynamic analysis of HIV-1 protease inhibitors. *Biochemistry* 1997, *36*, 6588–6596.
6. Velazquez-Campoy, A., Todd, M. J., Freire, E. HIV-1 Protease inhibitors: enthalpic versus entropic optimization of the binding affinity. *Biochemistry* 2000, *39*, 2201–2207.
7. Searle, M. S., Williams, D. H. The cost of conformational order: entropy changes in molecular associations. *J. Am. Chem. Soc.* 1992, *114*, 10690–10697.
8. Williams, D. H., Westwell, M. S. Aspects of weak interactions. *Chem. Soc. Rev.* 1998, *27*, 57–63.
9. Williams, D. H., Stephens, E., O'Brien, D. P., Zhou, M. Understanding noncovalent interactions: ligand binding energy and catalytic efficiency from ligand-induced reductions in motion within receptors and enzymes. *Angew. Chem. Int. Ed.* 2004, *43*, 6596–6616.
10. Davis, A. M., Teague, S. J. Hydrogen bonding hydrophobic interactions, and failure of the rigid receptor hypothesis. *Angew. Chem. Int. Ed.* 1999, *38*, 736–749.
11. Böhm, H. J. The development of a simple empirical scoring function to estimate the binding constant for a protein–ligand complex of known three-dimensional structure. *J. Comput.-Aided Mol. Des.* 1994, *8*, 243–256.
12. Gilli, P., Ferretti, V., Gilli, G. Enthalpy–entropy compensation in drug-receptor binding. *J. Phys. Chem.* 1994, *98*, 1515–1518.
13. Velazquez-Campoy, A., Luque, I., Freire, E. The application of thermodynamic methods in drug design. *Thermochim. Acta* 2001, *380*, 217–227.
14. Weiland, G. A., Minnemann, K. P., Molinoff, P. B. Fundamental difference between the molecular interactions of agonists and antagonists with the β-adrenergic receptor. *Nature* 1979, *281*, 114–117.
15. Miklavc, A., Kocjan, D., Mavri, J., Koller, J., Hadzi, D. On the fundamental difference in the thermodynamics of agonist and antagonist interactions with β-adrenergic receptors and the mechanism of entropy driven binding. *Biochem. Pharmacol.* 1990, *40*, 663–669.
16. Borea, P. A., Dalpiaz, K., Varani, K., Guerra, G., Gilli, G. Binding thermodynamics of adenosine A2a receptor ligands. *Biochem. Pharmacol.* 1995, *49*, 461–472.
17. Gilli, P., Gilli, G., Borea, P. A., Varani, K., Scatturin, A., Dalpiaz, A. Binding thermodynamics as a tool to investigate the mechanisms of drug-receptor interactions: thermodynamics of cytoplasmic steroid/nuclear receptors in comparison with membrane receptors. *J. Med. Chem.* 2005, *48*, 2026–2035. and references therein.
18. Hruby, V. J., Li, G., Hacksell-Luevano, C., Shenderovich, M. Design of peptides, proteins and peptidomimetics in chi space. *Biopolymers* 1997, *43*, 219–266. and references cited therein.
19. Leach, A. R., Lewis, R. A. A racing–bracing approach to computer-assisted ligand design. *J. Comp. Chem.* 1994, *15*, 233–240.
20. Nicklaus, M. C., Wang, S., Driscoll, J. S., Milne, G. W. A. Conformational changes of small molecules binding to proteins. *Bioorg. Med. Chem.* 1995, *3*, 411–428.
21. Charton, M., Motoc, I. Steric effects in drug design. *Top. Curr. Chem.* 1983, *114*, 1–6.
22. Veber, D. F., Johnson, S. R., Cheng, H.-Y., Smith, B. R., Ward, K. W., Kopple, K. D. Molecular properties the influence the oral bioavailability of drug candidates. *J. Med. Chem.* 2002, *45*, 2615–2623.
23. Lipinski, C. A., Lombardo, F., Dominy, B. W., Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 1997, *23*, 4–25.
24. Gougat, J., Ferrari, B., Sarran, L., Planchenault, C., Poncelet, M., Maruani, J., Alonso, R., Cudennec, A., Tiziano Croci, T., Fabio Guagnini, F., Katalin Urban-Szabo, K., Martinolle, J. P., Soubrié, P., Finance, O., Le Fur, G. SSR240612 [(2*R*)-2-[((3*R*)-3-(1,3-Benzodioxol-5-yl)-3-{[(6-methoxy-2-naphthyl)sulfonyl]amino}propanoyl)amino]-3-(4-{[2*R*,6*S*)-2,6-dimethylpiperidinyl]methyl}phenyl)-*N*-isopropyl-*N*-methylpropanamide Hydrochloride], a new nonpeptide antagonist of the bradykinin B1 receptor: biochemical and pharmacological characterization. *J. Pharmacol. Exp. Ther.* 2004, *309*, 661–669.
25. D'Amico, D. C., Aya, T., Human, J., Fotsch, C., Chen, J. J., Biswas, K., Riahi, B., Norman, M. H., Willoughby, C. A., Hungate, R., Reider, P. J., Biddlecome, G., Lester-Zeiner, D., Van Staden, C., Johnson, E., Kamassah, A., Arik, L., Wang, J., Viswanadham, V. N., Groneberg, R. D., Zhan, J., Suzuki, H., Toro, A., Mareska, D. A., Clarke, D. E., Harvey, D. M., Burgess, L. E., Laird, E. R., Askew, B., Ng, G. Identification of a nonpeptidic and conformationally restricted bradykinin B1 receptor antagonist with anti-inflammatory activity. *J. Med. Chem.* 2007, *50*, 607–610.
26. Shannon, C. E. Prediction and entropy. *Bell Sys. Tech. J.* 1951, *16*, 50–64.
27. Viswanadham, V. N., Mattice, W. L. Assessment of bond rotation interdependence in polymer chains: an information theory approach. *Macromolecules* 1987, *20*, 685–688.
28. Calderone, C. T., Williams, D. H. An enthalpic component in cooperativity: the relationship between enthalpy, entropy, and noncovalent structure in weak associations. *J. Am. Chem. Soc.* 2001, *123*, 6262–6267.
29. Watanabe, M., Kazuta, Y., Hayashi, H., Yamada, S., Matsuda, A., Shuto, D. Stereochemical diversity-oriented conformational restriction strategy. Development of potent histamine H3 and/or H4 receptor antagonists with an imidazolylcyclopropane structure. *J. Med. Chem.* 2006, *49*, 5587–5596.
30. Ali, S. M., Tedford, C. E., Gregory, R., Handley, M. K., Yates, S. L., Hirth, W. W., Phillips, J. G. Design, synthesis, and structure–activity relationships of acetylene-based histamine H3 receptor antagonists. *J. Med. Chem.* 1999, *42*, 903–909.
31. Kazuta, Y., Hirano, K., Natsume, K., Yamada, S., Kimura, R., Matsumoto, S.-i., Furuichi, K., Matsuda, A., Shuto, S. Cyclopropane-based conformational restriction of histamine. (1*S*,2*S*)-2-(2-Aminoethyl)-1-(1*H*-imidazol-4-yl)cyclopropane, a highly selective agonist for the histamine H3 receptor, having a *cis*-cyclopropane structure. *J. Med. Chem.* 2003, *46*, 1980–1988.
32. Ohmori, Y., Yamashita, A., Tsujita, R., Yamamoto, T., Taniuchi, K., Matsuda, A., Shuto, S. A method for designing conformationally restricted analogues based on allylic strain: synthesis of a novel class of noncompetitive NMDA receptor antagonists having the acrylamide structure. *J. Med. Chem.* 2003, *46*, 5326–5333.
33. Schreiber, S. L. Target-oriented and diversity-oriented organic synthesis in drug discovery. *Science* 2000, *287*, 1964–1969.
34. Burke, M. D., Schreiber, S. L. A planning strategy for diversity-oriented synthesis. *Angew. Chem. Int. Ed.* 2004, *43*, 46–58.
35. Daly, J. W., Garraffo, H. M., Spande, T. F., Decker, M. W., Sullivan, J. P., Williams, M. Alkaloids from frog skin: the discovery of epibatidine and the potential for developing novel non-opioid analgesics. *Nat. Prod. Rep.* 2000, *17*, 131–135.
36. Abe, H., Arai, Y., Aoyagi, S., Kibayashi, C. Synthesis of conformationally constrained spirodihydrofuropyridine analogues of epibatidine. *Tetrahedron Lett.* 2003, *44*, 2971–2973.
37. Brieaddy, L. E., Mascarella, S. W., Navarro, H. A., Atkinson, R. N., Damaj, M. I., Martin, B. R., Carroll, F. Y. Synthesis of bridged analogs of epibatidine 3-Chloro-5,7,8,9,9a,10-hexahydro-7,10-methanopyrrolo[1,2-*b*]-2,6-naphthyridine and 2-chloro-5,5a,6,7,8,

10-hexahydro-5,8-methanopyrrolo[2,1-*b*]-1,7-naphthyridine. *Tetrahedron Lett.* 2001, *42*, 3795–3797.

38. Hirashima, S., Suzuki, T., Ishida, T., Noji, S., Yata, S., Ando, I., Komatsu, M., Ikeda, S., Hashimoto, H. Benzimidazole derivatives bearing substituted biphenyls as hepatitis C virus NS5B RNA-dependent RNA polymerase inhibitors: structure–activity relationship studies and identification of a potent and highly selective inhibitor JTK-109. *J. Med. Chem.* 2006, *49*, 4721–4736.
39. Ikegashira, K., Oka, T., Hirashima, S., Noji, S., Yamanaka, H., Hara, Y., Adachi, T., Tsuruha, J.-I., Doi, S., Hase, Y., Noguchi, T., Ando, I., Ogura, N., Ikeda, S., Hashimoto, H. Discovery of conformationally constrained tetracyclic compounds as potent hepatitis C virus NS5B RNA polymerase inhibitors. *J. Med. Chem.* 2006, *49*, 6950–6953.
40. Meunier, J. C., Molereau, C., Toll, L., Suaudeau, C., Moisand, C., Alvinerie, P., Butour, J. L., Guillemot, J. C., Ferrara, P., Monsarrat, B., Mazarguil, H., Vassart, G., Parmentier, M., Costentin, J. Isolation and structure of the endogenous agonist of opioid receptor-like ORL1 receptor. *Nature* 1995, *377*, 532–535.
41. Kawamoto, H., Ozaki, S., Itoh, Y., Miyaji, M., Arai, S., Nakashima, H., Kato, T., Ohta, H., Iwasawa, Y. Discovery of the first potent and selective small molecule opioid receptor-like (ORL1) antagonist: 1-[(3*R*,4*R*)-1-cyclooctylmethyl-3- hydroxymethyl-4-piperidyl]-3-ethyl- 1,3-dihydro-2*H*-benzimidazol-2-one (J-113397). *J. Med. Chem.* 1999, *42*, 5061–5063.
42. Jong, L., Zaveri, N., Toll, L. The design and synthesis of a novel quinolizidine template for potent opioid and opioid receptor-like (ORL1, NOP) receptor ligands. *Bioorg. Med. Chem. Lett.* 2004, *14*, 181–185.
43. Thomas, J. B., Mascarella, S. W., Rothman, R. B., Partilla, J. S., Xu, H., McCullough, K. B., Dersch, C. M., Cantrell, B. E., Zimmerman, D. M., Carroll, F. I. Investigation of the N-substituent conformation governing potency and receptor subtype-selectivity in (+)-(3*R*,4*R*)-dimethyl-4-(3-hydroxyphenyl)- piperidine opioid antagonists. *J. Med. Chem.* 1998, *41*, 1980–1990.
44. Le Bourdonnec, B., Goodman, A. J., Michaut, M., Ye, H.-F., Graczyk, T. M., Belanger, S., Herbertz, T., Yap, G. P. A., DeHaven, R. N., Dolle, R. E. Elucidation of the bioactive conformation of the *N*-substituted *trans*-3,4-dimethyl-4-(3-hydroxyphenyl)piperidine class of -opioid receptor antagonists. *J. Med. Chem.* 2006, *49*, 7278–7289.
45. Gao, Y., Wei, C.-Q., Burke, T. R., Jr. Olefin metathesis in the design and synthesis of a globally constrained Grb2 SH2 domain inhibitor. *Org. Lett.* 2001, *3*, 1617–1620.
46. Furet, P., Gay, B., Caravatti, G., Garcia-Echeverria, C., Rahuel, J., Schoepfer, J., Fretz, H. Structure-based design and synthesis of high affinity tripeptide ligands of the Grb2-SH2 domain. *J. Med. Chem.* 1998, *41*, 3442–3449.
47. Martin, S. F. Preorganization in biological systems: Are conformational constraints worth the energy?. *Pure Appl. Chem.* 2007, *79*, 193–200.
48. Benfield, A. P., Teresk, M. G., Plake, H. R., DeLorbe, J. E., Millspaugh, L. E., Martin, S. F. Ligand preorganization may be accompanied by entropic penalties in protein–ligand interactions. *Angew. Chem. Int. Ed.* 2006, *45*, 6830–6835.
49. Martin, S. F., Dorsey, G. O., Gane, T., Hillier, M. C., Kessler, H., Baur, M., Mathae, B., Erickson, J. W., Bhat, T. N., Munshi, S., Gulnik, S. V., Topol, I. A. Cyclopropane-derived peptidomimetics. Design, synthesis, evaluation, and structure of novel HIV-1 protease inhibitors. *J. Med. Chem.* 1998, *41*, 1581–1597.
50. Martin, S. F., Austin, R. E., Oalmann, C. J., Baker, W. R., Condon, S. L., DeLara, E., Rosenberg, S. H., Spina, K. P., Stein, H. H., Cohen, J., Kleinert, H. D. 1,2,3-trisubstituted cyclopropanes as conformationally restricted peptide isosteres: application to the design and synthesis of novel renin inhibitors. *J. Med. Chem.* 1992, *35*, 1710–1721.

Chapter 18

Homo and Heterodimer Ligands: the Twin Drug Approach

Jean-Marie Contreras and Wolfgang Sippl

I. INDRODUCTION
II. HOMODIMER AND SYMMETRICAL LIGANDS
A. Symmetry in nature
B. Homodimers as receptors ligands
C. Homodimers as enzyme inhibitors
D. Homodimers as DNA ligands
E. Homodimers of pharmacological interest
III. HETERODIMER AND DUAL ACTING LIGANDS
A. Hybrid molecules as ligands of two different receptors
B. Hybrids as enzymes inhibitors
C. Hybrids acting at one receptor and one enzyme
D. Other examples of dual acting drugs
IV. BINDING MODE ANALYSIS OF IDENTICAL AND NON-IDENTICAL TWIN DRUGS
A. Identical and non-identical twin drugs interacting with two adjacent binding sites located on the same macromolecule
B. Identical twin drugs interacting with two similar binding sites located on different monomers of the same macromolecule
C. Identical and non-identical twin drugs interacting with two different binding sites located on different macromolecules
V. CONCLUSION
REFERENCES

There is in living organisms, human beings and societies a balance between these main two forces, between creative asymmetry, imagination, or revolution and cooperative symmetry gic or order; between Dionysios and Apollon

Jean-Pierre Changeux

I. INTRODUCTION

During the study of the structure-activity relationships (SARs) of a lead compound, the combination of two pharmracological entities in a single compound could be considered as a promising drug design strategy. Drugs containing two pharmacophoric groups covalently bounded are called twin drugs. Numerous terms have appeared recently in the literature such as "dual, dimeric, bivalent, hybrid, mixed or multiple" associated with the terms "ligands, inhibitors, activators, modulators or antagonists." Reviews discussing the interest of designing multiple ligands have been published during the last years.[1–10] In this chapter, we will use mainly the term of twin drug and will focus on the combination of only two (identical or non-identical) pharmacological entities (Figure 18.1). The multi-target approach (more than two pharmacophores targeted) will be discussed in another chapter (Chapter 28).

The association of two identical pharmacophoric entities will generate an "identical twin drug" which is equivalent to a homodimer derivative. A compound, where two different pharmacological entities are bounded, is called a "non-identical twin drug" or heterodimer. The first design strategy is equivalent to a duplication/dimerization process of an active compound or lead. The aim of this approach is the production of a more potent and/or more selective drug

Copyright © 2008, Elsevier Inc.
All rights reserved.

compared to the single entity. The second strategy consists of an association of two different pharmacophores. In this case, the new compound will possess both initial pharmacological activities. This approach could be an advantage when the two targeted enzymes or receptors are involved in the same disease or disorder. The heterodimer drug will produce a synergic effect by modulating simultaneously the two biological targets.

The administration of twin drugs can be favorable compared to the two separated drugs. The new entity will have its own pharmacokinetic property (absorption, distribution and metabolism) and pharmacodynamic property. These properties will be more predictable compared to the administration of two separated drugs. This aspect represents the main advantage of designing dual acting drugs in addition to the beneficial therapeutic combination of the two active principles. The twin drug must express both activities in an appropriate balance: a stoichiometric association of diazepam (2–20 mg per day) with aspirin (200–2,000 mg per day) would be nonsense. A twin drug where the two different pharmacophores are released after its administration will be considered as a prodrug of the two different entities (Chapter 38). The linker (e.g. polymethylene, polyamine) or the covalent bond (e.g. amide, ester function) present in the twin drug should resist to the metabolic process.

FIGURE 18.1 Identical and non-identical twin drugs.

Combination of identical or non-identical pharmacophores can be classified according to the connection modes used between the two entities. The combination could be achieved by the means a linker or not (single bond) or according to an overlap mode. The linker group can be a polymeric chain (usually a methylenic chain), an aromatic or an heteroaromatic ring and in some cases an non-aromatic cycle. Pharmacophores can be overlapped when a common structural motif (i.e. a ring or a chemical function) is identified in the two different drugs. Three examples of twin drugs are given in Figure 18.2. Duplication of aspirin led to the identical twin drug diaspirin where the connection mode is a non-linker mode. Tacrine, an acetylcholinesterase (AChE) inhibitor, was dimerized by the mean of a linker (polymethylenic chain) leading to *bis*-tacrine derivatives.[11,12] Salicyclic acid and paracetamol structure can be merged (overlap mode) to give the non-identical twin drug acetaminosalol.

Non-identical twin drugs are also named dual acting drugs or hybrids because of the different pharmacological responses targeted by the two pharmacophoric moieties. The design of dual acting drugs, called the symbiotic approach,[13] can be realized accordingly to two strategies (Figure 18.3). The first strategy combines two non-identical

FIGURE 18.2 Combination modes for twin drugs.

selective pharmacophores into a hybrid molecule as illustrated by the sulfonamidic derivative. The associative synthesis of a chlorobenzenesulfonamide with an indole derivative through a methylenic linker generates a dual β-blocker and diuretic agent.[14] The two pharmacological entities could be easily identified in the conjugated derivative. A biphenyl motif was used to merge an angiotensin II receptor (AT_1) antagonist and an endothelin-1 receptor (ET_A) antagonist.[15] Structural elements of initial selective ligands are still recognizable in the hybrid molecule. The second strategy starts with a lead compound found to exhibit already both activities. A rational optimization will lead to an intrinsically dual acting drug such as the histamine (H_1) and platelet activating factor (PAF) antagonist benzocycloheptapyridinylene piperidine.[16] In this case, it is impossible to attribute the structural features of the molecule responsible of each biological activity. The degree of pharmacophore overlap is correlated with the molecular size of the twin drug and the structure complexity. With a low molecular weight, it would be difficult to clearly identify the structural elements necessary for both activities.

For the last 20 years drug design strategies were driven by the traditional concept: one disease–one target–one ligand approach. Identification of a biological target responsible of a disease has led to the design of potent and selective ligands or inhibitors. But in most cases, diseases involve multiple and complex systems where more than one biological target must be modulated. Some studies[17] have shown that simultaneous and moderate inhibition or activation of several targets is more efficient than the use of selective and potent drug. During the last years, combinatorial therapy (cocktail of several drugs) has been used intensively to treat diseases, such as cancer and AIDS. Thus, the development and the use of ligands that could modulate simultaneously multiple biological targets represent a promising approach for the treatment of complex disorders. This topic will be reviewed in details in Chapter 28.

In the present chapter, examples will be reported in details to illustrate the "twin drug approach" (i.e. combination of only two pharmacophores). The design of twin drugs will be classified into two differents parts (Figure 18.4).

FIGURE 18.4 Twin drug design.

FIGURE 18.3 The symbiotic approach.

Homodimer and symmetrical ligands will be discussed in a first time. Homodimer ligands result from the dimerization of a single pharmacophoric unit whereas symmetrical drug could be obtained after an optimization process starting from an initial symmetrical active compound. The second part will focus on heterodimer ligands and dual acting drugs. Heterodimer ligands are prepared by association of two biologically active moieties for different biological targets. Dual acting drugs possess intrinsically two biological activities that could not be correlated with structural features. Such derivatives are obtained after optimization of a lead compound that possess initially both activities. The binding mode of identical and non-identical twin drugs with macromolecular structures will also be discussed in the last part. Several examples will be presented where design of twin drugs has been guided by crystallographic and molecular modeling studies.

II. HOMODIMER AND SYMMETRICAL LIGANDS

Dimerization of a biologically active molecule represents an alternative approach in the optimization process of a lead compound. Generally, the duplication of the pharmacophore leads to an equivalent or more active derivative which exhibits different selectivity profile and pharmacokinetic properties. Enzyme inhibition will be improved when using homodimer inhibitors. Antagonists and agonists of specific receptors could be transformed respectively into agonists and antagonists of another receptor subtype. Several examples will be reported and discussed below.

A. Symmetry in nature

Nature is efficient in producing compounds with a high degree of symmetry which allows to reduce information and complexity levels.[18] Natural symmetry is observed for the assembly of macromolecules (oligomers)[19] like for HIV protease, hemoglobin and insulin. The aggregation of insulin monomers to hexamers in presence of zinc affords a macromolecular complex with high degree of symmetry (C3 symmetry).[20] DNA, by means of its symmetrical double-stranded structure, determines the cell's morphology and function. These well-organized macromolecular systems constitute binding sites for smaller molecules including water and ions. Symmetrical natural compounds present generally a C2 symmetry axis (Figure 18.5) like the alkaloids lobelanine (treatment for drug addicts), sparteine (Grave's disease treatment) or isochondrodendrine, and the anticoagulant dicumarol and antispermatogenic gossypol.[21,22] Some examples of C3 symmetrical compounds are known. Valinomycin,[23] a cyclic peptide lactone antibiotic, is a highly selective K^+-carrier. It consists of a cyclic trimer containing L-valine, D-α-hydroxyisovaleric acid (Hyi), D-valine and L-lactate. The C3 symmetrical agent enterobactin is a cyclic lactonic *N*-acylated serine trimer.

B. Homodimers as receptors ligands

Identical twin drugs have shown increasing potencies and/or modified selectivity profiles as receptor ligands, when compared to their corresponding single drug. Several receptors of biogenic amines (catecholamines), quaternary ammoniums (acetylcholine, PAF-acether), and peptides (angiotensin, endothelin) belong to the well-known class of G-protein-coupled receptors (GPCR). They present one subunit with

FIGURE 18.5 Natural compounds presenting a C_2 symmetry axis.

seven transmembrane spanning domains, three extracellular and three intracellular loops, which are coupled to G-proteins. Duplication of drugs within this series of ligands has been efficient in several cases.

Piperidine-based dimer ligands were prepared and evaluated[24] for their inhibiting activity for dopamine (DA) and serotonin (5-HT) transporters (Figure 18.6). Two *trans*-piperidine units were linked by a pentamethylene spacer to

FIGURE 18.6 Duplication of monoamine receptors ligands.

FIGURE 18.7 Ligands of acetylcholine receptor.

give a homodimer ligand. The dimerization process showed an increase of activity on both DA transporter (DAT) and 5-HT transporter (SERT). Several β_1 selective adrenoreceptor antagonists have been designed by duplication of the well-known oxprenolol.[25] The phenyloxypropanolamine pharmacophore was dimerized with a methylene linker to lead to the symmetrical *bis*-oxprenolol. According to the length of the linker, selective β_2 or β_1 antagonists could be obtained.

Methoctramine, a selective M_2 antagonist, is an useful probe for characterizing muscarinic acetylcholine receptor (mAChR) subtypes. The selectivity toward M_2 receptors was improved by modifying the terminal moiety of methoctramine.[26] The replacement of the 2-methoxybenzyl group by the hydrophobic moiety of pirenzepine leads to a very potent M2 antagonist whereas pirenzepine is known as a selective M_1 antagonist (Figure 18.7). The duplication of pirenzepine pharmacophore leads to a new antagonist with a different selectivity profile. Recently, studies[27,28] have shown that the length of the polymethylene chain of methoctramine polyamine derivatives could be important to convert muscarinic antagonist into selective nicotinic antagonist. Thus, starting from a common pattern, symmetrical twin drugs acting as antagonists at different receptors, have been designed. Epibatidine, an alkaloid known for its high affinity to nicotinic acetylcholine receptors (nAChRs) was dimerized.[29] The bivalent ligands obtained were evaluated on different nAChR subtypes and showed nanomolar binding affinities. These derivatives are weak partial agonists on a specific nAChR subtype whereas epibatidine is a full agonist.

Several studies have previously reported that dimerization of peptidergic receptor ligands can result in an increase in affinity, potency, and/or metabolic resistance.[30–32] The design of non-peptidergic antagonists of bradykinin B_2 receptor, potential therapeutic agents in the treatment of inflammation and pain, led to a symmetrical *bis*-phosphonium salt[33] (Figure 18.8). The 10 Å-distance between the two positive charges is in good agreement with that found between guanidinium cations of Arg (1) and Arg (9) of bradykinin.

In the case of enkephalines, dimerization has shown better analgetic properties when compared to their monomeric counterparts (Figure 18.9). The increase in potency and the selectivity profile depends on the length of the methylenic chain.[34] Dimerization of the oxymorphamine pharmacophore led to a bivalent ligand which showed a greater potency to opioid receptors than the monomer unit. The spacer length was important for the δ receptors selectivity.[35] In another example, the design of a bridged opioid dimer led to a κ-selective opioid receptor antagonist such as norbinaltorphimine.[36,37] In this case, the two identical morphonic units were linked by the mean of a pyrrole ring.

The potent calcium channel antagonist (blocker) nitrendipine is an effective antihypertensive agent used for the treatment of cardiovascular diseases. The design of symmetrical *bis*-1,4-dihydropyridine derivatives (BDHP) (Figure 18.10) showed an increase of the binding activity of about 10 times.[38] The length of the linker seems to have no significant contribution on the level on binding activity suggesting that the second 1,4-dihydropyridine moiety does not interact with another binding site.

Melatonin is responsible for the regulation of mammalian circadian rhythms and reproductive functions. This neurohormone acts on two different melatonin receptors (MT), namely MT_1 and MT_2. These receptors are present in different parts of the body (brain, kidney, etc.) and their physiological roles are not well known. Selective MT ligands (agonists and antagonists) are useful tools for the characterization of MT. Selective MT_1 ligands were designed (Figure 18.11) starting from the non-selective agonist agomelatine.[39] Two agomelatine moieties were linked by a polymethylenic chain with variable length. A selective MT_1 ligand was obtained which behaved as an antagonist. In this case, dimerization of a non-selective MT agonist led to a selective MT_1 antagonist.

The synthesis and design of homodimeric ligands was also achieved in the search of new peroxisome proliferator-activated receptor (PPAR) agonists. PPARs belong to the superfamily of nuclear hormone receptors (steroid, thyroid and retinoid receptors) and act as transcription factors of specific genes. PPARα activators (agonists) such as fibrates are used to lower triglycerides and to moderately raise HDL-cholesterol level. Patients with type 2 diabetes are treated with PPARγ activators (glitazones) in order to

FIGURE 18.8 Bradykinin antagonists.

FIGURE 18.9 Twin drugs for opoid receptors.

FIGURE 18.10 Other symmetrical receptor ligands.

FIGURE 18.11 Melatonin receptor ligands.

decrease glucose and insulin levels. More recently, PPARδ agonists have showed an important increase of HDL-cholesterol level associated with a decrease of triglycerides. The design of combined triple agonists represent a therapeutic interest for diabetic patients. Duplication of non-selective PPAR agonists led to a PPARα, γ and δ agonist with a different profile compared to the monomer derivative.[40] The monomer was dimerized according an overlap mode by using the biphenyl moiety as the linker (Figure 18.12).

FIGURE 18.12 PPARs ligands.

C. Homodimers as enzyme inhibitors

The symmetrical arrangement of enzymes into homodimers or tetramers defines the active site of the enzyme in a highly symmetrical fashion. Thus, symmetrical inhibitors will correspond generally to the binding site of the enzyme.

HIV reverse transcriptase (HIV RT) and protease (HIV PR) are essential for the maturation and production of infectious viral particles. A combined inhibition of HIV RT and HIV PR is capable of reducing the viral load in blood patients.[41] These enzymes, which exist respectively as an heterodimer and an homodimer for HIV RT and HIV PR, are well characterized: more than 170 structures of HIV PR and its complexes with various inhibitors have been solved by protein crystallography techniques.[42,43] Thus, a dipalmitoylated derivative of 2,7-naphthalene disulfonic acid demonstrated a micromolar activity for both HIV-1 and HIV-2 RT (Figure 18.13).[44] Symmetrical nature of HIV PR was used in the search of novel anti-HIV drugs that would embody the predicted characteristic of the active site. The design of inhibitors of HIV PR has led to symmetrical compounds, which can be divided into two groups: (a) pseudosymmetrical compounds, like derivatives A 74,704[45] and L 700,417 which contain asymmetric atoms in close proximity to the inhibitor two-fold axis; (b) fully C2-symmetrical inhibitors like the cyclic urea [46] and the diol derivatives [47] (Figure 18.13).

The protein kinase C (PKC) enzyme family plays a pivotal role in the signal transduction pathways of hormones, neurotransmitters and other endogenous substances. The 11 PKC isoenzymes identified are involved in the activation of many cellular functions. PKCs possess a C-terminal catalytic region with a serine/threonine kinase function and N-terminal regulatory region. The regulatory region is subdivided into an activator binding domain C1 and a Ca^{2+} binding domain C2. Selectivity and classification of PKC isoenzyme are based on the structural and functional differences of this regulatory domain. During the search of potent activators of PKC, homodimer ligands were prepared and evaluated.[48] Dimer ligands showed a 200-fold higher affinity for PKCα than the monomer unit (Figure 18.14). The spacer optimal length was 14 carbon atoms. These bivalent ligands seem to interact with both C1a and C1b activator binding domain of PKCα or with the C1 domains of two adjacent PKCα. Dequalinium (DECA) analogs with longer and saturated linkers exhibited enhanced potency for inhibition of PKCα. The presence of a two-point contact on the enzyme by DECA analogs explain the potency and the selectivity of such compounds.[49]

Matrix metalloproteinases (MMPs) represent important targets for the development of new potential anticancer agents. MMP-2, MMP-9 and MMP-14 in particular play a significant role in metastatic tumor dispersion and angiogenesis. Potent and selective inhibitors of MMPs would be very useful in tumors study. Duplication of a monomeric MMP inhibitor was achieved (Figure 18.15) based on the observation that the Cbz aminoethyl side chain could be exposed to the solvent environment.[50] The addition of another drug entity would allow to interact with an adjacent active site or with other MMPs proteins. Two inhibitor entities were linked by the mean of a flexible spacer that could potentially interact with some enzyme regions. Whereas the new dimeric compound showed a lower activity and a similar selectivity compared to the monomer, it appeared to generate less cytotoxicity effects on cells.

Inhibitors of factor Xa (FXa), a serine protease involved in the cascade coagulation, would allow to control the coagulation process and bleeding problems. Discovery of orally

FIGURE 18.13 HIV reverse transcriptase and protease inhibitors.

FIGURE 18.14 Protein kinase inhibitors.

FIGURE 18.15 MMPs inhibitors.

FIGURE 18.16 Various enzyme inhibitors.

active non-peptidic inhibitors of FXa present a therapeutic interest. Recently, a potent and selective FXa inhibitor (Figure 18.16) was reported, the bisamidinobenzylidenecycloheptanone (BABCH).[51] This symmetrical derivative was designed from non-peptidic inhibitors and showed a potent and selective (compared to thrombin and trypsin) inhibitory activity on FXa.[52] The replacement of the cycloheptanone central scaffold of the BABCH derivative by a substituted pyridine core led to potent and more selective FXa inhibitors. Inhibitors of enzymes involved in the polyamine metabolic pathway have been designed as potential antitumor and antiparasitic agents. S-Adenosylmethionine decarboxylase (SAMDC) is a rate-limiting enzyme of polyamine biosynthesis.[53] The *bis*-guanylpyridine analog was designed from the active amidinohydrazone moiety (Figure 18.16). The symmetrical derivative was found to be a potent and selective SAMDC inhibitor.

The design of bifunctional AChE inhibitors was achieved in order to obtain potent and selective derivatives. Tacrine, an AChE inhibitor used for the treatment of Alzheimer disease (AD) patients, has been dimerized leading to the *bis*-tacrine derivative [12] (Figure 18.17). The length of the methylene chain was optimized in order to obtain a potent and selective homodimer. *Bis*-tetrahydroaminacrine

FIGURE 18.17 AChE inhibitors.

FIGURE 18.18 DNA ligands.

or bis(7)-tacrine showed a simultaneous interaction with the active site and the peripheral site[54] (allosteric site) of the enzyme resulting in an improvement of potency and selectivity.[11]

D. Homodimers as DNA ligands

The DNA molecule is the primary target of many antitumor agents. Small molecules, which bind to DNA by intercalation, require polycyclic systems in their structure for efficient binding. Because of the symmetrical arrangement of the helical double strand, symmetry is found in the structure of DNA ligands. Pentamidine and the *bis*-amidinobenzimidazoles (Figure 18.18) bind to the minor groove of DNA and show higher affinity for AT-base pairs rich regions.[55] Compounds with an even number of methylenes connecting benzimidazole rings have a higher affinity for DNA than those with an odd number of methylenes. The mono-intercalator DACA, a mixed topoisomerase I/II inhibitor with a cytotoxic activity on tumor cell lines, was dimerized using an aminopolymethylenic chain.[56] Substitution of the bis(acridine-4-carboxamides) derivatives was investigated. Analogs with substituents at 5-position showed superior potencies in a panel of cell lines compared to the monomer unit.

E. Homodimers of pharmacological interest

Treatment of malaria is less effective due to a resistance to chloroquine, the most useful antimalarial drug. Chloroquine resistance involve several mechanisms that are not completely understood. Dimerization of chloroquinoline derivatives was achieved to bypass this multidrug-resistant mechanism. Several bisquinolines such as piperaquine[57] showed activity against chloroquine-resistant malaria (Figure 18.19). In a recent study, *bis*-aminoacridine derivatives with different linker were evaluated for their antiparasitic activity.[58] The activity profile of these compounds was strongly dependent on the nature and the length of the connecting linker between the heterocyclic rings.

The search of cationic cholinergic agents has led to numerous twin drugs (Figure 18.20). The bis-quaternary ammonium salts hexamethonium and decamethonium are

FIGURE 18.19 Aminoacridine twin drugs.

FIGURE 18.20 Cholinergic twin drugs.

potent blockers in ganglia and in neuromuscular junction, respectively. Other neuromuscular blocking agents such as succinyl and sebacyl dicholines can be regarded as pure acetylcholine twin drugs.

III. HETERODIMER AND DUAL ACTING LIGANDS

Heterodimers and dual acting drugs exert their dual action on two different biological targets. These targets could be receptors (GPCRs, receptor subtypes in the same family or different receptor families), enzymes or a combination of both of them. The association of two physiological effects is aimed to obtain a synergic response in the treatment of a disease or a disorder. Hybrid molecules could result from the association of two distinct active principles (associative synthesis) or from a compound with an intrinsic dual acting profile. In the first case, the two pharmacophoric entities are linked and recognizables whereas in the second case it is often difficult to identify the chemical part of the molecule responsible of a biological activity. In this last case, the design of dual acting drugs will be often based on structural modifications of one of the two pharmocophoric entities by incorporating SAR important elements of the other pharmacophore.

A. Hybrid molecules as ligands of two different receptors

GPCRs possess physical, biochemical and structural similarities. Thus, selectivity toward biogenic amines such as noradrenaline (NA), serotonin (5-HT), dopamine (DA) and histamine (H) depends on a typical Asp interaction (conserved amino acid localized in the third transmembranar helix) and additional binding interactions. Because pharmacophores of all these ligands are similar, the control of their selectivity constitutes an important challenge for medicinal chemists. Thus, it may be of interest to synthesize hybrid drugs that bind potently to different GPCRs as an agonist, antagonist or mixed agonist/antagonist.

During the search of antipsychotic agents without side effects (i.e. extrapyramidal side effects), 5-HT_2 receptor

FIGURE 18.21 Dopaminergic and serotonergic dual acting drug.

antagonists such as ritanserin have shown a decrease of the negative symptoms of schizophrenia. So, it has been proposed that a combined administration of a 5-HT_2 antagonist and a D_2 antagonist could be efficient in the treatment of schizophrenic patients. SAR study on bridged γ-carbolines, reported for its potent affinity for 5-HT_2 and moderate affinity for D_2 receptors, was achieved and led to the design of a compound with equipotent and nanomolar affinity for both receptors [59] (Figure 18.21). The efficacious antipsychotic agent ziprasidone was also designed from a naphthylpiperazine derivative with moderate 5-HT_{2A} and D_2 antagonisms.[60,61] A typical SAR study allowed to optimize both DA and serotonine affinities with less extrapyramidal side effects.

Positive inotropic drugs such as arpromidine have been developed by combining in a same molecule histamine H_1 antagonistic and H_2 agonistic properties[62] (Figure 18.22). Association of a weak and partial H_2 agonist with a weak H_1 antagonist led to heterodimer which is a very potent H_2 agonist (100-fold increase) and H_1 antagonist 10 times more active than the monomeric pharmacophore. The H_1 antihistaminic drug loratadine presents a weak PAF antagonistic property. Taking into account the physiological importance of PAF in asthma, it was of therapeutic potential interest to antagonize by a single molecule the action of both mediators.[16] Optimization of this intrinsically dual acting drug was achieved and led to a derivative with a better PAF antagonistic activity and a similar H_1 binding level. Thromboxane A_2 (TXA_2) is also implicated in the pathophysiological conditions of asthma. Therefore efforts have been made to design TXA_2 receptor antagonists. The symbiotic approach concept led to the recent discovery of dibenzoxepin derivatives. The antiallergic and H_1 antagonist KW-4994 was reported to possess also a weak TXA_2 antagonizing activity. Successful modifications led to the design of dual TXA_2 and H_1 antagonists such as KF 15766.[63]

The octapeptide hormone angiotensin II is involved in vascular smooth muscle contraction and release of other endogenous substances. Since angiotensin II receptor subtypes (AT_1 and AT_2) are presents in various proportions in many tissues and organs, dual antagonists may constitute efficient pharmacological tools. Starting from losartan (Figure 18.23), a selective AT_1 antagonist, and PD 123317, a selective

FIGURE 18.22 Histaminergic hybrid drugs.

AT_2 antagonist, dual and non-selective AT_1 and AT_2 antagonists were designed.[64] Structures of losartan and PD 123317 were merged by using common features (in blue) and the new scaffold was then optimized by a traditional SAR study leading to BIBS 222 derivative. A potent and orally active AT_1 antagonist, L-159,093, was modified to enhance binding affinity for AT_2 receptor.[65] After optimization, a potent and balanced AT_1/AT_2 antagonist (L-159,689) was obtained.

More recently, several studies in animals have reported that simultaneous blockade of AT_1 and ET_A should be very efficient in the treatment of hypertension and cardiovascular diseases such as heart failure.[66,67] The biphenylsulfonamide BMS-193884 designed as a potent and selective ET_A antagonist share the same biphenyl framework as irbesartan, a AT_1 antagonist (Figure 18.24). By merging key structural elements from these two derivatives, Bristol-Myers Squibb scientists obtained a hybrid which was optimized to give DARA 7 derivative.[15,68] DARA 7 shown a balanced activity at AT_1 and ET_A receptors and reduced blood pressure elevations in rats.

Dual acting drugs can also result from the combination of ligands belonging to completely distinct pharmacophores. As activation of substance P (SP) and adenosine (A_1) receptors produce the same effect (e.g. hypotension and analgesia), it was of therapeutic interest to combine in a single compound these properties. A xanthine derivative, an A_1 adenosine receptor antagonist, was linked with the pentapeptide terminal part of substance P to give a conjugate (Figure 18.25) with similar affinity for both A_1 and SP receptors.[69] The coupling was achieved by the means of an amino acid. The heterodimer ligand obtained will be a useful tool for SP pathways and adenosine action study.

FIGURE 18.23 Angiotensine receptor ligands.

Peroxisome proliferator-activated receptors[70] (PPARs) are members of the superfamily of nuclear receptor that includes receptors for steroid, retinoid and thyroid hormones. The demonstration that PPARα and PPARγ were the receptors through which, respectively, the fibrate (lipid lowering activity) and the glitazone (insulin sensitization) drugs mediate their biological activity has led to the design of a new generation of dual acting drugs. Dual PPARα/γ agonists appear well suited for the treatment of hyperglycemia with prevention of cardiovascular disease in type 2 diabetes. The aryloxazole derivative (Figure 18.26) is a typical example of the combination of the fenofibric acid and the thiazolidinedione derivative.[71] Fenofibric acid, a PPARα agonist, was merged with a selective PPARγ thiazolidinone derivative; the hybrid derivative obtained was substituted with a phenyl group to increase potency. Further SAR studies led to the design of a more potent dual acting PPARα and PPARγ agonist such as muraglitazar. On the other hand, optimization of series of α-substituted phenylpropanoic acid derivatives [72] afforded dual PPARα/PPARδ agonists.

B. Hybrids as enzymes inhibitors

As observed with receptors, enzymatic systems can be subdivided into enzyme families, and each enzymatic type presents several isoforms. Thus, for pharmacological and therapeutic purposes, it may be of interest to combine in a same molecule structural characteristics for inhibition of two different isoenzymes, two enzymes belonging to the same family, or two enzymes for which inhibitors are showing pharmacophore similarities.

FIGURE 18.24 Dual angiotensin and ET_A ligands.

FIGURE 18.25 Substance P and adenosine hybrid ligand.

FIGURE 18.26 Dual PPAR agonists.

Inhibitors of cyclooxygenase (COX-2) and 5-lipoxygenase (5-LO), enzymes involved in the biosynthesis of prostaglandins and leukotrienes, are being studied as non-steroidal antiinflammatory agents with improved safety profile. Dual acting compounds with both inhibiting activities represent potential treatment for patients suffering from arthritis and other inflammatory disorders. The thiazolone CI-1004 (Figure 18.27) was identified as a equipotent dual COX/5-LO inhibitor.[73] Compared to KME-4 compound, this dual acting drug showed non-ulcerogenic water-soluble and orally active antiinflammatory properties. Recently, combination of the pharmacophores of selective COX-2 and 5-LO inhibitors was achieved.[74] The new conjugate derivative showed a potent COX-2/5-LO inhibition and a high COX-2 selectivity.

The inactivation of the endogenous opioid peptide enkephaline is one of the physiological role of neutral endopeptidase (NEP). It has been suggested that simultaneous inhibition of the related metallopeptidases angiotensin I-converting enzyme (ACE) and NEP might be advantageous in the treatment of congestive heart failure or hypertension. Thiorphan, the well-known inhibitor of NEP, has dual NEP–ACE inhibiting properties, but it is hundred times less potent as an ACE inhibitor than as NEP inhibitor (Figure 18.28). A rigid benzazepinone was designed as PheLeu mimetic and showed a potent dual NEP/ACE inhibition.[75] Efforts were achieved in this dual inhibitors series and led to the design of the bicyclic thiazepinone omapatrilat.[76] This compound showed equipotent and dual NEP/ACE inhibition and demonstrated an excellent blood pressure lowering in animals. Omapatrilat has been developed for the treatment of hypertension and congestive heart failure.

Nucleoside reverse transcriptase inhibitors (NRTIs) such as zidovudine (AZT) or zalcitabine (ddC) have proven highly efficacy when they are used in combination with HIV PR inhibitors. Because of increasing adverse side effects, research has been focused on non-nucleoside reverse transcriptase inhibitors (NNRTIs), inhibitors of a non-competitive binding site. NRTI–NNRTI combination therapy will exhibit a synergic activity and have a greater efficacy. Thus, conjugates containing both nucleoside analog component and non-nucleoside type inhibitor (Figure 18.29) were designed and showed micromolar anti-HIV activity.[77] The heterodimer derivative did not exhibit synergic effect

FIGURE 18.27 Dual COX/5-LO inhibitors.

FIGURE 18.28 Dual NEP/ACE inhibitors.

suggesting that the individual pharmacophores do not bind simultaneously.

Administration of AChE inhibitors represent again a promising approach for treating Alzheimer's disease since the correlation between this enzyme and amyloid formation has been demonstrated. A potent and selective AChE inhibitor[78] was designed by merging the two different AChE inhibitors huperzine A and tacrine (Figure 18.30). The

FIGURE 18.29 Heterodimer HIV inhibitors.

FIGURE 18.30 AChE hybrid inhibitors.

hybrid derivative showed a improved inhibiting activity with a moderate selectivity (AChE versus Butyrylcholinesterase) profile. *Bis*-interacting galanthamine ligands[79] were prepared by using different methylenic linkers with a phthalimide moiety and the centrally active inhibitor galanthamine. Crystallographic structure of AChE and biochemical studies allowed to identify clearly two binding sites on the enzyme: an active site, located at the bottom of a deep and narrow gorge, and a peripheral site, located at the opening of the gorge (for details see Section IV.A.1.). Combination of the phthalimide moiety, known to interact with the peripheral site, with galanthamine derivative (active site inhibitor) led to a conjugate that binds to both active and peripheral sites.

In addition to the deficiency of the cholinergic neurotransmission observed in AD and correlated with cognitive function disorders, AD patients also suffer from depression and anxiety. Serotonin transporters (SERT) inhibitors are used to treat these symptoms. The design of dual inhibitors (AChE and SERT) was achieved to obtain a better therapeutic effect.[80] After hybridization of rivastigmine, a marketed AChE inhibitor, and fluoxetine, a potent SERT inhibitor, a new series of dual AChE and SERT inhibitors was design (Figure 18.31).[81,82] After SAR studies, a compound ((R)-analog) with potent inhibitory effect on both enzymes was obtained.[83]

C. Hybrids acting at one receptor and one enzyme

Hybrid molecules acting simultaneously on a receptor and on an enzyme may produce potent synergistic effects. An illustration is given by the example of derivatives interfering with TXA_2, a powerful inducer of platelet aggregation

FIGURE 18.31 Dual SERT/AChE inhibitors.

FIGURE 18.32 TXA_2 hybrid drugs.

and vascular smooth muscle contraction. The inhibition of TXA_2 synthase (TxS) and the selective blockade of TXA_2 receptors (TxR) have been pursued as alternative therapeutic strategies to prevent the thrombotic action of TXA_2. Thus, TxS inhibiting and TxR antagonistic properties have been combined in a single molecule such as samixogrel[84] starting from isbogrel, a synthase inhibitor, and daltroban, a TXA_2 receptor antagonist (Figure 18.32). Because samixogrel

FIGURE 18.33 Hybrid drugs with synergic effects.

showed a moderate plasma levels after oral administration (low solubility in aqueous solution) it was then optimized into the guanidine derivative terbogrel which exhibits potent *in vivo* antithrombotic effects.[85]

Depending on the physiological hypothesis, both targets may belong to different systems. Dihydropyridines such as nifedipine are known to be calcium receptor (Ca^{2+}) antagonists. These drugs are commonly used for treatment of patients with cardiovascular diseases (hypertension, myocardial infarction). Combination of Ca^{2+} antagonism and TxS inhibition might induce an increase of the therapeutic efficacy for particular pathologies where both enhanced TXA_2 synthesis and cellular calcium overload are involved. The imidazol part of dazoxiben, a TxS inhibitor, was merged with the core structure of nifedipine[86] to led FEC 24265 (Figure 18.33). The hybrid derivative showed a relatively more favorable *in vivo* pharmacological profile. Platelet activating factor (PAF) is involved in inflammation process and related to pathologies such as ischemia, thrombosis and asthma. The association of PAF antagonist and TxS inhibitor would procure beneficial therapeutic treatment. Ridogrel, a potent TxS inhibitor, was directly linked to the PAF antagonist, E-6123.[87] The heterodimer ligand expressed dual and equipotent PAF affinity and TxS inhibiting activity and showed activity after oral administration.

Novel antihypertensive agents possessing both β-blocking and angiotensin-converting enzyme (ACE) inhibiting properties represent beneficial approach for the treatment of elevated blood pressure. Hybridization of the β-blocker pindolol and the ACE inhibitor enelapril led to the derivative BW-A575C which expresses both activities (Figure 18.34).[88] This dual β-blocker/ACE inhibitor drug offer therapeutic promise in hypertension and congestive heart failure.

D. Other examples of dual acting drugs

Combined treatment is necessary in the long-term treatment of essential hypertension. β-Blocker and diuretic properties in a same molecule would present a great interest for hypertension management. Few attempts to synthesize hybrid molecules by combining the structures of a β-adrenoreceptor antagonist and a diuretic were described (Figure 18.35). A hybrid sulfonamide was achieved by linking the β-blocker propranolol derivative with the 2-chlorobenzene sulfonamide moiety of mefruside.[14,89]

FIGURE 18.34 Dual β-blocker/ACE inhibitors.

FIGURE 18.35 Hypotensive hybrid drugs.

Dual acting antibacterial drugs were designed by linking quinolones (Figure 18.36), such as ciprofloxacin, to cefotaxime. The hybrid ligand was optimized and demonstrated potent activity against a broad spectrum of Gram-positive and Gram-negative bacteria.[90]

In the search of new and efficient antidepressants, dual acting drugs with selective serotonin (5-HT) reuptake (SSR) inhibition property and 5-HT_{1A} receptor antagonism were designed (Figure 18.37). A combination of the SSR inhibitor duloxetine and an arylpiperazine derivative, a class of derivatives known to have high affinity for 5-HT_{1A} receptor (e.g. NAN-190), led to the hybrid benzothiophene piperazine, a potential class of antidepressant with a dual mechanism of action.[91] In the same manner, blockade of terminal 5-$HT_{1B/1D}$ receptors by selective antagonists would in theory prevents the initial decrease. Thus, coadministration of SSR inhibitors and 5-$HT_{1B/1D}$ antagonists would lead to a large increase of 5-HT extracellular concentration and would be efficient in the treatment of depressive disorders. Coupling of a selective inhibitor of 5-HT reuptake (e.g. the indolylpiperidine derivative) with GR 127935, the first selective 5-$HT_{1B/1D}$ antagonist reported, allowed to obtain urea derivative showing both 5-HT reuptake inhibition and 5-$HT_{1B/1D}$ antagonism *in vitro*.[92]

IV. BINDING MODE ANALYSIS OF IDENTICAL AND NON-IDENTICAL TWIN DRUGS

Bioactive molecules that affect one target only might not always affect complex biological networks in the desired way even if they are able to perfectly interact with one target protein. Single targets in complex biological networks might have alternative routes that are sometimes different enough not to respond to the same molecule, and many networks are robust and prevent major changes in their outputs. [93,94] Therefore developing molecules acting on multiple targets has become an innovative approach in drug design.[4,17]

The huge number of protein crystal structures available today and the increasing reliability of modeling approaches have improved our understanding on the structural and energetic aspects of protein–ligand interaction. Among the 45,000 (July 2007) solved protein structures stored in the Protein Data Bank, PDB (www.pdb.org/pdb) several complexes are available where identical and non-identical twin drugs have been cocrystallized together with their target proteins. These complexes shed light on the binding mode of twin drugs and give important hints how to rationally design multivalent ligands.[3]

The term twin drug or bivalent ligand refers to the phenomenon in which a biological structure that contains

FIGURE 18.36 Antibacterial hybrid drug.

FIGURE 18.37 Antidepressant hybrid drugs.

multiple binding sites binds simultaneously to multiple ligands. This can lead to significant enhancements in binding strength for weakly binding molecules. Prominent examples of multiple ligands include, for example, binding by antibodies and lectins, and the receptor–ligand interactions involved in cell adhesion or viral attachment. The before introduced nomenclature (identical or non-identical twin drugs) is generally based on the structure of the ligands. However, classification can also be made on the basis of the target binding sites. Twin drugs, either identical or non-identical, can bind in different ways to their macromolecular target structures. Twin drug molecules which contain flexible linker regions are supposed to bind to (a) the same target protein to adjacent binding pockets; (b) to similar binding pockets located on different monomers of the same protein (i.e. a bivalent ligand is able to cross two monomers of the same target protein); or (c) to dissimilar binding pockets located on different proteins. The analysis of solved protein–ligand complexes in the PDB provides the basis for this classification. However, the question, that arises, is whether the observed binding mode represents an artefact from the crystallization process or has biological relevance.

To illustrate the above introduced classification, several examples will be presented where the development of twin drugs has been accompanied or guided by crystallographic, biophysical or molecular modeling studies.

A. Identical and non-identical twin drugs interacting with two adjacent binding sites located on the same macromolecule

In this part we discuss the binding of symmetrical and non-symmetrical twin drugs to a non-symmetrical binding site.

1. *Acetylcholinesterase inhibitors*

The three-dimensional structure of AChE from *Torpedo californica* revealed that the active site lies at the bottom of a deep and narrow gorge (20 Å) lined by the rings of several conserved aromatic amino acids.[95] At the "anionic subsite" of the active site, modeling suggested that the quaternary amino group of acetylcholine binds to the indole side chain of the conserved residue, Trp84, as was subsequently demonstrated for several complexes with AChE. The complexes of AChE with *bis*-quaternary ligands (decamethonium and BW284C51) led to the assignment of Trp279 as the major element of a second binding site, near the top of the active-site gorge, named "peripheral binding site," 14 Å away from the active site. These structural assignments were supported by a large body of biochemical studies, involving site-directed mutagenesis, which confirmed the importance of aromatic amino acid residues in AChE.[96] In an effort to improve drug potency and selectivity, the twin drug strategy was applied to the development of dual-site acting AChE inhibitors.[97,98] Several identical tacrine dimers, for example bis(5)-tacrine and bis(7)-tacrine, were synthesized and evaluated after computational studies predicted weak affinity of tacrine for the peripheral binding site residues (Trp279 and Tyr70) in conjunction with high affinity of tacrine for the catalytic anionic site (Trp84 and Phe330).[12,99] The increased inhibitory activity of the bivalent tacrine molecules relative to tacrine was attributed to dual-site binding.

The determination of the crystal structures of bis(5)-tacrine and bis(7)-tacrine (Figure 18.38) gave the structural basis for the observation that bis(7)-tacrine is an optimal inhibitor (*Torpedo californica* IC_{50} = 1.5 nM).[99] It forms favorable sandwich type stacking interactions in both with Trp84 (anionic site) and Trp279 (peripheral binding site) with minimal protein rearrangement. Bis(5)-tacrine which is less potent (IC_{50} = 28 nM), has only a one-sided favorable interaction in the peripheral site and induces a conformational change in the protein backbone near the acyl binding pocket. The structural changes to the native enzyme observed in the bis(5)-tacrine AChE complex showed that the selection of the optimal linker length has a dramatic influence on the inhibitory activity. These results underline the problem of current structure-based drug design approaches, where the protein structure is considered to be rigid and conformational changes are not taken into account when carrying out ligand docking.

Besides tacrine also other already known "anionic site" AChE inhibitors were used to develop AChE twin drugs.[100] The development of identical and non-identical twin drugs was guided by docking studies which indicated affinity of huperzine and tacrine for both the catalytic and peripheral binding sites in AChE.[97] The modeling studies also demonstrated beneficial hydrophobic effects imparted by the alkylene linker to the peripheral site ligand.[101] Further biochemical studies of AChE revealed that the peripheral binding site at the mouth of the gorge, was implicated in promoting aggregation of the beta-amyloid (Aβ) peptide responsible for the neurodegenerative process in AD.[102] This feature of AChE initiated the development of dual-site binding inhibitors, in hopes of increasing AChE inhibition potency, and protecting neurons from Aβ toxicity.[103] The reported increased affinity for AChE of bivalent (−)-galanthamine, tacrine, hupyridone, and huperzine derivatives, and of a tacrine/propidium heterodimeric ligands, supported the design of inhibitors with dual-site binding properties (Figure 18.39).[104] The (−)-bis(10)-hupyridone inhibitors shows a inhibitory activity of 2.4 nM (*Torpedo californica* AChE) which is more than 200-fold higher compared to the monomeric AChE inhibitor huperzine A.

2. *Bisubstrate inhibitors*

Another strategy employs symmetrical and asymmetrical bivalent ligands designed to bind at the cofactor and substrate

(a)

(b)

FIGURE 18.38 (a) AChE complexed with bis(7)-tacrine (left, colored orange) and bis(5)-tacrine (right, colored green). The molecular surface of the binding pocket is colored according the electrostatic potential (red = negative potential, blue = positive potential). The bis(7)-tacrine has the appropriate linker length to favorably interact with both binding sites (IC_{50} = 1.5 nM), whereas the bis(5)-tacrine can only bind in an unfavorable protein conformation (IC_{50} = 28 nM); (b) Molecular structures of the two cocrystallized AChE inhibitors.[99]

(a)

FIGURE 18.39 (a) Interaction of (−)-bis(10)-hupyridone (left side) and (−)-bis(12)-hupyridone (right side) with AChE. The two tryptophane residues of the active and peripheral binding sites are colored magenta. Hydrogen bonds between the inhibitor and the residues of the enzyme are shown as green line.

FIGURE 18.39 (Continued) (b) Molecular structures of the two cocrystallized AChE inhibitors.[104]

(a)

FIGURE 18.40 (a) Interaction of the bisubstrate inhibitor RM65 (magenta) and SAM (green) with the PRMT1 binding site. The symmetrical inhibitor blocks the cofactor and the substrate binding site of PRMT1; (b) Molecular structures of the bisubstrate PRMT1 inhibitor and SAM.[106]

sites of an enzyme, thus eliciting competitive inhibition. This type of ligands are called bisubstrate inhibitors. The bisubstrate concept has led to the development of compounds with powerful therapeutic properties. Mupirocin (pseudomonic acid-A) is a femtomolar inhibitor of bacterial isoleucyl-tRNA synthetase and is one of the most widely used topical antibiotic.[105] Other asymmetrical bisubstrate inhibitors have been discovered for GNC5-related *N*-acetyltransferases[106] and protein kinases.[107]

Jung *et al*. applied this approach to develop inhibitors for a histone modifying enzyme – the histone arginine methyltransferase PRMT1.[108] An identical PRMT1 inhibitor with cellular activity was obtained which blocks both the cofactor S-adenosylmethionine (SAM) and the substrate (histone) binding site simultaneously (Figure 18.40). The binding mode of RM65 is shown in Figure 18.40 in comparison with the cofactor SAM. Although the binding site is asymmetric, the symmetrical ligand shows favorable interactions with both substrate and cofactor binding site.

B. Identical twin drugs interacting with two similar binding sites located on different monomers of the same macromolecule

In this part we discuss the binding of symmetrical twin drugs to symmetrical binding sites located on different protein monomers.

1. *Sirtuin inhibitors*

NAD^+-dependent histone deacetylases (sirtuins) are enzymes which cleave off acetyl groups from lysine residues in histones but also other proteins.[109] Reversible acetylation level is an important factor in the regulation of the activity of such proteins. Potent selective sirtuin inhibitors are interesting tools for the investigation of the biological functions of those enzymes and may be future drugs for the treatment of cancer.[110,111] The crystal structure of the sirtuin subtype Sirt5 in complex with the identical twin drug suramin revealed that two protein monomers are linked by one molecule of suramin (Figure 18.41).[112] Sirt5, which is a monomer in solution, also was found to dimerize in solution upon suramin binding, as confirmed by size-exclusion chromatography. The monomer–monomer interface is mostly non-polar and there are no direct hydrogen bonds between the two monomers, thus the dimeric structure of Sirt5 is mainly stabilized by the bound suramin molecule itself. Both in the crystal structure as well as in solution, suramin acts as a linker resulting in dimerization of Sirt5. The simultaneous binding of one molecule of suramin at the surfaces of two monomers was also observed in the crystal structure of the suramin-myotoxin II complex. For sirtuins, however, this finding might introduce a new class of inhibitors that not only bind specifically into the active site but also function as linker molecules, thus limiting enzyme mobility and accessibility. In an effort to identify more potent and selective sirtuin inhibitors the suramin structure was modified.[113] The derived derivatives were found to be potent sirtuin inhibitors, with high activity for the Sirt1 subtype. Interestingly, not only the bivalent suramin molecule was found to block Sirt1 (297 nM), but also compound NF154 (Figure 18.41a) – which resembles one half of the suramin structure – is a potent Sirt1 inhibitor (525 nM).[113] Docking studies showed that NF154 interacts in a similar way with the sirtuin binding pocket. However, cross-linking of the two sirtuin monomers is not possible (Figure 18.41b).

2. *Glutathione S-transferase inhibitors*

Glutathione *S*-transferases (GSTs)1 catalyze the conjugation of the nucleophilic tripeptide glutathione (GSH, γ-Glu-Cys-Gly) to structurally diverse hydrophobic electrophiles. Among the electrophilic substrates for GSTs are alkylating agents used in cancer chemotherapy. GSTs are known to be overexpressed in malignant tissues suggesting that they may play a role in acquired resistance to antitumor agents[114] Therefore, the coadministration of potent, selective GST inhibitors as adjuvants to chemotherapy has emerged as a possible strategy to restore the drug sensitivity of resistant cells.[115] The geometry of the GST dimer, with its two identical active sites at opposite ends of a solvent-accessible intersubunit cleft, presents an opportunity to design symmetrical twin drugs that occupy both active sites simultaneously[116] The symmetrical inhibitors developed by Atkins *et al.* were found to bind to the protein dimer possessing two equivalent active sites in close proximity (about 20 Å apart).[117]

The available crystal structure of GST in complex with an anthrachinone sulfonate derivative (Figure 18.42) was used as a basis for the development of novel potent inhibitors. The authors used a related molecule, the Uniblue A derivative shown in Figure 18.43, to design twin drug molecules and to analyze the free energy of binding by isothermal titration calorimetry (ITC) measurements.[108] Because of the proximity of the two active sites of the GST dimer and their location at opposite ends of the solvent-accessible

(a) (b)

FIGURE 18.41 (a) Cross-linking of two Sirt5 monomers (cyan and green) by suramin (magenta). Hydrogen bonds are shown as orange line; (b) Binding mode of suramin and NF154.

FIGURE 18.41 (Continued) (c) Molecular structures of suramin and NF154.

FIGURE 18.42 Ribbon diagram of the complex with the anthrachinone sulfonate moiety and the GST dimer. The anthrachinone sulfonate moieties, as observed in the X-ray structure, are shown in capped sticks. The connection of these moieties by appropriate linkers (indicated by the red arrow) resulted in potent symmetrical GST inhibitors.[108]

intersubunit cleft, the authors reasoned that bivalent inhibitors could be designed to simultaneously occupy both binding sites. The appropriate linker length to bridge both monomers was guided by docking and molecular modeling studies. Whereas the monomer Uniblue A derivative showed an IC_{50} of 5000 nM, the twin drug is more than 100-fold more active (IC_{50} 44 nM). The ITC data supported the increased affinity of the bivalent inhibitors versus the monomeric analogs as observed in the inhibition experiments. The observed difference in ΔH between the two inhibitors is very close to the expected change for a bivalent inhibitor in which the linker does not interact significantly with the protein. The ΔH value for the twin drug molecule was found to be twice as large as for the monomeric inhibitor. On the other site a decreased entropy is observed for the bivalent inhibitor. The decrease in entropy upon binding of the bivalent inhibitor was assumed to arise from immobilization of the flexible linker, which contains 10 rotatable bonds. The restriction of rotational freedom within the linker upon binding offsets any entropic gain of limiting the translational entropy of the binding element in the bivalent compound. In conclusion, the basis for the improved affinity of the identical twin drug is significantly more favorable binding enthalpy, which is partially offset by a less favorable binding entropy.

3. *G-protein-coupled receptor ligands*

G-protein-coupled receptors (GPCRs) are membrane proteins that are characterized by a common seven helix transmembrane motif. The crucial role GPCRs play in many biological processes and the availability of selective small molecule GPCR ligands explain why GPCRs are among the most important of all target families.[118] In the last decades, increasing evidence has become available that GPCRs,

FIGURE 18.43 Molecular structures of GST inhibitors.

FIGURE 18.44 GPCR dimerization and binding of a twin drug molecule. The bivalent ligand bridges two independent GPCR binding sites.

upon activation, dimerize to its active form and subsequently produce its biological action. GPCRs assemble in the cell membrane as either homodimers or heterodimers.[119] Studies showing that GPCR heterodimerization can modify the receptor pharmacology have sparked an interest in the development of drugs that selectively target receptor heterodimers. One approach to target a pair of GPCRs has been to synthesize and use twin molecules targeting the two receptor binding sites on the homo or heterodimer simultaneously.[120] The rationale for employing the twin drug approach for GPCRs stems from the possibility that bivalent ligands may be capable of bridging independent receptor binding pockets on a receptor dimer resulting in a thermodynamically more favorable interactions than a monovalent binding of two ligands. This type of interaction is shown schematically in Figure 18.44.

A major breakthrough in the understanding of the GPCR family was achieved in 2000, when the crystal structure of bovine rhodopsin was resolved, and for the first time detailed structural insights were gained.[121] The structure of bovine rhodopsin was used by various groups for the generation of homology models of GPCRs. These models were used to guide the design of bivalent ligands addressing two GPCR binding sites. Portoghese *et al.* reported several identical twin molecules with varying linker length designed to investigate pharmacodynamic and organizational features of opioid receptors.[35,122,123] The twin drug approach has been shown to be generally applicable to other GPCRs. So far bivalent ligands for adenosine,[124] dopamine,[125] gonadotropin releasing hormone,[126] melatonine,[39] muscarine,[127] opioid, serotonin[128] and vasopressin receptors[129] have been reported.

Recently carried out molecular modeling studies on GPCR dimer models provided first information about the distance between the two individual binding pockets.[130] The evaluation of the minimal length between the two spacer attachment points revealed a minimal distance of approximately 22 Å.[131] Modeling with shorter spacers of (14 atoms) gave unfavorable highly extended conformation of the flexible linker, whereas a more favorable and extended conformation was obtained for longer spacers. The observed optimal distance is in good agreement with the data from the known bivalent ligands (examples are given in Figure 18.45).

C. Identical and non-identical twin drugs interacting with two different binding sites located on different macromolecules

1. GPCR (*heterodimer*) *ligands*

In the previous part we have discussed the binding of identical twin drugs to GPCR homodimers. However, it was recently also shown that many monomers from different GPCRs are able to interact with each other resulting in the formation of heterodimers.[118] An increasing number of studies point toward a role for GPCR heterodimerization

FIGURE 18.45 Molecular structure of identical twin drugs for GPCR homodimers.

in modulating receptor pharmacology and suggest that heterodimers could represent a functional unit. Thus, the design of selective bivalent ligands might be a promising strategy. A few examples have been published so far, where heterodimer selective ligands were found to have lesser side effects because of their greater selectivity. Portoghese *et al.* synthesized a series of bivalent ligands for opioid receptor dimers, consisting of a μ receptor agonist pharmacophore and a δ receptor antagonist pharmacophore separated by spacers of variable lengths.[35,120,121] The antinociceptive activity of the bivalent ligands was superior to that achieved by the coadministration of individual opioid receptor ligands.

The future analysis of the binding of such ligands to heterodimers and homodimers, guided by structure-based modeling on GPCR dimers, will provide insights into the molecular determinants required for selective occupancy and/or activation of heterodimers.[132]

V. CONCLUSION

Drugs combining two pharmacophores in a single molecule have been described in numerous domains of medicinal chemistry. Historically, they resulted from empirical structural modifications, but today rational design of homo and hetero ligands may involve the knowledge of the structure of the protein which contains these binding sites. Already, the literature contains several rational approaches to the discovery of identical and non-identical twin drugs. One problem which has to be further addressed is the physicochemical property of twin drug molecules. More sophisticated design strategies and computational tools will certainly be needed for that. Greater application of structure-based approaches and pharmacophore modeling will facilitate the design of molecules showing desired activities and optimal pharmacological profiles. Computer-based methods that can rapidly search for similar binding sites will help to predict off-site effects. Protein X-ray crystallography has revealed in several cases a high degree of symmetry resulting from the existence of dimeric (C-2), trimeric (C-3) or tetrameric protein assemblies. Simultaneous interaction of identical twin drugs (symmetrical binding sites) has to be associated with the increasing potency observed in many cases. The dimer could also show an improved selectivity profile (e.g. selectivity between different isoforms of an enzyme) compared to the initial monomer. However, combining in a same molecule two non-identical pharmacophores leads to a new compound, which may not bind simultaneously to each of the considered binding site. Recent findings in molecular pharmacology, molecular biology, enzymology and physiology will help to select pertinent pairs of targets involved in different pathologies. The search for a global synergic effect would be the goal in this approach.

For therapeutic purposes the earlier search of selective drugs is replaced by the design of non-selective dual drugs with tuning of their selectivity profile. The design of dual acting drugs is more complex but more challenging than the conventional design of a compound with a single activity. Today, the concept of "one target–one ligand" is still the major approach for drug development in pharmaceutical industry but the design of ligands, able to act on different targets simultaneously, represents probably an efficient alternative in the treatment of complex diseases or disorders. A recent review has reported and analyzed literature examples of multiple ligands combinations.[2] Even if successful examples have been reported in this chapter, some disadvantages may exist in the using of the twin drug approach: (i) Combining two pharmacophore components in a single molecule may lead to an inactive compound. A good knowledge of SAR data within each pharmacophore (types of interaction, steric hindrance-sensitive

regions, local hydrophilic and hydrophobic areas), and the choice of the linker (nature, position of the linkage) are critical for the success of the approach. (ii) The hybrid showed the awaited pharmacological profile, but the attempt failed because of non-predicted pharmacodynamic and toxicological problems. (iii) The balanced potency of the dual acting drugs has to be carefully evaluated. Design of agonist/antagonist hybrids has to take in consideration that drugs with antagonistic activity on receptors usually have to be given in concentrations significantly lesser than those needed for agonists (affinities in the nM and μM range respectively). In a similar manner, design of hybrids combining a receptor ligand and an enzyme inhibitor should take into account both efficacies, and particularly kinetic properties of the considered enzyme. However, the approach is workable and successful attempts have been obtained in cardiovascular domain, particularly for the treatment of hypertension, and in central nervous system research. This approach should be applied for the treatment of diseases that need restoration of the dopaminergic or cholinergic balance. In spite of this, numerous recent works dealing with the design of twin drugs acting in various systems have been reported in this chapter, and account for the increasing interest of this approach in drug design.

In a general manner, medicinal chemists should take into account the use of the twin drug approach as soon as they get a lead compound that need to be optimized. Dimerization of a lead or association of two different pharmacophores must be considered during the primary exploration of SARs as well as isosteric replacement, homology and conformational restriction are used during this process.

REFERENCES

1. Contreras, J. M., Bourguignon, J. J. Identical and non-identical twin drugs. In *The Practice of Medicinal Chemistry* (Wermuth, C. G., Ed.). Academic Press: London, 2003, pp. 251–273.
2. Morphy, R., Kay, C., Rankovic, Z. From magic bullets to designed multiple ligands. *Drug Discov. Today* 2004, *9*, 641–651.
3. Morphy, R., Rankovic, Z. Designed multiple ligands. An emerging drug discovery paradigm. *J. Med. Chem.* 2005, *48*, 6523–6543.
4. Csermely, P., Agoston, V., Pongor, S. The efficiency of multi-target drugs: the network approach might help drug design. *Trends Pharmacol. Sci.* 2005, *26*, 178–182.
5. Mencher, S. K., Wang, L. G. Promiscuous drugs compared to selective drugs (promiscuity can be a virtue). *BMC Clin. Pharmacol.* 2005, *5*, 3.
6. Sams-Dodd, F. Target-based drug discovery: is something wrong?. *Drug Discov. Today,* 2005, *10*, 139–147.
7. Frantz, S. Drug discovery: playing dirty. *Nature,* 2005, *437*, 942–943.
8. Morphy, R., Rankovic, Z. The physicochemical challenges of designing multiple ligands. *J. Med. Chem.* 2006, *49*, 4961–4970.
9. Espinoza-Fonseca, L. M. The benefits of the multi-target approach in drug design and discovery. *Bioorg. Med. Chem.* 2006, *14*, 896–897.
10. Morphy, R., Rankovic, Z. Fragments, network biology and designing multiple ligands. *Drug Discov. Today,* 2007, *12*, 156–160.
11. Pang, Y.-P., Quiram, P., Jelacic, T., Hong, F., Brimijoin, S. Highly potent, selective, and low cost *bis*-tetrahydroaminacrine inhibitors of acetylcholinesterase. *J. Biol. Chem.* 1996, *271*, 23646–23649.
12. Hu, M. K., Wu, L. J., Hsiao, G., Yen, M. H. Homodimeric tacrine congeners as acetylcholinesterase inhibitors. *J. Med. Chem.* 2002, *45*, 2277–2282.
13. Baldwin, J. J., Lumma, W. C., Lundell, G. F., Ponticello, G. S., Raab, A. W., Engelhardt, E. L., Hirschmann, R., Sweet, C. S., Scriabine, A. Symbiotic approach to drug design: antihypertensive β-adrenergic blocking agents. *J. Med. Chem.* 1979, *22*, 1284–1290.
14. Cecchetti, V., Fravolini, A., Schiaffella, F., Tabarrini, O., Bruni, G., Segre, G. *o*-Chlorobenzenesulfonamidic derivatives of (aryloxy)propanolamines as β-blocking/diuretic agents. *J. Med. Chem.* 1993, *36*, 157–161.
15. Murugesan, N., Gu, Z., Fadnis, L., Tellew, J. E., Baska, R. A., Yang, Y., Beyer, S. M., Monshizadegan, H., Dickinson, K. E., Valentine, M. T., Humphreys, W. G., Lan, S. J., Ewing, W. R., Carlson, K. E., Kowala, M. C., Zahler, R., Macor, J. E. Dual angiotensin II and endothelin A receptor antagonists: synthesis of 2′-substituted *N*-3-isoxazolyl biphenylsulfonamides with improved potency and pharmacokinetics. *J. Med. Chem.* 2005, *48*, 171–179.
16. Piwinski, J. J., Wong, J. K., Green, M. J., Ganguly, A. K., Billah, M. M., West, R. E., Jr, Kreutner, W. Dual antagonists of platelet activating factor and histamine. Identification of structural requirements for dual activity of *N*-acyl-4-(5,6-dihydro-11*H*-benzo[5,6]cyclohepta-[1,2-*b*]pyridin-11-ylene) piperidines. *J. Med. Chem.* 1991, *34*, 457–461.
17. Stephenson, V. C., Heyding, R. A., Weaver, D. F. The "promiscuous drug concept" with applications to Alzheimer's disease. *FEBS Lett.* 2005, *579*, 1338–1342.
18. Changeux, J. P. Remarks on the symmetry and cooperative properties of biological membranes. In *Symmetry and Function of Biological Systems at the Macromolecular Level* (Engström, A., Strandberg, B., Eds), Almqvist and Wiksell: Stockholm, 1969, pp. 235–256.
19. Blundell, T., Sewell, T., Turnell, B. Symmetry in the structure and organization of proteins. In *Structural Studies on Molecules of Biological Interest* (Dodson, G., Glusker, J. P., Sayre, D., Eds), Oxford University Press: London, 1981, pp. 390–403.
20. Sidddle, K. The insulin receptor. In *Receptor Subunits and Complexes* (Burgen, A., Barnard, E. A., Eds), Cambridge University Press: Cambridge, 1992, pp. 261–351.
21. Ariëns, E. J. Drug design: a general introduction. In *Drug Design* (Ariëns, E. J., Ed.), Vol. 1. Academic Press: New York, 1971, pp. 1–270.
22. Bell, M. R., Batzold, F. H., Winneker, R. C. Chemical control of fertility. In *Annual Reports in Medicinal Chemistry* (Bailey, D. M., Ed.), Vol. 21. Academic Press: San Diego, 1986, pp. 169–177.
23. Neupert-Laves, K., Dobber, M. The crystal structure of a K+ complex of valinomycin. *Helv. Chim. Acta.* 1975, *58*, 432–442.
24. Tamiz, A. P., Bandyopadhyay, B. C., Zhang, J., Flippen-Anderson, J. L., Zhang, M., Wang, C. Z., Johnson, K. M., Tella, S., Kozikowski, A. P. Pharmacological and behavioral analysis of the effects of some bivalent ligand-based monoamine reuptake inhibitors. *J. Med. Chem.* 2001, *44*, 1615–1622.
25. Kierstead, R. W., Faraone, A., Mennona, F., Mullin, J., Guthrie, R. W., Crowley, H. Beta-1-selective adrenoceptor antagonists. Blocking activity of a series of binary (aryloxy) propanolamines. *J. Med. Chem.* 1983, *26*, 1561–1569.
26. Melchiorre, C., Bolognesi, M. L., Chiarini, A., Minarini, A., Spampinato, S. Synthesis and biological activity of some methoctramine-related tetraamines bearing a 11-acetyl-5,11-dihydro-6*H*-pyrido[2,3-*b*][1,4]-benzodiazepin-6-one moiety as antimuscarinics: a second generation of highly selective M_2 muscarinic receptor antagonists. *J. Med. Chem.* 1993, *36*, 3734–3737.
27. Rosini, M., Budriesi, R., Bixel, M. G., Bolognesi, M. L., Chiarini, A., Hucho, F., Krogsgaard-Larsen, P., Mellor, I. R., Minarini, A., Tumiatti, V., Usherwood, P. N. R., Melchiorre, C. Design, synthesis, and biological evaluation of symmetrically and unsymmetrically substituted methoctramine-related polyamines as muscular nicotinic

receptor noncompetitive antagonists. *J. Med. Chem.* 1999, *42*, 5212–5223.
28. Rosini, M., Bixel, M. G., Marucci, G., Budriesi, R., Krauss, M., Bolognesi, M. L., Minarini, A., Tumiatti, V., Hucho, F., Melchiorre, C. Structure-activity relationships of methoctramine-related polyamines as muscular nicotinic receptor noncompetitive antagonists. 2. Role of polymethylene chain lengths separating amine functions and of substituents on the terminal nitrogen atoms. *J. Med. Chem.* 2002, *45*, 1860–1878.
29. Wei, Z. L., Xiao, Y., Kellar, K. J., Kozikowski, A. P. Synthesis and pharmacological characterization of bivalent ligands of epibatidine at neuronal nicotinic acetylcholine receptors. *Bioorg. Med. Chem. Lett.* 2004, *14*, 1855–1858.
30. Roth, R. A., Cassell, D. J., Morgan, D. O., Tatnell, M. A., Jones, R. H., Schüttler, A., Brandenburg, D. Effects of covalently linked insulin dimers on receptor kinase activity and receptor down regulation. *FEBS Lett.* 1984, *170*, 360–364.
31. Fauchère, J. C., Rossier, M., Capponi, A., Vallotton, M. B. Potentiation of the antagonistic effect of $ACTH_{11-24}$ on steroidogenesis by synthesis of covalent dimeric conjugates. *FEBS Lett.* 1985, *183*, 283–286.
32. Chino, N., Yoshizawa-Kumagaye, K., Noda, Y., Watanabe, T. X., Kimura, T., Sakakibara, S. Synthesis and biological properties of antiparrallel and parallel dimers of human α-human atrial natriuretic peptide. *Biochem. Biophys. Res. Comm.* 1986, *141*, 665–672.
33. Salvino, J. M., Seoane, P. R., Douty, B. D., Awad, M. A., Dolle, R. E., Houck, W. T., Faunce, D. M., Sawutz, D. G. Design of potent nonpeptide competitive antagonists of the human bradykinin B_2 receptor. *J. Med. Chem.* 1993, *36*, 2583–2584.
34. Costa, T., Wüster, M., Herz, A., Shimohigashi, Y., Chen, H.-C., Rodbard, D. Receptor binding and biological activity of bivalent enkephalins. *Biochem. Pharmacol.* 1985, *34*, 25–30.
35. Portoghese, P. S., Larson, D. L., Sayre, L. M., Yim, C. B., Ronsisvalle, G., Tam, S. W., Takemori, A. E. Opioid agonist and antagonist bivalent ligands. The relationship between spacer length and selectivity at multiple opioid receptors. *J. Med. Chem.* 1986, *29*, 1855–1861.
36. Lin, C.-E., Takemori, A. E., Portoghese, P. S. Synthesis and κ-opioid antagonist selectivity of a norbinaltorphimine congener. Identification of the address moiety required for κ-antagonist activity. *J. Med. Chem.* 1993, *36*, 2412–2415.
37. Portoghese, P. S. Bivalent ligands and the message-address concept in the design of selective opioid receptor antagonists. *Trends Pharmacol. Sci.* 1989, *10*, 230–235.
38. Joslyn, A. F., Luchowski, E., Triggle, D. J. Dimeric 1,4-dihydropyridines as calcium channel antagonists. *J. Med. Chem.* 1988, *31*, 1489–1492.
39. Descamps-Francois, C., Yous, S., Chavatte, P., Audinot, V., Bonnaud, A., Boutin, J. A., Delagrange, P., Bennejean, C., Renard, P., Lesieur, D. Design and synthesis of naphthalenic dimers as selective MT1 melatoninergic ligands. *J. Med. Chem.* 2003, *46*, 1127–1129.
40. Sauerberg, P., Bury, P. S., Mogensen, J. P., Deussen, H. J., Pettersson, I., Fleckner, J., Nehlin, J., Frederiksen, K. S., Albrektsen, T., Din, N., Svensson, L. A., Ynddal, L., Wulff, E. M., Jeppesen, L. Large dimeric ligands with favorable pharmacokinetic properties and peroxisome proliferator-activated receptor agonist activity *in vitro* and *in vivo*. *J. Med. Chem.* 2003, *46*, 4883–4894.
41. De Clerq, E. Toward improved anti-HIV chemotherapy: therapeutic strategies for intervention with HIV infections. *J. Med. Chem.* 1995, *38*, 2491–2517.
42. Appelt, K. Crystal structures of HIV-1 protease-inhibitor complexes. *Persp. Drug Discov. Des.* 1993, *1*, 23–48.
43. Bone, R., Vacca, J. P., Anderson, P. S., Holloway, M. K. X-ray crystal structure of the HIV protease complex with L-700,417, an inhibitor with pseudo C_2 symmetry. *J. Am. Chem. Soc.* 1991, *113*, 9382–9384.
44. Tan, G. T., Wickramasinghe, A., Verma, S., Singh, R., Hughes, S. H., Pezzuto, J. M., Baba, M., Mohan, P. Potential anti-AIDS naphtalenesulfonic acid derivatives. Synthesis and inhibition of HIV-1 induced cytopathogenesis and HIV-1 and HIV-2 reverse transcriptase activity. *J. Med. Chem.* 1992, *35*, 4846–4853.
45. Erickson, J. W., Neidhardt, D. J., VanDrie, J., Kempf, D. J., Wang, X. C., Norbeck, D. W., Plattner, J. J., Rittenhouse, J. W., Turon, M., Wideburg, N., Kohlbrenner, W. E., Simmer, R., Helfrich, R., Paul, D. A., Knigge, M. Design, activity, and 2.8 angström crystal structure of a C_2 symmetric inhibitor complexed to HIV-1 protease. *Science* 1990, *249*, 527–533.
46. Hulten, J., Bonham, N. M., Nillroth, U., Hansson, T., Zuccarello, G., Bouzide, A., Aqvist, J., Classon, B., Danielson, U. H., Karlen, A., Kvarnstrom, I., Samuelsson, B., Hallberg, A. Cyclic HIV-1 protease inhibitors derived from mannitol: synthesis, inhibitory potencies, and computational predictions of binding affinities. *J. Med. Chem.* 1997, *40*, 885–897.
47. Alterman, M., Bjorsne, M., Muhlman, A., Classon, B., Kvarnstrom, I., Danielsson, H., Markgren, P.-O., Nillroth, U., Unge, T., Hallberg, A., Samuelsson, B. Design and synthesis of new potent C_2-symmetric HIV-1 protease inhibitors. Use of the l-mannaric acid as a peptidomimetic scaffold. *J. Med. Chem.* 1998, *41*, 3782–3792.
48. Sridhar, J., Wei, Z. L., Nowak, I., Lewin, N. E., Ayres, J. A., Pearce, L. V., Blumberg, P. M., Kozikowski, A. P. New bivalent PKC ligands linked by a carbon spacer: enhancement in binding affinity. *J. Med. Chem.* 2003, *46*, 4196–4204.
49. Qin, D., Sullivan, R., Berkowitz, W. F., Bittman, R., Rotenberg, S. A. Inhibition of protein kinase Cα by dequalinium analogues: dependence on the linker length and geometry. *J. Med. Chem.* 2003, *43*, 1413–1417.
50. Rossello, A., Nuti, E., Catalani, M. P., Carelli, P., Orlandini, E., Rapposelli, S., Tuccinardi, T., Atkinson, S. J., Murphy, G., Balsamo, A. A new development of matrix metalloproteinase inhibitors: twin hydroxamic acids as potent inhibitors of MMPs. *Bioorg. Med. Chem. Lett.* 2005, *15*, 2311–2314.
51. Shaw, K. J., Guilford, W. J., Dallas, J. L., Koovakkaat, S. K., McCarrick, M. A., Liang, A., Light, D. R., Morrissey, M. M. (Z,Z)-2,7-Bis(4-amidinobenzylidene)cycloheptan-1-one: identification of a highly active inhibitor of blood coagulation factor Xa. *J. Med. Chem.* 1998, *41*, 3551–3556.
52. Phillips, G., Davey, D. D., Eagen, K. A., Koovakkat, S. K., Liang, A., Ng, H. P., Pinkerton, M., Trinh, L., Whitlow, M., Beatty, A. M., Morrissey, M. M. Design, synthesis, and activity of 2,6-diphenoxypyridine-derived factor Xa inhibitors. *J. Med. Chem.* 1999, *42*, 1749–1756.
53. Stanek, J., Caravatti, G., Capraro, H.-G., Furet, P., Mett, H., Schneider, P., Regenass, U. *S*-Adenosylmethionine decarboxylase inhibitors: new aryl and heteroaryl analogues of methylglyoxal bis(guanylhydrazone). *J. Med. Chem.* 1993, *36*, 46–54.
54. Taylor, P., Lappi, S. Interaction of fluorescence probes with acetylcholinesterase. The site and specificity of propidium binding. *Biochemistry* 1975, *14*, 1989–1997.
55. Fairley, T. A., Tidwell, R. R., Donkor, I., Naiman, N. A., Ohemeng, K. A., Lombardy, R. J., Bentley, J. A., Cory, M. Structure, DNA minor groove binding, and base pair specificity of alkyl- and aryl-linked bis(amidinobenzimidazoles) and bis(amidinoindoles). *J. Med. Chem.* 1993, *36*, 1746–1753.
56. Gamage, S. A., Spicer, J. A., Atwell, G. J., Finlay, G. J., Baguley, B. C., Denny, W. A. Structure-activity relationships for substituted bis(acridine-4-carboxamides): a new class of anticancer agents. *J. Med. Chem.* 1999, *42*, 2383–2393.
57. Vennerström, J. L., Ellis, W. Y., Ager, A. L., Jr, Anderson, S. L., Gerena, L., Milhous, W. K. Bisquinolines. 1. *N,N*-bis(7-Chloroquinolin-4-yl) alkanediamines with potential against chloroquine-resistant malaria. *J. Med. Chem.* 1992, *35*, 2129–2134.

58. Girault, S., Grellier, P., Berecibar, A., Maes, L., Mouray, E., Lemière, P., Debreu, M.-A., Davioud-Charvet, E., Sergheraert, C. Antimalarial, antitrypanosomal, and antileishmanial activities and cytotoxiciy of bis(9-amino-6-chloro-2-methoxyacridines): influence of the linker. *J. Med. Chem.* 2000, *43*, 2646–2654.
59. Mewshaw, R. E., Silverman, L. S., Mathew, R. M., Kaiser, C., Sherrill, R. G., Cheng, M., Tiffany, C. W., Karbon, E. W., Bailey, M. A., Borosky, S. A., Ferkany, J. W., Abreu, M. E. Bridged γ-carbolines and derivatives possessing selective and combined affinity for 5-HT_2 and D_2 receptors. *J. Med. Chem.* 1993, *36*, 1488–1495.
60. LoweIII, J. A., Seeger, T. F., Nagel, A. A., Howard, H. R., Seymour, P. A., Heym, J. H., Ewing, F. E., Newman, M. E., Schmidt, A. W., Furman, J. S. *et al* 1-Naphthylpiperazine derivatives as potential atypical antipsychotic agents. *J. Med. Chem.* 1991, *34*, 1860–1866.
61. Howard, H. R., LoweIII, J. A., Seeger, T. F., Seymour, P. A., Zorn, S. H., Maloney, P. R., Ewing, F. E., Newman, M. E., Schmidt, A. W., Furman, J. S., Robinson, G. L., Jackson, E., Johnson, C., Morrone, J. 3-Benzisothiazolylpiperazine derivatives as potential atypical antipsychotic agents. *J. Med. Chem.* 1996, *39*, 143–148.
62. Buschauer, A. Synthesis and *in vitro* pharmacology of arpromidine and related phenyl(pyridylalkyl)guanidines, a potential new class of positive inotropic drugs. *J. Med. Chem.* 1989, *32*, 1963–1970.
63. Ohshima, E., Takami, H., Harakawa, H., Sato, H., Obase, H., Miki, I., Ishii, A., Ishii, H., Sasaki, Y., Ohmori, K., Karasawa, A., Kubo, K. Dibenz[*b,e*]oxepin derivatives: novel antiallergic agents possessing thromboxane A_2 and histamine H_1 dual antagonizing activity 1. *J. Med. Chem.* 1993, *36*, 417–420.
64. Zhang, J. C., Entzeroth, M., Wienen, W. Characterization of BIBS 39 and BIBS 222: two new non peptide angiotensine II receptor antagonists. *Eur. J. Pharmacol.* 1992, *218*, 35–41.
65. de Laszlo, S. E., Quagliato, C. S., Greenlee, W. J., Patchett, A. A., Chang, R. S., Lotti, V. J., Chen, T. B., Scheck, S. A., Faust, K. A., Kivlighn, S. S. *et al* A potent, orally active balanced affinity angiotensin II AT1 antagonist and AT2 binding inhibitor. *J. Med. Chem.* 1993, *36*, 3207–3210.
66. Ikeda, T., Ohta, H., Okada, M., Kawai, N., Nakao, R., Siegl, P. K., Kobayashi, T., Miyauchi, T., Nishikibe, M. Antihypertensive effects of a mixed endothelin-A- and -B-receptor antagonist, J-104132, were augmented in the presence of an AT1-receptor antagonist, MK-954. *J. Cardiovasc. Pharmacol.* 2000, *36*, S337–S341.
67. Bohlender, J., Gerbaulet, S., Kramer, J., Gross, M., Kirchengast, M., Dietz, R. Synergistic effects of AT(1) and ET(A) receptor blockade in a transgenic, angiotensin II-dependent, rat model. *Hypertension*, 2000, *35*, 992–997.
68. Murugesan, N., Tellew, J. E., Gu, Z., Kunst, B. L., Fadnis, L., Cornelius, L. A., Baska, R. A., Yang, Y., Beyer, S. M., Monshizadegan, H., Dickinson, K. E., Panchal, B., Valentine, M. T., Chong, S., Morrison, R. A., Carlson, K. E., Powell, J. R., Moreland, S., Barrish, J. C., Kowala, M. C., Macor, J. E. Discovery of *N*-isoxazolyl biphenylsulfonamides as potent dual angiotensin II and endothelin A receptor antagonists. *J. Med. Chem.* 2002, *45*, 3829–3835.
69. Jacobson, K. A., Lipkowski, A. W., Moody, T. W., Padgett, W., Pijl, E., Kirk, K. L., Daly, J. W. Binary drugs: conjugates of purines and a peptide that bind to both adenosine and substance P. *J. Med. Chem.* 1987, *30*, 1529–1532.
70. Willson, T. M., Brown, P. J., Sternbach, D. D., Henke, B. R. The PPARs: from orphan receptors to drug discovery. *J. Med. Chem.* 2000, *43*, 527–550.
71. Brooks, D. A., Etgen, G. J., Rito, C. J., Shuker, A. J., Dominianni, S. J., Warshawsky, A. M., Ardecky, R., Paterniti, J. R., Typhonas, J., Karanewsky, D. S., Kauffman, R. F., Broderick, C. L., Oldham, B. A., Montrose-Rafizadeh, C., Winneroski, L. L., Faul, M. M., McCarthy, J. R. Design and synthesis of 2-methyl-2-(4-[2-(5-methyl-2-aryloxazol-4-yl)ethoxy]phenoxy)propionic acids: a new class of dual PPARα/γ agonists. *J. Med. Chem.* 2001, *44*, 2061–2064.
72. Kasuga, J., Yamasaki, D., Araya, Y., Nakagawa, A., Makishima, M., Doi, T., Hashimoto, Y., Miyachi, H. Design, synthesis, and evaluation of a novel series of alpha-substituted phenylpropanoic acid derivatives as human peroxisome proliferator-activated receptor (PPAR) alpha/delta dual agonists for the treatment of metabolic syndrome. *Bioorg. Med. Chem.* 2006, *14*, 8405–8414.
73. Unangst, P. C., Connor, D. T., Cetenko, W. A., Sorenson, R. J., Kostlan, C. R., Sircar, J. C., Wright, C. D., Schrier, D. J., Dyer, R. D. Synthesis and biological evaluation of 5-[[3,5-bis(1,1-dimethylethyl)-4-hydroxyphenyl]methylene]oxazoles, -thiazoles, and imidazoles: novel dual 5-lipoxygenase and cyclooxygenase inhibitors with antiinflammatory activity. *J. Med. Chem.* 1994, *37*, 322–328.
74. Barbey, S., Goossens, L., Taverne, T., Cornet, J., Choesmel, V., Rouaud, C., Gimeno, G., Yannic-Arnoult, S., Michaux, C., Charlier, C., Houssin, R., Henichart, J. P. Synthesis and activity of a new methoxytetrahydropyran derivative as dual cyclooxygenase-2/5-lipoxygenase inhibitor. *Bioorg. Med. Chem. Lett.* 2002, *12*, 779–782.
75. Flynn, G. A., Beight, D. W., Mehdi, S., Koehl, J. R., Giroux, E. L., French, J. F., Hake, P. W., Dage, R. C. Application of a conformationally restricted phe-leu dipeptide mimetic to the design of a combined inhibitor of angiotensin I-converting enzyme and neutral endopeptidase 24.11. *J. Med. Chem.* 1993, *36*, 2420–2423.
76. Robl, J. A., Sun, C. Q., Stevenson, J., Ryono, D. E., Simpkins, L. M., Cimarusti, M. P., Dejneka, T., Slusarchyk, W. A., Chao, S., Stratton, L., Misra, R. N., Bednarz, M. S., Asaad, M. M., Cheung, H. S., Abboa-Offei, B. E., Smith, P. L., Mathers, P. D., Fox, M., Schaeffer, T. R., Seymour, A. A., Trippodo, N. C. Dual metalloprotease inhibitors: mercaptoacetyl-based fused heterocyclic dipeptide mimetics as inhibitors of angiotensin-converting enzyme and neutral endopeptidase. *J. Med. Chem.* 1997, *40*, 1570–1577.
77. Pontikis, R., Dollé, V., Guillaumel, J., Dechaux, E., Note, R., Nguyen, C. H., Legraverend, M., Bisagni, E., Aubertin, A.-M., Grierson, D. S., Monneret, C. Synthesis and evaluation of "AZT-HEPT", "AZT-pyridinone", and "ddC-HEPT" conjugates as inhibitors of HIV reverse transcriptase. *J. Med. Chem.* 2000, *43*, 1927–1939.
78. Camps, P., El Achab, R., Gorbig, D. M., Morral, J., Munoz-Torrero, D., Badia, A., Banos, J. E., Vivas, N. M., Barril, X., Orozco, M., Luque, F. J. Synthesis, *in vitro* pharmacology, and molecular modeling of very potent tacrine-huperzine A hybrids as acetylcholinesterase inhibitors of potential interest for the treatment of Alzheimer's disease. *J. Med. Chem.* 1999, *42*, 3227–3242.
79. Mary, A., Renko, D. Z., Guillou, C., Thal, C. Potent acetylcholinesterase inhibitors: design, synthesis, and structure-activity relationships of *bis*-interacting ligands in the galanthamine series. *Bioorg. Med. Chem.* 1998, *6*, 1835–1850.
80. Van der Schyf, C. J., Geldenhuys, W. J., Youdim, M. B. Multifunctional drugs with different CNS targets for neuropsychiatric disorders. *J. Neurochem.* 2006, *99*, 1033–1048.
81. Kogen, H., Toda, N., Tago, K., Marumoto, S., Takami, K., Ori, M., Yamada, N., Koyama, K., Naruto, S., Abe, K., Yamazaki, R., Hara, T., Aoyagi, A., Abe, Y., Kaneko, T. Design and synthesis of dual inhibitors of acetylcholinesterase and serotonin transporter targeting potential agents for Alzheimer's disease. *Org. Lett.* 2002, *4*, 3359–3362.
82. Toda, N., Tago, K., Marumoto, S., Takami, K., Ori, M., Yamada, N., Koyama, K., Naruto, S., Abe, K., Yamazaki, R., Hara, T., Aoyagi, A., Abe, Y., Kaneko, T., Kogen, H. Design, synthesis and structure-activity relationships of dual inhibitors of acetylcholinesterase and serotonin transporter as potential agents for Alzheimer's disease. *Bioorg. Med. Chem.* 2003, *11*, 1935–1955.
83. Toda, N., Tago, K., Marumoto, S., Takami, K., Ori, M., Yamada, N., Koyama, K., Naruto, S., Abe, K., Yamazaki, R., Hara, T., Aoyagi, A., Abe, Y., Kaneko, T., Kogen, H. A conformational restriction approach to the development of dual inhibitors of acetylcholinesterase and serotonin transporter as potential agents for Alzheimer's disease. *Bioorg. Med. Chem.* 2003, *11*, 4389–4415.

84. Soyka, R., Heckel, A., Nickl, J., Eisert, W., Muller, T. H., Weisenberger, H. 6,6-Disubstituted hex-5-enoic acid derivatives as combined thromboxane A_2 receptor antagonists and synthetase inhibitors. *J. Med. Chem.* 1994, *37*, 26–39.
85. Soyka, R., Guth, B. D., Weisenberger, H. M., Luger, P., Muller, T. H. Guanidine derivatives as combined thromboxane A_2 receptor antagonists and synthase inhibitors. *J. Med. Chem.* 1999, *42*, 1235–1249.
86. Cozzi, P., Carganico, G., Fusar, D., Grossoni, M., Menichincheri, M., Pinciroli, V., Tonani, R., Vaghi, F., Salvati, P. Imidazol-1-yl and pyridin-3-yl derivatives of 4-phenyl-1,4-dihydropyridines combining Ca^{++} antagonism and thromboxane A_2 synthase inhibition. *J. Med. Chem.* 1993, *36*, 2964–2972.
87. Fujita, M., Seki, T., Inada, H., Shimizu, K., Takahama, A., Sano, T. Approach to dual-acting platelet activating factor (PAF) receptor antagonist/thromboxane synthase inhibitor (TxSI) based on the link of PAF antagonists and TxSIs. *Bioorg. Med. Chem. Lett.* 2002, *12*, 341–344.
88. Allan, G., Cambridge, D., Hardy, G. W., Follenfant, M. J. BW A575C a chemically novel agent with angiotensin converting enzyme inhibitor and beta-adrenoceptor-blocking properties. *Br. J. Pharmacol.* 1987, *90*, 609–615.
89. Cecchetti, V., Schiaffella, F., Tabarrini, O., Zhou, W., Fravolini, A., Goi, A., Bruni, G., Segre, G. Symbiotic approach to drug design: *N*-[(4-Chloro-3-sulfamoylbenzamido)ethyl]propanolamine derivatives as beta-adrenergic blocking agents with diuretic activity. *Eur. J. Med. Chem.* 1991, *26*, 381–386.
90. Albrecht, H. A., Beskid, G., Christenson, J. G., Deitcher, K. H., Georgopapadakou, N. H., Keith, D. D., Konzelmann, F. M., Pruess, D. L., Wei, C. C. Dual-action cephalosporins incorporating a 3′-tertiary-amine-linked quinolone. *J. Med. Chem.* 1994, *37*, 400–407.
91. Martinez-Esparza, J., Oficialdegui, A.-M., Perez-Silanes, S., Heras, B., Orus, L., Palop, J.-A., Lasheras, B., Roca, J., Mourelle, M., Bosch, A., Del Castillo, J.-C., Tordera, R., Del Rio, J., Monge, A. New 1-aryl-3-(4-arylpiperazin-1-yl)propane derivatives, with dual action at 5-HT_{1A} serotonin receptors and serotonin transporter as a new class of antidepressants. *J. Med. Chem.* 2001, *44*, 418–428.
92. Matzen, L., van Amsterdam, C., Rautenberg, W., Greiner, H. E., Harting, J., Seyfriend, C. A., Bottcher, H. 5-HT reuptake inhibitors with 5-HT1B/1D antagonistic activity: a new approach toward efficient antidepressants. *J. Med. Chem.* 2000, *43*, 1149–1157.
93. Ocampo, M. T., Chaung, W., Marenstein, D. R., Chan, M. K., Altamirano, A., Basu, A. K., Boorstein, R. J., Cunningham, R. P., Teebor, G. W. Targeted deletion of mNth1 reveals a novel DNA repair enzyme activity. *Mol. Cell. Biol.* 2002, *22*, 6111–6121.
94. Papp, B., Pal, C., Hurst, L. D. Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast. *Nature,* 2004, *429*, 661–664.
95. Sussman, J. L., Harel, M., Frolow, F., Oefner, C., Goldman, A., Toker, L., Silman, I. Atomic structure of acetylcholinesterase from *Torpedo californica*: a prototypic acetylcholine-binding protein. *Science*, 1991, *253*, 872–879.
96. Silman, I., Sussman, J. L. Acetylcholinesterase: "classical" and "non-classical" functions and pharmacology. *Curr. Opin. Pharmacol.* 2005, *5*, 293–302.
97. Du, D. M., Carlier, P. R. Development of bivalent acetylcholinesterase inhibitors as potential therapeutic drugs for Alzheimer's disease. *Curr. Pharm. Des.* 2004, *10*, 3141–3156.
98. Sippl, W., Contreras, J. M., Parrot, I., Rival, Y. M., Wermuth, C. G. Structure-based 3D QSAR and design of novel acetylcholinesterase inhibitors. *J. Comput. Aided Mol. Des.* 2001, *15*, 395–410.
99. Rydberg, E. H., Brumshtein, B., Greenblatt, H. M., Wong, D. M., Shaya, D., Williams, L. D., Carlier, P. R., Pang, Y. P., Silman, I., Sussman, J. L. Complexes of alkylene-linked tacrine dimers with *Torpedo californica* acetylcholinesterase: Binding of Bis5-tacrine produces a dramatic rearrangement in the active-site gorge. *J. Med. Chem.* 2006, *49*, 5491–5500.
100. Greenblatt, H. M., Guillou, C., Guenard, D., Argaman, A., Botti, S., Badet, B., Thal, C., Silman, I., Sussman, J. L. The complex of a bivalent derivative of galanthamine with torpedo acetylcholinesterase displays drastic deformation of the active-site gorge: implications for structure-based drug design. *J. Am. Chem. Soc.* 2004, *126*, 15405–15411.
101. Barril, X., Orozco, M., Luque, F. J. Predicting relative binding free energies of tacrine-huperzine A hybrids as inhibitors of acetylcholinesterase. *J. Med. Chem.* 1999, *42*, 5110–5119.
102. De Ferrari, G. V., Canales, M. A., Shin, I., Weiner, L. M., Silman, I., Inestrosa, N. C. A structural motif of acetylcholinesterase that promotes amyloid beta-peptide fibril formation. *Biochemistry,* 2001, *40*, 10447–10457.
103. Piazzi, L., Rampa, A., Bisi, A., Gobbi, S., Belluti, F., Cavalli, A., Bartolini, M., Andrisano, V., Valenti, P., Recanatini, M. 3-(4-[[Benzyl (methyl)amino]methyl]phenyl)-6,7-dimethoxy-2H-2-chromenone (AP2238) inhibits both acetylcholinesterase and acetylcholinesterase-induced beta-amyloid aggregation: a dual function lead for Alzheimer's disease therapy. *J. Med. Chem.* 2003, *46*, 2279–2282.
104. Wong, D. M., Greenblatt, H. M., Dvir, H., Carlier, P. R., Han, Y. F., Pang, Y. P., Silman, I., Sussman, J. L. Acetylcholinesterase complexed with bivalent ligands related to huperzine a: experimental evidence for species-dependent protein-ligand complementarity. *J. Am. Chem. Soc.* 2003, *125*, 363–373.
105. Parenti, M. A., Hatfield, S. M., Leyden, J. J. Mupirocin: a topical antibiotic with a unique structure and mechanism of action. *Clin. Pharm.* 1987, *6*, 761–770.
106. Poux, A. N., Cebrat, M., Kim, C. M., Cole, P. A., Marmorstein, R. Structure of the GCN5 histone acetyltransferase bound to a bisubstrate inhibitor. *Proc. Natl. Acad. Sci. USA* 2002, *99*, 14065–14070.
107. Parang, K., Cole, P. A. Designing bisubstrate analog inhibitors for protein kinases. *Pharmacol. Ther.* 2002, *3*, 145–157.
108. Spannhoff, A., Machmur, R., Heinke, R., Trojer, P., Bauer, I., Brosch, G., Schule, R., Hanefeld, W., Sippl, W., Jung, M. A novel arginine methyltransferase inhibitor with cellular activity. *Bioorg. Med. Chem. Lett.* 2007, *17*, 4150–4153.
109. North, B. J., Verdin, E. Sirtuins: Sir2-related NAD-dependent protein deacetylases. *Genome Biol.* 2004, *5*, 224–234.
110. Schafer, S., Jung, M. Chromatin modifications as targets for new anticancer drugs. *Arch. Pharm. (Weinheim)* 2005, *338*, 347–357.
111. Heltweg, B., Gatbonton, T., Schuler, A. D., Posakony, J., Li, H., Goehle, S., Kollipara, R., Depinho, R. A., Gu, Y., Simon, J. A., Bedalov, A. Antitumor activity of a small-molecule inhibitor of human silent information regulator 2 enzymes. *Cancer Res.* 2006, *66*, 4368–4377.
112. Schuetz, A., Min, J., Antoshenko, T., Wang, C. L., Allali-Hassani, A., Dong, A., Loppnau, P., Vedadi, M., Bochkarev, A., Sternglanz, R., Plotnikov, A. N. Structural basis of inhibition of the human NAD^+-dependent deacetylase SIRT5 by suramin. *Structure,* 2007, *15*, 377–389.
113. Trapp, J., Meier, R., Hongwiset, D., Kassack, M. U., Sippl, W., Jung, M. Structure-activity studies on suramin analogues as inhibitors of NAD(+)-dependent histone deacetylases (sirtuins). *Chem. Med. Chem.* 2007, 2(10), 1419–1431.
114. Tew, K. D. Glutathione-associated enzymes in anticancer drug resistance. *Cancer Res.* 1994, *54*, 4313–4320.
115. Lyttle, M. H., Hocker, M. D., Hui, H. C., Caldwell, C. G., Aaron, D. T., Engqvist-Goldstein, A., Flatgaard, J. E., Bauer, K. E. Isozyme-specific glutathione-S-transferase inhibitors: design and synthesis. *J. Med. Chem.* 1994, *37*, 189–194.
116. Oakley, A. J., Lo Bello, M., Ricci, G., Federici, G., Parker, M. W. Evidence for an induced-fit mechanism operating in pi class glutathione transferases. *Biochemistry,* 1998, *37*, 9912–9917.
117. Lyon, R. P., Hill, J. J., Atkins, W. M. Novel class of bivalent glutathione S-transferase inhibitors. *Biochemistry* 2003, *42*, 10418–10428.

118. Drews, J. Genomic sciences and the medicine of tomorrow. *Nat. Biotechnol.* 1996, *14*, 1516–1518.
119. Milligan, G. G protein-coupled receptor dimerization: function and ligand pharmacology. *Mol. Pharmacol.* 2004, *66*, 1–7.
120. Rozenfeld, R., Decaillot, F. M., IJzerman, A. P., Devi, L. A. Heterodimers of G protein-coupled receptors as novel and distinct drug targets. *Drug Discov. Today*, 2006, *3*, 437–443.
121. Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M., Miyano, M. Crystal structure of rhodopsin: A G protein-coupled receptor. *Science*, 2000, *289*, 739–745.
122. Daniels, D. J., Kulkarni, A., Xie, Z., Bhushan, R. G., Portoghese, P. S. A bivalent ligand (KDAN-18) containing delta-antagonist and kappa-agonist pharmacophores bridges delta2 and kappa1 opioid receptor phenotypes. *J. Med. Chem.* 2005, *48*, 1713–1716.
123. Bhushan, R. G., Sharma, S. K., Xie, Z., Daniels, D. J., Portoghese, P. S. A bivalent ligand (KDN-21) reveals spinal delta and kappa opioid receptors are organized as heterodimers that give rise to delta(1) and kappa(2) phenotypes. Selective targeting of delta-kappa heterodimers. *J. Med. Chem.* 2004, *47*, 2969–2972.
124. Jacobson, K. A., Xie, R., Young, L., Chang, L., Liang, B. T. A novel pharmacological approach to treating cardiac ischemia. Binary conjugates of A1 and A3 adenosine receptor agonists. *J. Biol. Chem.* 2000, *275*, 30272–30279.
125. Abadi, A. H., Lankow, S., Hoefgen, B., Decker, M., Kassack, M. U., Lehmann, J. Dopamine/serotonin receptor ligands, part III[1]: synthesis and biological activities of 7, 7′-alkylene-*bis*-6, 7, 8, 9, 14, 15-hexahydro-5H-benz[d]indolo[2, 3-g]azecines – application of the bivalent ligand approach to a novel type of dopamine receptor antagonist. *Arch. Pharm. (Weinheim)* 2002, *335*, 367–373.
126. Bonger, K. M., van den Berg, R. J., Heitman, L. H., AP, I. J., Oosterom, J., Timmers, C. M., Overkleeft, H. S., van der Marel, G. A. Synthesis and evaluation of homo-bivalent GnRHR ligands. *Bioorg. Med. Chem.* 2007, *15*, 4841–4856.
127. Rajeswaran, W. G., Cao, Y., Huang, X. P., Wroblewski, M. E., Colclough, T., Lee, S., Liu, F., Nagy, P. I., Ellis, J., Levine, B. A., Nocka, K. H., Messer, W. S., Jr Design, synthesis, and biological characterization of bivalent 1-methyl-1,2,5,6-tetrahydropyridyl-1,2,5-thiadiazole derivatives as selective muscarinic agonists. *J. Med. Chem.* 2001, *44*, 4563–4576.
128. Soulier, J. L., Russo, O., Giner, M., Rivail, L., Berthouze, M., Ongeri, S., Maigret, B., Fischmeister, R., Lezoualc'h, F., Sicsic, S., Berque-Bestel, I. Design and synthesis of specific probes for human 5-HT4 receptor dimerization studies. *J. Med. Chem.* 2005, *48*, 6220–6228.
129. Chini, B., Manning, M. Agonist selectivity in the oxytocin/vasopressin receptor family: new insights and challenges. *Biochem. Soc. Trans.* 2007, *35*, 737–741.
130. Fotiadis, D., Jastrzebska, B., Philippsen, A., Muller, D. J., Palczewski, K., Engel, A. Structure of the rhodopsin dimer: a working model for G-protein-coupled receptors. *Curr. Opin. Struct. Biol.* 2006, *16*, 252–259.
131. Russo, O., Berthouze, M., Giner, M., Soulier, J. L., Rivail, L., Sicsic, S., Lezoualc'h, F., Jockers, R., Berque-Bestel, I. Synthesis of specific bivalent probes that functionally interact with 5-HT_4 receptor dimers. *J. Med. Chem.* 2007, 50(18), 4482–4492.
132. Filizola, M., Wang, S. X., Weinstein, H. Dynamic models of G-protein coupled receptor dimers: indications of asymmetry in the rhodopsin dimer from molecular dynamics simulations in a POPC bilayer. *J. Comput. Aided Mol. Des.* 2006, *20*, 405–416.

Chapter 19

Application Strategies for the Primary Structure–Activity Relationship Exploration

Camille G. Wermuth

I. INTRODUCTION
II. PRELIMINARY CONSIDERATIONS
III. HIT OPTIMIZATION STRATEGIES
A. Some information about the target is available
B. No information about the target is available
C. The predominant objective is potency
D. The predominant objective is the establishment of SARs
E. The predominant objective consists of analog design
IV. APPLICATION RULES
A. Rule number one: the minor modifications rule
B. Rule number two: the biological logic rule
C. Rule number three: the structural logic's rule
D. Rule number four: the right substituent choice
E. Rule number five: the easy organic synthesis (EOS) rule
F. Rule number six: eliminate the chiral centers!
G. Rule number seven: the pharmacological logic rule
REFERENCES

"Le bon sens est la chose du monde la mieux partagée:
car chacun pense en être si bien pourvu, que ceux même qui
sont les plus difficiles à contenter en tout autre chose,
n'ont point coutume d'en désirer plus qu'ils n'en ont."

"Common sense is the worldly thing which is the best shared:
as each of us thinks to be so well provided with,
that even those who are the most difficult to satisfy in any other thing,
don't want to desire more of it than they already have."

René Descartes (1596–1650)[1]

I. INTRODUCTION

When confronted to a new lead structure or when, for patent reasons, he has to enlarge the protection perimeter around newly discovered structures, the medicinal chemist may be daunted by the immensity of his task. Effectively the possibilities of molecular variations around the lead structure are immense and *a priori* the synthesis of several thousand potential analogs can be envisaged. The aim of this chapter is to provide some guidelines and strategies rendering easier and more efficacious the decision on which compounds to prepare and which ones to reject. The proposed guidelines derive essentially from common-sense reasoning, feature that may explain why they are often forgotten. In addition to his personal experience, the author was inspired by the articles of Messer,[2] Cavalla,[3] Craig,[4] and Austel.[5]

II. PRELIMINARY CONSIDERATIONS

Before considering the different possibilities of molecular variation presented in the previous chapters (homology, bioisosterism, conformational restriction, optical isomerism, ring system modifications, and synthesis of twin drugs, etc.), one has to decide what kind of general strategy should be applied. Depending on the lead structure's size and on its degree of complexity, the strategy may involve a simplification (disjunctive approach), conservation of the

Copyright © 2008, Elsevier Ltd
All rights reserved.

FIGURE 19.1 Neostigmine and rivastigmine are the result of disjunctive approach applied to physostigmine.

TABLE 19.1 Drugs Resulting from Disjunctive Manipulations

Lead	Derivative
Cocaine	Procaine
Tubocurarine	Decamethonium
Morphine	Morphinanes Benzomorphanes Phenylpiperidines
Atebrine	Chloroquine
Asperlicine	Benzodiazepine analog
Phylloquinone (vitamin K_1)	Menadione (vitamin K_3)
Triampterene	Amiloride
Cimetidine	Roxatidine
Somatostatine	Simplified peptide
Bothrops jaracaVenin	Teprotide Captopril

same level of complexity (analogical approach) or enlargement through additional elements (conjunctive approach).

Simplification of the original lead compound is especially appropriate for natural substances. This approach, known as the *disjunctive approach*[6] (see Chapter 16) consists of a molecular dissection that deletes functions, structural elements, or cycles. Classical examples of disjunctive approaches are found in the pruning of the acetylcholinesterase inhibitor physostigmine to yield neostigmine and, later one, rivastigmine (Figure 19.1) or the change from somatostatin to a simplified hexapeptide.[7]

Other examples of disjunctive approaches are collected in Table 19.1. The main result of the methodology is the identification of the portions of the molecule that are essential for the expected biological activity and of those that are not.

TABLE 19.2 Drug Analogs Possessing a Similar Size than the Model Compound

Initial drug	Analog
Chlorpromazine	Thioridazine
Imipramine	Amitryptyline
Propranolol	Pindolol
Furosemide	Bumetamide
Enalapril	Perindopril
Cimetidine	Ranitidine
Pravastatine	Fluindostatine

Conservation of the lead compound's degree of complexity proceeds usually through isosteric exchanges or functional inversions and can be considered as being the *analogical approach*[8] (Table 19.2).

Finally, when additional moieties are grafted to the molecule, one speaks of *conjunctive approaches*.[6] They can also consist in the attachment of additional structural elements as in the association of two separate drugs (associative synthesis, non-symmetrical twin drugs, the symbiotic approach[9] or in the duplication of the parent drug (symmetrical twin drugs).

The change of the γ-aminobutyric acid (GABA)B receptor agonist CGP 27 492 to the GABAB receptor antagonist CGP 54 062 (Figure 19.2) represents a typical example of conjunctive approach resulting from the attachment of additional structural elements.[10] A similar case is provided by the design of the H2 receptor agonist impromidine.[11]

Examples of drugs resulting from *associative synthesis* (non-symmetrical twin drugs) and of *duplication* of the parent drug (symmetrical twin drugs) are listed in Tables 19.3 and 19.4. A more detailed study is presented in Chapter 18.

III. HIT OPTIMIZATION STRATEGIES

The strategy of hit optimization will heavily depend on the amount of information available at the start of the study. Particularly, if some knowledge of the 3D structure of the target is available, the synthesis program can take the corresponding information immediately into account. This will also be the case if some earlier structure–activity relationship (SAR) studies are available. However, in the most frequent cases, the target is new and original. In such a situation all the initial SAR exploration rests in the hands of the medicinal chemist. The purpose of this chapter is to let him benefit from a more than 30 years experience and to help him in the choice of the most appropriate strategy.

FIGURE 19.2 Conjunctive approach in drug design. The attachment of two benzylic groups and a hydroxyl in S-configuration changes a $GABA_B$ receptor agonist into a $GABA_B$ antagonist[10] similarly the H_2-histaminergic agonist impromidine is the result of a conjunctive approach applied to histamine.[11]

TABLE 19.3 Non-symmetrical Twin Drugs

Drug No. 1	Drug No. 2	Twin drug
Cafeine	Amphetamine	Fenethylline
Aspirine	Paracetamol	Benorylate
Clofibric acid	Nicotinic acid	Etofibrate
Hydrazinopyridazine	β-blocker	Prizidilol
Pindolol	Captopril	BW-B385C[12]

FIGURE 19.3 In a rather complex molecule, the simple shift of a nitrogen atom to another position of the diazine ring atom abolishes almost completely the affinity for the CRF receptor.

TABLE 19.4 Symmetrical Twin Drugs

Drug	Activity
Bialamicol	Anti-amebic
Ethambucol	Tuberculostatic
Probucol	Antihyperlipoproteinemic
Thiamine disulfide	Vitamin
Dicumarol	Anticoagulant
Netropsin	DNA binding agent
Succinylcholine	Skeletal muscle relaxant

A. Some information about the target is available

The first point to consider is to ascertain if the hit (or the lead) that has to be optimized is relevant to a computer-aided design. If the X-ray 3D structure of the target protein is already described, it becomes possible to match the different candidate molecules with the target structure and eliminate those which evidently are too bulky or which possess an inadequate geometry. Similar situation occurs if a non-experimental 3D model of the target is available. For more details on computer-assisted drug design (CADD) see chapters 10, 28, 29 and 30. An important point which is always to be kept in mind is that CADD has limitations. Thus the two corticotropin-releasing factor (CRF) antagonists of Figure 19.3 are structurally very close positional isomers. However, one of them shows nanomolar affinity for the CRF receptor whereas the other one is practically inactive.[13] No evident CADD explanation is available for this behavior.

In the absence of any target structure some information can be gathered in comparing the hit structure with the endogenous ligand (if known) or with other structures showing affinities for the same target (if known). If several compounds exist which are recognized by the target it becomes possible to practice the active analog approach (Chapters 28 and 29).

TABLE 19.5 Astra-Zeneca "Generic Lead Target Profile" for Progressing HTS Hits to Leads

Parameter	Values
Potency	100 nM
Molecular weight	<450
C log P	<3
Log D	<3
Solubility	>10 μg/mL
Clear SAR around the potential lead	
Structure must provide patent opportunities	
Selectivity – Use PanLabs/Cerep batteries	
Need *in vivo* biological validation	
HERG screening	
Early toxicity *in vivo* screening	
Metabolism	
P450 inhibition IC_{50}	10 μM for 5 major isoenzymes
Rat hepatocyte intrinsic clearance	<14 μL/min/mg
Human microsome intrinsic clearance	<23 μL/min/mg
Rat IV clearance	<35 mL/min/mg
Volume	>0.5 L/kg
$t_{1/2}$	>0.5 h
Pharmacokinetics	
Rat PO bioavailability	>10%
Plasma protein binding	<99.5

B. No information about the target is available

When one has to deal with a real new target, the only way to progress is to practice enough syntheses to identify the molecular features that are favorable and those that are detrimental to the activity. Such molecular variation programs can be practiced in several manners. Usually the predominant parameter wanted first is potency, but other qualities of the future drug molecules are relevant of drug optimization. One can mention selectivity, satisfactory ADME (absorption, distribution, metabolism, and excretion) and toxicity profiles, optimal physicochemical properties such as chemical stability, water solubility, and absence of polymorphs, finally the compounds must be patentable. Table 19.5, due to Baxter *et al.*[14] summarizes the different criteria practiced at the Astra-Zeneca company.

C. The predominant objective is potency

One of the most fruitful strategies revealing the features associated with high potency consist in what we call the *topological exploration of the lead compound*. In this approach, the possible modifications of the molecule are considered from the four cardinal points: the south, the north, the west and the east, and from the center of the molecule (Figure 19.4).

FIGURE 19.4 Scheme of the topological exploration of the pirenzepine molecule.

As an illustration, let us assume that our starting lead compound is the muscarinic antagonist pirenzepine. We can then consider successively the different modification sites.

1. *East side modifications*

In the present example, the east side modifications concern mainly in changes at the level of the pyridine ring, and the questions that should be answered are the following ones (Figure 19.5):

- Is the pyridine nitrogen necessary? Can the pyridine ring be replaced by a phenyl ring (a)?
- Can the pyridine nitrogen be displaced to other positions (b, c, d)?
- What is the influence of substituents on the pyridine ring (e)? Can it be substituted by various functions (associated with typical electronic, steric, or lipophilic changes)?
- Can the pyridine ring be changed to other aromatic heterocycles: pyridazine (f), pyrimidine (g), pyrazine (h), triazines (i), thiazoles (j), etc?

2. *North side modifications*

On the north side of pirenzepine one can consider the NH and the C═O groups separately or together, as an amide function (Figure 19.6).

- Can the NH group be substituted (a)? (Possible role as hydrogen bond donor).
- Can the NH group be replaced by a CH_2 group (b) (or any other possible bioisosteric group)?

FIGURE 19.5 East side modifications on the pirenzepine molecule.

FIGURE 19.6 North side modifications on the pirenzepine molecule.

- Is the carbonyl group necessary? Can it be changed to a CH_2 group (d)?
- Can the amide be replaced by any bioisostere such as described in Chapter 15 (c, e, f)?

3. *West side modifications*

Due to the almost symmetrical structure of pirenzepine, essentially the same kind of modifications can be applied than that suggested for the east side.

For both east and west sides, "benzo-splitting" (see Chapter 16) can be considered (Figure 19.7).

4. *South side modifications*

They concern the changes made on the basic side chain, a huge amount of possibilities exist on different levels:

On the piperazine ring (Figure 19.8)

– Can the *N*-methyl group be replaced by higher alkyl, aralkyl or aryl groups? (a)
– Can it be replaced by homopiperazine (b), by piperidine (c and c), by ring-opened diamines (e)?
– Can it be substituted (g and h) or bridged (i)?
– Can it be replaced by some vague bioisosteric equivalent such as a guanidino group (*N*-methyl piperazine was used as guanidine substitute in the design of thrombine inhibitors)?

On the carboxamido function (Figure 19.9)

– Can the carbonyl group be reduced to a CH_2 group (a)?
– Can the carboxamido function be replaced by a carbon–carbon double bond (b)?
– Can it be included in a bioisosteric and constraint ring system (c)?

5. *Center modifications*

The literature available for the tricyclic psychotropic drugs proposes a great number of possibilities of variations of the potentially applicable to the present central diazepinone ring (Figure 19.10).

They include ring contraction (a), ring extension (b), ring bridging (c), and changes in the nature and number of the heteroatoms.

Taken together, all these modifications suggest the synthesis of an impressive number of analogs of a given hit structure and hopefully open the way to some more potent original analogs.

D. The predominant objective is the establishment of SARs

When one deals with a hitherto unknown active compound, he or she wants to acquire information about the atoms and/or the functionalities involved in the interaction with the target and about the nature of the chemical bonds they achieve in the drug–target complex. In the examples below the hit compound is imaginary. First, a topological attribution of the functional groups similar to that of the previous paragraph is achieved (Figure 19.11). Second, a case by case discussion of each functional feature is undertaken.

FIGURE 19.7 "Benzo-splitting" applied to the east and the west side of pirenzepine.

FIGURE 19.8 Variations on the *N*-methylpiperazine ring.

FIGURE 19.10 Modifications of the central ring of pirenzepine.

FIGURE 19.9 Variations on the side chain carboxamido group of pirenzepine.

FIGURE 19.11 An imaginary hit compound serves as demonstration molecule. The different sites (1–10) of the molecule are potentially able to influence the drug–receptor interaction.

If the carbonyl group of the molecule is considered, one can assume that, either it interacts by exchanging hydrogen bonds with its target protein or it plays an architectural role. This latter is due to the planar structure imposed by the sp^2 character of its carbon atom. In order to decide between the two alternatives, some key substances have to be prepared and tested (Figure 19.12).

The synthesis of the thiocarbonyl analog leads to a compound with potency similar to that of the carbonyl derivative but which is unable to make hydrogen bonds. If it is still active, it means that hydrogen bonding at this region of the molecule is not necessary for activity. If it is no more active, the ketonic carbonyl can be kept. Possibly the carbonyl group can even be replaced by the corresponding ketoxime, a better hydrogen bond former. The reduction of the carbonyl group to the two enantiomeric alcohols transforms the planar sp^2 carbon atom into pyramidal sp^3 carbons. These compounds may be active if the carbonyl group plays a conformational role.

The presence of an α-methylene group permits the enolization of the hit molecule, this tautomery is blocked through the insertion of a spiro cyclopropyl ring (Figure 19.13). Result of a formal oxidation yields the secondary alcohol. Double bond possibilities are α-methylene ($X = CH_2$), the diketone ($X = O$) and the corresponding ketoxime ($X = N\text{-}OH$).

Further investigations concern the different substituents beared by the hit molecule (Figure 19.14):

1. The methyl group can it be replaced by a hydrogen? By a larger group (e.g. isopropyl)?
2. The methoxy substituent can it be replaced by an isoelectronic substituent or by another hydrogen bond acceptor?

FIGURE 19.12 The ketonic carbonyl function of the hit compound can play an electronic role in forming hydrogen bonds or an architectural role in maintaining the oxygen atom coplanar with nearest three carbon atoms.

FIGURE 19.13 The α-methylene group allows enolization of the hit compound. This is no more possible with the α-cyclopropyl analog. Various oxidative functionalities are possible.

3. Role of the chlorine: electronic? lipophilic? metabolic?
4. Nitrogen substituent: desmethyl? higher alkyl? acyl? carbamate? urea?

E. The predominant objective consists of analog design

In the losartan molecule, the substituted imidazole moiety is attached to the typical tetrazolyl-diphenyl unit (Figure 19.15). In practicing analog synthesis, the Novartis scientists conserved unchanged this latter part of the molecule but tried to prepare a bioisosteric equivalent of the substituted imidazole possessing similar interaction possibilities. The lipophilic *n*-butyl chain was maintained, the CN dipole was replaced by a CO dipole, and the ensemble chlorine substituent plus two imidazolic carbon atoms was replaced by an isolipophilic isopropyl group. Finally, as it was known that *in vivo* the primary alcoholic group of losartan was oxidized into the corresponding carboxylic group, the carboxylic function was directly introduced in valsartan.

FIGURE 19.14 Investigations about the roles played by the various substituents of the hit molecule.

IV. APPLICATION RULES

It would certainly be tiresome and beyond the possibilities of a medicinal chemistry team to be forced to prepare all the compounds imagined by means of "paper" chemistry. The following rules aim to codify precisely the use of all the strategies and to increase their efficacy in establishing priorities and selection rules.

A. Rule number one: the minor modifications rule

This rule[8] can be defined as the priority given to the design of analogs that are close to the lead structure and that result from only minor changes. Minor changes are achieved by very simple organic reactions such as hydrogenations, hydroxylations, methylations, acetylations, racemate resolutions, changes in substituents and isosteric replacements. The modification can produce either an increase in potency or an increase in selectivity or even sometimes the suppression of unwanted toxic or side effects (Table 19.6).

A simple change in the aliphatic chain length can abolish *mutagenic* properties. Thus, in a series of muscarinic M_1 partial agonists (Figure 19.16), the compound with a dimethylene side chain is potent (IC_{50} [^{3}H]-pirenzepine = 3 nM), but mutagenic. The higher trimethylenic homolog is less potent (IC_{50} [^{3}H]-pirenzepine = 15 nM), but safe in terms of mutagenicity.[16]

In some instances, very slight changes such as hydrogenation or dehydrogenation can induce dramatic changes in the activity profile of drug molecules. Examples are found in the imidazoline I_3 receptor ligands which act on

FIGURE 19.15 The angiotensin agent valsartan results from a rational analog design.

TABLE 19.6 Minor Modifications

Original compound	Modified compound	Result[a]
Ergotamine	Dihydroergotamine	Increase in potency as α-adrenergic antagonist decrease in toxicity
Chlorothiazide	Hydrochlorothiazide	20-fold increase in potency
Chloroquine	Hydroxychloroquine	Decrease in toxicity
Morphine	Codeine	Change in activity profile (analgesic → antitussive)
Carbachol	Bethanechol	Increase in selectivity (exclusively muscarinic)
Imipramine	Desmethylimipramine	Change in activity profile (noradrenergic → serotonergic)
Tolbutamide	Chlorpropamide	Longer duration of action (5–7 h → 24–48 h)
Racemic amphetamine	Dexamphetamine	Less cardiovascular side effects

[a] Taken from Goodman and Gilman[15]

FIGURE 19.16 The mutagenicity of the original dimethylenic muscarinic M_1 partial agonist could be abolished in changing the dimethylene to a trimethylene side chain.

the insulin secretion. The imidazolinic compound efaroxan acts as an agonist,[17] the corresponding imidazole acts as an antagonist[18] (Figure 19.17). A similar passage is found for the passage from the agonistic benzofuranic compound 2-BFI[19] to its antagonistic dihydro derivative.[20]

The minor modifications rule, even supported by prestigious results, is largely unrecognized. Making use of ordinary chemistry, it is not always accepted with enthusiasm by organic chemists. The very simple reactions that are involved do not add much to their fame and they are more fascinated by the challenge of a total synthesis, especially of a natural substance bearing many chiral centers. Seen from a practical point of view, priority has nevertheless to be given to this principle. Its simplicity of implementation, and especially the spectacular results that it brings, militate in its favor.

B. Rule number two: the biological logic rule

The second rule of application rests on the earliest possible utilization of biochemical data. Indeed, even when a medicinal chemist ignores all of the biochemistry of the substance that he studies, he earns to consider the former under the biochemical angle which will provide him with a good number of subjects to think about and may sometimes allow anticipation of the behavior of the molecule. In particular, biological activity may be rationalized if it stems from the chemical or physicochemical properties of the series.

Very general properties can be foreseen in so far as functions or moieties present in the structure can suggest interference with a biological system; for example, hydrazines or hydroxylamines and pyridoxal-containing coenzymes, complexing agents and metallic coenzymes, electron donors or acceptors and oxido-reduction coenzymes, tensioactive amphiphilic substances and production of hemolysis in erythrocytes. In a more precise manner, the alkylating properties of compounds such as the nitrogen mustards, the nitrosoureas and the mitomycins, or the

FIGURE 19.17 Hydrogenations and dehydrogenations can induce switches from agonist to antagonist profiles.

cross-linking properties of *cis*-platinum derivatives relate to their anticancer activity. The activity of anthracyclines, ellipticine, and the anthracenediones has been shown to be due to intercalation in the double helix of DNA. The anticoccidial polyether ionophores are potent complexing agents for mono- and divalent cations.

The biological action of a compound is also readily explicable if it *mimics a natural substrate or mediator*. This is the case for enzymes with inhibitors (angiotensin I and captopril), suicide substrates (GABA and vigabatrin), antimetabolites (*p*-aminobenzoic acid and *p*-aminobenzene-sulfonamide) and for receptors with agonists (acetylcholine and muscarine), antagonists (GABA and gabazine) and uptake inhibitors (GABA and nipecotic acid). The analogy with the endogenous substance can be sometimes very vague as seen with quaternary ammonium compounds that all present more or less affinity for the cholinergic receptors, or with purines, that often are recognized by the phosphodiesterases.

The pathways of drug *metabolism* follow some general rules (see Chapter 33) and the metabolites of a given substance can, at least qualitatively, be imagined in advance. As a consequence, various measures can be taken to favor or, conversely, to slow down the biodegradation. Some chemical groupings are more than others prone to yield unwanted *toxic metabolites* (see Chapter 33). Among the best known are the aromatic nitro, nitroso, azo and amino compounds, the bromoarenes, the hydrazines and the hydroxylamines, and the polyhalogenated aliphatic or aromatic compounds. A proposed explanation of the arylamine toxicities (the so-called para effect) is their facile oxidation to an electrophilic quinonic system followed by addition of thiol nucleophiles. A process that models well-known hapten formation reactions.[21] Finally, if the active principle is an acid or a base, the choice of the salifying counterion has also to follow some selection criteria (see Chapter 37). Oxalates and nitrates for example are not very popular whereas hydrochlorides represent a satisfactory compromise.

C. Rule number three: the structural logic's rule

This rule implies that, as soon as some structural data are available (intercharge distances, E or Z conformations, axial or equatorial substituent orientations, misoriented substituents, etc.), they have to be fed back into drug design. When dealing with enzymes or receptors of unknown structure, one route to such information consists in comparing already known active compounds, recognized by the same molecular target, and to deduce the important stereoelectronic features associated with potency and selectivity. This approach is referred to as pharmacophore identification or receptor mapping (see Chapters 28 and 29). Initially presented by Marshall *et al.*[22] it has some predictive merit[23,24] and, at least avoids unnecessary syntheses of *a priori* inactive compounds. In practice the most efficient methods to succeed consist of steady comings and goings between synthetic and computer chemistry in order to achieve the ideal interplay between intuition and computer assistance.

A structural guide is also available for drugs designed to bind to the neurotransmitters of the central nervous system (CNS). On the basis of a comparison of the crystal structure of recognized representative compounds from each of eight major CNS active drug classes, Andrews and Lloyd[25,26] identified a common structural basis, essentially characterized by an aromatic plane distant by about 5 Å from a nitrogen moiety. The explanation for this finding resides in the biochemical origin of the neurotransmitters[24]. Most of them are of the arylethylamine type, as a result from the decarboxylation of aromatic amino acids such as DOPA or histidine.

Another example is given by acetylcholine which can be considered as bioisosteric with GABA (Figure 19.18), this property explains the observation that a compound such as the GABAA receptor antagonist bicuculline is recognized by both the GABAA and the nicotinic receptors.[24]

D. Rule number four: the right substituent choice

Half of all the existing drugs contain easily substituted aromatic rings. The replacement, in such rings, of a hydrogen by a substituent (alkyl, halogen, hydroxyl, nitro, cyano, alkoxy, amino, carboxylate, etc.) can dramatically modify the intensity, the duration, and perhaps even the nature of the pharmacological effect (see Chapters 20 and 22). It becomes therefore of prime importance to proceed to the optimal choice of substituents *so as to explore with the*

FIGURE 19.18 Similar intercharge distances between the protonated nitrogen and the carbonyl dipole exist in bicuculline, GABA, acetylcholine, and carbachol.

smallest set possible the 3D space formed by lipophilic, electronic, and steric parameter coordinates (Figure 19.19).

The right substituent choice minimalizes the number of test compounds that have to be synthesized to insure a significant space volume. This point represents a 3D extension of the Craig plot discussed by Craig[4] and by Austel.[5] In this context, the decision tree proposed by Topliss[27] allows a fast identification of the substituents associated with the highest potency. Application examples of the Topliss scheme are discussed by Martin and Dunn.[28]

E. Rule number five: the easy organic synthesis (EOS) rule

Synthesis of new compounds is a costly and lengthy process; therefore any measure able to render it more efficacious is welcome. Thus for example, when the decision is taken to prepare a given set of compounds, why not first prepare those whose synthesis is the easiest? In the same line of thought why not prepare first compounds for which intermediates are commercially available?

A particular recommendation is to synthesize heterocycles. In statistics established in 1982 on 1,522 drug molecules and which can still be considered as valid, Kleemann and Engel[29] highlighted the fact that, among the synthetic drugs 62% contained at least one heterocyclic ring, the percentage within natural compounds being even higher (77%). Indeed, heterocycles present many advantages: (1) They allow the insertion of elements capable to give interactions there that the carbocycles do not give. (2) They allow a greater number of combinations. It becomes therefore easier to be original. (3) They represent rigid analogs of endogenous substances that themselves are often nitrogenous metabolites of amino acids. (4) Often their facile synthesis permits the preparation of large series.[30] One of the major problems when dealing with isosteric or bioisosteric replacements in heterocyclic systems is the selection of the *a priori* most promising candidate among several dozens of possible rings. A simple clue, which reflects the dipolar moment, can be given by knowledge and comparison of the *boiling points* of the basic heterocycles (see chapter 15).

F. Rule number six: eliminate the chiral centers!

Although optical isomerism is discussed in Chapter 26, some practical considerations on chiral molecules are appropriate here. Nowadays, it is well accepted that racemates and both enantiomers are usually three different pharmacological entities and that it requires extensive pharmacological, toxicological, and clinical pharmacological research before it can be decided whether it is advantageous to use racemates or enantiomers in clinical practice. According to Soudijn,[30] these research efforts could be reduced to about one-third when drugs without centers or planes of asymmetry could be developed with the same or higher affinity.

Effectively *asymmetry is far from being an absolute requisite for activity!* The alkaloid morphine possesses five chiral centers, on the other hand its synthetic derivative fentanyl is devoid of any asymmetric center but nonetheless belongs to the most potent analgesics known. Usually, chiral centers are eliminated in creating symmetry (see Chapter 26, Figures 26.9 and 26.10). A typical example of this process is the design of non-chiral immunosuppressive 2-aminopropane-1,3-diols starting from the natural compound myriocin[31] (Figure 19.20).

In some instances, the chiral centers can be at least partially eliminated. This is the case for the synthetic analogs of the hydroxymethyl-glutaryl coenzyme A (HMG-CoA) reductase inhibitor mevinolin. Mevinolin itself (see Chapter 18) has seven asymmetric centers but the five chiral centers contained in the hexahydronaphtalene ring system

FIGURE 19.19 3D space formed by lipophilic, electronic, and steric coordinates.

FIGURE 19.20 Suppression of chiral centers through introduction of symmetry.[31]

are unnecessary for HMG-CoA inhibition. The second generation of mevinolin analogs, retains only two of the initial seven chiral centers.[32]

When one nevertheless has to deal with chiral centers, why not first prepare the racemic compound and start with an enantioselective synthesis only if an interesting activity is found? This latter point is tricky, because many people believe that two enantiomers might happen to antagonize each other. They refer to the numerous examples published in the literature.[30,33–35] In reality two optical isomers *are never antagonists at comparable dosages*. This comes from the space-relationship required for the interaction with the receptor site that is only slightly altered by passing from S to R forms, or vice versa. If one of the enantiomers achieves the optimal fit to the receptor site in exchanging the highest number of non-covalent linkages, its antipode gives rise to only weaker interactions, even under the most favorable conditions (see Chapter 17).

From a practical point of view, this absence of *stoichiometric* antagonism entails two consequences: (a) if a racemic mixture does not show any activity, it is useless to carry out the separation of the two antipodes, and (b) a racemic mixture usually has the average potency of both constituents, thus, the maximal benefit one can achieve in resolving racemic mixtures is to increase the potency to twice that of the racemate.

G. Rule number seven: the pharmacological logic rule

We already insisted in Chapter 2 on the fact that a correctly performed pharmacological study must satisfy certain criteria (relationship between dose and effect, presentation of the confidence limits, comparison with a reference compound, and determination of the time of the peak action). On the chemical side it is also extremely important to provide the pharmacologists with reference compounds published by the competitors laboratories. Even if it is felt tedious and time consuming to resynthesize an already described compound, the operation is always worthwhile and sometimes surprising. How often a good looking published molecule, for which attractive activities are claimed, looses much of its charm once it is reinvestigated by one's own team! An other point that may contribute to increasing the credibility of your work, is the so-called *a contrario* probe. In other words, when your own SAR studies allow identification of the molecular features associated with high activity, proceed, of course to the synthesis of the most interesting representatives, but also prepare at least one compound that, according to your results, should be inactive.

REFERENCES

1. Descartes, R. *Discours de la méthode, 1ére partie 1637*. Librairie Larousse: Paris, 1972. p. 27
2. Messer, M. Traditional or pragmatic research. In *Drug Design: Fact or Fantasy?* (Jollès, G., Wooldridge, K. R. H., Eds). Academic Press: London, 1984, pp. 217–224.
3. Cavalla, J. F. Drug design valuable for refining an active drug. In *Decision Making in Drug Research* (Gross, F., Ed.). Raven Press: New York, 1983, pp. 165–172.
4. Craig, P. N. Guidelines for drug and analog design. In *The Basis of Medicinal Chemistry* (Wolff, M. E., Ed.), Vol. 1. Wiley-Interscience: New York, 1980, pp. 331–348.
5. Austel, V. Features and problems in practical drug design. In *Steric Effects in Drug Design* (Charton, M., Motoc, I., Eds), Vol. 114. Lange & Springer: Berlin, 1984, pp. 8–19. Topics in Current Chemistry
6. Schueller, F. W. *Chemobiodynamics and Drug Design*. McGraw-Hill: New York, 1960.
7. Freidinger, R. M., Veber, D. F. Design of novel cyclic hexapeptide somatostatin analogs from a model of the bioactive conformation. In *Conformationally Directed Drug Design* (Vida, J. A., Gordon, M., Eds). American Chemical Society: Washington, DC, 1984, pp. 169–187.
8. Wermuth, C. G. Modifications chimiques des médicaments en vue de l'amélioration de leur action. *Agressologie* 1966, *7*, 213–219.
9. Nicolaus, B. J. R. Symbiotic approach to drug design. In *Decision Making in Drug Research* (Gross, F., Ed.). Raven Press: New York, 1983, pp. 173–186.
10. Froestl, W., Furet, P., Hall, R. G., Mickel, S. J., Strub, D., Sprecher, G. V., Baumann, P. A., Bernasconi, R., Brugger, F., Felner, A., Gentsch, C., Hauser, K., Jaeckel, J., Karlsson, G., Klebs, K., Maître, L., Marescaux, C., Moser, P., Pozza, M. F., Rihs, G. $GABA_B$ antagonists: novel CNS-active compounds. In *Perspectives in Medicinal Chemistry* (Testa, B., Kyburz, E., Fuhrer, W., Giger, R., Eds). VHC: Weinheim, 1993, pp. 259–272.
11. Durant, G. J., Duncan, W. A. M., Ganellin, C. R., Parsons, M. E., Blakemore, R. C., Rasmussen, A. C. Impromidine (SK&F 92 676) is a very potent and specific agonist for histamine H_2 receptors. *Nature* 1978, *276*, 403–405.
12. Hardy, G. W., Allan, G. BW-B3895C. *Drugs Fut.* 1988, *13*, 204–206.
13. Chen, S. *et al* CRF antagonists. *J. Med. Chem.* 1996, *39*, 4358–4360.
14. Baxter, A., Bennion, C., Bent, J., Boden, K., Brough, S., Cooper, A., Kinchin, E., Kindon, N., McInally, T., Mortimore, M., Roberts, B., Unitt, J. Hit-to-lead studies: the discovery of potent, orally bioavailable triazolethiol CXCR2 receptor antagonists. *J. Med. Chem.* 2003, *13*, 2625–2628.
15. Goodman-Gilman, A., Rall, T. W., Nies, A. S., Taylor, P. *Goodman and Gilman's The Pharmacological Basis of Therapeutics*, 8th Edition. Pergamon Press: New York,1990.
16. Kan, J. -P. Sanofi-Synthelabo Recherche, Montpellier, France, personal communication, January 17, 1997.
17. Chan, S. L., Pallett, A. L., Morgan, N. G. Clotrimazole and efaroxan stimulate insulin secretion by different mechanisms in rat pancreatic islets. *Naunyn-Schmiedebergs Arch. Pharmacol.* 1997, *356*, 763–768.
18. Chan, S. L., Atlas, D., James, R. F., Morgan, N. G. The effect of the putative endogenous imidazoline receptor ligand, clonidine-displacing substance, on insulin secretion from rat and human islets of Langerhans. *Br. J. Pharmacol.* 1997, *120*, 926–932.
19. Morgan, N. G., Chan, S. L., Mourtada, M., Monks, L. K., Ramsden, C. A. Imidazolines and pancreatic hormone secretion. *Ann. NY Acad. Sci.* 1999, *881*, 217–228.
20. Brown, C. A., Chan, S. L., Stillings, M. R., Smith, S. A., Morgan, N. G. Antagonism of the stimulatory effects of efaroxan and glibenclamide in rat pancreatic islets by the imidazoline, RX801080. *Br. J. Pharmacol.* 1993, *110*, 1017–1022.
21. Sanner, M. A., Higgins, T. J. Chemical Basis for Immune Mediated Idiosyncratic Drug Hypersensitivity. In *Ann. Rep. Med. Chem.*

(Bristol, J. A., Ed.), Vol. 26. Academic Press: San Diego, CA, 1991, pp. 181–190.

22. Marshall, G. R., Barry, C. D., Bosshard, H. E., Dammkoehler, R. A., Dunn, D. A. The conformational parameter in drug design: the active analog approach. In *Computer-Assisted Drug Design* (Olson, E. C., Christoffersen, R. E., Eds). American Chemical Society: Washington, DC, 1979, pp. 205–226.
23. Marshall, G. R., Cramer, D. R., III Three-dimensional structure–activity relationships. *Trends Pharmacol. Sci.* 1988, *9*, 285–289.
24. Wermuth, C. G., Langer, T. Pharmacophore identification. In *3D QSAR in Drug Design – Theory, Methods, and Applications* (Kubinyi, H., Ed.). ESCOM: Leiden, 1993, pp. 117–136.
25. Andrews, P. R., Lloyd, E. J. A common structural basis for CNS drug action. *J. Pharm. Pharmacol.* 1983, *35*, 516–518.
26. Lloyd, E. J., Andrews, P. R. A common structural model for central nervous system drugs and their receptors. *J. Med. Chem.* 1986, *29*, 453–462.
27. Topliss, J. G. Utilization of operational schemes for analog synthesis in drug design. *J. Med. Chem.* 1972, *15*, 1006–1011.
28. Martin, Y. C., Dunn, W. J. Examination of the utility of the Topliss schemes for analog synthesis. *J. Med. Chem.* 1973, *16*, 578–579.
29. Kleemann, A., Engel, J. *Pharmazeutische Wirkstoffe- Synthese, Patente, Anwendungen (Preface)*. Georg Thieme Verlag: Stuttgart, 1982.
30. Soudijn, W. Advantages and disadvantages in the application of bioactive racemates or specific isomers as drugs. In *Stereochemistry and Biological Activity of Drugs* (Ariëns, E. J., Soudijn, W., Timmermans, P. B. M. W. M., Eds). Blackwell Scientific Publications: Oxford, 1983, pp. 89–102.
31. Kiuchi, M., Adachi, K., Kohara, T., Minoguchi, M., Hanano, T., Aoki, Y., Mishina, T., Arita, M., Nakao, N., Ohtsuki, M., Hoshino, Y., Teshima, K., Chiba, K., Sasaki, S., Fujita, T. Synthesis and immunosuppressive activity of 2-substituted 2-aminopropane-1,3-diols and 2-aminoethanols. *J. Med. Chem.* 2000, *43*, 2946–2961.
32. Baader, E., Bartmann, W., Beck, G., Bergmann, A., Granzer, E., Jendrella, H., Kerekjarto, B.v., Kesseler, K., Krause, R., Paulus, E., Schubert, W., Wess, G. Rational approaches to enzyme inhibitors: new HMG-CoA reductase inhibitors. In *Trends in Drug Research* (Claassen, V., Ed.). Elsevier: Amsterdam, 1990, pp. 49–71.
33. Towart, R., Wehninger, E., Meyer, H. Effects of unsymmetrical ester substituted 1,4-dihydropyridine derivatives and their optical isomers on contraction of smooth muscle. *Naunyn-Schmiedeberg's Arch. Pharmacol.* 1981, *317*, 183–7185.
34. Lotti, V. J., Taylor, D. A. α2-Adrenergic agonist and antagonist activity of the respective (–)- and (+)-enantiomers of 6-ethyl-9-oxaergoline. *Eur. J. Pharmacol.* 1982, *85*, 211–215.
35. Hof, R. P., Rüegg, U. T., Hof, A., Vogel, A. Stereoselectivity at the calcium channel: opposite action of the enantiomers of a 1,4-dihydropyridine. *J. Cardiovasc. Pharmacol.* 1985, *7*, 689–693.

Part IV

Substituents and Functions: Qualitative and Quantitative Aspects of Structure–Activity Relationships

Han van de Waterbeemd

Section Editor

Chapter 20

Substituent Groups

Patrick Bazzini and Camille G. Wermuth

I. INTRODUCTION
II. METHYL GROUPS
A. Effects on solubility
B. Conformational effects
C. Electronics effects
D. Effects on metabolism
E. Extensions to other small alkyl groups
III. EFFECTS OF UNSATURATED GROUPS
A. Vinyl series
B. Allylic series
C. Acetylenic series
D. Cyclenic equivalents of the phenyl ring
IV. EFFECTS OF HALOGENATION
A. The importance of the halogens in the structure–activity relationship
B. Usefulness of the halogens and of cognate functions
V. EFFECTS OF HYDROXYLATION
A. Effects on solubility
B. Effects on the ligand–receptor interaction
C. Hydroxylation and metabolism
VI. EFFECTS OF THIOLS AND OTHER SULFUR-CONTAINING GROUPS
A. Drugs containing thiol
B. Drugs containing oxidized sulfides
C. Drugs containing thiocyanate or thiourea
VII. ACIDIC FUNCTIONS
A. Effects on solubility
B. Effects on biological activity
VIII. BASIC GROUPS
IX. ATTACHMENT OF ADDITIONAL BINDING SITES
A. To increase lipophilicity
B. To achieve additional interactions
REFERENCES

Fifty percent of the currently used drugs contain at least one aromatic ring that can be matter of substitution

John Taylor[1]

I. INTRODUCTION

The replacement, in an active molecule, of a hydrogen atom by a substituent (alkyl, halogen, hydroxyl, nitro, cyano, alkoxy, amino, carboxylate, etc.) or a functional group can deeply modify the potency, the duration, perhaps even the nature of the pharmacological effect. Structure–activity relationship studies implying substituent modifications represent therefore a common practice in medicinal chemistry, all the more since half of the existing drugs contain easy-to-substitute aromatic rings. The perturbations brought by the substituent can affect various parameters of a drug molecule such as its partition coefficient, its electronic density, its steric environment, its bioavailability and pharmacokinetics, and, finally its capacity to establish direct interactions between the substituent and the receptor or the enzyme.

In reality it is impossible to modify only one alone of these five parameters. Thus for example, the replacement of a hydrogen atom by a methyl group is going simultaneously to play on the five parameters listed above. Nevertheless, through a careful selection of the adequate substituent, it is possible to vary one of the considered parameters in a dominant manner.

To illustrate the repercussions on the biological activity resulting from substituent effects, we will study

Copyright © 2008, Elsevier Ltd
All rights reserved.

successively the effects of methyl groups, of unsaturated groups, and of halogen substitution. More shortly, hydroxy groups, thiols and acidic or basic functions will be discussed. A final section deals with the attachment of large lipophilic additional binding moieties.

II. METHYL GROUPS

In this section, we show how a methyl group, so often considered as chemically inert, is able to alter deeply the pharmacological properties of a molecule. We will envisage successively effects on the solubility, conformational effects, electronic effects and effects on the bioavailability and the pharmacokinetics. In the last paragraph, we will present some replacement possibilities of the methyl groups by related groups and extend the study to some larger alkyl groups.

A. Effects on solubility

As a rule, the grafting of one or several methyl groups on an active molecule renders the former more lipophilic and therefore less soluble in water. However, in some particular cases, grafting one or several methyl groups to a molecule results in an increase of the water solubility by mechanisms such as the increase of hydrophobic bonding possibilities or diminution of the crystal lattice energy.

1. *Increase in lipophilicity*

Normally one expects that methyl groups increase the lipophilicity. Indeed, the $\log P$ (logarithm of the partition coefficient P between n-octanol and water) is 2.69 for toluene, compared to $\log P = 2.13$ for benzene.[2] More generally, the passage of (M)-H to (M)-CH_3 gives place to a positive increment of 0.52 in Hansch constants calculations (see Chapter 23).

The increase in lipophilicity due to methylation can drastically modify the bioavailability of the drug, and thus its efficacy. Tyrosine dimethylation of the synthetic opioid DPDPE, for example (Figure 20.1), produces an increase in affinity both for the δ-opioid and the μ-opioid receptors. The *in vivo* analgesic activity increases also in comparison with the non-methylated DPDPE, notably as a result of enhanced bioavailability.[3]

Another example of correlation between lipophilicity and activity is found in a series of imidazolinediones (Table 20.1). The introduction of a single methyl group (compound 1 and 2) increases the lipophilic character of the molecule, and its ability to displace rimonabant (SR-141716A), the specific antagonist of CB_1 cannabinoid receptors.[4]

2. *Hydrophobic interactions*

As stated above, the usual result of the methyl group addition to a given molecule is to augment its lipophilicity: there are, however, exceptions to this rule, especially when the grafting of one or several methyl groups can render the molecule more compact (more "globular"). A good illustration of this effect is provided by aliphatic alcohols.[6] As expected, one observes that the increase in lipophilicity when passing from n-butanol to n-pentanol is accompanied by a decrease of the water solubility. However, 2-pentanol, and even more neopentanol, although possessing one methyl more than n-butanol, are less lipophilic, that means more soluble in water (Table 20.2).

Similarly, one can observe, in Table 20.1 (compound 3), that the imidazolinedione is less lipophilic when R2 is an isopropyl than when R2 is an n-propyl.

How to explain this anomaly? It has to be simply attributed to an entropic effect.[7] In aqueous solution the particle is imprisoned in a three-dimensional network (a cluster) of structured water molecules. On the other hand, a smaller amount of structured water molecules is needed to create a cluster around a compact molecule than around an extended one (Figure 20.2). This new structural arrangement is energetically favorable.

For the same reason some aromatic rings bearing protonated basic side chains are more soluble than anticipated.

Compound	R	Receptor binding K_i (nM)	
		μ	δ
DPDPE	H	2018	17.7
DMT-DPDPE	Me	58.3	1.8

FIGURE 20.1 Structure–activity of DPDPE and its dimethyl-analog (Tyr-D-Pen-Gly-Phe-D-Pen).

TABLE 20.1 Structure, Affinity to CB_1 Cannabinoids Receptors and Lipophilicity of Imidazolinediones[4]

Compound	R1	*n*	R2	Percentage of displacement at 10 μM	Lipophilicity (log P)*
1	H	2	N-morpholine	<5	2.16
	CH_3	2	N-morpholine	24	3.49
2	H	5	—CH_3	36	5.37
	CH_3	5	—CH_3	47	6.48
3	H	0	—$(CH_2)_2$—CH_3	<20	3.84
	H	0	—$CH(CH_3)_2$	<5	2.89

*The lipophilicity was calculated using the CLIP method[5].

TABLE 20.2 Solubility in Water at 20°C of *n*-Butanol **1**, Isobutanol **2**, *tert*-Butanol **3**, *n*-Pentanol **4**, 2-Pentanol **5** and Neopentanol **6**[6]

No.	Compound	Solubility (g/100 g H_2O)
1		8.2
2		5
3		Miscible
4		2.4
5		4.9
6		12.2

In these derivatives, the chain folds in such a way that the cationic head becomes placed under the aromatic ring and can establish a typical donor–acceptor interaction with the π cloud of the aromatic ring (folding effect).[8]

FIGURE 20.2 A lesser amount of structured water molecules is needed to wrap a compact molecule (2,2,3-trimethylbutane) than to wrap an extended one (*n*-heptane).

3. *Crystal lattice cohesion*

Greater water solubility can also result from a decrease of the crystal lattice energy, the methyl groups hindering the various intermolecular interactions (hydrogen bonds, dipole–dipole bonds, etc.). In the antibacterial sulfonamide series, the substitution of the pyrimidine ring of sulfadiazine by one, then two, methyl groups causes an increase in solubility (Table 20.3).[9] *A priori*, one would expect why the methyl substituted derivatives are less soluble. This for the double reason that they show increased lipophilicity and that they are less dissociated than the parent molecule.

Indeed the inductive character of the methyl groups disfavors ionization and the non-ionized form of a molecule is always less soluble than the corresponding ionized form. Despite this unfavorable electronic effect, sulfamidine is approximately five times more soluble than sulfadiazine. Similarly, the grafting of only one methyl group to the

TABLE 20.3 Increased Water Solubility Caused by Insertion of Methyl Groups[9]

R1	R2	Drug	pK (acidic)	Percent ionized at pH 5.2	Solubility pH 5.2–37°C (M)
H	H	Sulfadiazine	6.5	3.9	0.0005
CH_3	H	Sulfamerazine	7.1	1.4	0.0013
CH_3	CH_3	Sulfamidine	7.4	0.7	0.0024

FIGURE 20.3 The presence on diphenhydramine of an *ortho*-methyl group prevents the side chain to adopt the favorable coplanar conformation as found in phenindamine.[11]

herbicide simazine provides atrazine that is 14 times more soluble in water.[10]

B. Conformational effects

The steric hindrance generated by a methyl group can create constraints and impose particular conformations that may be favorable or unfavorable for ligand–receptor interactions.

Harms and Nauta have studied the effects of methyl substitution on the aromatic ring of the spasmolytic diphenhydramine.[11] The presence of a methyl in *para* position corresponds to a 3.7-fold increase in antihistaminic activity compared to the non-substituted derivative (Figure 20.3). Conversely, the presence of a methyl in *ortho* position inactivates the molecule (one fifth of the activity of the non-substituted derivative).

The explanation proposed by the authors is as follows: the methyl group in *ortho* position would prevent the side chain to adopt the usual "antihistaminic" conformation such as found for phenindamine, for example. Curiously the *ortho–ortho′*-disubstituted analog of diphenhydramine shows local anesthetic properties (40 times those of diphenhydramine).

In the same way, the activity of histamine on the H_1 receptor is directly correlated with its conformation. That is why the presence of a methyl group in position 4, by modifying the orientation of the imidazole ring relative to the side chain, decreases dramatically the potency of 4-methylhistamine (Table 20.4) on H_1 receptor (400-fold less than histamine) whereas its potency on the H_2 receptor is almost unchanged.[12] Indeed, a different chemical property is needed at the H_2 receptor: namely the tautomeric property of the imidazole ring to act as a proton-transfer agent (see Section II.C.).

In steroids the two angular methyl groups in position 18 and 19 stand on the surface and form a screen above the β face. This entails selective attacks on the rear face (α face) of the molecule.[14]

Besides, the presence of the methyl in position 18 imposes a preferential conformation to the methylketone chain placed in position 17.[14]

The antihypertensive imidazoline clonidine (Figure 20.4; R1 = R2 = Cl) and its analogs activate norepinephrine

TABLE 20.4 H_1 and H_2 Receptor Agonist Activity of Histamine Derivatives[13]

R1	R2	Derivative	H_1 receptor*	H_2 receptor*
H	Me	4-Methylhistamine	0.23	39.0
Me	H	2-Methylhistamine	16.5	2.0

*Agonist activity relative to histamine (=100).

FIGURE 20.4 The restricted rotation resulting from *o*- and *o′*-substitution imposes a quasi-perpendicular orientation of the imidazolinic ring toward the phenyl ring.[15]

as well as specific receptors of the central nervous system (CNS). The maximal activity in this series is always observed when both the *ortho* positions are substituted (R1 = R2 = methyl, chlorine or ethyl, etc.). This situation implies a restrained rotation of the atropisomery type and the impossibility for the two cycles to lay in a coplanar situation. Correspondingly, the geometry of the molecule becomes close to that of the norepinephrine.[15]

An interesting example of the effects induced by the presence or the absence of a simple methyl is found in two non-peptide angiotensin AT_1 ligands (Table 20.5). Both bind with high affinity to the AT_1 and stimulate phosphatidylinositol turnover. However, L-162,782 acts as a powerful partial agonist with a bell-shaped dose–response of 64% of the maximal level reached with angiotensin II. Compound L-162,389 only gave a response of 6% which characterized it an antagonist. The authors speculated that the receptor/ligand complex of these compounds after binding is able to change between inactive and active conformations.[16]

C. Electronics effects

The methyl group and, more generally all alkyl groups, are the only substituents acting by an inductive electron-donating effect. All the other groups are electron donors by mesomeric effects. This means that the methyl and the alkyls are electron donors in any environment while a basic group, dimethylaminoethyl for example, will be a mesomeric donor in basic or neutral medium, but will become strongly electron attracting by protonation in gastric medium (pH ≈ 2). Table 20.6, taken from Chu,[17] presents some numerical values for substituents commonly met in medicinal chemistry.[18] Hansch's π constant accounts for the contribution of lipophilicity, Hammett's σ constants reflect the electronic effects and the molar refraction (MR) is related to the volume of the substituent. The table illustrates clearly the dramatic change in Hammett's σ parameter when passing from a free amino group ($\sigma = -0.66$) to a protonated one ($\sigma = +0.60$).

A practical consequence is that it is always judicious to include a methyl (or an alkyl) group in a structure–activity relationships (SAR) study. Thus in any series of R-substituted molecules, when one wants to vary R, the methyl group is generally chosen as a representative of an electron-donating group, the second substituent being chosen among the electron attractors (Cl, CN, NO_2, CF_3, etc.).

As we have seen previously (Table 20.4), the activity of histamine on H_1 receptor is profoundly altered by methylation.

This is also the case concerning the H_2 receptor. To be active on this receptor, the aqueous solution of histamine derivatives must be in a monocationic form, which is considered most likely to be the physiologically active form. We now know that in aqueous solution at pH 7.4, the prevalent tautomer of histamine is the N_3H (Figure 20.5).[19,20] Indeed, the imidazole ring must not be protonated to act as a proton-transfer agent.[12] However, the presence of a methyl group on the imidazole ring, by its electron-donating effect, has a noticeable effect on the degree of protonation of the imine nitrogen N1, giving to it a higher proton affinity. Thus, at pH 7.4, there are more diprotonated species (the non-active form) in solution for 2-methylhistamine and 4-methylhistamine than for histamine (99.85%, 87.7% and 3.4%, respectively). This explains their relative weak activity compared to histamine (Table 20.4).[13]

TABLE 20.5 Dual Agonist and Antagonist Nonpeptide Ligands of the Angiotensine AT1 Receptor[16]

Compound	R	IC_{50} human AT1 wild-type (nM)	Maximal response obtained during stimulation
L-162,389	H	3.97 ± 0.69	5.8 ± 1.3%
L-162,782	Me	24.6 ± 2.3	64 ± 3%

Note: AII = 100%

TABLE 20.6 Some Common Aromatic Substituent Constants[17]

Group	π_{para}	σ	MR
H	0.00	0.00	1.03
CH_3	0.56	−0.17	5.65
CF_3	0.88	0.54	5.02
Cl	0.71	0.23	6.03
OH	−0.67	−0.37	2.85
OCH_3	−0.02	−0.27	7.87
NH_2	−1.23	−0.66	5.42
NH_3^+	–	0.60	–
NO_2	−0.28	0.78	7.36
CN	−0.57	0.66	6.33
CO_2H	−0.32	0.45	6.93
$COCH_3$	−0.55	0.50	11.18

FIGURE 20.5 Prevalent species in aqueous solution at pH 7.4.[13]

D. Effects on metabolism

Seen from the metabolic point of view the methyl group plays a particularly important role. Three possibilities are currently met: (a) the methyl group is oxidized, (b) the methyl group is shifted and (c) the methyl group is not (or only slightly) attacked and can then serve as blocking group.

1. *Oxidation of the methyl group*

The oxidation of the methyl group begins generally with the formation of the hydroxymethyl analog and continues usually until the carboxyl step. This is observed for simple compounds like camphor or 2-methyl-pyridine but also for drugs like tolbutamide and alpidem, explaining the relatively short half-life of these latter compounds.

Sometimes the oxidation of the methyl group gives rise to an active metabolite, contributing thus to a reasonable half-life to the drug (Figure 20.6).[21]

The grafting of a methyl group, especially on aromatic rings, represents often a good mean of detoxification. It is rapidly oxidized to an inactive and easy-to-eliminate carboxylic group. When the grafted chains are longer than methyl, the attack takes place rather at the benzylic position, at position ω-1 or on ramifications (Figure 20.7).[22]

Angular methyl groups of steroids are usually resistant to metabolic oxidation, probably in relation with a local steric hindrance.

2. *The methyl group is shifted*

A methyl group, when grafted on a nitrogen or sulfur atom, can transform this latter in an "onium," able to act as methyl donor. In living organisms the usual suppliers of methyl rests are choline and methionine. Methionine is first activated *in vivo* by combination with adenosine to yield *S*-adenosyl-methionine (SAM; Figure 20.8).

More generally, any *S*- or *N*-methylated drug can *a priori* constitute a methyl donor. On the other hand, when the methyl (or alkyl) rest is linked to a good leaving group, as found for alkyl sulfates or sulfonates such as methyl sulfate or busulfan, alkylating reagents are produced and there exists a huge risk of carcinogenicity.

FIGURE 20.6 Metabolic pathway of tolderodine in mice and dogs, leading to its active metabolite 5-HM.[21]

FIGURE 20.7 Privileged oxidative attacks of long chains.

FIGURE 20.8 Methyl donors and alkylating compounds: choline **1**, methionine **2**, *S*-adenosylmethionine **3**, dimethyl sulfate **4** and butane-1,4-diol bis-methane sulfonate **5** (busulfan).

FIGURE 20.9 The methylation of the ene-diol function of ascorbic acid leads to a chemically stable, but pharmacologically inactive compound.

3. *The methyl serves to block a reactive function*

A reactive function, such as an active hydrogen belonging to a hydroxyl, thiol or amino, can be masked by methylation. Methyl groups can thus serve to protect sensitive functionalities from metabolic hydroxylation.

The ene-diol function is essential to the antioxidant properties of vitamin C, it is therefore not surprising that its methylation leads to an inactive compound (Figure 20.9).

As such, the endogenous peptides, methionine- and leucine enkephalin are inactive by the oral route. Starting from Met-enkephalin, Roemer *et al.* prepared a less vulnerable analog of methionine enkephalin (Tyr-D-Ala-Gly-*N*-Me-Phe-Met(O)-ol), with prolonged parenteral and oral analgesic activity[23] (Figure 20.10). Several modifications were needed: replacement of glycine by the unnatural D-alanine, *N*-methylation of the Gly-Phe amide bond, oxidation of methionine to the sulfoxide, and reduction of the C-terminus to the corresponding alcohol.

For other examples starting from peptide leads, see the excellent reviews of Plattner and Norbeck[24] and of Fauchère.[25]

FIGURE 20.10 Less vulnerable analog of Met-enkephalin.[23]

In steroids the 6α-position (e.g. prednisolone, Figure 20.11) is a position that is normally hydroxylated. Grafting a methyl in this place prevents its hydroxylation. Halogens (particularly fluorine) suit even better because they are not sensitive at all to oxidative attacks.

Allopregnanolone (3α-hydroxy-5α-pregnan-20-one) has anticonvulsant properties. However, this steroid is not orally active and its first metabolite is the hormonally active 3-keto-5α-pregnan-20-one. Ganaxolone is a 3β-methylated analog of allopregnanolone. Its 3β-methylation has two effects: (i) it prevents the metabolic attack of the 3α-hydroxy function, (ii) it enhances the bioavailability of the pregnane steroids.[26] Ganaxolone is presently in phase IIb study as antiepileptic (Figure 20.11).[27]

FIGURE 20.11 Protection of steroids against metabolic attacks.

BOX 20.1 The Morphine family[28]

Morphine was first isolated in 1804 by both Seguin and Courtois. But it is the German pharmacist Sertürner who first published its results in 1805. He called this white powder "morphium," because its effects point out the Greek god of dreams Morpheus. Structure–activity relationships of morphine analogs illustrate well how simple methyl group can change the pharmacological profile.

- The replacement of the methyl by an hydrogen reduces the potency. This can be explained by the fact that the secondary NH obtained in compound 2 is more polar and therefore crosses less easily the blood-brain barrier (BBB).
- Codeine (compound 3), which differs from morphine by a single methyl, is used for treating moderate pain, coughs and diarrhea. The *in vitro* assays on isolated receptors indicate that codeine should be 1,000 times less active than morphine. But when codeine is given orally to patients, it is only five times less active. This difference between *in vitro* and *in vitro* is due to a demethylation of codeine in the liver, the removal of the R1 methyl group leading to morphine.
- In compound 4, heterocodeine exhibits the highest potency of this table. One could conclude that a substitution on R2 increases activity, but we must remember that these tests have generally been done *in vitro*. In reality, the improvement in activity is more due to the modification of the pharmacokinetic properties of the compound than to its intrinsic affinity for the receptor. Indeed, the methylation of the OH gives a more lipophilic drug, with better the BBB crossing and thus, higher morphine concentration in the CNS.
- And now a few words about the acetylated analogs of morphine: if 3-acetylmorphine has a weaker activity (proof of the importance of the free phenolic OH) than morphine, 6-acetylmorphine and diamorphine (heroin) have an increased activity. For both, it is due to the fact that they are less polar than morphine, and enter the brain more quickly and in higher concentration. But whereas 6-acetylmorphine acts immediately on the receptor, the R1 acetyl group of diamorphine has to be removed before. This explains the difference of activity between compounds 6 and 7.

Compound	Name	R1	R2	R3	Analgesic activity compared to morphine
1	Morphine	H	H	Me	100
2	Normorphine	H	H	H	25
3	Codeine	Me	H	Me	20
4	Heterocodeine	H	Me	Me	500
5	3-Acetylmorphine	Ac	H	Me	<10
6	6-Acetylmorphine	H	Ac	Me	400
7	Diamorphine	Ac	Ac	Me	200

FIGURE 20.12 Structure of 1,5-benzodiazepine active on cholecystokinin receptors.[29]

TABLE 20.7 Aromatic Lipophilic, Electronic and Steric Descriptors for Some Current Aliphatic Rests[30]

Group	π (Hansch)	σ (Hammett)	E_s (Taft)
Methyl	0.50	−0.17	0.00
Isopropyl	1.30	−0.19	−1.08
Cyclopropyl	1.20[a]	−0.30[b]	–
Cyclobutyl	1.80	−0.20	−0.67
Tertiobutyl	1.98	−0.30	−2.46
Cyclopentyl	2.14	−0.20	−1.12
Cyclohexyl	2.51	−0.15	−1.40

[a]Taken from Ref. [31].
[b]Taken from Ref. [32].

E. Extensions to other small alkyl groups

The methyl group is the prototype of a saturated aliphatic substituent with lipophilic and electron-donor inductive effect. In some instances, it can advantageously be replaced by related groups bringing either symmetry, or more lipophilicity, or an increased inductive effect. We remind below some possibilities.

Keep in mind, however, that modifications, even by alkyl groups, can dramatically affect the properties of a substance, as found some 1,5-benzodiazepine where hydrogen and methyl substituents yield weak anatagonists of cholecystokinin whereas ethyl, propyl and butyl substituents yield agonists (Figure 20.12).[29]

1. Numerical values

The values reported in Table 20.7,[30] allow the comparison of some characteristic alkyl groups in aromatic substitution. One will note the comparable bulkiness (E_s) of the isopropyl and cyclopentyl groups, while the tertiobutyl group is far more voluminous. Furthermore, it is remarkable to observe that the electron-donor effect of the cyclopentyl group is superior to that of the cyclohexyl group.

2. Gem-dimethyl and spiro-cyclopropyl

Gem-dimethyl and spiro-cyclopropyl are useful to render a carbon atom quaternary and therefore resistant to metabolic attacks.

Starting form retinoic acid (RA), the active vitamin A metabolite, one of the first constrained analogs, TTNPB, was synthesized (Figure 20.13). The incorporation of a second *gem*-dimethyl group into the tetrahydronaphthalene ring aimed to block its oxidation of the ring. The resulting compound is 10-fold more potent in the tracheal organ culture (TOC) than all-*trans*-RA[33] but also 10,000 more toxic…[34] However, other less toxic retinoids, are used in chemoprevention. Bexarotene in combination with chemotherapeutic agents has demonstrated interesting results with advanced non-small cell lung cancer. Two phase III trials are currently under way to fully characterize the role of bexarotene in the treatment of these cancers.[35]

Gem-dimethyl can also constitute solutions to introduce symmetry into a chiral center, or to protect a close and sensitive function, as in the case of gemeprost, an analog of prostaglandin E_1 used in medical abortion, where the *gem*-dimethyl groups at C-16 protect the alcohol moiety at C-15 from rapid metabolic oxidation.[36]

3. Isopropyl and cyclopropyl

The cyclopropyl rest is less bulky than the isopropyl group for a maximal electron-donor effect. This electronic effect is involved when the cyclopropyl group of efavirenz, a non-nucleoside reverse transcriptase inhibitor, interacts with the aromatic ring of tyr181 via a π-aryl interaction which is presumably favorable to binding.[37]

The lipophilic effect of cyclopropyl explains why abacavir, a nucleoside reverse transcriptase inhibitor, has an improved absorption in the CNS compared to diaminopurine dioxolane (DAPD) (Figure 20.14).[38]

For a review on cyclopropane derivatives in medicinal chemistry, see Cussac *et al.*[40]

4. The cyclopentyl group

The cyclopentyl group creates the maximal inductive effect for a relatively reasonable bulkiness. It is often a good filling of a hydrophobic pocket as illustrated for the cAMP phosphodiesterase inhibitor rolipram (Figure 20.15). The inhibitory activity toward type IV cAMP-phosphodiesterase is increased 10 times when the *meta*-methoxy group is replaced by a *meta*-cyclopentyl group (rolipram).[41]

FIGURE 20.13 *Gem*-dimethyl derivatives.

FIGURE 20.14 Reverse transcriptase inhibitors.[39]

FIGURE 20.15 Structures of rolipram and of its dimethoxy analog.[41]

Presumably the cyclopentyl rest fills, in an optimal manner, a hydrophobic pocket of the active site of the enzyme. The cyclopentyl group has also proven advantageous in replacing a *gem*-dimethyl in a series of inhibitors of acyl-CoA-cholesterol acyltransferase, which is an enzyme implied in the absorption of the alimentary cholesterol.[42]

III. EFFECTS OF UNSATURATED GROUPS

The introduction of an unsaturated group (vinyl, ethynyl, allyl, etc.) in a drug molecule generally entails one or several of the following consequences.[43,44]

1. *Existence of electronic effects*: the unsaturated rests behave as electron attractors through inductive effects. Furthermore, direct interactions of donor–acceptor type are possible thanks to the π electron cloud surrounds present in multiple bonds.
2. Possibility of existence of a geometrical isomery (e.g. *cis–trans* geometric isomery).
3. *Possibility of activation through conjugation*: the association of several unsaturated functions in conjugated position (dienes, enynes, enones, enolides, polyunsaturated derivatives) renders the corresponding molecules very reactive. It facilitates especially the addition of biological nucleophiles and notably of thiols.

FIGURE 20.16 Medicines containing a vinyl group.

4. *Facilitation of the metabolism*: The unsaturated element constitutes often the vulnerable site of the molecule, that will be attacked first (e.g. by formation of an epoxide that evolves into a diol that, on its turn, can undergo oxidative cleaving), but this is not always the case. Therefore, one should pay attention to the problems posed by the formation of these metabolites (aldehydes, carboxylic acids …): they can also be biologically active.
5. Increase of the narcotic power and the toxicity in comparison with the corresponding saturated compound. Ethylene, acetylene, trichlorethylene, divinyl oxide and, by extension, cyclopropane are examples of unsaturated narcotics.

With regard to their classification, we will distinguish four series of unsaturated derivatives: the vinyl series, the allyl series, the acetylenic series, and the ring-unsaturated derivatives that are bioisosteric to aromatic rings.

A. Vinyl series

Beside active substances containing actual vinyl rests, this series comprises molecules containing substituted vinyl groups as well as cyclopropyl groups.

1. *Vinyl groups*

They are not excessively used in medicinal chemistry. Divinyl oxide is an excellent general anesthetic but it polymerizes easily and forms peroxides. Stabilization of the compound is usually achieved by addition of 0.01% of *N*-phenyl α-naphtylamine. On the other hand, compounds such as kainic acid, vinylbital, quinine, 17α-vinyl-testosterone, compound SKF 100 047, and vigabatrin (Figure 20.16) are perfectly stable vinyl derivatives.

TABLE 20.8 Inhibitory Effects of Nucleoside Derivatives on Human T-lymphoblast Molt/4F and MT4 Cells[46]

Nucleoside	R	Molt/4F IC_{50} (μM)	MT4 IC_{50} (μM)
Uridine	Vinyl	14 ± 0.12	11
	Ethynyl	>372	286 ± 78
Adenosine	Vinyl	6.5 ± 0.1	15 ± 5
	Ethynyl	25 ± 0.1	22 ± 7

Nowadays, several literature reports about the interest of vinyl groups. Indeed, the introduction of alkyl, akenyl and alkynyl groups into purine and pyrimidine nucleosides is of great interest for their potential activity.[45] As an example, the introduction of vinyl and ethynyl on uridine and adenosine shows a interesting inhibitory effect toward human tumor cell lines (Table 20.8). According to the authors, these groups induce the opposite conformation of the glycoside bond if compared to the natural nucleoside, which made them more cytostatic than the natural ones.[46]

Rancourt *et al.*[47] have shown some interesting results with aminocyclopropane carboxylic acid (ACCA) derivatives

FIGURE 20.17 Recent studies on vinyl-containing compounds.[47,48]

FIGURE 20.18 Tranylcypromine represents a stable substitute of the enamine aminostyrene.

as inhibitors of the hepatitis C virus NS3 protease. The replacement of hydrogen by ethyl group affords a modest 3-fold improvement of potency (from $IC_{50} = 14$ to 4.8 μM). But the modification of the ethyl group to a vinyl group allow once again a 7-fold increase (from $IC_{50} = 4.8$ to 0.63 μM) (Figure 20.17).This difference may be due in a part to a beneficial electronic interaction between the π-electrons of the vinyl group and those of the phenyl group of the near Phe154.

The incorporation of vinyl groups into antibacterial C12 ketolides has a favorable impact on the pharmacokinetic and pharmacodynamic properties: increase of the lung to plasma AUC ratio, of the bioavailability, and half-life in plasma and lung. These properties directly impact on the *in vivo* potent efficacy in rat lung infection models.[48]

2. *Cyclopropyl groups*

Cyclopropyl rings can constitute interesting substitutes for vinyl rests when they are too fragile or when they give place to unwanted isomeries or tautomeries. Thus tranylcypromine, an antidepressant acting by inhibition of the monoamine-oxidases (MAO), is a stable compound, while its ethylenic analog no longer is (Figure 20.18). A supplementary advantage in the use of cyclopropanic analogs comes from their fixed stereochemistry: there is no spontaneous conversion from *cis* to *trans* isomer as frequently observed with ethylenic derivatives.

B. Allylic series

All allylic derivatives are relatively hepatotoxic and irritant. Allylic alcohol itself serves to create experimental hepatic lesions that allow testing hepatoprotecting drugs. We will envisage three categories of allylic derivatives: *C*-allyl derivatives, *N*-allyl derivatives, and *O*-and *S*-allyl derivatives which often possess alkylating properties.

1. *C-allyl derivatives*

They present the double advantage to be lipophilic (rapid onset) and to give place to fast biodegradation (short duration of action). However, they often conserve the intrinsic hepatotoxicity of the allyl group.[49] Allobarbital is a sedative-hypnotic that is no longer used; allylestrenol acts as a pure progestative hormone and alprenolol is a β-blocker (Figure 20.19).

Acetamidoeugenol is an intravenous anesthetic of ultra-short duration of action, it has been withdrawn because it provokes irritations and lesions of the vascular wall. Acetamidoeugenol is oxidized very rapidly *in vivo* into the corresponding aryl-acetic acid. This observation was the basis of the synthesis of another intravenous short-acting anesthetic such as propanilide.[50]

Compound AL-438, a non-steroidal selective glucocorticoid modulator, is a potent anti-inflammatory agent (in the same range as prednisolone). This 5-allyl compound

FIGURE 20.19 *C*-allyl derivatives.

FIGURE 20.20 Nalorphine and cognate derivatives.[52]

has low side effects compared to other glucocorticoid treatment with no increase of glucose level, and no effects over bone formation rates. Compound AL-438 is considered as a lead compound, and optimization is still ongoing.[51]

2. N-*allyl derivatives*

The replacement in morphine, and in some of its simplified analogs, of the *N*-methyl group by a *N*-allyl group (and, later on, by some related groups)[52] has constituted a decisive step in the study of opiate analgesics. Indeed this modification had for the first time achieved the passage of morphinic receptor agonists to the corresponding antagonists (Figure 20.20).

Aloxidone and albutoïne are anticonvulsivant *N*-allyl derivatives of hydantoin and thiohydantoin (Figure 20.21). The dibenzazepine azapetine is an α-adrenergic blocking agent used as peripheral vasodilator.

3. O- *and* S-*allyl derivatives*

Several of these compounds are listed in the Merck Index. The β-blocking oxprenolol, the arylacetic analgesic-anti-inflammatory drug aclofenac and the fungicide enilconazole are *O*-allyl derivatives. Penicillin O and penicillin S are both *S*-allyl derivatives.

Alkylating allyl derivatives: When the allyl rest bears a good leaving group, it generates easily the allylic cation. This cation is stabilized by mesomery, and is an excellent electrophile. Many natural compounds can release allylic alcohols. A first example is found in allicine, the antibacterial principle of garlic, which results from the action of alliinase on alliine (Figure 20.22). Several varieties of senecio (*Senecio vulgaris* L., Compositae, *Senecio platyphyllus* etc.) contain alkaloids derived from pyrrolizidine: senecionine, seneciphylline, etc. These substances provoke hepatic cancers, notably in cattle. Here too, the alkylating properties are to put to the account of the allylic structure

FIGURE 20.21 *N*-allyl derivatives of hydantoin, thiohydantoin and dibenzazepine.

FIGURE 20.22 Alkylating allylic derivatives.

(Figure 20.22)[53] It is reasonable to admit the formation of allylic carbamates also in mitomycines A, B and C which are antimitotic drugs used in cancerology. Presumably the allyl function is created by an elimination reaction involving the departure of an acetalic hydroxy or methoxy group.

C. Acetylenic series

Acetylenic groups are used for their electronic effects, as equivalents of aromatic rings and to impose structural constraints.

1. *Electronic effects*

The acetylene function exerts an electron-attracting effect. This effect can be reinforced by substitution of the acetylenic hydrogen. Ethynyl compounds are essentially found among the light sedative-hypnotic drugs (CNS-depressing effect of the unsaturated derivatives) where, most of the acetylenic alcohols are used as carbamic esters: meparfynol, ethinamate, etc (Figure 20.23).

The acetylenic CH can act as an hydrogen bond donor (like for the antiparkinsonian rasagiline), and thus allow an additional interaction with the receptor.[54]

In steroid series, the presence of ethynyl groups in position 17α provides orally active compounds then metabolized, by hydration of the triple bond, to 17α-methylketones (Figure 20.24).

Because of the potential biologic activity of ethynyl group, it should not be systematically avoided during SAR exploration, even if mono-substituted ethynyl derivatives can sometimes be highly cytotoxic.[55]

2. *Aromatic ring equivalents*

Thanks to their π electron clouds and to their small volume, ethynyl groups can sometimes function as bioisosteres of aromatic rings and give similar donor–acceptor interactions.

An X-ray structure of ABT-279 (Figure 20.25), an inhibitor of dipeptidyl peptidase-IV (DPP-IV), bound in the human active site shows an interesting feature: Tyr547 and Phe357 of DPP-IV formed a narrow tunnel between their

FIGURE 20.23 Ethynylated sedatives.

FIGURE 20.24 Ethynylated steroids.

FIGURE 20.25 Acetylenic groups as aromatic ring equivalents.

phenyl rings that accommodates the ethynyl group. This suggests a π–π stacking effect between the triple bond and the aromatic systems of the receptor. Further investigations showed good pharmacokinetics and excellent preclinical safety profile of the ethynyl compound, allowing ABT-279 to be selected as a candidate for clinical evaluation in humans for the treatment of type 2 diabetes.[56]

Lee *et al.* also supposed a stacking effect for a 6-ethynyl quinazoline. This additional interaction with the receptor increased five times the potency of this compound when compared to the 6-chloro analog. However, this acetylenic analog is reactive and rapidly metabolized, giving for example Michael addition in a biological system, and thus led the authors not to test further this compound.[57]

Another example of the rapid metabolization of ethynyl groups with sodium methohexital which is used as an injectable barbituric for very short anaesthesias.

3. *Structural constraints*

In inserting an acetylenic function between two carbon atoms, one achieves a structure with four "on-line" atoms representing a rigid entity with a distance of 4.2 Å between the two extreme atoms (Figure 20.26). This kind of arrangement is found in the cholinergic agonist oxotremorine and in the γ-aminobutyric acid (GABA) analog 4-amino-tetrolic acid. This compound is recognized, like GABA itself, by the enzyme GABA-transaminase for which it acts as an inhibitor.[58]

FIGURE 20.26 Rigidity and extension imposed by a triple bond.

FIGURE 20.27 Cyclenic equivalents of the phenyl ring.

A structural constraint is also present in the antimycotic terbinafine.[59] SAR exploration led to a simple pharmacophore model A in which the *trans*-enyne group achieves the best length spacer between the lipophilic domain (L2) and the polar group (P).[60]

D. Cyclenic equivalents of the phenyl ring

The cyclohexenyl ring and, to a lesser extent the cyclopentenyl and cycloheptenyl rings can possibly replace a phenyl ring. This is the case for the barbiturics cyclobarbital and heptabarbital which are entirely comparable to phenobarbital. From the metabolic point of view, the cyclohexenyl ring is oxidized in position α to the double bond to produce the corresponding cyclohexenone (Figure 20.27).

Another example comes from the benzodiazepine series where tetrazepam can be compared to diazepam. However, in this case a slight difference in the activity profile exists: tetrazepam is less sedative, hypnotic and anticonvulsant than diazepam, on the other hand it has more muscle relaxant and analgesic effects, therefore come its indication in visceral and particular pain. From the chemical point of view, the cyclohexenic double bond is introduced in a rather unexpected manner by means of a radicalar rearrangement of a *N*-chloroamide (Figure 20.28).[61]

FIGURE 20.28 Cyclenic equivalents in the benzodiazepine series.

No.	Hypotensive activity ED_{20} (mg/kg)
1	0.01
2	3.00

Affinity for the benzodiazepine receptor

R	IC_{50} (μM)
o-Cl	70.00
m-Cl	3.90
p-Cl	0.56

FIGURE 20.29 (a) The *ortho–ortho′* substitution in clonidine maintains the planes of the aromatic rings in a perpendicular position to each other and (b) The *ortho*–chloro isomer of the benzodiazepine ligand CGS 9896 has a 125-fold lower affinity than the parent molecule.

IV. EFFECTS OF HALOGENATION

Presently one drug out of three is a halogenated derivative and halogens are found in drugs belonging to practically all therapeutic classes. That has not always been the case, indeed, in the past medicines were mostly of natural origin and the natural substance chemistry is relatively deprived of halogenated substances. Halogen-containing drugs have entered in usage only after 1820. The first organic halogenated drugs have mainly been used for their depressive action on the CNS: production of a general anesthesia with chloroform, sedation or hypnosis with chloral and bromural. From 20th century onward a regular growth of the number of halogenated drugs had been observed and it became explosive at the end of the World War II. Even halogenated drugs from natural origin became available, such as chlortetracycline or chloramphenicol, and also substances from marine origin or from fermentation broths.

A. The importance of the halogens in the structure–activity relationship

1. *Steric effects*

The obstruction of a molecule by means of halogen substitution can impose certain conformations or mask certain functions. In the case of clonidine the bulky halogen atoms prevent the free rotation and maintain the planes of the aromatic rings in a perpendicular position to each other (Figure 20.29a).[15]

Also in a series of benzodiazepine receptor ligands derived from CGS 9896, strong steric effects are described by Fryer *et al.* (Figure 20.29b).[62] Indeed, the *ortho* and the *para* isomers can be considered as having the same lipophilicity and very similar electronic effects. Thus the reduction of binding affinity of the *ortho*-chloro compound is attributed to the steric effect.

TABLE 20.9 van der Waals Radii[63]

	Radii in Å		Radii in Å
H	1.20	O	1.52
F	1.47	S	1.80
Cl	1.75	CH_3	1.80
Br	1.85	CF_3	2.20
I	1.98	$CH(CH_3)_2$	2.20

The halogen Van der Waals radii give a good idea of the size of each atom or groupment. In Table 20.9, the values proposed by Bondi[63] show that the size of fluorine is close to the size of oxygen. Chlorine, bromine, sulfur and methyl group are close to each other, and due to the length of the C—F bond (1.39 Å), trifluoromethyl is equal to isopropyl.

But there is a difference between the steric size of an atom or a groupment (the absolute size) and its steric effect (the volume in which its electronic effects are sensible).

2. *Electronic effects*

The electronic effects of the halogens are to ascribe to their inductive electron-attracting properties. These latter are maximal for chlorine and bromine, less marked for iodine, and very weak for fluorine. The mesomeric donor effect of the halogen atoms is usually not involved in biological medias. The influences of halogens on the potency of MAO inhibition and of dopamine uptake blockade *in vitro* are shown in Figure 20.30. The choice of the optimal substituent allows noticeable gains in potency, compared to the parent molecule.[64,65]

Progressive mono- and di-substitution of diazepam-related benzodiazepinones enhances the affinity for the mitochondrial benzodiazepine receptor (MBR) by a factor of 233 (Figure 20.31).[66]

Monoamine oxidase inhibition

X	IC_{50} (nM)
H	1,200
Br	200
CF_3	100
SO_2CF_3	27

[^{3}H] dopamine uptake

R1	R2	IC_{50} (nM)
CH_3O	CH_3O	2876
H	Cl	115
Cl	Cl	75

FIGURE 20.30 Influence of halogenated substituents on MAO inhibition potency *in vitro*.

Compound	Substitution pattern 2′	4′	7	IC_{50} nM MBR
Ro 5-3464	H	H	H	700
Ro 5-5115	H	Cl	H	54
Diazepam	H	H	Cl	72
Ro 5-6900	Cl	H	Cl	11
Ro 5-4864	H	Cl	Cl	3

FIGURE 20.31 Chlorine effects in the benzodiazepine series.[66]

Human NK_1 receptor binding: IC_{50} (nM)

FIGURE 20.32 Successive *N*-acetylation and CH3 → CF3 replacement achieve a 958-fold increase in affinity.[66]

N-acylation of L-tryptophan benzyl esters, followed by replacement of the 3,5-dimethyl groups by their trifluoromethyl analogs, achieved an almost 1,000-fold increase in potency in a series of substance P receptor antagonists (Figure 20.32).[66]

3. *Electrostatic similitude*

In certain active molecules, the role of the fluorine or chlorine atoms is not apparent at first glance. For example, two compounds such as *m*-trifluoromethyl-phenylethylamine and 5-hydroxy-tryptamine, which are chemically different show many pharmacological similarities. In this case, the explanation lays in the similitude of the electrostatic potential maps (Figure 20.33).

Conversely two closely related pyrazoloquinolines, compounds CGS 8216 and its *para*-chloro analog CGS 9896 present a totally opposed activity profile on the same benzodiazepine receptor.[67] A dramatic effect resulting from chlorine substitution is also found in the change from β-phenyl-GABA to β-(*p*-chlorophenyl)-GABA.[68]

4. *Hydrophobic effects*

The predominantly lipophilic influence of halogen substitution is seen in the classical cases of the halocarbon

anaesthetics,[69] the halogenophenol antiseptics and the halogenated insecticides (Figure 20.34).

For these compounds there is a direct correlation between biological activity and certain physicochemical parameters such as partition coefficient, surface tension or vapor pressure. The accumulation of halogen atoms favors the passage of the biomembranes and the access to the CNS. Hansch has defined a parameter, π, representing the hydrophobic contribution of a substituent. It measures the individual participation of each group to the partition coefficient of the molecule (see Chapter 23). Table 20.10 shows π values for several substituent groups on bis-substituted phenyl rings.

As one can see, even if the participation of fluorine is modest compared to hydrogen, introducing halogens like chlorine (Figure 20.35),[69] bromine or a trifluoromethyl group induce a important lipophilic contribution (larger than the contribution of a methyl group).[70]

FIGURE 20.33 Electrostatic potential maps of *m*-trifluoromethylphenylethylamine and 5-hydroxy-tryptamine (Generated with DS Viewer Pro 6.0).

TABLE 20.10 Selected Values of Hansch's π Parameter[70]

X	π	X	π
H	0.00	OH	−0.67
F	0.14	OCH_3	−0.02
Cl	0.71	OC_2H_5	0.38
Br	0.86	$COCH_3$	−0.55
CH_3	0.56	CF_3	0.88
C_2H_5	1.02	NH_2	−1.23

Compound	log *P*	Compound	log *P*
$C_2H_5—O—C_2H_5$	0.89	$CHBr_2Cl$	2.16
$CHCl_3$	1.97	$CHBr_3$	2.40
$CHBrCl_2$	2.00	$CF_3CHBrCl$	2.30

The anaesthetic activity increases in the same order

FIGURE 20.34 Compounds in which the halogens play essentially a lipophilic role.

FIGURE 20.35 Chlorophenols experimental log *P* values.[69]

5. *Reactivity of the halogens*

In terms of bond strength, all C-halogen bonds are weaker than the C—H bond (Table 20.11), except for the C—F bond, due to the high electronegativity of the fluorine and an orbital size similar to that of carbon.[71]

B. Usefulness of the halogens and of cognate functions

Depending on their physical properties and their reactivity, the derivatives of fluorine, chlorine, bromine and iodine present various degrees of usefulness (Table 20.12).

1. *The case of fluorine*

With chlorine, it is one of the most used halogens in medicinal chemistry (around 20% of the used medicines and 28% of the agrochemicals include fluorine) when they are attached to a non-activated carbon atom.

For a long time, fluorine has been considered as a bioisostere of hydrogen, which is no more true. Indeed, fluorine is very different from hydrogen, with a van der Waals radius comparable to that of oxygen, it induces an increase in lipophilicity and its electronegativity is the highest in the periodic classification.

The difference in electronegativity between fluorine and carbon creates a large dipole moment in this bond. This dipole may contribute to the molecule's ability to be engaged in intermolecular interactions. Fluorine is able to participate to hydrogen bonds with the hydrogen of water. These bonds are weaker than those obtained with oxygen, but they are still strong enough to contribute to the binding of fluoroaromatic compounds to active site and/or receptor (Figure 20.36).[73]

The importance of electrostatics in the interaction of aromatics fluorine with cations and hydrogen bond donors can be visualized using electrostatic potential surfaces (Figure 20.37).[73] In monofluorobenzene, the potential of the fluorine is concentrated on the unique fluorine present, whereas in polyfluorobenzene the negative charge is spread over several fluorine atoms. For this reason, monofluorobenzene may give stronger interactions.

Another major use of fluorine is to block metabolically sensitive positions of a molecule. When the fluorine is placed in an activated position (e.g. R—CO—Cl → R—CO—CF_3), like for the antineoplastic valrubicin it can advantageously replaced by a CF_3 group.

TABLE 20.11 Atomic Radii and Characteristics of Carbon-Halogen Bonds[72]

Atomic radius (Å)	Bond*	Interatomic distance (Å)	Bond strength (kcal/mol)
H: 0.29	C–H	1.14	93
F: 0.64	C–F	1.45	114
Cl: 0.99	C–Cl	1.74	72
Br: 1.14	C–Br	1.90	59
I: 1.33	C–I	2.12	45

*In aliphatic series.

TABLE 20.12 Some Aromatic Substituent Constants for Halogens and Equivalent Functions[17]

Group	π	σ	MR
H	0.00	0.00	1.03
F	0.14	0.06	0.92
Cl	0.71	0.23	6.03
Br	0.86	0.23	8.88
I	1.12	0.18	13.94
CF_3	0.88	0.54	5.02
CH_3	0.56	−0.17	5.65
CN	−0.57	0.66	6.33
SO_2CF_3	0.55	0.93	12.86
SCF_3	1.44	0.50	13.81
SCN	0.41	0.52	13.40

FIGURE 20.36 Example of fluoroaromatic compounds in which an intermolecular participation of the fluorine has been proposed.[73]

2. *The case of chlorine*

A chlorine substituent produces simultaneously an increase in lipophilicity, an electron-attracting effect and a metabolic obstruction. The discovery of bicalutamide, a non-steroidal selective androgen receptor modulators (SARM) gave birth to a series of analogs such as compound A and B (Figure 20.38). Chen *et al.* have found that, compared to compound A, compound B has an higher affinity, an improved half-life, a smaller volume of distribution and a lower clearance.[74] The improved half-life and lower clearance might be explained by the ability of the two halogens to prevent metabolism.[75]

3. *The case of bromine*

Bromine is the less used halogen, and when it serves, it is usually incorporated as a bromo-aryl. The reproach against bromine is to generate reactive alkylating intermediates, more easily than chlorine or fluorine. Therefore it can confer, in a long-term treatment, toxic potentialities to the molecule that bears it. This was the case for the anti-inflammatory-analgesic drug bromfenac sodium, withdrawn from the US market due to reports of hepato-toxicity.[76]

However, other bromoaromatic compound have been marketed, like the anxiolytic bromazepam, or brimonidine used to lower pressure in the eyes in patients who have glaucoma and ocular hypertension (Figure 20.39).

FIGURE 20.37 Electrostatic potential surface of fluorophenyls, view from the front and the edge. *Source*: Taken from Ref. [73].

4. *The case of iodine*

Although even lesser tolerated than bromine, iodine is indispensable to the treatment of certain thyroidal deficiencies. Administrated by internal route, iodine and iodine derivatives trigger either acute hypersensitivity reactions (larynx oedema, cutaneous hemorrhages, fever, arthralgies, etc.) or chronic reactions (iodism). In addition to its use in certain dysfunctions of the thyroid gland, iodine presents to specific uses: covalent iodine derivatives serve as radiological contrast substances and ^{131}iodine (half-life: 8 days) is used as radioactive tracing agent.

5. *Extensions-cognate groups*

Chlorine, trifluoromethyl, cyano or azido groups are more or less bioisosteres. Other possible candidates are: SCN, SCF_3, SO_2CF_3 and $CH{=}CF_2$ (see Chapter 15).

V. EFFECTS OF HYDROXYLATION

The substitution of OH for H affects biological activity profoundly, as in the conversion of ethane to ethanol or

FIGURE 20.38 Structure and binding affinity to androgen receptor of non-steroidal SARMs.[74]

benzene to phenol. Simple alcohols have narcotic, and simple phenols have bacteriostatic properties. Polyfunctional compounds can act as chelating or complexing agents.

A. Effects on solubility

The introduction of an alcoholic or a phenolic hydroxy group into an active molecule changes the partition coefficient toward more hydrophilicity and renders the molecule more water soluble (see Figure 20.40). The value of the Hansch π constant for an aromatic hydroxy group is –0.67. That means that to compensate the loss in lipophilicity due to the monohydroxylation of an active compound, it is necessary to attach at an approriate place of the molecule a chlorine atom ($\pi = 0.71$) for example.

An interesting parameter has to be noted about the hydrophilicity of trihydroxybenzenes, and by extention to other polyhydroxylated compound. The fact that the hydroxyl groups of 1,2,3-trihydroxybenzene, and in a lesser extent those of 1,2,4-trihydroxybenzene, are able to form intramolecular hydrogen bonds leave these hydroxyl less available to generate hydrogen bonds with water, making them more hydrophobic than 1,3,5-trihydroxybenzene.

B. Effects on the ligand–receptor interaction

For some hydroxylated drugs like morphine, dopamine, haloperidol, γ-hydroxybutyrate, serotonine or most of the steroids, the hydroxy group is an essential element for hydrogen bonding with the receptor. For others the attachment of a hydroxy group can result in potency changes. Examples are found in hycanthone, which is 10 times more active against schistosomes than lucanthone,[77] or in hydroxylated minaprine analogs which show a 10 times better affinity for M_1 muscarinic receptors than the parent drugs.[78] One spectacular example of hydrogen bonding with hydroxyl (or amino) group has been show by Cappelli *et al.* (Figure 20.41).[79] By the simple replacement of an hydrogen by an hydroxyl group, they increase 750-fold the binding affinity of the compound allowing it to show sub-picomolar affinity!

C. Hydroxylation and metabolism

As a rule metabolic hydroxylation of an active compound represents a detoxication (phase I) mechanism. It results

FIGURE 20.39 Bromoaryl compounds.

FIGURE 20.40 Experimental log *P* for various hydroxyaryl compounds.[69]

R	IC_{50} (nM)
H	0.075
NH_2	0.0017
OH	0.0001

FIGURE 20.41 Structure and binding affinity of 3-quinolinecarboxamides to human endogenous NK1 receptor in UC11MG cells.[79]

generally from a first-pass effect and can be followed or not by a conjugation reaction (see Chapters 31 and 32). Classical examples of drugs detoxified through hydroxylation are paracetamol, oxyphenbutazone and hydroxychloroquine. Other important reactions of hydroxy compounds, whether alcoholic or phenolic, are based on their capacity to accept activated groups through the action of group-transferring enzymes (methylation, sulfation, phosphorylation, glycosylation, etc.).

VI. EFFECTS OF THIOLS AND OTHER SULFUR-CONTAINING GROUPS

Thiol and disulfide groups are of wide occurrence in natural products. They are found in small molecules such as lipoic acid, glutathione and thiamine as well as in cysteine-containing peptides and proteins (hormones, enzymes, antibiotics). In all these substances, thiol and disulfide groups are clearly associated either with high chemical reactivity or with consolidation of peptide and protein architecture. Being too reactive, the thiol and the disulfide groups are normally not used in medicinal chemistry as substituents in quantitative structure–activity relationships (QSAR) studies. Occasionally, methylthio substitution on aromatic rings is practiced, but even then, the obtained thioethers are very reactive. They are easily converted to sulfoxides and *vice versa* (see Chapter 33).

A. Drugs containing thiol

Drugs containing thiol groups are mainly used for the strong affinity that the thiolate anion presents toward heavy metals. This is the case for thiol-containing angiotensin-converting enzyme inhibitors which bind to a zinc-containing enzyme (see Chapter 8, and Ganellin and Roberts[80]). Captopril, a marketed drug used in the treatment of arterial hypertension and congestive heart failure, was obtained after manipulation of succinylproline. However, some adverse reactions due to the mercapto group conduct to the replacement of the thiol group by a carboxyl group, leading to another potent angiotensin-converting enzyme, enalapril.

Presently, other captopril analogs, like zofenopril (a prodrug of which the cleavage of the thioester led the active compound zofenoprilat) are still marketed. They are able to coordinate Zn(II) and act as free radical scavenger. Their lipophilicity allows them to attain high lever in heart tissue and then they can be used as cardioprotective drugs (Figure 20.42).[81]

The heavy-metal chelating properties of thiols were taken advantage of in the design of dimercaprol ("British Anti-Lewisite," BAL) as counter poison of the arsenical war gas lewisite (Figure 20.43). Today dimercaprol is used to treat poisoning by compounds of gold, mercury, antimony and arsenic. The toxic nature of the heavy metals is masked and chelate is stable enough to be excreted as such in the urine.

Penicillamine (D-β,β-dimethylcysteine) is an effective chelator of copper, mercury, zinc and lead which promotes the excretion of these metals in the urine. It is clinically used in patients with Wilson's disease, with rheumatoid arthritis and with heavy-metal intoxications.[82] Ziram and ferbam are the zinc and the iron salts respectively of dimethyldithiocarbamic acid (Figure 20.44). They are widely used as selective fungicides in agriculture. Pyrithione (1-hydroxy-2(1*H*)-pyridinethione), as its zinc salt is used in dermatology as antiseborrheic.

B. Drugs containing oxidized sulfides

The sulfoxide (S═O) and the sulfone (O═S═O) functions are very polar and usually confer mediocre CNS bioavailability. But this characteristic should not be a barrier to the use of this type of derivatives. As an example, the triptan class of compounds does generally have poor brain penetration data, but these characteristics are not a good guide to central activity, especially with very potent drugs such as the triptans (in the nanomolar range on 5-$HT_{1B/1D}$ receptors). Indeed, in contrast to most other CNS agents that are antagonists, triptans are agonist and so will require only low fractional receptor occupancy to exert central effects.[83] Another example: sulfonamides presenting moderate brain penetration (~10%) are known, like the 5-HT_6 receptor antagonist SB-271046 (Figure 20.45).[84]

Many other sulfone containing drugs are marketed, the major part for antibacterial use, but also as cardioprotective agents, antihypertensive, analgesic…

C. Drugs containing thiocyanate or thiourea

Covalently bound thiocyanates are rather unusual. Recent examples are the corticotropin-releasing factor (CRF) antagonist NPC 22009[85] and 4-phenoxy-phenoxyethyl thiocyanate (Figure 20.46) which acts as inhibitor of the sterol biosynthesis in *Trypanosoma cruzi*,[86] the protozoon agent responsible for Chagas disease. Sometime thiocyanates are

FIGURE 20.42 Antihypertensive drugs.

FIGURE 20.43 Chelating properties of thiol derivatives.

FIGURE 20.44 Ziram and pyrithione-zinc are chelated salts.

used as counteranions for the preparation of pharmaceutically acceptable salts.[87]

Heterocyclic thioureas such as 6-propylthiouracil, methimidazole and carbimidazole (Figure 20.47) are used as antithyroid drugs. They inhibit the formation of thyroid hormones; one of the presumed mechanisms is the inhibition of the iodine incorporation into the tyrosyl residues of thyroglobulin. It was proposed that the iodine atom is bound to a protein as a sulfenyl iodide. The thioureas may act in establishing covalent —S—S— bonds and in displacing iodine as HI.[88]

Some thioureas are also used in short-acting anesthetic compounds, like thialbarbital, thiamylal and thiopental. These compounds are used to induce anesthesia before the injection of anesthetic products and are now replaced slowly by propofol (Figure 20.48). More generally, thiourea and its simpler aliphatic derivatives, as well as thioamides, produce goiter and have to be avoided in drug design. A natural compound, L-5-vinyl-2-thiooxazolidone (goitrin) is responsible for the goiter of cattle eating turnips or cruciferous plants.[89]

Besides their affinity for metallic ions, thiol groups have other characteristics such as ability to interconvert to disulfides through redox reactions, to add to conjugated double bonds and to form complexes with the pyridine nucleotides by nucleophilic attack at the 4-position of the pyridine ring.

FIGURE 20.45 Compounds acting on CNS receptors.

FIGURE 20.46 Covalently bound thiocyanates.[85,86]

FIGURE 20.47 Heterocyclic thioureas with antithyroid effects.

VII. ACIDIC FUNCTIONS

The prototypical representatives of the group are the carboxylic acids. However, a huge number of bioisosteres such as sulfonic or phosphonic acids, tetrazoles or 3-hydroxyisoxazoles are available (see Chapter 15). In addition functions like esters, amides, peptides, aldehydes, primary alcohols and related functions can work as prodrugs or bioprecursors (see Chapter 36).

A. Effects on solubility

The introduction of an acidic group into a biologically active compound which contains no such group already has as consequence essentially a solubilizing effect. This effect can even be enhanced through salt formation (see Chapter 37). Carboxylic acids are often highly ionized at the physiological pH values and this is even more the case for sulfonic acids. As a rule, strong and highly ionized acids cannot cross the biological membranes which are permeable only to non-dissociated molecules or in some case to molecules whose charge is very distributed; they are then subject to a rapid clearance from the animal body. However, once absorbed, they can establish strong ionic interactions with the basic amino acids, especially with lysine, contained in the blood serum albumine, or the enzyme and receptor proteins.

A visible effect of the solubility modifications happening after introduction of a carboxylic group is found in the history of antihistaminic compounds (Figure 20.49). The first generation of antihistaminic drugs, in which hydroxyzine, were lipophilic compounds. They were able to cross the BBB, and had a sedating action because they were not P-glycoprotein (P-gp) substrate (which means they were not considered as xenobiotics and pumped out of the brain). Nowadays, hydroxyzine is still used as anxiolytic.

The second generation of antihistaminic compounds are less lipophilic, thanks to the replacement of the hydroxy group by a carboxylic acid. They are also P-gp substrate, which limits CNS exposure.[90,91]

FIGURE 20.48 Heterocyclic thioureas with anesthetic effects, and goitrin.

FIGURE 20.49 Antihistaminic compounds.

B. Effects on biological activity

Changes in biological activity distinguish the sulfonic from the carboxylic acids. Broadly speaking, the sulfonic acids as a class are generally not biologically active. Exceptions are certain complex dyes or trypanocides (Trypan blue, suramin, etc.) and sulfonic amino acids such as taurine and hypotaurine for which an active transport mechanism exists. For carboxylic acids the situation depends on whether the carboxylic function is introduced in small or large molecules.

1. *In small molecules*

The introduction of a carboxylic group changes fundamentally the biological activity. Very often the initial biological activity is destroyed and the toxicity of the parent compound is reduced. In a series of cyproheptadine analogs the replacement of a chlorine substituent on a benzo ring by a carboxylic group resulted in a 4,000-fold loss in affinity for the spiperone-labeled dopamine receptor.[92] Conversely, the presence of the carboxylic group can sometimes create the conditions necessary for activity (Figure 20.50).

FIGURE 20.50 Introduction of a carboxylic group in a small molecule can destroy the activity originally present or, conversely, create the conditions necessary for activity.

2. *In large molecules*

High pharmacological activity is maintained despite the presence of the carboxylic group. Examples are the anti-inflammatory arylacetic acids, the prostaglandins, cromolin and the related anti-asthmatics, and finally the β-lactam antibiotics. In these drugs, the relative weight of the carboxyl is notably smaller. An illustrative example is found in a series of acids and esters derived from coformycin and acting as AMP deaminase inhibitors (Figure 20.51). In this series, the free acid is clearly more potent than the corresponding ethyl ester.[93]

The discovery of losartan (Figure 20.52), and the numerous analogs obtained after (candesartan, eprosartan, irbesartan, olmesartan, tasosartan, telmisartan, valsartan …)[94,95] is an example of manipulation of acidic groups.

Starting with the benzyl-imidazole (compound 1), which binds to angiotensin II (AII) receptor but exhibits a very poor potency, Duncia *et al.*[94] supposed that a second acidic group (compound 2) would provide, at physiological pH, a negative charge in the area of space where Tyr-4 and Asp-1 residues reside. This substitution increased 10-fold the potency of the compound. Another 10-fold increase has been obtained by the introduction of a second phenyl ring between the two first rings (compound 3). The replacement of the carboxylic acid group by a more lipophilic acid isostere (compound 6) gave some oral activity, but without comparing the effect obtained by intravenous injection. Finally, a greater oral activity was obtained with the tetrazole substituent, due in part to the greater oral bioavailability, and rendered possible the development of the compound losartan (compound 7).[95]

Antitumor sulfonamides targeting G1 phase provide an *a contrario* example. Despite the relatively large size of the molecule, the presence of a carboxylic function completely abolishes the *in vitro* antiproliferative activity[96] whereas the corresponding carboxamide and sulfonamide are highly active (Figure 20.53).

AMP deaminase

R	K_i (μM)
H	0.79
Et	3.80

FIGURE 20.51 The free carboxylic acid is more active than the corresponding ethyl ester.

3. *Other carboxylic functions*

With carboxyl-derived functions such as esters and amides, the initial activity of the drug, lost in introducing the carboxyl group, is often regained. Amides, ureides, hydantoins and barbiturates share CNS-depressing properties and are frequently indispensable elements of sedative, tranquillizing and anticonvulsant drugs. Nitriles as substituents are often comparable to chlorine atoms, but sometimes more toxic.

VIII. BASIC GROUPS

The basic groups met in medicinal chemistry are the amines, the amidines, the guanidines and practically all nitrogen-containing heterocycles. Basic groups are polar and one

Compound	R	IC_{50} (nM)
1	H	150
2	COOH	1.6

Compound	R	Est. pK_a	IC_{50} (nM)
3	COOH	5	0.23
4	$COONHSO_2Ph$	8.44	0.14
5	$NHCOCF_3$	9.5	6.3
6	$NHSO_2CF_3$	4.5	0.083
7	Tetrazole	5–6	0.019

FIGURE 20.52 Genesis of losartan.[94,95]

would expect that highly ionized bases (especially quaternary ammonium salts) would resemble the sulfonic acids and show limited activity due to their mediocre membrane permeability. In practice bases with pK_a values superior to 10 have very limited chance to reach the CNS.

As seen for the acidic groups, the introduction of a basic group into a biologically active compound which contains no such group already has as consequence essentially a solubilizing effect. This effect can also be enhanced through salt formation (see Chapter 37). In drug–protein interactions the classical counter-anions of organic bases are the aspartic and the glutamic carboxylates.

The biological activity of amines and basic heterocycles is immense and justifies the adage: "no biological activity without nitrogen." Steroids, prostaglandins, non-steroidal anti-inflammatory drugs are, of course, exceptions. Primary amines demonstrate often less specific effect than secondary or tertiary amines. Acylation deactivates the amines strongly as does the introduction in some other place of the molecule of a carboxylic or sulfonic group (formation of zwitterions: bipolar ions). Diamines and polyamines are usually more active than monoamines. Aromatic amines are always more hazardous than aliphatic amines and form toxic metabolites, examples are 2-naphtylamine, benzidine, aniline (see Chapter 33). They are easy to detoxicate by introducing a carboxyl group, as evidenced by the change from aniline to the non-toxic *p*-aminobenzoic acid.

IX. ATTACHMENT OF ADDITIONAL BINDING SITES

A. To increase lipophilicity

Many polar active molecules, selected through *in vitro* screening tests are unable to cross the biomembranes and their bioavailability is particularly low. Attaching a very lipophilic moiety can sometimes help to overcome this drawback.

A typical example is given by the development of the anticonvulsant drug thiagabine[91] starting from nipecotic acid (Figure 20.54). Cyclic amino acids such as nipecotic acid and guvacine have been shown to inhibit GABA uptake. However, these small amino acids do not readily cross the BBB and thus limit their potential clinical usefulness. A considerable improvement has been the discovery by Yunger *et al.*[98,99] of compound SKF 89976A, a *N*-(4,4-diphenyl-3-butenyl) substituted nipecotic acid. *In vitro* this compound is approximatily 20 times more potent than the parent amino acid as inhibitor of [^{3}H] GABA uptake. Moreover, it is active orally as anticonvulsant in mice and rats with ED_{50}'s around 8 mg/kg.

Further developments demonstrated that the 4,4-diphenyl-3-butenyl moiety could be replaced by ether-type analogs[100,101] and that the attachment of a lipophilic 3,3-diphenylpropyl side chain can take place at the carbon atom at the 6-position of the amino acid.[102] However in this latter case, despite an reasonable *in vitro* activity (IC_{50} = 100 nM), no *in vivo* activity is observed. Final optimization of SKF 89976A led to the bis-thiophene tiagabine.[97] In a similar way, ω,ω-diphenyl-alkyl chains (butterflies) were also attached to L-glutamic acid to yield glutamate metabotropic m3 receptor-selective agonists.[103]

B. To achieve additional interactions

The fixation of large aromatic substituents such as 2-naphtyl and 3,3-diphenylpropyl to the low-efficacy partial agonist 4-PIOL transform this series in powerful $GABA_A$ receptor antagonists. This improvement is due to the substantial size of the cavity and the strong interactions between the aromatics and the binding cavity of the $GABA_A$ receptor (Figure 20.55).[104]

Histaprodifen, a 3,3-diphenylpropyl substituted analog of histamine, is a potent histamine H_1 receptor agonists. The increase in potency between histamine and histaprodifen may be due to a new orientation of the imidazole ring caused by the space filling substitution at the 2-position. This new orientation allows an additional interaction with the H_1 receptor.[105]

X	Antiproliferative activity IC_{50} (μg/mL) Colon 38 murine adenocarcinoma	P388 murine leukemia
COOH	>100	>100
$CONH_2$	0.19	0.87
SO_2NH_2	0.10	0.45

FIGURE 20.53 The antitumor activity of the carboxamide and of the sulfonamide is completely abolished for the corresponding free carboxylic acid.[93]

FIGURE 20.54 Lipophilic derivatives of nipecotic acid and of guvacine.

FIGURE 20.55 Large aromatic substituents attached to 5-(4-piperidyl)-3-isoxazolol.

Instead of ensuring high lipophilicity, these aralkyl groups serve to achieve additional interactions with the target macromolecule. This is typically the case for the angiotensin converting enzyme inhibitor enalaprilat. The exchange in captopril of the thiol function for a carboxylic group as ligand for the enzyme zinc atom entails an important decrease in activity (Chapter 6). This decrease could be compensated by the attachment of a phenethyl moiety.

Several marketed compounds are using the "butterfly" chains with different spacer length, with substituents on the phenyl rings (more often fluorine), and/or substituents onto the carbon bearing the two phenyls (alkyl, amide, hydroxyl, nitrile …).

REFERENCES

1. Taylor, J. The topliss approach-opinion of an industrial scientist on QSAR methods. *Lecture Given at the Louis Pasteur University*. Stasbourg: New york, 1984. March 8.
2. Rekker, R. F., Mannhold, R. *Calculation of Drug Lipophilicity*. Verlag VCH: Weinheim, 1992.
3. Witt, K. A., Davis, T. P. CNS drug delivery: opioid peptides and the blood-brain barrier. *AAPS J.* 2006, *8*(1), E76–E88.
4. Ooms, F., Wouters, J., Oscari, O., Happaerts, T., Bouchard, G., Carrupt, P. A., Testa, B., Lambert, D. M. Exploration of the pharmacophore of 3-alkyl-5-arylimidazolidinediones as new CB(1) cannabinoid receptor ligands and potential antagonists: synthesis, lipophilicity, affinity, and molecular modelling. *J. Med. Chem.* 2002, *45*(9), 1748–1756.
5. Gaillard, P., Carrupt, P. A., Testa, B., Boudon, A. Molecular lipophilicity potential, a tool in 3D QSAR: method and applications. *J. Comput. Aided Mol. Des.* 1994, *8*(2), 83–96.
6. Ginnings, P., Baum, R. Aqueous solubilities of the isomeric pentanols. *J. Am. Chem. Soc.* 1937, *59*, 1111–1113.
7. Nemethy, G. Hydrophobic interactions. *Angew. Chem. Int. Ed. Engl.* 1967, *6*(3), 195–206.
8. Hansch, C., Anderson, S. M. The effect of intramolecular hydrophobic bonding on partition coefficients. *J. Org. Chem.* 1967, *32*, 2583–2586.
9. Gilligan, D., Plummer, N. Comparative solubilities of sulfadiazine, sulfamerazine and sulfamethazine and their N4-acetyl derivatives at varying pH levels. *Proc. Soc. Exp. Biol. Med.* 1943, *53*, 142–145.
10. Albert, A. *Selective Toxicity*. Chapman & Hall: London, 1979, p 41.

11. Harms, A. F., Nauta, W. T. The effects of alkyl substitution in drugs- I. Substituted dimethylaminoethyl benzhydryl ethers. *J. Med. Pharm. Chem.* 1960, *2*, 57–77.
12. Durant, G. J., Ganellin, C. R., Parsons, M. E. Chemical differentiation of histamine H1- and H2-receptor agonists. *J. Med. Chem.* 1975, *18*(9), 905–909.
13. Reggio, P., Topiol, S., Weinstein, H. Molecular determinants for the agonist activity of 2-methylhistamine and 4-methylhistamine at H2-receptors. *J. Med. Chem.* 1986, *29*(11), 2412–2415.
14. Velluz, L. La stucture nor-stéroïde. In *Actualités Pharmacologiques* (Hazard, R., Cheymol, J., Eds). Masson & Cie: Paris, 1961, pp. 221–243.
15. Wermuth, C. G., Schwartz, J., Leclerc, G., Garnier, J. P., Rouot, B. Communication préliminaire: conformation de la clonidine et hypothèses sur son interaction avec un récepteur α-adrénergique. *Eur. J. Med. Chem.* 1973, *8*, 115–116.
16. Perlman, S., Costa-Neto, C. M., Miyakawa, A. A., Schambye, H. T., Hjorth, S. A., Paiva, A. C., Rivero, R. A., Greenlee, W. J., Schwartz, T. W. Dual agonistic and antagonistic property of nonpeptide angiotensin AT1 ligands: susceptibility to receptor mutations. *Mol. Pharmacol.* 1997, *51*(2), 301–311.
17. Chu, K. C. The quantitative analysis of structure–activity relationships. In *The Basis of Medicinal Chemistry/Burger's Medicinal Chemistry* (Wolf, M. E., Ed.), Vol. I. John Wiley & Sons: New York, 1980, pp. 393–418.
18. van de Waterbeemd, H., el Tayar, N., Carrupt, P. A., Testa, B. Pattern recognition study of QSAR substituent descriptors. *J. Comput. Aided Mol. Des.* 1989, *3*(2), 111–132.
19. Ganellin, C. R. Conformation of histamine derivatives. 3. A relationship between conformation and pharmacological activity. *J. Med. Chem.* 1973, *16*(6), 620–623.
20. Paiva, T. B., Tominaga, M., Paiva, A. C. Ionization of histamine, *N*-acetylhistamine, and their iodinated derivatives. *J. Med. Chem.* 1970, *13*(4), 689–692.
21. Andersson, S. H., Lindgren, A., Postlind, H. Biotransformation of tolterodine, a new muscarinic receptor antagonist, in mice, rats, and dogs. *Drug Metab. Dispos.* 1998, *26*(6), 528–535.
22. Herz, J. E., Swaminathan, S., Pinkerton, F. D., Wilson, W. K., Schroepfer, G. J., Jr. Inhibitors of sterol synthesis. A highly efficient and specific side-chain oxidation of 3 beta-acetoxy-5 alpha-cholest-8(14)-en-15-one for construction of metabolites and analogs of the 15-ketosterol. *J. Lipid Res.* 1992, *33*(4), 579–598.
23. Roemer, D., Beuscher, H. H., Hill, R. C., Pless, J., Bauer, W., Cardinaux, F., Closse, A., Hauser, D., Huguenin, R. A synthetic enkephalin analogue with prolonged parenteral and oral activity. *Nature* 1977, *268*, 547–549.
24. Plattner, J. J., Norbeck, D. W. Obstacles to drug development from peptide leads. In *Drug Discovery Technologies* (Clark, C. R., Moos, W. H., Eds). Ellis Horwood Limited: Chichester, 1990, pp. 92–126.
25. Fauchère, J. L. Elements for the rational design of peptide drugs. In *Adv. Drug Res.* (Testa, B., Ed.), Vol. 1. Academic Press: London, 1986, pp. 29–69.
26. Gee, K. W., McCauley, L. D., Lan, N. C. A putative receptor for neurosteroids on the GABAA receptor complex: the pharmacological properties and therapeutic potential of epalons. *Crit. Rev. Neurobiol.* 1995, *9*(2-3), 207–227.
27. Carter, R. B., Wood, P. L., Wieland, S., Hawkinson, J. E., Belelli, D., Lambert, J. J., White, H. S., Wolf, H. H., Mirsadeghi, S., Tahir, S. H., Bolger, M. B., Lan, N. C., Gee, K. W. Characterization of the anticonvulsant properties of ganaxolone (CCD 1042; 3alpha-hydroxy-3beta-methyl-5alpha-pregnan-20-one), a selective, high-affinity, steroid modulator of the gamma-aminobutyric acid(A) receptor. *J. Pharmacol. Exp. Ther.* 1997, *280*(3), 1284–1295.
28. Patrick, G. L. The opium analgesics. In *An Introduction to Medicinal Chemistry*, 1st Edition. Oxford University Press: Oxford, 1995, pp. 250–254.
29. Aquino, C. J., Armour, D. R., Berman, J. M., Birkemo, L. S., Carr, R. A., Croom, D. K., Dezube, M., Dougherty, R. W., Jr., Ervin, G. N., Grizzle, M. K., Head, J. E., Hirst, G. C., James, M. K., Johnson, M. F., Miller, L. J., Queen, K. L., Rimele, T. J., Smith, D. N., Sugg, E. E. Discovery of 1,5-benzodiazepines with peripheral cholecystokinin (CCK-A) receptor agonist activity. 1. Optimization of the agonist "trigger". *J. Med. Chem.* 1996, *39*(2), 562–569.
30. Tute, M. S. Principles and practice of Hansch analysis: a guide to structure–activity correlation for the medicinal chemist. In *Advan. Drug Res.* (Harper, N. J., Simmonds, A. B., Eds), Vol. 6. Academic Press: London, 1971, pp. 1–77.
31. Hansch, C., Anderson, S. M. The structure–activity relationship in barbiturates and its similarity to that in other narcotics. *J. Med. Chem.* 1967, *10*, 745–753.
32. Martin, Y. C. *Quantitative Drug Design: A Critical Introduction*. Marcel Dekker, Inc.: New York, 1978.
33. Dawson, M. I., Hobbs, P. D., Derdzinski, K., Chan, R. L., Gruber, J., Chao, W. R., Smith, S., Thies, R. W., Schiff, L. J. Conformationally restricted retinoids. *J. Med. Chem.* 1984, *27*(11), 1516–1531.
34. Lindamood, C., III, Cope, F. O., Dillehay, D. L., Everson, M. P., Giles, H. D., Lamon, E. W., McCarthy, D. J., Sartin, J. L., Hill, D. L. Pharmacological and toxicological properties of arotinoids SMR-2 and SMR-6 in mice. *Fundam. Appl. Toxicol.* 1990, *14*(1), 15–29.
35. Rigas, J. R., Dragnev, K. H. Emerging role of rexinoids in non-small cell lung cancer: focus on bexarotene. *Oncologist*, 2005, *10*(1), 22–33.
36. Lednicer, D., Mitscher, L. A., Georg, G. I. *The Organic Chemistry of Drug Synthesis*, Vol. 4. John Wiley & Sons: New York, 1990, p. 11.
37. Rizzo, R. C., Wang, D., Tirado-Rives, J., Jorgensen, W. L. Validation of a model for the complex of HIV-1 reverse transcriptase with sustiva through computational resistance profiles. *J. Am. Chem. Soc.* 2000, *122*(51), 12898–12900.
38. Ray, A. S., Yang, Z., Chu, C. K., Anderson, K. S. Novel use of a guanosine prodrug approach to convert 2′,3′-didehydro-2′,3′-dideoxyguanosine into a viable antiviral agent. *Antimicrob. Agents Chemother.* 2002, *46*(3), 887–891.
39. Bai, D. L., Tang, X. C., He, X. C. Huperzine A, a potential therapeutic agent for treatment of Alzheimer's disease. *Curr. Med. Chem.* 2000, *7*(3), 355–374.
40. Cussac, M., Pierre, J. L., Boucherle, A., Favier, F. Intérêt des dérivés du cyclopropane en Chimie Thérapeutique. *Ann. Pharm. Fr.* 1975, *33*, 513–529.
41. Marivet, M. C., Bourguignon, J. J., Lugnier, C., Mann, A., Stoclet, J. C., Wermuth, C. G. Inhibition of cyclic adenosine-3′,5′-monophosphate phosphodiesterase from vascular smooth muscle by rolipram analogues. *J. Med. Chem.* 1989, *32*, 1450–1457.
42. Trivedi, B. K., Holmes, A., Stoeber, T. L., Blankley, C. J., Roark, H. W., Picard, J. A., Shaw, M. K., Essenburg, A. D., Stanfield, R. L., Krause, B. R. Inhibitors of acyl-CoA:cholesterol acyltransferase. 4. A novel series of urea ACAT inhibitors as potential hypocholesterolemic agents. *J. Med. Chem.* 1993, *36*, 3300–3307.
43. Sexton, W. A. *Chemical Constitution and Biological Activity*, 3rd Edition. E. & F. N. Spon Ltd.: London, 1963. p. 103
44. Craig, P. N. Guidelines for drug and analog design. In *The Basis of Medicinal Chemistry-Burger's Medicinal Chemistry, 4th ed.* (Wolf, M. E., Ed.), Vol. I. John Wiley & Sons: New York, 1980, pp. 331–348.
45. Hobbs, J. B. Purine and pirimidine targets. In *Comprehensive Medicinal Chemistry* (Sammes, P. G., Ed.). Pergamon Press: Oxford, 1990, pp. 299–331.
46. Manfredini, S., Baraldi, P. G., Bazzanini, R., Marangoni, M., Simoni, D., Balzarini, J., De Clercq, E. Synthesis and cytotoxic activity of 6-vinyl- and 6-ethynyluridine and 8-vinyl-and 8-ethynyladenosine. *J. Med. Chem.* 1995, *38*(1), 199–203.
47. Rancourt, J., Cameron, D. R., Gorys, V., Lamarre, D., Poirier, M., Thibeault, D., Llinas-Brunet, M. Peptide-based inhibitors of the hepatitis C virus NS3 protease: structure–activity relationship at the C-terminal position. *J. Med. Chem.* 2004, *47*(10), 2511–2522.

48. Burger, M. T., Lin, X., Chu, D. T., Hiebert, C., Rico, A. C., Seid, M., Carroll, G. L., Barker, L., Huh, K., Langhorne, M., Shawar, R., Kidney, J., Young, K., Anderson, S., Desai, M. C., Plattner, J. J. Synthesis and antibacterial activity of novel C12 vinyl ketolides. *J. Med. Chem.* 2006, *49*(5), 1730–1743.
49. Browning, E. *Toxicity and Metabolism of Industrial Solvents*. Elsevier: New York, 1965, pp. 377–331, pp. 401–411
50. Scholtan, W., Sy, L. Kolloid-chemische Eigenschaften eines neuen Kurznarkotikums. Teilgewicht der Mizelle und Verteilungsgleichgewicht des Wirkstoffes. (Colloidal chemical properties of a new short-acting anesthetic. Particle weight of the micelle and distribution equilibrium of the active compound.). *Arzneimit.-Forsch.* 1966, *16*, 679–691.
51. Rosen, J., Miner, J. N. The search for safer glucocorticoid receptor ligands. *Endocr. Rev.* 2005, *26*(3), 452–464.
52. Milne, G. M., Jr., Johnson, M. R. Narcotic antagonists and analgesics. In *Annu. Rep. Med. Chem.* (Clarke, F. H., Ed.), Vol. 11. Academic Press: New York, 1967, pp. 23–32.
53. Culvenor, C. C. J., Dann, A. T., Dick, A. T. Alkylation as the mechanism by which the hepatotoxic pyrrolizidine alkaloids act on cell nuclei. *Nature*, 1962, *195*, 570–573.
54. Sterling, J., Herzig, Y., Goren, T., Finkelstein, N., Lerner, D., Goldenberg, W., Miskowski, I., Molnar, S., Rantal, F., Tamas, T., Toth, G., Zagyva, A., Zekany, A., Finberg, J., Lavian, G., Gross, A., Friedman, R., Razin, M., Huang, W., Krais, B., Chorev, M., Youdim, M. B., Weinstock, M. Novel dual inhibitors of AChE and MAO derived from hydroxy aminoindan and phenethylamine as potential treatment for Alzheimer's disease. *J. Med. Chem.* 2002, *45*(24), 5260–5279.
55. Cristofoli, W. A., Wiebe, L. I., De Clercq, E., Andrei, G., Snoeck, R., Balzarini, J., Knaus, E. E. 5-Alkynyl analogs of arabinouridine and 2′-deoxyuridine: cytostatic activity against herpes simplex virus and varicella-zoster thymidine kinase gene-transfected cells. *J. Med. Chem.* 2007, *50*(12), 2851–2857.
56. Madar, D. J., Kopecka, H., Pireh, D., Yong, H., Pei, Z., Li, X., Wiedeman, P. E., Djuric, S. W., Von Geldern, T. W., Fickes, M. G., Bhagavatula, L., McDermott, T., Wittenberger, S., Richards, S. J., Longenecker, K. L., Stewart, K. D., Lubben, T. H., Ballaron, S. J., Stashko, M. A., Long, M. A., Wells, H., Zinker, B. A., Mika, A. K., Beno, D. W., Kempf-Grote, A. J., Polakowski, J., Segreti, J., Reinhart, G. A., Fryer, R. M., Sham, H. L., Trevillyan, J. M. Discovery of 2-[4-{{2-(2*S*,5*R*)-2-cyano-5-ethynyl-1-pyrrolidinyl]-2-oxoethyl]amino]-4-methyl-1-piperidinyl]-4-pyridinecarboxylic acid (ABT-279): a very potent, selective, effective, and well-tolerated inhibitor of dipeptidyl peptidase-IV, useful for the treatment of diabetes. *J. Med. Chem.* 2006, *49*(21), 6416–6420.
57. Lee, S. J., Konishi, Y., Yu, D. T., Miskowski, T. A., Riviello, C. M., Macina, O. T., Frierson, M. R., Kondo, K., Sugitani, M., Sircar, J. C. *et al*. Discovery of potent cyclic GMP phosphodiesterase inhibitors. 2-Pyridyl- and 2-imidazolylquinazolines possessing cyclic GMP phosphodiesterase and thromboxane synthesis inhibitory activities. *J. Med. Chem.* 1995, *38*(18), 3547–3557.
58. Beart, P. M., Uhr, M. L., Johnston, G. A. R. Inhibition of GABA transaminase activity by 4-aminotetrolic acid. *J. Neurochem.* 1972, *19*, 1855–1861.
59. Stutz, A., Petranyi, G. Synthesis and antifungal activity of (E)-*N*-(6,6-dimethyl-2-hepten-4-ynyl)-*N*-methyl-1-naphthalenemethanamine (SF 86-327) and related allylamine derivatives with enhanced oral activity. *J. Med. Chem.* 1984, *27*(12), 1539–1543.
60. Nussbaumer, P., Leitner, I., Mraz, K., Stutz, A. Synthesis and structure–activity relationships of side-chain-substituted analogs of the allylamine antimycotic terbinafine lacking the central amino function. *J. Med. Chem.* 1995, *38*(10), 1831–1836.
61. Schmitt, J. Sur un nouveau myorelaxant de la classe des benzodiazépines: le tétrazepam. *Eur. J. Med. Chem.* 1967, *2*, 254–259.
62. Fryer, R. I., Zhang, P., Rios, R., Gu, Z.-Q., Basile, A. S., Skolnick, P. Structure–activity relationship studies at the benzodiazepine receptor (BZR): a comparison of the substituent effects of pyrazoloquinolinone analogs. *J. Med. Chem.* 1993, *36*, 1669–1673.
63. Bondi, A. van der Waals volumes and radii. *J. Phys. Chem.* 1964, *68*(3), 441–451.
64. Taylor, J. B., Kennewell, P. D. *Introductory Medicinal Chemistry*. Ellis Horwood Ltd.: Chichester, 1981, p. 89
65. Newman, A. H., Allen, A. C., Izenwasser, S., Katz, J. Novel 3α-(diphenylmethoxy)tropane analogs: potent dopamine uptake inhibitors without cocaine-like behavioral profiles. *J. Med. Chem.* 1994, *37*, 2258–2261.
66. MacLeod, A. M., Merchant, K. J., Cascieri, M. A., Sadowski, S., Ber, E., Swain, C. J., Baker, R. *N*-acyl-L-tryptophan benzyl esters: potent substance P receptor antagonists. *J. Med. Chem.* 1994, *36*, 2044–2045.
67. Gee, K. W., Yamamura, H. I. A novel pyrazoloquinoline that interacts with brain benzodiazepine receptors: characterization of some *in vitro* and *in vivo* properties of CGS 9896. *Life Sci.* 1982, *30*, 2245–2252.
68. Bowery, N. G., Bittiger, H., Olpe, H.-R. *$GABA_B$ receptors in mammalian function*. John Wiley & Sons: Chichester, 1990.
69. Katritzky, A. R., Dobchev, D. A., Fara, D. C., Hur, E., Tamm, K., Kurunczi, L., Karelson, M., Varnek, A., Solov'ev, V. P. Skin permeation rate as a function of chemical structure. *J. Med. Chem.* 2006, *49*(11), 3305–3314.
70. Hansch, C., Leo, A., Hoekman, D. *Exploring QSAR: Hydrophobic, Electronic and Steric Constants*, In: American Chemical Society (Ed.), ACS Professionnal Reference Books, Washington, DC, 1995.
71. Lide, D. R. *CRC Handbook of Chemistry and Physics*, 86th Edition. CRC Press: Boca-Raton, FL, 2005.
72. Buu-Hoï, N. P. Les dérivés organiques du fluor d'intérêt pharmacologique. In *Progress in Drug Research* (Jucker, E., Ed.), Vol. 3. Birkhäuser Verlag: Basel, 1961, pp. 9–74.
73. Razgulin, A. V., Mecozzi, S. Binding properties of aromatic carbon-bound fluorine. *J. Med. Chem.* 2006, *49*(26), 7902–7906.
74. Chen, J., Hwang, D. J., Bohl, C. E., Miller, D. D., Dalton, J. T. A selective androgen receptor modulator for hormonal male contraception. *J. Pharmacol. Exp. Ther.* 2005, *312*(2), 546–553.
75. Wu, D., Wu, Z., Yang, J., Nair, V. A., Miller, D. D., Dalton, J. T. Pharmacokinetics and metabolism of a selective androgen receptor modulator in rats: implication of molecular properties and intensive metabolic profile to investigate ideal pharmacokinetic characteristics of a propanamide in preclinical study. *Drug Metab. Dispos.* 2006, *34*(3), 483–494.
76. Anonymous. Bromfenac Sodium, Duract^R. *Drugs Fut.* 1998, *23*, 1234–1235.
77. Rosi, D., Peruzzotti, G., Dennis, E. W., Berberian, D. A., Freele, H., Archer, S. A new, active metabolite of "miracil D". *Nature*, 1965, *208*, 1005–1006.
78. Wermuth, C. G. Aminopyridazines – an alternative route to potent muscarinic agonists with no cholinergic syndrome. *Il Farmaco* 1993, *48*, 253–274.
79. Cappelli, A., Giuliani, G., Pericot Mohr Gl, G., Gallelli, A., Anzini, M., Vomero, S., Cupello, A., Scarrone, S., Matarrese, M., Moresco, R. M., Fazio, F., Finetti, F., Morbidelli, L., Ziche, M. A non-peptide NK1 receptor agonist showing subpicomolar affinity. *J. Med. Chem.* 2004, *47*(6), 1315–1318.
80. Ganellin, C. R., Roberts, S. M. *Medicinal Chemistry, the Role of Organic Chemistry in Drug Research*, 2 Edition. Academic Press: London, 1993.
81. Avendano Lopez, M. C. Drug optimization by molecular manipulation. *Meth. Find. Exp. Clin. Pharmacol.* 2002, *24*(Suppl. A), 9–11.
82. Klaassen, C. D. Heavy metals and heavy-metal antagonists. In *Goodman and Gilman's The Pharmacological Basis of Therapeutics* (Goodman Gilman, A., Rall, T. W., Nies, A. S., Taylor, P., Eds). Pergamon Press: New York, 1990, pp. 1592–1614.
83. Evans, D. C., O'Connor, D., Lake, B. G., Evers, R., Allen, C., Hargreaves, R. Eletriptan metabolism by human hepatic CYP450

enzymes and transport by human P-glycoprotein. *Drug Metab. Dispos.* 2003, *31*(7), 861–869.

84. Bromidge, S. M., Brown, A. M., Clarke, S. E., Dogson, K., Gager, T., Grassam, H. L., Jeffrey, P. M., Joiner, G. F., King, F. D., Middlemiss, D. N., Moss, S. F., Newman, H., Riley, G., Routledge, C., Wyman, P. 5-Chloro-*N*-(4-methoxy-3-piperazin-1-yl-phenyl)-3-methyl-2-benzothiophenesulfonamide (SB-271046): a potent, selective, and orally bioavailable 5-HT_6 receptor antagonist. *J. Med. Chem.* 1999, *42*, 202–205.
85. Abreu, M. E., Rzeszotarski, W., Kyle, D. J., Hiner, R. Preparation of oxopyrazolinyl thiocyanate and bis(oxopyrazolinyl)disulfides as corticotropin-releasing factor antagonist. *US Patent 5063245 A*, to Nova Pharmaceutical Corporation, 1991.
86. Szajnman, S. H., Yan, W., Bailey, B. N., Docampo, R., Elhalem, E., Rodriguez, J. B. Design and synthesis of aryloxyethyl thiocyanate derivatives as potent inhibitors of *Trypanosoma cruzi* proliferation. *J.Med.Chem.* 2000, *43*, 1826–1840.
87. Stahl, P. H., Wermuth, C. G. *Pharmaceutical Salts*. Helvetica Chimica Acta-John Wiley & Sons: Zürich, 2002.
88. Jirousek, L., Pritchard, E. On the chemical iodination of tyrosine with protein sulfenyl iodide and sulfenyl periodide derivatives. *Biochim. Biophys. Acta* 1971, *243*, 230–238.
89. Haynes, R. C., Jr. Thyroid and Antithyroid Drugs. In *Goodman and Gilman's The Pharmacological Basis of Therapeutics* (Goodman Gilman, A., Rall, T. W., Nies, A. S., Taylor, P., Eds), 8 Edition. Pergamon Press: New York, 1990, pp. 1361–1383.
90. Montoro, J., Sastre, J., Bartra, J., del Cuvillo, A., Davila, I., Jauregui, I., Mullol, J., Valero, A. L. Effect of H1 antihistamines upon the central nervous system. *J. Invest. Allergol. Clin. Immunol.* 2006, *16*(Suppl. 1), 24–28.
91. Chen, C., Hanson, E., Watson, J. W., Lee, J. S. P-glycoprotein limits the brain penetration of nonsedating but not sedating H1-antagonists. *Drug Metab. Dispos.* 2003, *31*(3), 312–318.
92. Remy, D. C., Britcher, S. F., King, S. W., Anderson, P. S., Hunt, C. A., Randall, W. C., Bélanger, P., Atkinson, J. G., Girard, Y., Rooney, C. S., Fuentes, J. J., Totaro, J. A., Robinson, J. L., Risley, E. A., Williams, M. Synthesis and receptor binding dtudies relevant to the neuroleptic activities of some 1-methyl-4-piperidylene-9-substituted-pyrrolo[2,1-*b*][3]benzazepine derivatives. *J. Med. Chem.* 1983, *26*, 974–980.
93. Bookser, B. C., Kasibhatla, S. R., Erion, M. D. AMP deaminase inhibitors. 4. Further N3-substituted coformycin aglycon analogues: N3-alkylmalonates as ribose 5′-monophosphate mimetics. *J. Med. Chem.* 2000, *43*(8), 1519–1524.
94. Duncia, J. V., Chiu, A. T., Carini, D. J., Gregory, G. B., Johnson, A. L., Price, W. A., Wells, G. J., Wong, P. C., Calabrese, J. C., Timmermans, P. B. The discovery of potent nonpeptide angiotensin II receptor antagonists: a new class of potent antihypertensives. *J. Med. Chem.* 1990, *33*(5), 1312–1329.
95. Carini, D. J., Duncia, J. V., Aldrich, P. E., Chiu, A. T., Johnson, A. L., Pierce, M. E., Price, W. A., Santella, J. B., III, Wells, G. J., Wexler, R. R. *et al*. Nonpeptide angiotensin II receptor antagonists: the discovery of a series of *N*-(biphenylylmethyl)imidazoles as potent, orally active antihypertensives. *J. Med. Chem.* 1991, *34*(8), 2525–2547.
96. Owa, T., Yoshino, H., Okauchi, T., Yoshimatsu, K., Ozawa, Y., Sugi, N. H., Nagasu, T., Koyanagi, N., Kitoh, K. Discovery of novel antitumor sulfonamides targeting G1 phase of the cell cycle. *J. Med. Chem.* 1999, *42*(19), 3789–3799.
97. Andersen, K. E., Braestrup, C., Gronwald, F. C., Jorgensen, A. S., Nielsen, E. B., Sonnewald, U., Sorensen, P. O., Suzdak, P. D., Knutsen, L. J. The synthesis of novel GABA uptake inhibitors. 1. Elucidation of the structure–activity studies leading to the choice of (*R*)-1-[4,4-bis(3-methyl-2-thienyl)-3-butenyl]-3-piperidinecarboxylic acid (tiagabine) as an anticonvulsant drug candidate. *J. Med. Chem.* 1993, *36*(12), 1716–1725.
98. Yunger, L. M., Fowler, P. J., Zarevics, P., Setler, P. E. Novel inhibitors of γ-aminobutyric acid (GABA) uptake: anticonvulsant actions in rats and mice. *J. Pharmacol. Exp. Ther.* 1984, *228*, 109–115.
99. Ali, F. E., Bondinell, W. E., Dandridge, P. A., Frazee, J. S., Garvey, E., Girard, G. R., Kaiser, C., Ku, T. W., Lafferty, J. J., Moonsammy, G. I., Oh, H.-J., Rush, J. A., Stler, P. E., Stringer, O. D., Venslavsky, J. W., Volpe, B. W., Yunger, L. M., Zircle, C. L. Orally active and potent inhibitors of γ-aminobutyric acid uptake. *J. Med. Chem.* 1985, *28*, 653–660.
100. Falch, E., Krogsgaard-Larsen, P. GABA uptake inhibitors. Syntheses and structure–activity studies on GABA analogues containing diarylbutenyl and diarylmethoxyalkyl *N*-substituents. *Eur. J. Med. Chem.* 1991, *26*, 69–78.
101. Pavia, M. R., Lobbestael, S. J., Nugiel, D., Mayhugh, D. R., Gregor, V. E., Taylor, C. P., Schwarz, R. D., Brahce, L., Vartanian, M. G. Structure–activity studies on benzhydrol-containing nipecotic acid and guvacine derivatives as potent, orally-active inhibitors of GABA uptake. *J. Med. Chem.* 1992, *35*, 4238–4248.
102. N'Goka, V., Schlewer, G., Linget, J. M., Chambon, J. P., Wermuth, C. G. GABA-uptake inhibitors: construction of a general pharmacophore model and successful prediction of a new representative. *J. Med. Chem.* 1991, *34*(8), 2547–2557.
103. Wermuth, C. G., Mann, A., Schoenfelder, A., Wright, R. A., Johnson, B. G., Burnett, J. P., Mayne, N. G., Schoepp, D. D. (2*S*,4*S*)-2-amino-4-(4,4-diphenylbut-1-yl)- pentane-1,5-dioic acid: a potent and selective antagonist for metabotropic glutamate receptors negatively linked to adenylate cyclase. *J. Med. Chem.* 1996, *39*(4), 814–816.
104. Frolund, B., Jorgensen, A. T., Tagmose, L., Stensbol, T. B., Vestergaard, H. T., Engblom, C., Kristiansen, U., Sanchez, C., Krogsgaard-Larsen, P., Liljefors, T. Novel class of potent 4-arylalkyl substituted 3-isoxazolol GABA(A) antagonists: synthesis, pharmacology, and molecular modelling. *J. Med. Chem.* 2002, *45*(12), 2454–2468.
105. Elz, S., Kramer, K., Pertz, H. H., Detert, H., ter Laak, A. M., Kühne, R., Schunack, W. Histaprodifens: Synthesis, Pharmacological *in Vitro* Evaluation and Molecular Modeling of a New Class of Highly Active and Selective Histamine H_1 Receptor Agonists. *J.Med.Chem.* 2000, *43*, 1071–1084.

Chapter 21

The Role of Functional Groups in Drug–Receptor Interactions

Laurent Schaeffer

I. INTRODUCTION
II. GENERAL PRINCIPLES
III. THE IMPORTANCE OF THE ELECTROSTATIC AND STERIC MATCH BETWEEN DRUG AND RECEPTOR
A. Electrostatic interactions
B. Steric interactions
C. Enthalpy/entropy compensation
IV. THE STRENGTHS OF FUNCTIONAL GROUP CONTRIBUTIONS TO DRUG–RECEPTOR INTERACTIONS
A. Measuring functional group contributions
B. The methyl group and other nonpolar substituents
C. The hydroxyl group and other hydrogen-bond forming substituents
D. Acidic and basic substituents
E. Practical applications for the medicinal chemist
F. Ligand efficiency
V. COOPERATIVE BINDING
REFERENCES

Alice remained looking thoughtfully at the mushroom for a minute, trying to make out which were the two sides of it; and, as it was perfectly round, she found this a very difficult question.

(Alice's Adventures in Wonderland, Lewis Carroll)

I. INTRODUCTION

An understanding of noncovalent interactions in ligand–receptor complexes is essential for the appreciation of drug action mechanisms as well as for rational drug design. The purpose of this chapter is to provide an overview of the physical and chemical factors which contribute most significantly to the strength of drug–receptor interactions.[1–3] The first part consists of a physical description of the influence of electrostatic and steric match on the various types of nonbonded drug–receptor interactions. The second part provides a more chemical interpretation, concentrating on the intrinsic strengths of individual functional group contributions to the affinity of drugs for their receptors. Some practical applications for the medicinal chemist will then be proposed. The concept of cooperative binding will be exposed in conclusion in order to underline the limitations of the pure additive calculation method based on average binding energies.

II. GENERAL PRINCIPLES

When a ligand and a receptor are sufficiently close, the ligand can diffuse up to and dock into its binding site on the receptor. This may first be mediated by the long-range electrostatic interactions between the ligand and the receptor and then strengthened by short-range hydrogen bonds and Van der Waals' interactions. Binding is accompanied by conformational changes ranging from modest shifts of a few atoms to movements of whole macromolecule domains. When the ligand is a drug, its biological activity is directly related to its affinity for the receptor, that is, the stability of the drug–receptor complex. The strength of this interaction is measured by its K_d, the dissociation constant for the complex at equilibrium:

$$K_d = \frac{[\text{drug}][\text{receptor}]}{[\text{complex}]} \tag{21.1}$$

Copyright © 2008, Elsevier Ltd
All rights reserved.

The smaller the K_d and the greater is the affinity of the drug for the receptor. This dissociation constant is related to the corresponding Gibbs free energy change which itself is composed of an enthalpic (ΔH) and entropic contribution ($T\Delta S$):[1,4]

$$\Delta G = -2.303\ RT \log\ K_d = \Delta H - T\Delta S \qquad (21.2)$$

Under physiological conditions (T = 310 K) this is approximated (in kJ/mol) by

$$\Delta G = -5.85 \log K_d \qquad (21.3)$$

The experimental measurement of the equilibrium constant thus provides a direct calculation of ΔG. Typically, K_d is in the range of 10^{-2} and 10^{-12} M meaning that the affinity of a ligand toward its receptor falls into an energy interval between −10 and −70 kJ/mol in aqueous solution.[5] A drug binding with a K_d of 10^{-9} M requires, for example, $(-5.85) \times (-9) = 52.6$ kJ/mol to dissociate from the receptor.

According to the equation (21.2), ligand–receptor interactions are characterized by enthalpy–entropy compensation in which one term favors and the other disfavors binding. While enthalpic contributions include electrostatic, hydrogen bond, and Van der Waals' interactions, entropic contributions arise from several sources. On one hand, the loss of flexibility upon binding has an important entropic cost, counterbalanced on the other hand by the displacement of ordered water molecules. This will be discussed in the next section as well as the various types of drug–receptor interactions.

III. THE IMPORTANCE OF THE ELECTROSTATIC AND STERIC MATCH BETWEEN DRUG AND RECEPTOR

What determines K_d or in other words how does the binding affinity relate to structural properties of a complex and its formation from separate, individually solvated species? What are the prerequisites that allow a receptor to bind a ligand tightly and selectively? On first glance, the most important requirement appears to be a good steric and electronic complementarity between receptor and ligand, usually described by Van der Waals and Coulomb interactions. Although most noncovalent interactions depend to some degree on both types of complementarity, we will separate them in the following discussion into those which are primarily electrostatic and those which are primarily steric.

A. Electrostatic interactions

Electrostatic interactions are the net result of the attractive forces between the positively charged nuclei and the negatively charged electrons of the two molecules. The attractive force between these opposite charges leads to three main bond types: charge–charge, charge–dipole, and dipole–dipole interactions. The strength of any electrostatic interaction can be calculated from equation (21.4), where q_i and q_j are two charges separated by a distance r_{ij} in a medium of dielectric constant ε. This equation applies equally to ionic interactions, where the charges q_i and q_j are integer values, or to polar interactions, in which the total energy is summed over the contributions calculated from the partial charges on all the individual atoms.

$$E = \frac{q_i \cdot q_j}{\varepsilon \cdot r_{ij}} \qquad (21.4)$$

1. *Charge–charge interactions or ionic bonds.*

According to the equation (21.4), the strength between two charges is inversely proportional to the distance separating them. Since the strengths of other noncovalent bonds are even more sharply dependent on distance than that of ionic bonds, ionic attraction frequently dominates the initial long-range interactions between drugs and receptors. A simple ionic interaction basically provides a ΔG of −20 kJ/mol. It also follows from equation (21.4) that the strengths of ionic interactions are crucially dependent on the dielectric constant ε of the surrounding medium. Indeed, in biological systems, charges are often separated by water or other molecules and dielectric micro-environments are variable, with less shielding of charges in regions of hydrocarbon sidechains and greater shielding in regions of polar sidechains. For instance, in hydrophobic pockets, like the interior of a protein molecule, the dielectric constant is around 4, whereas in bulk-phase water the corresponding value is 80. In other environments, intermediate values are appropriate, for example, for interactions occurring near the surface of a protein a ε value of 28 is commonly used. Charge–charge interactions between biological systems and drugs are possible insofar as ionic species are strongly present in biomacromolecules and drugs at physiological pH. Cationic environments are provided by protonation of basic groups such as the amino acid side chains lysine, arginine, and, to a much lesser extent, histidine (Table 21.1). On the other hand, acidic groups, such as the carboxylic acid side chains of glutamic acid and aspartic acid, are deprotonated to give anionic groups.

Concerning drugs, both cationic and anionic compounds are commonly used (protonated basic side chain, protonated aza heterocycles, deprotonated carboxylic acids, enolic species, and acidic sulfonamides) (Figure 21.1).

The antihypertensive drug captopril is an example of a molecule that participates in an ionic bond with the Lys1087 residue of the angiotensin-converting enzyme (ACE) receptor (Figure 21.2).

2. *Charge–dipole and dipole–dipole interactions*

Molecules composed of atoms of different electronegativities have usually an asymmetric distribution of electrons, which produces electronic dipoles. These dipoles within a cell or in aqueous medium can be attracted by a close-by ion, establishing so-called charge–dipole interactions. A permanent dipole can also interact with another permanent dipole leading to a dipole–dipole interaction. In a recent review, Diederich *et al.* has listed these orthogonal multipolar interactions among which C—X···C═O (X═halogen), C═O···C═O, C≡N···C═O, S═O···C═O, C—OH···C═O, and H_2O···C═O.[8] Because the charge of a dipole is less than that of an ion, charge–dipole and dipole–dipole interactions are weaker than ionic bonds. They are nevertheless key contributors to the overall strengths of drug–receptor interactions, since they occur in any molecule in which electronegativity differences between atoms result in significant bond, group or molecular dipole moments. The key differences between ionic and dipolar interactions relate to their dependence on distance and orientation (Table 21.2).

Indeed, while steric effects are of little importance in ionic interactions, stricter geometric requirements apply to dipolar interactions, which may be either attractive or repulsive, depending on the orientation of the dipole moments. In most cases, these interactions provide a ΔG of −5 to −30 kJ/mol. The insomnia drug zafirlukast gives an illustration of dipolar interactions (Figure 21.3).

TABLE 21.1 Main Ionizable Groups in Proteins and Nucleic Acids[6,7]

Chemical function	Charge	pK_a
Fully or almost fully ionized groups at pH 7.4		
Carboxyl α (terminal COOH)	−	1.8–2.4
Carboxyl β (Asp)	−	3.7
Carboxyl γ (Glu)	−	4.3
Primary phosphoryl	−	0.7–1.0
Secondary phosphoryl	−	5.9–6.0
α-Ammonium (Lys)	+	10.5
Guanidinium	+	12.5
Partially ionized groups at pH 7.4		
Sulfhydryl (Cys)	−	8.2
Imidazolium (His)	+	6.0
α-Ammonium (terminal peptide NH_2)	+	7.5–10.3
N Amidic (Glu)	+	0.1
N Amidic (Asp)	+	8.8
Nonionized groups at pH 7.4		
Phenolic hydroxyl (Tyr)	−	10.0
Heteroaromatic hydroxyl (uracyl, thymine, guanine)	−	9.2–9.8
Osidic hydroxyl	−	12.3–12.6
Amino residue (adenine, guanine, cytosine)	+	3.3–4.6

3. *Inductive interactions*

The electric field generated by a charged molecule or a molecule with a permanent dipole can induce a dipole in a second

FIGURE 21.2 Captopril as example of a charge–charge interaction with the ACE receptor.

FIGURE 21.1 Cationic and anionic drugs.

TABLE 21.2 Types of Noncovalent Interactions

Type of interaction	Energy and dependency on distance	Example
Charge–charge Longest-range force, nondirectional	ΔG of −20 to −40 kJ/mol 1/r	
Charge–dipole Depends on orientation of dipole	ΔG of −12 to −20 kJ/mol $1/r^2$	
Dipole–dipole Depends on mutual orientation of dipoles	ΔG of −4 to −12 kJ/mol $1/r^3$	
Charge-induced dipole Depends on polarizability of molecule in which dipole is induced	ΔG of −2 to −10 kJ/mol $1/r^4$	
Dipole-induced dipole Depends on polarizability of molecule in which dipole is induced	ΔG of −2 kJ/mol $1/r^6$	
Dispersion Involves mutual synchronization of fluctuating charges	ΔG of −2 to −4 kJ/mol $1/r^6$	
Hydrogen bond Charge attraction + partial covalent bond	ΔG of −4 to −30 kJ/mol	

molecule that is located nearby in space. The strength of the interaction depends on the dipole moment of the first molecule and the polarizability of the second. That means when an electron-donating molecule (or group) comes into contact with an electron-withdrawing molecule (or group), the donor may transfer some of its charge to the acceptor, forming a charge–transfer complex. In the case of an intramolecular redistribution of charge, this will be referred to as an induced polarization, whereas a redistribution of charge between two molecules is described as a charge–transfer interaction. In either case, the resulting interactions are always attractive and strongly dependent on the distance separating the two molecules as well as on the difference between the ionization potential of the donor and the electron affinity of the acceptor. Donor groups contain π-electrons, such as alkenes, alkynes, and aromatic moieties with electron-donating substituents, or groups presenting nonbonded electrons pairs (O, N, S). Acceptor groups contain electron-deficient π-orbitals, such as alkenes, alkynes, and aromatic moieties having electron-withdrawing substituents,

FIGURE 21.3 Zaleplon as example of ion–dipole and dipole–dipole interactions.

and weakly acidic protons. Groups on receptors acting as electron donors are for instance the aromatic ring of tyrosine or the carboxylate group of aspartate. Cystein is an example of an electron-withdrawing group whereas histidine, tryptophan, and asparagine are both electron donors and acceptors. In general, this type of interaction can contribute as much as 2–3 kJ/mol.

An interesting example of the importance of inductive interactions is the calculation by Bajorath *et al.* on the binding of folate (Figure 21.4) and dihydrofolate to dihydrofolate reductase. This revealed a shift in net charge equivalent to half an electron from the pteridine ring to the glutamate moiety on binding to the enzyme, with the major change in density being focussed on the bonds that are catalytically reduced.[9]

4. *Hydrogen bonds*

Hydrogen bonds are specific, short range, and directional nonbonded interactions. They occur between a hydrogen atom bound covalently to an electronegative atom (usually N, S or O), and an additional electronegative atom (Table 21.3). Distances of 2.5–3.2 Å between hydrogen-bond donor X and Y and X−H···Y angles of 130–180° are typically found.[10] Their strength is optimal when the three concerned atoms are aligned and when the H donor tends to point directly at the acceptor electron pair. As a result of its electrostatic nature, the strength of a hydrogen bond depends also on its microscopic environment and on the local dielectric constant ε of the surrounding medium (Coulombic interaction energy is proportional to ε^{-1}). Therefore, buried hydrogen bonds are regarded as more important for protein–ligand interactions than those formed in solvent-exposed regions.[11] The free energy for hydrogen bonding can be between −4 and −30 kJ/mol, but usually is in the range of −12 to −20 kJ/mol. Binding affinities increase by about one order of magnitude per hydrogen bond.

FIGURE 21.4 Folate structure.

TABLE 21.3 Potential Hydrogen Bond Donor and Acceptor Groups Classified According to Their Strength of Interaction[12]

	Donor*	Acceptor
Very strong	N^+H_3, X^+–H, F–H	CO_2^-, O^-, N^-, F^-
Strong	O–H, N–H, Hal–H	O=C, O–H, N, S=C, F–H, Hal^-
Weak	C–H, S–H, P–H, M–H	C=C, Hal–C, π, S–H, M, Hal–M, Hal–H, Se

*X is any atom, Hal is any of the lighter halogens, and M is a transition metal

Although their strength is weaker than ionic or covalent bonds, they are in general the predominant contribution to the specificity of molecular recognition.[13,14]. They help also to determine the conformation and folding ways of numerous macromolecules.

The double helical structure of DNA, for example, is due largely to hydrogen bonding between the base pairs, which link one complementary strand to the other and enable replication (Figure 21.5). Hydrogen bonds are also essential to maintain the structural integrity of α-helix and β-sheet conformations of peptides and proteins. By causing the macromolecules to fold into a specific shape, the hydrogen bonds contribute to the apparition of their biochemical functions.

In drug design, hydrogen bonds are exploited to obtain specificity, which is achieved not only through favorable short-range directionally specific interactions, but also because ligand–receptor arrangements that leave bonding capacity unsatisfied are disfavored. The number of hydrogen bonds in a drug molecule may be limited by requirements on polarity for absorption and permeation. The Lipinski rule-of-five, for example, suggests that compounds with more than 5 hydrogen bond donors or more than 10 hydrogen bonds acceptors are more likely to have poor absorption or permeation characteristics.

Among the numerous examples of drug–receptor interactions through hydrogen bonds, the antibiotic vancomycin

FIGURE 21.5 Hydrogen bonds between DNA base pairs.

FIGURE 21.6 Crystal structure of a short peptide L-Lys-D-Ala-D-Ala (bacterial cell wall precursor – in green) bound to the antibiotic vancomycin (in blue) through 5 hydrogen bonds.[15]

is especially interesting because it binds selectively with peptides having a terminal D-Ala-D-Ala moiety in bacterial cell through five hydrogen bonds (Figure 21.6). Vancomycin is lethal to the bacteria since once it has bound to these particular peptides they are unable to be used to construct the bacteria's cell wall.

5. *Cation–π interactions*

Such interaction occurs between a cation and the large, permanent quadrupole moment of an aromatic ring. The interaction energy depends on both the nature of the π-system and the nature of the cation. The importance of cation-π interactions has been first recognized by Ma and Dougherty.[16] In fact, cation-π interactions play a key role for molecular recognition in biological receptors. They have been considered in such diverse systems as acetylcholine receptors (nicotinic, muscarinic, and Ach esterase), K^+ channels, the cyclize enzymes of steroid biosynthesis, and enzymes that catalyze methylation reactions involving S-adenosylmethionine.[17] A remarkable example is given by the strong complexation of 7-methyl-GTP (dissociation constant $K_d \approx 1.1 \times 10^8\ M^{-1}$, $\Delta G = -45\,kJ/mol)^{-1}$ by a messenger RNA 5′-cap-binding protein, the eukaryotic translation initiation factor eIF4E. The cationic nucleobase in this complex is sandwiched at Van der Waals distance (*ca.* 3.5 Å) between two tryptophan side chains (Figure 21.7).[18]

The study of the factor Xa, a serine protease from the blood coagulation cascade, has conducted Diederich *et al.*[19] to observe that the aromatic box formed by the side chains of Phe174, Tyr99, and Trp215 in the S4-pocket is a very effective onium binding site. By comparing the affinity of the quaternary ammonium ion (±)-**1** to its tert-butyl analog, the free enthalpy increment for cation-π interactions in this box was determined as $\Delta G = -2.8 \times 4.18\,kJ/mol$ (Figure 21.8).

6. *Arene–arene interactions*[17,20]

Despite their weak and poorly directional character, π–π interactions have been recognized to play an important role

FIGURE 21.7 Partial view of the X-ray crystal structure (PDB code: 1L8B) of the messenger RNA 5′-cap-binding protein eIF4E bound to 7-methyl-GTP, which shows the sandwiching of the cationic nucleobase between the side chains of Trp102B and Trp56B.

FIGURE 21.8 Schematic representation of the complex between factor Xa and the tricyclic inhibitor (±) − 1.

in molecular recognition. Burley and Petsko indeed demonstrated in a study involving 34 proteins that on average 60% of aromatic side chains (Phe, Trp, Tyr) are involved in aromatic–aromatic interactions.[21] It is nowadays recognized that London dispersion interactions are the major source of stabilization energy between two aromatic molecules; however, the electrostatic component associated with the quadrupole moment of the aromatic ring is an influential factor in determining the geometry of the interaction. Assessment of binding interactions in aqueous solution is complicated by additional hydrophobic effects leading to apolar complexation or to intramolecular hydrophobic collapse.

Three lowest energy arrangements are commonly involved in π–π interactions (Figure 21.9) among which the T-shaped structure is the predominant one. The highest edge-to-face attraction is observed when an electron-attracting substituent renders the interacting H atom more acidic (higher positive partial charge) and when an electron-donating substituent increases the basicity (π-electron density) of the interacting π system. The other important arrangement of aromatic rings besides the edge-to-face contact is the parallel alignment. In this case two aromatic partners, one bearing strong electron-donor and the other strong electron-acceptor groups, form parallel stacking

complexes in solution, with the geometry largely determined by molecular orbital interactions (charge–transfer complexes). The term "polar/π" interaction was introduced to emphasize the importance of the electrostatic term in π-stacking. Since attractive electrostatic interactions between atoms of opposite partial charges often overcome the repulsion between close-shell π clouds, π–π stacking interactions are also abundant between heterocyclic π systems. A nice example is provided by the ternary complex of the anticancer drug 1843U89 and dUMP formed at the active site of thymidylate synthase (Figure 21.10).[22]

B. Steric interactions

As the lock-and-key model suggests, shape complementarity is very important for ligand–receptor binding and specificity. While electrostatic interactions are the dominant interactions involving polar molecules, there are also strong interactions between nonpolar molecules, particularly at short intermolecular distances.

FIGURE 21.9 Proposed lowest energy structures of the benzene dimer. *d*: distance between planes, d′: lateral offset.

1. *Dispersion forces*

Van der Waals or London dispersion forces are the universal forces responsible for attractive interactions between nonpolar molecules. The occurrence of these short-range interactions is due to the fact that any atom will, at any given instant, be likely to possess a finite dipole moment as a result of the movement of electrons around the nuclei. When molecules are approaching each other, the temporary dipoles of one molecule induce opposite dipoles in the other approaching molecules, thus resulting in a net attractive force. Although the individual interactions between pairs of atoms are relatively weak (about 2 kJ/mol), the total contribution to binding from dispersion forces can be very significant if there is a close fit between drug and receptor. The quality of the steric match is thus the dominant factor in nonpolar interactions.

2. *Short-range repulsive forces*

The short-range repulsive forces resulting from the overlap of the electron clouds of any two molecules increase exponentially with decreasing internuclear separation. The balance between these repulsive interactions and the dispersion forces thus determines both the minimum and the most favorable nonbonded separation between any pair of

FIGURE 21.10 Heterocyclic π stacking between dUMP and the anticancer drug 1843U89 bound at the active site of thymidylate synthase (PDB code: 1TSD).

atoms. The equilibrium distance can be determined from crystal data, and is equivalent to the sum of the Van der Waals radii of the two interacting atoms. For nonpolar molecules this balance between the attractive dispersion forces and the short-range repulsive forces is generally defined in terms of the Buckingham (6-exp) potential given in equation (21.5) or the alternative Lennard–Jones 6–12 potential, given in equation (21.6).

$$E = \frac{Ae^{-Br}}{r^d} - \frac{C}{r^6} \tag{21.5}$$

$$E = \frac{A^r}{r^{12}} - \frac{C}{r^6} \tag{21.6}$$

3. *Conformational energy*

While intramolecular interactions within the drug molecule are the primary factor in determining the lowest energy conformation of the unbound drug, intermolecular interactions with the receptor also have a significant effect on conformation. If the bound conformation of a flexible molecule is also its lowest energy conformation, there is no conformational energy cost involved in binding. If, on the other hand, the optimal interaction between drug and receptor requires a higher-energy conformation, this energy difference will reduce the apparent strength of the interaction between the two molecules.

C. Enthalpy/entropy compensation

Consider the formation of a specific noncovalent bond (e.g. **A**···**B** for the transformation **A + B** → **A**···**B**). An increase in its strength (which corresponds to an increasing negative contribution to ΔH, and a more favorable binding process) will be accompanied by an increasing restriction in the relative motion of **A** and **B** in **A**···**B** (which corresponds to a negative contribution to ΔS, and so unfavorable to binding). This opposing interplay between enthalpy and entropy is known as enthalpy/entropy compensation, and is a fundamental property of noncovalent interactions.[23] It can be interpreted that an enhancement of intermolecular binding is accompanied by a loss in degrees of freedom of mobility and vice versa. The two effects can be traded off against each other because the strength of noncovalent bonds is, at room temperature, comparable to the thermal energies that oppose them. The enthalpy/entropy compensation is of particular importance for the prediction of receptor–ligand interactions: whereas the individual enthalpic and entropic contributions can vary over large ranges, the total change in free enthalpy is frequently close to zero. As a consequence, small relative errors in the prediction of ΔH and ΔS can have significant influence on ΔG. This concept is less important in the study of covalent bonds, which are typically too strong to be effectively opposed by thermal motions at room temperature.

1. *Hydrophobic interactions*

The hydrophobic effect is the property that nonpolar molecules tend to self-associate in the presence of aqueous solution. This short-range attractive interaction is due to both enthalpic and entropic effects. It describes the energetic preference of nonpolar molecular surfaces to interact with other nonpolar molecular surfaces and thereby to displace water molecules from the interacting surfaces. When a nonpolar molecule is surrounded by water, stronger than normal water–water interactions are formed around the solute molecule to compensate for the weaker interactions between solute and water.[24] This results in an increasingly ordered arrangement of water molecules around the solute and thus negative entropy of dissolution. The decrease in entropy is roughly proportional to the nonpolar surface area of the molecule. The association of two such nonpolar molecules in water reduces the total nonpolar surface area exposed to the solvent, thus reducing the amount of structured water, and therefore providing a favorable entropy of association. The enthalpic contribution to hydrophobic interactions is due to the water molecules occupying lipophilic binding sites and which are consequently unable to form hydrogen bonds with the receptor. Their release from hydrophobic pocket let them to form strong hydrogen bonds with the bulk water. As for Van der Waals forces, hydrophobic interactions are individually weak (0.1 to 0.2 kJ/mol for every square angstrom of solvent-accessible hydrocarbon surface),[25] but the total contribution of hydrophobic bonds to drug–receptor interactions is substantial. Similarly, the overall strength of the hydrophobic interaction between two molecules is very dependent on the quality of the steric match between the two molecules. If this is not sufficiently close to squeeze all of the solvent from the interface, a substantial entropy penalty must be paid for each of the trapped water molecules. Hydrophobic interactions are also regarded as the main driving force for conformational changes of the receptor upon ligand binding. This induced fit can be viewed as a "collapse" of the receptor about the ligand.[26] As an extreme case the binding of trifluoroperazine (Figure 21.11) to Ca^{2+} calmodulin induces a

FIGURE 21.11 Trifluoperazine, a ligand of Ca^{2+}-calmodulin that induce an adaptation of the protein binding pocket.

conformational change of the protein from an extended to a compact form.[27]

2. *Translational and rotational entropy*

The transformation of two mobile molecules into one mobile complex results in the loss of translational and rotational entropy. Indeed, by binding to its receptor, a drug molecule is losing three translational and three rotational degrees of freedom replaced by six vibrational degrees of freedom in the complex.[28] The resulting entropy change is dependent on the relative tightness of the complex which is formed. For a typical ligand–protein interaction, the estimated change in free energy resulting from the loss of entropy on binding (at 310 K) ranges from 12 kJ/mol for a very weak interaction to 60 kJ/mol for a tightly bound complex.[29]

3. *Conformational entropy*

In the case of flexible drug molecules there is a further entropy loss due to the conformational restriction, which accompanies binding. Based on the observed entropy changes accompanying cyclization reactions, the extent of this entropy loss is estimated[30] at 5–6 kJ/mol per internal rotation, although the actual figure again depends on the overall strength of the interaction between the drug and the receptor. In the case of rigid analogs there is no such loss of conformational entropy on binding. Provided that they offer a good steric and electrostatic match to the receptor, rigid analogs should therefore have a free energy advantage relative to more flexible drugs. To optimize entropic contributions, compounds are usually designed to be relatively rigid with few rotatable bonds. Conformational flexibility is however important in biomolecular binding processes. A recent conformational analysis of drug-like ligands binding to proteins shows that many ligands do not bind in a minimum energy conformation. Energetically unfavorable conformational rearrangements can be tolerated in some cases without penalizing the tightness of binding.[31] On the other hand, small-scale motions including bond stretching, bond angle bending, and dihedral angle variations are able to reduce slightly the receptor affinity for its ligand. The timescale of these motions is around 10^{-12} seconds and the amplitude is less than 1 Å. For enzyme catalysis, for example, it was shown that movement of less than 1 Å can alter catalytic rates by several orders of magnitude.

IV. THE STRENGTHS OF FUNCTIONAL GROUP CONTRIBUTIONS TO DRUG–RECEPTOR INTERACTIONS

The total free energy of interaction between a drug and its receptor provides a measure of the strength of the association between the two molecules, but tells us little or nothing about the overall quality of their match. Does the observed binding reflect a composite of interactions between every part of the drug and its receptor, or is it a case of one or two strong interactions contributing sufficient energy to disguise an otherwise mediocre fit? Is the observed increase in interaction energy resulting from the addition of a new functional group consistent with what might have been anticipated? To answer these questions we need some means of estimating the individual functional group contributions to drug–receptor interactions.

A. Measuring functional group contributions

When cooperativity is ignored, contributions of ΔG values to the total free energies of binding may be added together. Approaches based on functional group additives equation (21.7) or the additivity of free enthalpy components equation (21.8) have frequently been applied to understand and predict protein–ligand interactions.[32] Pioneering studies in this field were performed by Andrews *et al.*[31], and Lau and Pettitt.[33]

$$\Delta G = \Delta G_{\mathrm{Me}} + \Delta G_{\mathrm{OH}} + \Delta G_{\mathrm{Ph}}(\mathrm{Ph}) + \cdots \qquad (21.7)$$

$$\Delta G = \Delta G_{\text{H-bridge}} + \Delta G_{\text{solvation}} + \Delta G_{\text{conformation}} + \cdots T\Delta S \qquad (21.8)$$

In order to have a brief overview of the methods used to predict the free energy of binding of a ligand to its receptor, we will describe and discuse here some of them.

As a first approximation, the free energy of binding can be defined in terms of the binding energies for the individual functional groups which make up the drug molecule according to equation (21.9)

$$\Delta G = T\Delta S_{t,r} + n_r E_r = \Sigma n_{\mathrm{x}} E_{\mathrm{x}} \qquad (21.9)$$

where $T\Delta S_{\mathrm{t,r}}$ is the loss of overall translational and rotational entropy associated with binding of the drug molecule, n_{r} is the number of internal degrees of conformational freedom lost on binding the drug molecule, and E_{r} is the energy equivalent of the entropy loss associated with the loss of each degree of conformational freedom on receptor binding.

1. *Intrinsic binding energy*

The final term in equation (21.9) is the sum of the binding energies E_{x} associated with each functional group X, of which there are n_{x} present in the drug. In the ideal case, when the specified functional group is aligned optimally and without strain with the corresponding functional group in the

FIGURE 21.12 Complex formation between a ligand (containing polar functionalities **A** and **B**) and a receptor (containing polar functionalities **D** and **E**) with exchange of four water molecules to the bulk solvent.

receptor, E_x is referred to as the intrinsic binding energy.[34] In other cases the term apparent binding energy is used.

It should be noted that each binding energy E_x is actually a combination of the various enthalpic and entropic interactions outlined above. These include the enthalpy of interaction between the functional group and its corresponding binding site on the receptor, the enthalpy changes associated with the removal of water of hydration from the functional group and its target site and the subsequent formation of bonds between the displaced water molecules, and the corresponding entropy terms associated with the displacement and subsequent bonding of water molecules (Figure 21.12).

It is apparent that these intrinsic binding energies may be regarded, at least approximately, as properties of the functional group that should be relatively independent of the groups to which the particular functional group is attached. Such intrinsic binding potentials might thus reasonably be used in an additive manner to provide an overall estimate of the drug–receptor interaction.

2. *Anchor principle*

It follows from equation (21.9) that the binding energy, E_x, due to the interaction between the receptor and a specific functional group, X, can be estimated by comparing the binding energies for pairs of compounds which differ only in the presence or absence of the specific functional group. This approach was first applied by Page[29] who referred to it as the "anchor principle." It is based on the premise that the difference in binding of a drug molecule with or without the particular functional group is due to factors associated solely with that group, that is, the binding energy E_x plus any degrees of conformational freedom lost specifically as a result of binding of group X. Other degrees of conformational freedom lost on binding and the loss of overall rotational and translational entropy associated with the remainder of the drug molecule (the anchor) are assumed to be unaffected by the presence or absence of X.

Similarly, the impact of a single amino acid substitution in the active site of an enzyme on transition-state stabilization, as determined by the change in either catalytic efficiency or inhibitor binding, provides a measure of the relative binding energy of the two side chains.

Clearly, the magnitude of the binding energies obtained using the anchor principle will vary widely with the quality of the interaction. If the functional groups are not properly aligned, as might reasonably be expected in many mutant proteins, a small or even repulsive interaction may result. Alternatively, the strength of the additional bond may be offset by a reduction in the strengths of the existing bonds. Under these circumstances the anchor principle will lead to an underestimate of the true bond strength.

3. *Average binding energy*

An alternative to the pair-by-pair approach inherent in the anchor principle was developed by Andrews *et al.*[31] who sought to average the contributions of individual functional groups to the observed binding energies of 200 ligand–protein interactions in aqueous solution. For this purpose, the average loss of overall rotational and translational entropy accompanying drug-receptor binding, $T\Delta S_{t,r}$ in equation (21.9), was estimated at 58.5 kJ/mol at 310 K. Regression analysis against n_r (obtained by counting the number of degrees of conformational freedom in each of the 200 ligand structures) and n_x (the number of occurrences of each functional group, X, in each of the 200 ligand structures) as the independent variables was then used to obtain average values of the binding energies associated with each functional group and for the loss of entropy associated with each degree of conformational freedom.

The results of this analysis showed that the loss of entropy associated with each internal rotation ΔG_r on receptor binding is equivalent to a reduction in the free energy of binding by average of 3 kJ/mol.

The corresponding binding energies obtained by the averaging process were C (sp^2 or sp^3) 3 kJ/mol; O, S, N, or halogen, 5 kJ/mol; OH and C═O, 10 and 14 kJ/mol, respectively; and CO_2^-, OPO_3^{2-}, and N^+, 34, 42 and 48 kJ/mol, respectively. Once again, it should be stressed that these are not intrinsic binding energies in the sense defined above. This would be the case only if each functional group in each drug in the series was optimally aligned with a corresponding functional group in the receptor. In fact, since every functional group of every drug was included in the analysis, the calculated values are averages of apparent

TABLE 21.4 Functional Group Contributions to Drug–Receptor Interactions (kJ/mol)

Functional group type	Technique employed to determine interaction energy		
	Anchor principle	Site-directed mutagenesis	Average energy
Nonpolar (per carbon atom)	12–14	1–3	3–6
H–bonding (uncharged)	16	2–6	5–14
H–bonding (charge-assisted)	20–42	15–19	
Charged (carboxyl, amine)	18–28	12–25	34–48
$T\Delta S_{t,r}$	12–60		58.5
ΔG_r (internal rotation)	5–6		3

TABLE 21.5 Average Values for the Parameters of Equation (21.10)[25]

Parameter	Physical process	Value (kJ/mol)
$\Delta G_{t,r}$	Energy cost of bimolecular association	+5.4
ΔG_r	Energy cost of restriction of an internal rotor	+1.4
ΔG_h	Benefit of the hydrophobic effect (per Å^2 of buried hydrocarbon)	−0.17 (Å^2)
ΔG_p	Benefit of making a neutral hydrogen bond of ideal geometry	−4.7
ΔG_{ionic}	Benefit of making an ionic hydrogen bond of ideal geometry	−8.3

binding energies, including those for some groups which may not interact with the receptor at all.[35] The calculated averages are thus almost certainly smaller than the corresponding intrinsic binding energies, although they follow expected trends in that charged groups lead to stronger interactions than polar groups, which in turn are stronger than nonpolar groups such as sp^2 or sp^3 carbons.

The apparent contributions of some functional groups and/or bond types to overall binding energies derived from the various studies reviewed above are summarized in Table 21.4. Also included are corresponding values used or suggested for the overall loss of rotational and translational entropy, $T\Delta S_{t,r}$, and the loss of conformational entropy resulting from restriction of free rotation, ΔG_r. These first attempts at a semiquantification of drug–receptor interactions in terms of costs and benefits were later refined by Williams *et al.*[36,37] who applied an additional term to account for the hydrophobic effect. They suggested that the magnitude of this effect is proportional to the surface area of the hydrocarbon that is removed from exposure to water upon formation of the complex. It can be estimated in terms of ΔG_h per unit of area A of hydrocarbon buried, which can readily be measured with the aid of computer graphics. It leads to the following equation:

$$\Delta G = \Delta G_{t,r} + n\Delta G_r + A\Delta G_h + \Sigma\Delta G_p \qquad (21.10)$$

ΔG is the observed free energy of binding, $\Sigma\Delta G_p$ ist he sum of the free energies of binding for all the polar interactions made in the binding site, and the other terms are as defined above. Böhm "trained" a variant of equation (21.10) with a set of 45 interactions of experimentally known binding constants from the association of ligands of small molecular weight with proteins through sets of known interactions.[38] He divided the original ΔG_p values into two groups: those involving ionic interactions (ΔG_{ionic}) and those involving hydrogen bonds formed between neutral entities (the term ΔG_p was retained). Since the modified form of the equation has only five types of ΔG contributions and the 45 binding sites involve different combinations of these five types of ΔG contributions, average values for them can be obtained. These values (Table 21.5) have proved very useful in the pharmaceutical industry.

It is interesting to note that the average value $\Delta G_{t,r} = +5.4$ kJ/mol is remarkably small, and represents only about one tenth of the maximum theoretical entropy loss corresponding to complete immobilization of the ligand. This small value presumably reflects, at least in part, the large residual motion that the drugs can exercise relative to the receptor to which they are bound. The average cost of restricting the rotation of an internal bond in the drugs (+1.4 kJ/mol) is slightly less than that found for the formation of crystals from neat liquids that contain internal rotors (2–3 kJ/mol). This finding probably reflects the fact that rotations are somewhat less restricted in these binding sites than they are in crystals. Most importantly, the application of the equation gives useful approximate binding constants in many cases.

B. The methyl group and other nonpolar substituents

The initial application of the anchor principle described by Page[29] related to data on the selectivity of amino acid-tRNA synthetases, from which he estimated intrinsic

FIGURE 21.13 Isoleucine **2** and desmethyl-isoleucine **3**.

binding energies for the methylene group in the range 12–14 kJ/mol. For example, the calculated binding energies of equation (21.3) for isoleucine **2** (Figure 21.13) and its desmethyl analog **3** to isoleucyl-tRNA synthetase are 29.7 and 15.9 kJ/mol, respectively, indicating that the methyl group contributes a total of 13.8 kJ/mol to the overall interaction.

This estimate, having been derived from observations on a highly selective enzyme–substrate interaction, is probably also approaching the intrinsic limit for the binding contribution of a methyl group. Indeed, according to the calculation of Williams[36], the burial of a methyl group ($\approx$25 Å^2) contributes only -4 kJ/mol. For longer hydrocarbon side chains the positive contribution due to dispersion forces and hydrophobic interactions tends to be offset by the loss of conformational entropy on binding. Thus, the "average" binding energy of 3 kJ/mol obtained by Andrews *et al.* for sp^2 and sp^3 carbon groups is identical to the "average" reduction in free energy of binding estimated for the loss of conformational freedom around a single bond.[31] Clearly, this effect will be greater in saturated hydrocarbon chains than their more conformationally constrained unsaturated or cyclic analogs.

C. The hydroxyl group and other hydrogen-bond forming substituents

The most extensive studies of hydroxyl group contributions to drug–receptor interactions are those of Wolfenden *et al.* on the contribution of hydrogen bonds formed by hydroxyl groups in transition-state analogs. In a series of 13 examples of paired ligands with and without hydroxyl groups, they used[39] the anchor principle to determine apparent binding energies for single hydroxyl groups ranging from 20–42 kJ/mol.

Thus, in comparing the binding of 1,6-dihydropurine ribonucleoside **4** (Figure 21.14) and its 6-hydroxy derivative **5** to adenosine deaminase, they observed a difference in binding energy of 41 kJ/mol.[40] The authors suggested that the 6-hydroxyl group, which has very limited freedom of movement, is likely to be in almost ideal alignment with the active site for forming a hydrogen bond. This conjecture has been verified by the determination of the crystal structure of the inhibitory complex between adenosine deaminase and 6-hydroxy-1,6-dihydropurine ribonucleoside, which showed that the 6-hydroxyl group interacts

FIGURE 21.14 1,6-Dihydropurine ribonucleoside **4** and its 6-hydroxy analog **5**.

with a zinc atom, with a protonated histidyl residue, and with an aspartic acid residue at the enzyme's active site.[41]

Once again, the data from active-site mutagenesis studies are less striking, but nevertheless reveal some very substantial hydrogen-bonding interactions. In Fersht's studies[42] on tyrosyl-tRNA synthetase, for example, hydrogen bonds between this enzyme and uncharged substrate groups contributed between 2 and 6 kJ/mol towards specificity, while hydrogen bonds to charged groups contributed between 15 and 19 kJ/mol, corresponding to a factor of 1,000 in specificity. These numbers are however higher than the average contributions determined by Böhm[38] (4.7 and 8.3 kJ/mol, respectively).

D. Acidic and basic substituents

Simple observations on the interactions of individual charged groups with appropriate enzymes may lead to an indication of their binding energies. The phosphate ion, for example, binds alkaline phosphatase[43] with a dissociation constant of 2.3×10^{-6} M, equivalent to a ΔG value of approximately 33 kJ/mol. Taking the most conservative estimate for the loss of rotational and translational entropy associated with this interaction, 12 kJ/mol for a loosely bound complex, equation (21.9) then gives a lower estimate for binding of the phosphate ion of 45 kJ/mol. If the same value of $T\Delta S_{t,r}$ is applied to the binding of oxalate ion to transcarboxylase,[44] for which the dissociation constant is 1.8×10^{-10} M (57 kJ/mol), equation (21.9) gives an apparent binding energy of 24 kJ/mol per carboxylate group after allowance for a minimal conformational entropy loss of 3 kJ/mol.

These figures are broadly consistent with the average values of Andrews *et al.*[31] which were in the range 34–48 kJ/mol for charged phosphate,[45] amine and carboxyl groups.

E. Practical applications for the medicinal chemist

1. *Assessing a lead compound*

Summation of the average contributions of individual binding groups, including allowance for conformational,

rotational and translational entropy terms as shown in equation (21.9), provides a simple back-of-the-envelope calculation of the strength of binding which might be expected for a drug forming a typical interaction with a receptor. This figure, when compared to the observed affinity of the drug for the target receptor, then gives a direct indication of the actual quality of the electrostatic and steric match between the drug and the receptor.

a. Binding is tighter than expected

If the observed binding of a drug to its receptor turns out to be substantially stronger than that calculated from equation (21.9) it is reasonable to expect that the drug structure offers a good fit to the receptor in a reasonably low-energy conformation. The structure should therefore provide an excellent starting point for the development of even more bioactive compounds.

A good example of this is biotin (Figure 21.15), which was the most extreme case of a positive deviation from the calculated "average" in the original set of 200 ligand–protein interactions studied by Andrews *et al.*[31]

$$\begin{aligned}\Delta G_{\text{av}} &= T\Delta S_{rt} + 5E_{\text{r}} + 8E_{\text{Csp}^3} + 2E_{\text{N}} \\ &\quad + E_{\text{S}} + E_{\text{C=O}} + E_{\text{COOH}} \\ &= -58.5 + 5(-3) + 8(3) + 2(5) \\ &\quad + 5 + 14 + 34 = 13.5\ \text{KJ/mol}\end{aligned} \tag{21.11}$$

$$\begin{aligned}\Delta G_{\text{obs}} &= -5.85 \log K_{\text{d}} \\ &= -5.85(-15) = 87.7\ \text{KJ/mol}\end{aligned} \tag{21.12}$$

Application of equation (21.9) to biotin (see above) gives an « average » binding energy of 13.5 kJ/mol, whereas substitution into equation (21.3) of the experimentally observed binding constant to the protein avidin ($10^{-15}\,\text{mol}^{-1}$) gives a binding energy of 87.7 kJ/mol. The difference of almost 74 kJ/mol implies an exceptionally good fit between biotin and the structure of the protein. It has since been established that this is indeed the case, with polarization of the biotin molecule by the protein actually leading to an ionic interaction where a neutral hydrogen-bonding interaction had been assumed.

FIGURE 21.15 Structure of biotin.

b. Binding is looser than expected

If the observed binding is significantly weaker than anticipated on the basis of an "average" energy calculation, the fit between the drug and the receptor is less than perfect. In some cases this will be because the match between drug and receptor is less a matter of hand and glove than of "square peg and round hole," and the only realistic option for the drug designer is to start again.

In other cases, simpler remedies may be followed:

1. The fit may be unsatisfactory because only part of the drug is interacting with the receptor. This situation applies particularly to large drug molecules (e.g. peptide hormones), for which selective pruning of unused parts of the structure may produce simpler compounds without loss of affinity;
2. The drug may be binding to the receptor in a comparatively high energy conformation. In this case the design of more rigid structures which are already fixed in the desired conformation will give an increase in binding energy equivalent to the conformational energy cost of binding the more flexible analog.

2. *Assessing the effectiveness of substituents*

Equally simple back-of-the-envelope calculations based on equation (21.9) can be used to predict the increase in binding energy which might be expected upon the addition of a functional group which is optimally aligned with a corresponding group in the receptor. This figure, when compared to the observed increase in affinity, gives direct feedback on whether or not the new group is actually performing the function anticipated in the design strategy.

An interesting example of how this approach can be used to assess the validity of a drug design hypothesis is provided by the receptor-based design of sialidase inhibitors as potential anti-influenza drugs. Starting from the knowledge[46] of the structurally invariant active site of influenza A and sialidases, von Itzstein *et al.*[47] postulated that substitution of the 4-hydroxyl group of the nonselective sialidase inhibitor 2-deoxy-2,3-didehydro-D-acetylneuraminic acid **6** (Figure 21.16) with a positively charged

FIGURE 21.16 Sialidase inhibitors: 2-deoxy-2,3-didehydro-D-*N*-acetyl neuraminic acid **7** and its **4**-guanidino analog **8**.

substituent would fill an occupied pocket lined with anionic residues. Synthesis and testing of the 4-guanidino analog **7** revealed a reduction in K_i from 10^{-6} to 10^{-10} mole/l, equivalent to an additional binding energy of 23 kJ/mol. Although not at the upper limit of the increments in binding energy anticipated for well-aligned ionic interactions on the basis of the data in Table 21.4, this figure is certainly consistent with the design hypothesis, as is borne out by the crystal structure of the complex.[47] This shows that the guanidino lies between two target carboxyl groups in the active site of the enzyme, although only one appears to be optimally placed for a strong interaction.

F. Ligand efficiency

The idea of ligand efficiency (LE) has recently emerged as a useful guide to optimize fragment and lead selection in the discovery process.[48,49] Preliminary work has been published by Kuntz *et al.*[50] This key contribution, where affinities were examined for a variety of ligands against many different targets, showed that ΔG tends to increase little with molecular mass when the ligand contains more than *ca.* 15 heavy atoms (HA). Later, Hopkins *et al.*[48] proposed to define LE as the binding free energy for a ligand divided by its number of HA:

$$\mathrm{LE} = -\Delta G/\mathrm{HA} \tag{21.13}$$

According to them, comparison of lead compounds on the basis of LE rather than the potency alone could be useful in deciding the potential for further optimization for particular hits and chemical scaffold. LE has since become an important concept in drug discovery partly due to the realization that large ligands have a decided disadvantage in terms of the molecular properties necessary for bioavailability.[51,52] LE is notably used when prioritizing the output from HTS or other screening strategies. It is helpful, particularly when trying to assess the relative value of fragments for follow up in fragment-based drug design.[53] To obtain a final compound with MW < 500 and ≈10 nM potency, LE needs to stay above 1.25 kJ/mol per HA.[48]

V. COOPERATIVE BINDING

Equations (21.9) and (21.10) are based on the assumption that contributions to binding energies can be partitioned in term of individual interactions, and that these individual binding energies are additive and independent of each other. However, in general, it is impossible to study one binding interaction in isolation from the others at an interface. In practise indeed, cooperativity between noncovalent interactions is observed.

Noncovalent interactions are said to interact with each other in a positively (or negatively) cooperative manner when the binding energy that is derived from their acting together is greater (or smaller) than would be derived from the sum of their acting separately. This concept developed by Williams *et al.*[25,54,55] is illustrated in Figure 21.17.

The consequence of a positive cooperativity is that a structural tightening occurs in the bound state with a benefit in enthalpy and a smaller cost in entropy. By studying the unusually strong reversible binding of biotin by avidin ($K_a \approx 10^{15}$ mol/l) and streptavidin ($K_a \approx 10^{13}$ mol/l), Houk *et al.*[56] observed that the five hydrogen bonds of the ligand–receptor complex act cooperatively (Figure 21.18), leading to stabilization that is larger than the sum of individual hydrogen-bonding energies. The charged aspartate is

FIGURE 21.17 Schematic representation of a receptor that binds ligands **X**, **Y**, and **Z** with affinities ΔG_X, ΔG_Y, and ΔG_Z, respectively. (a) Binding of **Z** results in a structure with an intermolecular distance d_0. (b) When **Y** and **Z** are connected by a rigid, strain-free linker (**Y–Z**) they bind to the receptor with positive cooperativity (ΔG_{Y-Z} more negative than $\Delta G_Y + \Delta G_Z$) and there is structural tightening ($d_1 < d_0$). (c) If **X** is connected to **Y–Z** by a rigid, strain-free linker to form **X–Y–Z** then further structural tightening will occur ($d_2 < d_1$) leading to a further cooperative enhancement. (d) The shorter linker between **Y** and **Z** does not allow both these binding interactions to occur with optimal geometry. **Y–Z** binds the receptor with negative cooperativity (ΔG_{Y-Z} more positive than $\Delta G_Y + \Delta G_Z$) and there is structural loosening ($d_3 > d_0$).

FIGURE 21.18 Schematic representation of streptavidin–biotin interactions.

the key residue that provides the driving force for cooperativity by greatly polarizing the urea of biotin. If the residue is removed, the network is disrupted.

Following the same principle, a negative cooperativity produces a decrease of the ligand–receptor interaction strength. This is illustrated by Figure 21.17d) where **Y** and **Z** are rigidly held in a conformation that does not allow both binding interactions to occur with the preferred geometry. This situation could be induced, for example, by introducing a linker between **Y** and **Z** that is too short. The "pull" of **Y** towards its preferred binding geometry will adversely affect the binding of **Z** by forcing it away from its preferred binding geometry, and vice versa (ΔG_{Y-Z} is less negative than $\Delta G_Y + \Delta G_Z$).

REFERENCES

1. Klebe, G., Dullweber, F., Böhm, H.-J. Thermodynamic models of drug-receptor interactions: a general introduction. In *Drug-Receptor Thermodynamics: Introduction and Applications* (Raffa, R. B., Ed.). Wiley: Chichester, 2001, pp. 83–104.
2. Silverman, R. B. Drug-receptor interactions. *The Organic Chemistry of Drug Design and Drug Action*. Elsevier Academic Press: Amsterdam, 2004, pp. 123–131.
3. Motiejunas, D., Wade, R. C. Structural, energetic, and dynamic aspects of ligand-receptor interactions. In *Comprehensive Medicinal Chemistry* (Taylor, J. B., Triggle, D. J., Eds), Vol. 4. Elsevier, 2007, pp. 193–213.
4. Di Cera, E. *Thermodynamic Theory of Site-Specific Binding Processes in Biological Macromolecules*. Cambridge University Press: Cambridge, 1995.
5. Böhm, H.-J., Klebe, G. What can we learn from molecular recognition in protein-ligand complexes for the design of new drugs?. *Angew. Chem. Int. Ed. Engl.* 1996, *35*, 2588–2614.
6. Goldstein, A., Aronow, L., Kalman, M. S. *Principles of Drug Action. The Basis of Pharmacology*. Hoeber Medical Division: New York, 1968, p. 884.
7. Dawson, R. M. C., Elliot, D. C., Elliot, W. H., Jones, K. M. *Data for Biochemical Research*. Oxford University Press: Oxford, 1969.
8. Paulini, R., Muller, K., Diederich, F. Orthogonal multipolar interactions in structural chemistry and biology. *Angew. Chem. Int. Ed. Engl.* 2005, *44*, 1788–1805.
9. Bajorath, J., Kraut, J., Li, Z. Q., Kitson, D. H., Hagler, A. T. Theoretical studies on the dihydrofolate reductase mechanism: electronic polarization of bound substrates. *Proc. Natl. Acad. Sci. USA* 1991, *88*, 6423–6426.
10. Jeffrey, G. A. *An Introduction to Hydrogen Bonding*. Oxford University Press: New York, 1997.
11. Stahl, M., Bohm, H. J. Development of filter functions for protein-ligand docking. *J. Mol. Graph. Model.* 1998, *16*, 121–132.
12. Williams, M. A., Ladbury, J. E. Hydrogen bonds in protein-ligand complexes. In *Protein-Ligand Interactions. From Molecular Recognition to Drug Design* (Böhm, H.-J., Schneider, G., Eds). Wiley-VCH: Weinheim, 2003, p. 139.
13. Fersht, A. *Enzyme Structure and Mechanism*. Freeman: New York, 1985.
14. Fersht, A. The hydrogen bond in molecular recognition. *Trends Biochem. Sci.* 1987, *12*, 301–304.
15. Knox, J. R., Pratt, R. F. Different modes of vancomycin and D-alanyl-D-alanine peptidase binding to cell wall peptide and a possible role for the vancomycin resistance protein. *Antimicrob. Agents Chemother.* 1990, *34*, 1342–1347.
16. Ma, J. C., Dougherty, D. A. The cation-π interaction. *Chem. Rev.* 1997, *97*, 1303–1324.
17. Meyer, E. A., Castellano, R. K., Diederich, F. Interactions with aromatic rings in chemical and biological recognition. *Angew. Chem. Int. Ed. Engl.* 2003, *42*, 1210–1250.
18. Niedzwiecka, A., Marcotrigiano, J., Stepinski, J., Jankowska-Anyszka, M., Wyslouch-Cieszynska, A., Dadlez, M., Gingras, A. C., Mak, P., Darzynkiewicz, E., Sonenberg, N., Burley, S. K., Stolarski, R. Biophysical studies of eIF4E cap-binding protein: recognition of mRNA 5′ cap structure and synthetic fragments of eIF4G and 4E-BP1 proteins. *J. Mol. Biol.* 2002, *319*, 615–635.
19. Schärer, K., Morgenthaler, M., Paulini, R., Obst-Sander, U., Banner, D. W., Schlatter, D., Benz, J., Stihle, M., Diederich, F. Quantification of cation-pi interactions in protein-ligand complexes: crystal-structure analysis of factor Xa bound to a quaternary ammonium ion ligand. *Angew. Chem. Int. Ed. Engl.* 2005, *44*, 4400–4404.
20. Di Fenza, A., Heine, A., Koert, U., Klebe, G. Understanding binding selectivity toward trypsin and factor Xa: the role of aromatic interactions. *Chem. Med. Chem.* 2007, *2*, 297–308.
21. Burley, S. K., Petsko, G. A. Weakly polar interactions in proteins. *Adv. Protein Chem.* 1988, *39*, 125–189.
22. Weichsel, A., Montfort, W. R. Ligand-induced distortion of an active site in thymidylate synthase upon binding anticancer drug 1843U89. *Nat. Struct. Biol.* 1995, *2*, 1095–1101.
23. Dunitz, J. D. Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions. *Chem. Biol.* 1995, *2*, 709–712.
24. Silverstein, K. A. T., Haymet, A. D. J., Dill, K. A. The strength of hydrogen bonds in liquid water and around nonpolar solutes. *J. Am. Chem. Soc.* 2000, *122*, 8037–8041.
25. Williams, D. H., Stephens, E., O'Brien, D. P., Zhou, M. Understanding noncovalent interactions: ligand binding energy and catalytic efficiency from ligand-induced reductions in motion within receptors and enzymes. *Angew. Chem. Int. Ed. Engl.* 2004, *43*, 6596–6616.
26. Davis, A. M., Teague, S. J. Hydrogen bonding, hydrophobic interactions, and failure of the rigid receptor hypothesis. *Angew. Chem. Int. Ed. Engl.* 1999, *38*, 736–749.
27. Vandonselaar, M., Hickie, R. A., Quail, J. W., Delbaere, L. T. Trifluoperazine-induced conformational change in Ca(2+)-calmodulin. *Nat. Struct. Biol.* 1994, *1*, 795–801.
28. Yu, Y. B., Privalov, P. L., Hodges, R. S. Contribution of translational and rotational motions to molecular association in aqueous solution. *Biophys. J.* 2001, *81*, 1632–1642.
29. Page, M. I. Entropy, binding energy, and enzymic catalysis. *Angew. Chem. Int. Ed. Engl.* 1977, *16*, 449–459.
30. Page, M. I., Jencks, W. P. Entropic contributions to rate accelerations in enzymic and intramolecular reactions and the chelate effect. *Proc. Natl. Acad. Sci. USA* 1971, *68*, 1678–1683.
31. Andrews, P. R., Craik, D. J., Martin, J. L. Functional group contributions to drug-receptor interactions. *J. Med. Chem.* 1984, *27*, 1648–1657.
32. Gohlke, H., Klebe, G. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angew. Chem. Int. Ed. Engl.* 2002, *41*, 2644–2676.
33. Lau, W. F., Pettitt, B. M. Selective elimination of interactions: a method for assessing thermodynamic contributions to ligand binding with application to rhinovirus antivirals. *J. Med. Chem.* 1989, *32*, 2542–2547.
34. Jencks, W. P. On the attribution and additivity of binding energies. *Proc. Natl. Acad. Sci. USA* 1981, *78*, 4046–4050.
35. Andrews, P. R., Kubinyi, H., Ed.)*3D QSAR in drug design: theory, methods and applications*. ESCOM Science Publishers B.V.: Leiden, 1993, pp. 13–40.
36. Williams, D. H., Cox, J. P. L., Doig, A. J., Gardner, M., Gerhard, U., Kaye, P. T., Lal, A. L., Nichols, I. A., Salter, C. J., Mitchell, R. C. Toward the semiquantitative estimation of binding constants. Guides

for peptide-peptide binding in aqueous solution. *J. Am. Chem. Soc.* 1991, *113*, 7020–7030.
37. Searle, M. S., Williams, D. H. The cost of conformational order: Entropy changes in molecular associations. *J. Am. Chem. Soc.* 1992, *114*, 10690–10697.
38. Böhm, H. J. The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure. *J. Comput. Aided Mol. Des.* 1994, *8*, 243–256.
39. Wolfenden, R., Kati, W. M. Testing the limits of protein-ligand binding discrimination with transition-state analogue inhibitors. *Acc. Chem. Res.* 1991, *24*, 209–215.
40. Kati, W. M., Wolfenden, R. Contribution of a single hydroxyl group to transition-state discrimination by adenosine deaminase: evidence for an "entropy trap" mechanism. *Biochemistry* 1989, *28*, 7919–7927.
41. Wilson, D. K., Rudolph, F. B., Quiocho, F. A. Atomic structure of adenosine deaminase complexed with a transition-state analog: understanding catalysis and immunodeficiency mutations. *Science* 1991, *252*, 1278–1284.
42. Fersht, A. R., Shi, J. P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M., Winter, G. Hydrogen bonding and biological specificity analysed by protein engineering. *Nature* 1985, *314*, 235–238.
43. Levine, D., Reid, T. W., Wilson, I. B. The free energy of hydrolysis of the phosphoryl-enzyme intermediate in alkaline phosphatase catalyzed reactions. *Biochemistry* 1969, *8*, 2374–2480.
44. Northrop, D. B., Wood, H. G. Transcarboxylase. VII. Exchange reactions and kinetics of oxalate inhibition. *J. Biol. Chem.* 1969, *244*, 5820–5827.
45. Hirsch, A. K., Fischer, F. R., Diederich, F. Phosphate recognition in structural biology. *Angew. Chem. Int. Ed. Engl.* 2007, *46*, 338–352.
46. Colman P. M., Krug, R. M., (Ed.) *The influenza virus*. Plenum Press: New York, 1989, pp. 175–218.
47. von Itzstein, M., Wu, W.-Y., Kok, G. B., Pegg, M. S., Dyason, J. C., Jin, B., Phan, T. V., Smythe, M. L., White, H. F., Oliver, S. W., Colman, P. M., Varghese, J. N., Ryan, D. M., Woods, J. M., Bethell, R. C., Hotham, V. J., Cameron, J. M., Penn, C. R. Rational design of potent sialidase-based inhibitors of influenza virus replication. *Nature* 1993, *363*, 418–423.
48. Hopkins, A. L., Groom, C. R., Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discov. Today* 2004, *9*, 430–431.
49. Abad-Zapatero, C., Metz, J. T. Ligand efficiency indices as guideposts for drug discovery. *Drug Discov. Today* 2005, *10*, 464–469.
50. Kuntz, I. D., Chen, K., Sharp, K. A., Kollman, P. A. The maximal affinity of ligands. *Proc. Natl. Acad. Sci. USA* 1999, *96*, 9997–10002.
51. Abad-Zapatero, C. Ligand efficiency indices for effective drug discovery. *Expert Opinion on Drug Discovery* 2007, *2*, 469–488.
52. Reynolds, C. H., Bembenek, S. D., Tounge, B. A. The role of molecular size in ligand efficiency. *Bioorg. Med. Chem. Lett.* 2007, *17*, 4258–4261.
53. Leach, A. R., Hann, M. M., Burrows, J. N., Griffen, E. J. Fragment screening: an introduction. *Mol. Biosyst.* 2006, *2*, 430–446.
54. Calderone, C. T., Williams, D. H. An enthalpic component in cooperativity: the relationship between enthalpy, entropy, and noncovalent structure in weak associations. *J. Am. Chem. Soc.* 2001, *123*, 6262–6267.
55. Williams, D. H., Davies, N. L., Zerella, R., Bardsley, B. Noncovalent interactions: defining cooperativity. Ligand binding aided by reduced dynamic behavior of receptors. Binding of bacterial cell wall analogues to ristocetin A. *J. Am. Chem. Soc.* 2004, *126*, 2042–2049.
56. DeChancie, J., Houk, K. N. The origins of femtomolar protein-ligand binding: hydrogen-bond cooperativity and desolvation energetics in the biotin-(strept)avidin binding site. *J. Am. Chem. Soc.* 2007, *129*, 5419–5429.

Chapter 22

Compound Properties and Drug Quality

Christopher A. Lipinski

I. INTRODUCTION
II. COMBINATORIAL LIBRARIES
A. Library design for HTS screens
B. Experimental synthesis success rate
C. Poor solubility and library design
D. Importance of the synthesis rate-determining step
E. If protocol development is rate determining
F. Poor ADME properties – business aspects
G. If library production is rate determining
H. Relative importance of ADME assays
III. CHEMISTRY CONTROL OF INTESTINAL PERMEABILITY
A. Improving permeability
B. Hydrogen bonding and permeability
C. Intramolecular hydrogen bonds
D. Permeability testing
IV. CHEMISTRY CONTROL OF AQUEOUS SOLUBILITY
A. The definition of poor solubility
B. Aqueous solubility and blunt SAR
C. Changing the pK_a
D. Improving aqueous solubility
V. *IN VITRO* POTENCY AND CHEMISTRY CONTROL
A. Lead complexity
VI. METABOLIC STABILITY
A. ADME computational models
B. Limitations of Caco-2 cell culture
C. Poor aqueous solubility and permeability assay noise
D. Physiologically relevant screening concentration
VII. ACCEPTABLE SOLUBILITY GUIDELINES FOR PERMEABILITY SCREENS
A. Batch-mode solubility prediction
REFERENCES

Our achievements speak for themselves. What we have to keep track of are our failures, discouragements, and doubts. We tend to forget the past difficulties, the many false starts, and the painful groping. We see our past achievements as the end result of a clean forward thrust, and our present difficulties as signs of decline and decay.

Eric Hoffer

I. INTRODUCTION

Compound properties required for an orally active drug have not changed in the last decade with respect to developability, that is, the probability that a compound will succeed to market approval.[1] Considerable literature supports the stability over time of the physicochemical properties of an approved drug. For example, the mean MWT of Food and Drug Administration (FDA) approved drugs is stable over 40 years at about 350 Da.[2] Developability stands in contrast to druggability, that is, the properties adequate for a compound to enter into clinical studies. With respect to druggability compound properties have markedly changed in recent years. A compound now, as in the past, must still have adequate solubility and permeability relative to its clinical potency to enter into clinical studies as an oral drug. However, what has changed in the last decade is the success of medicinal chemists in pushing the envelope of physicochemical properties of compounds entering into phase I. It is now abundantly clear from published ligand profiles that some very

Copyright © 2008, Elsevier Ltd
All rights reserved.

interesting biological targets are barely druggable[3] from a "Rule-of-5"[4] (Ro5) perspective. The new challenge to medicinal chemistry is the divergence between druggability and developability. It is now possible, albeit with considerable effort and resources, to discover non-Ro5 ligands that have the credentials to enter phase I against difficult targets. However, there is no evidence, at this time, that the historically very stable profiles of approved drugs are changing as a result of the interest in the newer more difficult targets. As a result of synthetic chemistry practice changes in the last decade we now increasingly see a dichotomy in medicinal chemistry. On the one hand, compounds, especially those intended for traditional high-throughput screening (HTS) still tend to be larger and more lipophilic or larger with more hydrogen-bonding functionality than in the previous decade. On the other hand, it is now abundantly clear that fragment screening is a viable alternative to HTS albeit with some limitations. The changes in HTS screening libraries continue to translate into poorer aqueous solubility and poorer intestinal permeability than in the previous era and the advent of fragment screening[5] now requires high aqueous solubilities in the high hundreds of micromolar to millimolar range. As a result, there is a greater divergence in compound properties between the early discovery stage/discovery–development interface compared to the later clinical stages than in the era covered by the previous edition of *The Practice of Medicinal Chemistry*. At the risk of repetition, the vast majority of recent chemistry changes are in the early discovery/discovery development interface stage and there are virtually no property changes at the later clinical development stages.

Accordingly, the discussions in the previous edition of this book on the properties of crystalline, well-characterized compounds intended for oral dosing which are entering the development phase, still remain valid. What is new is the phenomena of early stage discovery compounds with properties frequently far removed from those historically optimal in the later clinical stages. This chapter focuses on the recent changes in synthetic chemistry practices, shows how these changes modulated by the target impact on compound properties, and provides guidance on how the pattern of early compound properties can be improved toward those described in the previous edition of *The Practice of Medicinal Chemistry*. Changes in medicinal chemistry synthetic practice can be broadly characterized into three main areas: (1) changes related to synthesis of one compound at a time and moderate output parallel synthesis, (2) the more radical changes related to combinatorial chemistry, and (3) the chemistry involved in generation of fragment screening libraries. This chapter primarily focuses on issues related to the design of combinatorial libraries and secondarily with the newer area of fragment libraries. A highly technical discussion in book chapter format tends to rapidly become out of date. Thus, the main focus of both chapter sections will be on general principles, which tend not to be well covered in primary journal technical publications.

II. COMBINATORIAL LIBRARIES

The design of combinatorial libraries with drug-like properties is a trade-off between efficient chemistry and therefore better numerical compound output on one side and increasingly difficult chemistry but better physicochemical properties on the other. This trade-off is not solely a technical issue but is strongly influenced by people attitudes and organizational timing issues. The author believes that there is a hierarchy of properties that need to be controlled in a quality library. Simply put, those properties that are under poorest chemistry control and that are the most difficult to fix in chemistry optimization should be given highest priority. In this context, poor chemistry control equates to chemistry structure–activity relationship (SAR) which is difficult to control. Avoiding chemistry functionality known to be associated with toxicity is an example of chemistry SAR that is under good control. Figure 22.1 shows a chemical structure that we call "the molecule from hell." It was created by my computational colleague Dr. Beryl Dominy and contains chemistry functionality that in a single molecule triggers 27 different filters for chemistry functionality known to be associated with mutagenic activity. A molecule like this is used to test that a computational filter will detect the undesirable chemistry functionality. An example of chemistry SAR that is under poor control is the SAR relating to aqueous solubility and gastrointestinal tract permeability. Physicochemical profiling of solubility and permeability has recently been reviewed. These properties are difficult but not impossible to control by medicinal chemistry. In a combinatorial chemistry setting, poor aqueous solubility is almost universally a problem and must be explicitly addressed by both computational and experimental intervention strategies.

A. Library design for HTS screens

Library design viewed as a computational process and the experimental implementation of a library design can result

FIGURE 22.1 The "molecule from Hell."

in quite different outcomes. In a library design, one might consider a core template which can be modified by a variety of substituent groups. The physicochemical profile of the library is determined by the chemical structure of the core and the range of physicochemical properties of the substituent groups. For example, there might only be one core in a design but many substituent groups. As a result the physicochemical envelope of the library is very much influenced both by the availability of the chemical pieces that will become the substituents and by the physicochemical properties of those substituents. If the experimental implementation (synthesis) of a design is 100% successful then the actual experimental physicochemical envelope profile is the same as in the library design. In actuality, the chemical synthesis is seldom 100% successful. Thus, what matters is whether there is any bias in the experimental synthesis success rate such that the experimental library differs significantly from the design library. *A priori*, the chemical synthesis success rate is usually not known so that, in general, libraries are over designed. More compounds are designed than will actually ever be synthesized.

Since the last edition of this chapter we have learned quite a bit about the relative importance and balance of cores (scaffolds) and substituent groups in successful combinatorial libraries. The general trend in automated chemistry has been toward the synthesis of smaller libraries. For example, a collection of 10 libraries each with a different core and each containing 100 members works better in lead discovery than a library of 1,000 members all containing the same core. It is now apparent based on a decade of experience in fragment screening that a great deal of the binding energy of ligand to target comes more from the core rather than the substituents.[6] This explains the greater screening success of the smaller libraries since each library probes a different core even though generally the chemistry effort to make the multiple smaller size libraries is more difficult than to make the single larger library.

B. Experimental synthesis success rate

The experimental synthesis success rate almost always biases the experimental library so that the physicochemical profile relative to aqueous solubility is significantly worse than that in the design library. The fewest chemistry problems are found in lipophilic substituent moieties lacking polar functionality. In almost all cases polar functionality is electron withdrawing so that reactions of a substituent moiety like reactive amination, acylation or nucleophilic substitution proceed more poorly. Blocking and deblocking of a polar group adds to the complexity and length of a synthesis. As a result polar reagents which require blocking and deblocking are experimentally selected against. Robotic pipettors perform poorly or not at all on slurries of precipitates so any factor that increases the insolubility of a reagent in an organic solvent will bias the library outcome.

C. Poor solubility and library design

Poor solubility in an organic solvent arises from two quite different factors: solvation energy and disruption of intermolecular crystal packing forces in the crystalline reagent. Solvation of a lipophilic reagent in an organic solvent is typically not a problem. However, disruption of intermolecular crystal packing forces is very much a problem in an organic solvent, especially if the reagent has a high melting point. This type of problem is most likely to be present in a reagent with polar hydrogen-bond acceptor/donor functionality. Thus, the reagent insolubility problem tends to bias a library toward a more lipophilic and hence more aqueous insoluble profile. To accommodate diversity considerations a range of substituent moieties is selected. A large structural range translates into a broad molecular weight distribution. The combination of reagent solubility and diversity considerations results in an experimental library that is biased toward higher lipophilicity and higher molecular weight relative to the design library. The bias occurs because high lipophilicity and high molecular weight are the worst combination of "Rule-of-5" parameters in terms of leading to poor aqueous solubility. We now know that poor aqueous solubility due to strong intermolecular crystal packing forces in the crystalline reagent is clinically by far the worst solubility cause. An analysis of clinically approved poorly aqueous soluble drugs shows that those reaching marketing approval are far more likely to be insoluble because of solvation, that is, they are large and lipophilic, the so-called greaseballs. By contrast those insoluble due to crystal packing, the so-called brick dust compounds are very unlikely to reach marketing approval.[7] Likely it is the considerably greater opportunity for formulation fixes for the greaseballs that makes these compounds more likely to reach the approval stage.

D. Importance of the synthesis rate-determining step

Effectiveness of ADME (absorption, distribution, metabolism, excretion) design implementation depends on whether chemistry protocol development or chemistry production is rate determining. If chemistry production is rate determining there will be excess validated protocols relative to library production. This means that protocols can be prioritized as to their ADME attractiveness and the least attractive protocols from an ADME perspective may never be translated into actual library production. However, protocol development and not library production is often the rate-determining step. This eventuality is unfortunate because there is an understandable reluctance to discontinue chemistry synthetic efforts because of a poor ADME experimental profile if considerable chemistry

effort has already been expended. Consider the following situation: the effort toward library production is 70% complete and the experimental ADME profile is poor. Would you discontinue completion of library synthesis because of poor ADME if 70% of the chemistry effort had already been completed? Thus a key issue becomes how much chemistry experimental effort takes place before exemplars are experimentally profiled in ADME screens.

E. If protocol development is rate determining

If protocol development is rate determining, the effectiveness of ADME experimental assays depends very much on how the early exemplars are synthesized. In theory, the most effective method would be to obtain a well-spaced subset of the library in an experimental design sense. A traditional, noncombinatorial, synthesis would accomplish this but would not fit in well with a combinatorial optimization process. A possible means around this problem is to institute some type of early automated clean-up of combinatorial exemplars from partially optimized reaction schemes. This is not a tidy solution because the most efficient process would be an automated clean-up on the entire library after the optimization process was complete.

The least effective method of providing samples for experimental ADME profiling is a late-stage selection from the optimized combinatorial libraries. It is least effective, not because of chemistry efficiency considerations. A late-stage selection from the optimized combinatorial libraries is actually very chemistry efficient. However, the inefficiency comes from the aspect of people. The data come too late to prevent poor ADME quality compounds from being made. The timing problem in obtaining combinatorial exemplars is one of the driving forces that makes computational ADME profiling so attractive.

F. Poor ADME properties – business aspects

There is a business aspect here for a company selling combinatorial libraries if the combinatorial libraries contain compounds experimentally verified to have poor ADME properties. What do you do with the poor ADME compounds? The market for screening compounds is split. Some customers are interested in using a small molecule as a tool or probe for a biological pathway – the chemical biology use. Other customers are interested in true drug discovery where good ADME is an absolute requirement. So, not all customers want all compounds to have good ADME properties, especially if there is a price break on the poorer ADME compounds. There is still a sizeable market (perhaps 30% but dwindling) for compounds perceived to be possibly active in an *in vitro* assay irrespective of ADME properties. The argument in chemical biology is that one cannot proceed anywhere without that first "lead." The absolute requirement in chemical biology is selectivity. So, one should try to get the "chemical biology lead" first, and worry about ADME later, or never if the work never proceeds to true drug discovery.

G. If library production is rate determining

If library production is rate determining then the effectiveness of experimental ADME assays becomes much simpler. Operationally, it is fairly straightforward to de-prioritize a library that profiles poorly in ADME assays if there are more protocols than can be translated into production. One simply executes the best protocols in an ADME sense and drops from production those that profile the worst. There is a clear message here if one really cares about ADME properties. The manning and effort should be greater at the protocol validation stage as opposed to the library production stage. This ADME-derived message is contrary to that which one receives at the vendor trade shows. The vendor trade shows tend to emphasize the production side. This is hardly surprising if there is much more money to be made in selling production hardware and technology as opposed to selling tools for protocol design and validation.

H. Relative importance of ADME assays

I believe that ADME assays are not all equal in terms of contributing to drug quality. ADME design is most important for those properties under the poorest chemistry control. Good chemistry control equates with chemists' ability to control SAR. Good control of SAR means that a chemist can make a small molecular change with a resultant large change in the measured property. Poor control of SAR means a loss of relationship between molecular structure and the measured property. I believe it is inadvisable to totally filter out compounds with poor properties if they can easily be fixed by chemistry. The goal in drug research is to discover inherently active compounds with appropriate ADME properties. There is little value in being so restrictive that few inherently active compounds will be discovered.

III. CHEMISTRY CONTROL OF INTESTINAL PERMEABILITY

Chemistry control of intestinal permeability via passive *trans*-cellular processes is poor. The good news is that, in general, poor permeability is not a problem in combinatorial libraries. One really has to go out of one's way to introduce enough polar functionality in a combinatorial compound to

make a compound impermeable via passive *trans*-cellular processes (the most common absorptive pathway). The bad news is that if a compound has really poor intestinal permeability there is virtually nothing to fix the problem in terms of pharmaceutical formulation technology. There is recent progress examining the effect of excipients in enhancing permeability as a function of compound charge state. The general pattern is that excipients enhance the permeability of weak acids but retard the permeability of weak bases. The acceptable excipient permeability fixes all have a common feature that the excipient does not significantly damage the gastrointestinal barrier. To this authors knowledge there are no case studies of success in improving permeability by a barrier breaking approach. Chemistry control of intestinal permeability is poor because, except for a few very specific exceptions, chemistry SAR is blunt.

A. Improving permeability

The best general guide for improving permeability is to reduce the polar surface area (PSA) or to reduce the sum of the hydrogen-bond accepting and donating moieties in the compound. Some authors recommend a PSA cutoff of about 120 Å. In my experience a compound with PSA of less than about 140–150 Å should have reasonable intestinal permeability unless something else is wrong. The "something else is wrong" includes excessive basicity (above pK_a 11.5) or excessive acidity (below pK_a 3) or log D at pH 6.5 (or 7) below about 0.0. However, excessive acidity may be compatible with acceptable permeability if the anion exhibits extensive charge delocalization, for example, a vinylogous enolic acid-like system. In cases like this, acceptable permeability can occur via a charge delocalized ion pair. The blunt SAR feature comes from the phenomenon that permeability improves only gradually as a physicochemical property is moved in a desired direction. One does not see a sudden improvement in permeability from a properly positioned methyl group as one might see with respect to *in vitro* SAR. Biological phenomena might be responsible for apparent poor permeability. For example, the compound is extensively metabolized by cytochrome P450 3A4 in the gut intestinal wall or the compound is actively effluxed by P-glycoprotein in the gut wall. Unfortunately, both these biological systems tend to show broad substrate specificity so the chemistry SAR can be blunt.

B. Hydrogen bonding and permeability

Permeability is intensely affected by the presence of an intramolecular hydrogen bond. In this area tight SAR can occur. Any structural change that causes an internal hydrogen bond to form (or that prevents one from forming) can have a dramatic effect on permeability. The effect of a single intramolecular hydrogen bond on increasing permeability is great and can easily be a factor of 10 or more. This type of structural effect is currently difficult to computationally predict with success so far reported in the prediction of permeabilities of cyclic peptides. Using bond order and structure one can predict the possibility that an intramolecular hydrogen bond is possible. However, to my knowledge, no existing commercially available, reasonably fast batch mode program can calculate whether an intramolecular hydrogen bond is likely to form in reality, that is, the hydrogen bond formation is energetically likely.

C. Intramolecular hydrogen bonds

Experimental assays can correctly predict the permeability enhancing effect of an intramolecular hydrogen bond. These assays can be cell-based as in Caco-2 cell culture assays or MDCK cell assays; or they can be nonbiological as in a traditional log P or log D, or a parallel artificial membrane permeation assay (PAMPA) assay. PAMPA assays have grown in popularity with a considerable body of literature attesting to their utility. Experimental assays work because they are sensitive to the actual hydrogen-bonding properties of a compound. There is considerable variation in the extent to which biological and nonbiological assays have been adopted in the pharmaceutical industry as a predictor for intestinal permeability. In my opinion, this may partly reflect differences in chemistry across organizations. An organization with many conformationally flexible compounds bearing hydrogen-bond donor and acceptor groups might be particularly likely to use a nonbiological type of experimental assay because of the need to identify the more permeable compounds due to intramolecular hydrogen bonding. Conversely, an organization with many heterocyclic compounds having limited possibilities for intramolecular hydrogen bonding might not see the need for this type of assay.

D. Permeability testing

There is no uniformity as to how permeability testing is carried out in the pharmaceutical industry. This suggests that there may not be a great deal of difference in the effectiveness of the various experimental and computational permeability prediction methods. The people factor can easily be as (or even more) important than purely technical factors. The goal in a permeability assay or calculation is to influence chemistry behavior, that is to direct chemistry synthesis toward more permeable compounds. Thus the best assay or calculation may be the one that chemists (for whatever reason) believe and act on. This means that issues such as capacity, ease of use, ease of interpretation and internal credibility can be the deciding factor for effectiveness. As previously discussed, PSA calculations can be used as a permeability filter. PSA calculations are not very accurate if: the compound is highly conformationally flexible; the compound is multiply charged and if intramolecular H-bond

possibilities exist. For these types of compounds an experimental permeability assay may be the only option.

IV. CHEMISTRY CONTROL OF AQUEOUS SOLUBILITY

Chemistry control of aqueous solubility is poor. The good news is that if a compound has poor aqueous solubility, methods do exist to fix the problem in terms of pharmaceutical formulation technology. However, these are always expensive in time and manning, and depending on the degree of the solubility problem, may have limited or no precedent in terms of existing approved products. So far, the preferred solution to poor solubility is to fix it in chemistry. Formulation fixes are a last resort. The bad news is that, in general, poor aqueous solubility is by far the most common ADME problem in combinatorial libraries. It takes no effort at all to introduce poor aqueous solubility into a combinatorial compound library. The combination of high molecular weight and high lipophilicity outside the "Ro5" limits is an almost certain guarantee of poor aqueous solubility. Based on our experimental screening, lipophilicity above the Ro5 log *P* limit of 5 by itself carries with it a 75% chance of poor aqueous solubility.

A. The definition of poor solubility

The definition of "poor solubility" has reduced to a smaller number in the combichem/HTS era. A classic pharmaceutical science textbook might have defined poor solubility as anything below a solubility of 1 g/mL at pH 6.5 (or pH 7). Currently, most drug researchers would be very excited by a solubility as high as 1 g/mL. In general, with average permeability and a projected clinical potency of 1 mg/kg, a drug needs a minimum aqueous solubility of 50–100 mg/mL to avoid the use of nonstandard solubility fixing formulation technology. We find the guidelines published by Pfizer's Curatolo (see Ref. [8]) on maximum absorbable dose to be an excellent guide for the combination of permeability, solubility and potency required in an orally active drug.

Figure 22.2 is a bar graph illustrating the minimum acceptable solubility as a function of compounds projected clinical potency and permeability in medicinal chemistry. The middle set of bars show that a compound has to have a minimum thermodynamic solubility of 52 mg/mL when the permeability is average (avg K_a) and the projected clinical potency is 1 mg/kg.

B. Aqueous solubility and blunt SAR

Chemistry control of aqueous solubility is poor because, except for a few very specific exceptions, chemistry SAR is blunt. In this respect, control of solubility like that of permeability is poor. Solubility due to excessive lipophilicity

FIGURE 22.2 Minimum acceptable drug solubility as a function of projected clinical potency (0.1–10 mg/kg) and intestinal permeability (low–avg–high K_a).

improves only gradually as the lipophilicity is moved in the desired downward direction. Trying to decrease lipophilicity by incorporating polar functionality may or may not work. The potential solubility improvement attendant on introducing polar hydrophilic functionality can easily be more than counterbalanced by a decrease in solubility due to increased intermolecular crystal packing forces arising from the new polar functionality. In our experimental solubility screening, about 60% of poor solubility is unrelated to excessive lipophilicity. This 60% of poor solubility arises from high crystal packing forces due to intramolecular interactions in the solid crystal state that make disruption of the crystal energetically costly. Thus the blunt SAR feature in solubility comes from the phenomenon that solubility improves only gradually (or not at all) as a physicochemical property (lipophilicity) is moved in a desired downward direction.

C. Changing the pK_a

Changing the pK_a of an acidic or basic group in a molecule so that more of the compound exists in the ionized form at physiological pH lowers log *D* (at about pH 7) and, in general, should improve aqueous solubility. The improvement in solubility is limited, however, if the solubility of the neutral form of the compound (the inherent solubility) is very low. The situation is worsened if the starting pK_a is far from 7. We find this to be a particular problem with weak bases. Weakly basic pyridines, quinolines, quinazolines and thiazoles seem to be frequent members of combinatorial libraries. Understanding the ionization behavior of drugs and how this property relates to oral absorption is extremely complex and likely beyond the capability (and interest) of many medicinal chemists. The reader is referred to an excellent recent review in this complex area.[8]

The extent of poor aqueous solubility may be experimentally underestimated in a combinatorial library. No

combinatorial library is purified by traditional crystallization. The vast majority of compounds purified by an automated process will be isolated in amorphous form. Compounds in an amorphous solid form exist in a much higher energy state than a true crystalline solid, and aqueous solubilities of amorphous solids are always higher than those of crystalline solids.[9] This phenomenon may only be recognized if there is a high degree of interest in a combinatorial compound. The combinatorial compound is scaled up and purified by crystallization. The newly crystallized compound can easily be an order of magnitude more insoluble, and hence more poorly absorbed than the original sample.

D. Improving aqueous solubility

Aqueous solubility can be improved by medicinal chemistry despite the blunt SAR feature and the pharmaceutical industries record has been quite successful in this respect. However, to improve solubility requires commitment to a combination of computational and experimental interventions and a real effort on the part of chemists to incorporate solubility information into synthesis design. The importance of rapid experimental feedback is particularly important given the current imprecision in computationally predicting poor solubility arising from crystal packing interactions. It is critical not to miss a serendipitous improvement in solubility attendant on a molecular change. Owing to the blunt SAR feature, the easiest way to improve solubility with respect to library design is to try to design the best solubility profile right at the start. Sonication/acoustic technology, both contact and noncontact, is a new development that experimentally improves solubility.[10] The methods are currently under active exploration and data is limited as to the duration of the increased solubilities. Likely sonication introduces enough energy to a precipitated crystalline material that the equivalent of the solubility of amorphous material can be attained. Equipment capable of running in HTS mode is now available from several vendors.

V. IN VITRO POTENCY AND CHEMISTRY CONTROL

In vitro potency has always been under excellent chemistry control; the hallmark of good control being tight chemistry SAR. With respect to combinatorial chemistry several exceptions should be noted. Compounds are often encountered as leads in HTS screening that can be characterized as "phony HTS leads." These types of compounds should be avoided at all cost as templates in combinatorial chemistry. A common attribute of these leads is that the chemistry SAR is flat and fuzzy if they are subjected to lead optimization. Large chemistry changes can be made with only very small changes in activity. Often these types of problems can be avoided by similarity searches on the initial apparent lead. A loss or gain of activity related to a small structural change over the initial lead is a good sign. A change in activity of a factor of 10 between two compounds differing by only a single methyl group is a classic example of good SAR. A flat SAR among analogs is not a good sign. Often these "leads" are not very active, perhaps in the low micromolar range. Sometimes the exact same compound appears active in multiple HTS screens. Alternatively, as we have observed, members of a well-defined structural series appear as apparent HTS actives but are not necessarily the exact same compound across different HTS screens. Some phony HTS "leads" are removable by compound quality filters and there are recent publications on structural features of these worthless frequent hitters. A sub-class of false positives in screens occurs because the compounds form colloidal aggregates in the 50–1,000 nm size range. The biological activity occurs due to the aggregate and not the discrete compound. This is why the medicinal chemistry optimization fails. Aggregate formation tends to occur more frequently at higher concentrations. Recent publications suggest the incidence could be as high as 15–20% of compounds at concentrations of 30 μM or higher. Very steep dose response curves are a tip off to a potential aggregate problem. Experienced medicinal chemistry "intuition" works quite well in avoiding "phony HTS leads" but it would clearly be advantageous to use computational filters for this problem. Software, even if it worked no better than chemists' intuition, would be advantageous from the people viewpoint. Biologists generally do not understand or appreciate chemists' exquisitely tuned sense of what constitutes a "good" chemical structure. Thus many hard feelings and miscommunication between chemists and biologists could be avoided by a computation that merely mimics chemists' structural intuition.

A. Lead complexity

The other major limitation on good *in vitro* potency control relates to the complexity of the apparent lead. Some apparent leads simply lack a critical structural feature and cannot be easily optimized by traditional medicinal chemistry means. These are most likely to be detected among very weak actives in traditional HTS screens and form the majority of actives in techniques such as the fragment screening method of SAR by nuclear magnetic resonance (NMR). The increasing probability of a missing critical piece as MWT decreases probably sets the lower size range for lead discovery libraries using traditional HTS. In fragment screening by NMR or X-ray the medicinal chemist receives extra target chemistry structural binding information beyond simply ligand affinity. This extra information allows the optimization starting from low MWT starting points, something that would not have been predicted a decade ago. In fragment screening, the affinity of the initially weak starting binders can increase surprisingly rapidly on structural changes. The theory behind this surprisingly steep SAR is now known.[11] It turns out that the translational and rotational entropic penalty

for binding is as a first approximation similar for a small and large ligand. This appears to be the basis for the success stories of converting an initial millimolar binder to a nanomolar affinity range biber with an SAR on about a 100 analogs.

Several publications provide an excellent perspective on lead generation libraries.[12,13] The general theme is that the properties of a lead must allow for the almost inevitable increase in molecular weight and lipophilicity that accompany *in vitro* activity optimization. Andrews' binding energy can be used as a rough indicator of functional group complexity, that is, whether the apparent lead has enough "stuff" on it to interact with a receptor target. We find no difference in the overall sum of Andrews' binding energy between phase II drugs and combinatorial libraries. However, we do find large differences in the density of functionalization. Phase II drugs are much more compact and densely functionalized than combinatorial compounds. This is easily seen by simply plotting the ratio of Andrews' binding energy to molecular weight for the members of the two types of libraries. The ratio is much larger for the phase II compounds. The same functionality is placed on a more compact smaller structure; hence the ratio is larger. Many drug discovery organizations are now using ligand efficiency (LE) rather than IC_{50} to describe compound *in vitro* activity. The general idea is that for comparable IC_{50} the smaller compound is the better. In one implementation of LE the IC_{50} is converted to a receptor molar affinity and then to binding affinity in kilocalorie per mole which is then divided by the number of heavy atoms in the molecule.[14] An LE of 0.3 kcal per heavy atom or better is frequently used as the cutoff for an efficient ligand. Operationally, this means that if LE is maintained during optimization the ligand affinity will reach 10 nM before the MWT exceeds 500 and oral absorption becomes problematic.

VI. METABOLIC STABILITY

Metabolic stability is generally on all lists of ADME filters. The chemistry control is highly situational in the sense that in some chemical series control is excellent, that is the chemistry SAR is very tight. A specific example might be the blocking of hydroxylation by a fluorine atom. In other cases the control is poor, for example, the biology target SAR dictates the presence of a metabolically unstable moiety. A specific example is the hydroxamic acid moiety, found in many early lipoxygenase inhibitors, which was readily metabolically converted to the parent carboxylic acid. Where the chemistry control is highly dependent on the chemistry context, I believe it is dangerous to implement exclusionary filters for highly probable metabolic events. For example, perhaps as many as 40–50% of compounds might be substrates to some extent for cytochrome P450 3A4-mediated oxidation. As this event is very probable, I think it is unwise to implement a blanket combinatorial exclusionary filter for cytochrome P450 3A4 substrates. What might make more sense would be to factor in the probability that the problem could be fixed with chemistry. Thus in a combinatorial library one might want to allow compounds with a single cytochrome P450 3A4 potentially metabolically unstable site but filter out those compounds with two or more sites of metabolic instability. One could envision a quite different use of the same information depending on the research stage. In lead optimization, one would want to use all available metabolic stability information because the immediate goal is to improve the compound's drug-like properties. However, in a biology lead-seeking step such as the preparation of a combinatorial library for an HTS screen one is much more interested in the lead generation process. Here, one would want to apply metabolic stability criteria in a looser sense taking into account the probability of a subsequent medicinal chemistry fix. Current computational predictions of cytochrome oxidation predict the position of metabolism.[15] This is very useful in chemically blocking the metabolism. However, current computational predictions do not predict rate of metabolism. Having a balanced metabolism across multiple CYP subtypes is a very desirable feature. For example, some drug discovery organizations will discontinue work on a chemical series if the metabolism is entirely due to a polymorphically expressed cytochrome like CYP2D6 because of the problem of high clinical exposure in patients deficient in CYP2D6.

A. ADME computational models

Computational models for ADME properties work best when the models are based on single mechanism experimental assays. Scientists approaching an ADME computational model are often influenced by their familiarity with therapeutic target computational models. A typical biological HTS screen consists of a single mechanism assay. For example, a compound is screened to determine whether it is an agonist or antagonist for a single receptor subtype. For the single mechanism screen more experimental data usually mean a better computational model, so with this history it is very easy to fall into the trap of believing that more data in any ADME screen will result in a better computational model. In a therapeutic target assay one does not deliberately mix targets. So, there is no history of what to expect if the experimental endpoint is due to multiple mechanisms. Suppose one were to deliberately mix half a dozen biological targets in a single HTS screen such that a hit on any of the targets gave a common analytical endpoint. Could one develop a computational model for the endpoint if the experimental response was based on half a dozen unrelated structure activity patterns? I think this would be very difficult, especially as the number of experimental data points were increased. However, this is exactly what occurs if one tries to build a computational model based on experimental data in a multi mechanism ADME assay. Even the experimental SAR optimization against multiple mechanisms in a non-ADME screening

sense can induce medicinal chemistry skepticism. This occurs with phenotypic screens where chemists are reluctant to believe that three orders of magnitude activity increase can be achieved against a mechanistically imprecise end point.

B. Limitations of Caco-2 cell culture

Our experience with trying to build computational models based on experimental permeability screening in Caco-2 cell culture illustrates the problem introduced by multiple mechanisms. We found that deviation from a single mechanism could arise either in the assay *per se* or could arise from the compounds that were screened in the assay. One aspect of the multiple mechanism problem is the presence of active multiple biological transport mechanisms for both enhancing and reducing absorption in cell culture assays. This issue is well documented and is outside the scope of this chapter.

C. Poor aqueous solubility and permeability assay noise

Poor aqueous solubility, a compound-related factor rather than an assay-related factor, has a major effect by introducing "noise" into absorption screening and thus has an effect on making computational model building very difficult. It must be stressed that the compound solubility factor virtually never appears as an explicit consideration in the published permeability literature. Compound sets are published that are used to validate *in vitro* cell-based absorption assays. Validation usually means obtaining an acceptable correlation between human fraction absorbed data and *in vitro* permeability data. The absorption data always include the experimentally well-controlled but compound number – limited human fraction – absorbed data that are used to define absorption ranges in the FDA bioavailability waiver guidelines.[16] This limited compound set is then supplemented with additional compounds chosen from published human absorption literature. In our own work we have been able to accumulate literature human fraction-absorbed data on a total of about 330 compounds. Larger datasets of up to about 1,000 compounds exist, which are based on published reference texts or intensive literature searches supplemented by detective work to differentiate the absorption and metabolism components in oral bioavailability. The hallmark of compounds with human absorption data is that they are well-behaved compounds from a "drug-like" viewpoint. The fraction absorbed is heavily biased to the high percentage absorbed range and the compounds are almost universally soluble in aqueous media. This simply reflects the compound quality filtering process that must be passed for a compound to enter the types of studies likely to generate human fraction-absorbed data. In short, literature compound permeability validation sets are completely appropriate and say a good deal about assay issues in a permeability screen, but they have almost no relevance to assay reproducibility issues related to poor compound solubility.

D. Physiologically-relevant screening concentration

Table 22.1 sets the stage for the types of solubility among currently synthesized compounds that are likely to be

TABLE 22.1 Solubility Ranges Among Currently Synthesized Compounds

μg mL^{-1}	μM (MWT 300)	μM (MWT 400)	μM (MWT 500)	μM (MWT 600)	Compounds
1	3.33	2.50	2.00	1.67	30% of Groton compounds are in this solubility range
3	10.00	7.50	6.00	5.00	
5	**16.67**	**12.50**	**10.00**	**8.33**	
10	33.33	25.00	20.00	16.67	10% of Groton compounds are in this solubility range
20	66.67	50.00	40.00	33.33	
30	100.00	75.00	60.00	50.00	
40	133.33	100.00	80.00	66.67	
50	**166.67**	**125.00**	**100.00**	**83.33**	**Solubility acceptable for 1 mg/kg^{-1} potency**
60	200.00	150.00	120.00	100.00	
70	233.33	175.00	140.00	116.67	
80	266.67	200.00	160.00	133.33	60% of Groton compounds are in this solubility range
90	300.00	225.00	180.00	150.00	
100	333.33	250.00	200.00	166.67	
200	666.67	500.00	400.00	333.33	
300	1000.00	750.00	600.00	500.00	FDA 1 mg/kg^{-1} solubility in 250 mL water
500	1666.67	1250.00	1000.00	833.33	
1000	3333.33	2500.00	2000.00	1666.67	

submitted to a permeability screen such as a Caco-2 assay. In this type of assay a variety of biological transporters are present that mediate both absorption and efflux. The movement of a compound through the Caco-2-polarized cell layer through the action of these transporters can be saturated if the drug concentration is high enough. Thus it is important to screen at a physiologically relevant concentration. If the dose is too low the permeability estimate will be too low because the importance of efflux transporters will be overestimated.

VII. ACCEPTABLE SOLUBILITY GUIDELINES FOR PERMEABILITY SCREENS

Table 22.1 maps the acceptable solubility ranges as defined by pharmaceutical science to the molar concentration range of biological screening. For an average potency compound of about 1 mg/kg, the screening dose in a Caco-2 screen should be somewhere in the range of 100 mM. This concentration is the minimum required for adequate absorption. However, pharmaceutical industry Caco-2 screening doses are typically 10–25 mM. This dose range is chosen for practical reasons. If the assays were run at 100 mM a high incidence of insoluble or erratically soluble compounds would be encountered. In Caco-2 screening in our Groton, USA laboratories, one-third of compounds screened at 10 mM are insoluble in an aqueous medium. When one-third of compounds screened in an assay are insoluble in aqueous media, assay reproducibility becomes a major issue. I think it is entirely reasonable to question the value of permeability screening of combinatorial libraries given their general tendency toward poor solubility.

A. Batch-mode solubility prediction

The reader can quickly realize whether poor solubility might be a confounding factor for permeability screening of a combinatorial library from a solubility calculation. In my experience, the existing batch-mode solubility calculation programs generate similar and quite reasonable solubility histogram profiles when run on thousands of compounds (although I would not trust numerical prediction results for small numbers of compounds). Experimental permeability screening (especially if it is manning intensive) might not be worthwhile because of the solubility noise factor if a significant fraction of the library is predicted to be insoluble at 10 mM (the low end of the typical screening concentration range).

In summary, poor aqueous solubility is the single physicochemical property that is most likely to be problematical in a combinatorial library. It can be avoided in part by intelligent use of batch-mode solubility calculations. The solubility problem is not simply a technical issue in library design. It is exacerbated by chemistry synthesis considerations and by the timing of the availability of combinatorial exemplars. Formulation fixes are often available for the greaseball compounds unless the solubility is extremely poor, but these should be avoided as much as possible. Poor permeability is seldom a problem in combinatorial libraries, but is disastrous if present since other than modest excipient effects that mimic a food effect effective permeability formulation fixes do not currently exist.

REFERENCES

1. Leeson, P. D., Davis, A. M. Time-related differences in the physical property profiles of oral drugs. *J. Med. Chem.* 2004, *47*, 6338–6448.
2. Blake, J. F. Identification and evaluation of molecular properties related to preclinical optimization and clinical fate. *Med. Chem.* 2005, *1*, 649–655.
3. Morphy, R. The influence of target family and functional activity on the physicochemical properties of pre-clinical compounds. *J. Med. Chem.* 2006, *49*, 2969–2978.
4. Lipinski, C. A., Lombardo, F., Dominy, B. W., Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 2001, *46*, 3–26.
5. Leach, A. R., Hann, M. M., Burrows, J. N., Griffen, E. J. Fragment screening: an introduction. In *Structure-Based Drug Discovery*, 2006, pp. 142–172.
6. Hajduk, P. J., Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat. Rev. Drug Discov.* 2007, *6*(3), 211–219.
7. Bergström, C. A. S., Wassvik, C. M., Johansson, K., Hubatsch, I. Poorly soluble marketed drugs display solvation limited solubility. *J. Med. Chem.* 2007. ACS ASAP.
8. Avdeef, A. Solubility of sparingly-soluble ionizable drugs. *Adv. Drug Deliv. Rev.* 2007, *59*(7), 568–590.
9. Singhal, D., Curatolo, W. Drug polymorphism and dosage form design: a practical perspective. *Adv. Drug Deliv. Rev.* 2004, *56*(3), 335–347.
10. Oldenburg, K., Pooler, D., Scudder, K., Lipinski, C., Kelly, M. High throughput sonication: evaluation for compound solubilization. *Combin. Chem. High Throughput Screening* 2005, *8*, 499–512.
11. Murray, C. W., Verdonk, M. L. The consequences of translational and rotational entropy lost by small molecules on binding to proteins. *J. Comput.-Aided Mol. Des.* 2002, *16*(10), 741–753.
12. Leeson, P. D., Davis, A. M., Steele, J. Drug-like properties: guiding principles for design – or chemical prejudice?. *Drug Discov. Today Tech.* 2004, *1*(3), 189–195.
13. Oprea, T. I., Allu, T. K., Fara, D. C., Rad, R. F., Ostopovici, L., Bologa, C. G. Lead-like, drug-like or "pub-like": How different are they? *J. Comput.-Aided Mol. Des.* 2007, *21*(1–3), 113–119.
14. Hopkins, A. L., Groom, C. R., Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discov. Today* 2004, *9*(10), 430–431.
15. Cruciani, G., Carosati, E., De Boeck, B., Ethirajulu, K., Mackie, C., Howe, T., Vianello, R. MetaSite: understanding metabolism in human cytochromes from the perspective of the chemist. *J. Med. Chem.* 2005, *48*(22), 6970–6979.
16. Lennernäs, H., Abrahamsson, B. The biopharmaceutics classification system. In *Comprehensive Medicinal Chemistry II*, Vol. 5, 2007, pp. 971–988.

Chapter 23

Quantitative Approaches to Structure–Activity Relationships

Han van de Waterbeemd and Sally Rose

I. INTRODUCTION TO QSAR
II. BRIEF HISTORY AND OUTLOOK
III. QSAR METHODOLOGY
A. Descriptors
B. Methods for building predictive models
C. Global and local models, and consensus modeling
D. Time-series behavior and autoQSAR
E. Experimental design
F. Inverse QSAR and multi-objective optimization
IV. PRACTICAL APPLICATIONS
A. Limitations and appropriate use
B. Examples
C. Library design, compound acquisition and profiling
D. HTS analysis
E. Software
REFERENCES

Never trust anything that can think for itself if you can't see where it keeps its brain.
J. K. Rowling, Harry Potter and the Chamber of Secrets, 1999.

I. INTRODUCTION TO QSAR

Medicinal chemists typically attempt to develop *qualitative* structure–activity relationships (SAR) for a synthesized series of compounds in a drug discovery project which elucidate target binding affinity and ADMET (absorption, distribution, metabolism, excretion and toxicity) properties in their quest for an optimal compound meeting clinical candidate criteria. *Quantitative* structure–activity relationships (QSAR) and quantitative structure–property relationships (QSPR), using various statistical and mathematical tools, are even more powerful tools to guide the chemist in the design process. Thus, QSAR is used to select better compounds to synthesize or buy that will increase our understanding of SAR and increase the likelihood of finding active compounds by doing the least possible *in vitro* or *in vivo* experiments. Some typical applications of QSAR include predicting whether a proposed new compound:

- will be active or inactive
- will have an activity of *x* nM
- will be selective
- contributes to new information/knowledge
- has appropriate ADMET properties
- is important to better understand SAR and helps to create a better predictive model.

The interface between the computational disciplines of molecular modeling, chemoinformatics and QSAR is rather blurred and often overlapping. Molecular modeling is used to model compounds in 3D, calculate their properties (often for use in QSAR) and understand their interaction with protein targets. Chemoinformatics[1,2] is about encoding chemical compounds, storing and searching databases, and analysis of such data. QSAR refers more to building and using predictive models to predict various drug-related properties.

In this chapter, only QSAR methods which use physicochemical or structural properties of molecules will be discussed, while in Chapter 29 so-called 3D-QSAR approaches will be presented. 3D-QSAR techniques, for example, comparative molecular field analysis (CoMFA), commonly

Copyright © 2008, Elsevier Ltd
All rights reserved.

define a grid around a molecule and use the value of a molecule's electronic, steric and lipophilic field at specific grid points as descriptors and relate activity to these descriptors using multivariate QSAR statistical techniques, commonly partial least squares (PLS). In these methods, accurate molecular superposition and choice of grid point spacing is crucial.

QSAR methods can be classified in several ways. One approach is to look at the nature of the method, supervised versus unsupervised, where supervised methods use the activity values to create a predictive model from the descriptors and unsupervised methods model molecular similarity from descriptors, but do not use the activity values in the derivation of the model. Another way is to look at the nature of the relationship between activity and descriptors; categorical versus continuous, or linear versus non-linear (Figure 23.1)

This chapter considers predictive models in general as well as what could be considered as truly "quantitative" models. Predictive models can be used in many stages of the discovery process. This ranges from modeling biological activity, analysis of the corporate compound collection, selection of diverse "drug-like" compounds for purchase and screening, selection of subsets to screen against a particular target, analyses of High-throughput screening (HTS) data, design of small synthesis libraries in hit-to-lead (HTL), design of focused lead optimization (LO) libraries, early prediction of ADME and toxicity/safety, to optimization of absorption, distribution, metabolism and excretion (ADME) properties.

The use of various methods and required prediction precision depends on the stage in the discovery process. In the early stages of drug discovery, unsupervised molecular similarity methods are useful for compound selection and simple, rapid filters are useful for library design and lead profiling. Supervised QSAR methods are widely applied during LO to assist the medicinal chemist in optimizing potency, ADME and selectivity and for generating new ideas. Commonly, this is applied in parallel with mechanistic (physiologically-based pharmacokinetics, PBPK) models, needed for pharmacokinetics (PK) optimization. These latter type of quantitative models typically require *in vitro* data input.

FIGURE 23.1 Structure–activity and structure–property relationships using data modeling techniques may provide the basis for understanding and prediction of biological activity and physicochemical features.

The success of predictive models depend on several factors including the user expectation.[3] Distinction also needs to made between a useful model (often the reality)[4] and a statistically perfect model (these rarely exist!).

In this chapter, we briefly review the history of QSAR, the types of descriptors, the key methodologies and various aspects of model building and prediction in QSAR. A few examples are given to exemplify appropriate use of the available tools.

II. BRIEF HISTORY AND OUTLOOK

First attempts to express quantitatively relationships between chemical structure and bioactivity go back to the beginning of the previous century.[5] However, these approaches were only widely applied when computers and relevant mathematical methods became available. The credit goes to Corwin Hansch and Toshio Fujita for introducing these quantitative methods to medicinal chemistry in the 1960s.[6–11] Initially, in their pioneering work, Hansch *et al.* focused their attention on the role of octanol/water partition coefficients ($\log P$) in drug transport processes that were thought to contribute crucially to the measured activities. As we know now, $\log P$ is the most predominant descriptor in many structure–activity correlation studies.

Early QSAR studies were mainly focused on analyzing the effect of aromatic substituents on activity in congeneric compound series using substituent constants to describe the steric, electronic and lipophilic characteristics of the substituents. Hansch and Fujita proposed a method to describe quantitatively relationships between biological activity and chemical descriptors.[6,9] This can be expressed as follows:

$$\text{Biological activity} = \text{function (molecular or fragmental descriptors)} \tag{23.1}$$

The Hansch–Fujita approach is also called the linear free-energy relationship (LFER) or extrathermodynamic approach, since most of the descriptors are derived from rate or equilibrium constants.

There are numerous examples of traditional Hansch QSAR studies in the literature.[10,11] Some include large sets of descriptors, while others explore just a few. If physicochemical descriptor values are not readily available, indicator variables, denoting presence or absence of a certain structural feature, may be of help. This is the basis of the Free-Wilson (FW) method.

The FW model was proposed in 1964 at the same time as the Hansch model, but is far less widely used.[9,12] It uses

indicator values, having a value of unity for the presence of a substructural feature, for example, a *para*-chloro substituent and zero for its absence, as sole parameters in a Hansch model-like regression equation. FW has also been named the additivity model. The combination of physicochemical properties with substructure indicators is often the best way to proceed.

Traditional Hansch analysis using multiple linear regression (MLR) suffers from several shortcomings. One of the problems is that one often has more variables than compounds and Hansch analysis ideally requires at least five times as many compounds as descriptors. Hansch analysis becomes unstable when correlated descriptors are used. Furthermore, there is often a need to consider correlations between chemical descriptors and several biological tests simultaneously.

PLS (partial least squares projections to latent structures) is a generalization of regression which is appropriate to treat these problems.[13,14] PLS can handle numerous and even collinear (inter-correlated) variables, and allows for a certain amount of missing data. An important part of a PLS data modeling study is the cross-validation (CV) of the results to determine the robustness of the model (see Section III.B.4.). PLS considers all independent descriptors together and calculates their modeling power, that is, their contribution to the regression. PLS is particularly useful when many descriptors are taken into consideration. Experience has shown that PLS often gives the most relevant and statistically significant results and should be the preferred default technique in QSAR correlation studies.

A good example of a comparative study using Hansch and FW, both using MLR, with a PLS analysis on the same data, is given in.[15]

Progress in chemometrics has made a number of new statistical techniques available, which are increasingly being used. This concerns both new supervised and unsupervised (or "pattern recognition") techniques. Chemometrics was defined about 25 years ago as the chemical discipline which uses mathematical, statistical and related techniques to design optimal measurement procedures and experiments, and to extract maximum relevant information from chemical data. The science of chemometrics has been developed to promote applications of statistics in analytical, organic and medicinal chemistry.

The current focus on early prediction of ADMET properties, as well as the analysis of HTS data, has led to a revival and extension of the use of QSAR technology in the pharmaceutical industry. Other applications can be found in metabonomics, the application of chemometrics to analytical spectral data to predict disease or the effect of compounds on metabolism. Many methods originally known in economics and artificial intelligence research are now also being used in QSAR/QSPR.

We see currently a change from academic QSAR, with models using 10–100 s of compounds, to industrial QSAR, with models built using 1,000–10,000 s of compounds. Comparative analyses have been done on large sets of QSARs to understand the role of the size of the data set and the descriptors used in the modeling.[16] Another example on the industrial scale are the studies of the effect of time-series behavior on QSAR model performance (see Section III.D.).[17]

An increased interest in molecular similarity (or dissimilarity) has had a beneficial impact on the design of combinatorial chemistry libraries and because of this the diversity of compounds in a library has generally increased. More recently attention has also focused on "drug-likeness" of libraries (see Section IV.C.). This has brought about a change in the types and increase in the number of descriptors which are used in QSAR studies and this is discussed below.

The recent development of pipelining tools such as Pipeline Pilot, Orange and Knime (see Box 23.3) has allowed QSAR developers to automate many predictive procedures. These tools give the user facile access to databases, structures and data analysis methods. They allow complex QSAR procedures to be predefined and subsequently executed by non-experts; so bringing predictive QSAR models directly to the medicinal chemist's work bench.

III. QSAR METHODOLOGY

A. Descriptors

1. *Biological endpoints*

When a compound interacts with a chemical or biological environment, we may define physicochemical properties, such as lipophilicity or ionization constants, biochemical properties, such as binding constants, and biological properties, such as activity or toxicity.

The targets of drug action are diverse and include, for example, membrane-bound receptors, ion channels, enzymes and DNA. Accordingly, biologists have developed a wide variety of biological and pharmacological test systems producing different kinds of data. Some are quite simple and accurate, for example, IC_{50} values as a measure of ligand activity or K_i values as a measure of ligand binding affinity, while others are more complex with large errors, for example, *in vivo* data. Activity may be expressed as a continuous measure, for example, IC_{50} or % inhibition at a specified concentration, or as categorical data such as active versus inactive, agonist versus antagonist, or as strong, medium and weak. Both continuous and categorical data can be used in QSAR studies. The proper choice of mathematical model to relate biological to physicochemical descriptor data depends on the quality and kind of data to be analyzed. Therefore the classical Hansch approach using MLR (see section III. B.2.a) is not suited for all purposes.

Important off-target biological endpoints include interactions with the hERG channel and metabolizing enzymes such as cytochrome P450 s (CYPs), and interactions with

transporters such as P-glycoprotein (P-gp). Metabolic stability measured *in vitro* using microsomes and hepatocytes is suitable for QSAR modeling. There is also considerable interest in modeling human PK data such as bioavailability, oral absorption, volume of distribution, clearance and brain uptake, among others.

2. Key physicochemical properties

The *lipophilicity* of a compound is often considered as an important design factor since it is related to processes such as absorption, brain uptake, volume of distribution and protein binding. This property is often expressed as the partition coefficient of the neutral species ($\log P$) or distribution coefficient ($\log D$, typically measured at pH 7.4). The 1-octanol/water system is often taken as the reference or standard for partition coefficients. However, other partitioning systems may give useful information too. It has also been found that the differences between $\log P$ values measured in two different solvent systems ($\Delta \log P$) may contain relevant information related to the H-bonding capacity of a compound.[18] However, tedious measurements of multi-solvent partitioning data are not required on a larger scale, since it has been demonstrated that the calculation of the polar surface area (PSA) of a compound is an adequate substitute reflecting the hydrogen bonding capacity of a molecule.[19] In practice, the lipophilicities of series of compounds are often measured by RP-HPLC,[20] and more recently by high-throughput shake-plate approaches. Within a series of closely related compounds $\log P/D$ values are correlated to $\log k_w$ values from RP-HPLC. Though this is often not the case for more diverse molecules. However, one should consider each lipophilicity scale as unique and reflecting a combination of the steric and H-bonding properties of a compound.[21,22]

It is also good to realize that properties like fragment lipophilicity contributions are additive properties, but may be very much dependent on the structural environment (Figure 23.2). Some substitutions may have a more dramatic effect than expected. Radioactive labeling with ^{125}I is quite common for biological studies. One should be aware, however, that aromatic iodination increases the $\log P$ of the compound by *ca.* 1 $\log P$ unit, and thus a different tissue distribution may result. An aromatic fluoro substituent has very little effect on the lipophilicity, but mainly serves in drugs to avoid oxydative biotransformation.

The aqueous *solubility* of a compound is another key property in drug discovery and poor solubility is often the cause of a series' demise. Aqueous solubility is inversely related to lipophilicity. High-throughput methods are now available to measure solubility,[23,24] but approaches to estimate solubility from computational models are less successful. The models are not sufficiently accurate to predict poor solubility with great precision. Moreover, solubility is influenced by many factors such as ionic strength, type of buffer, crystal packing, etc., all properties which are difficult to model and consistently measured data is not available from the literature. Nevertheless, solubility predictions can help projects to avoid making many compounds of poor solubility and guide the project into the right ballpark.

FIGURE 23.2 Impact of substitution on the lipophilicity of a compound depends on its structural environment.

Many drugs have ionizable groups at physiological pH. Therefore *ionization constants* (pK_a) are important properties for medicinal chemists to modulate properties they influence such as lipophilicity and solubility; which in turn affect absorption and distribution. The presence of charged groups may also be essential for activity against certain targets.

3. Other descriptors

Molecular structures may be considered at different levels, each containing certain types of information.[21,25] The simplest representation is the empirical chemical formula, while a highly complex representation is a molecular electrostatic potential (MEP) representation on the van der Waals surface which includes both steric and electronic information.

Molecular properties can be divided into various categories. There are experimental and calculated properties; properties that can be calculated from a 1D, 2D or 3D representation of molecules; whole molecule or molecular fragment (or substituent) properties; pure or composed (derived) properties; and intrinsic properties or those that depend on the environment. Properties generally have a quantitative value, but may also be represented by a count of the number of occurrences, for example, number of H-bond donors.

FIGURE 23.3 A molecule can be broken into fragments in a variety of ways. The presence (1) or absence (0) of a fragment can be used as a descriptor or properties can be calculated for the individual fragments. Substituent constants obtained from a look-up table are classically used to define the properties of substituents (shown in red).

Intrinsic properties are directly related to the structure without considering any interaction, such as molecular weight. Chemical descriptors may contain structural (also called global, whole or full molecule) information, for example, dipole moment or molecular shape descriptors, or local information for substructural parts of the molecule, for example, a local dipole or bulk at a certain substitution site of the molecule (Figure 23.3). Classically, substituent constants were used to parameterise substituents on an aromatic or aliphatic system, which described their electronic, steric and hydrophobic effects compared to a hydrogen substituent. Substituent constants were measured on a standard test system, but could be applied to a wide variety of core structures using a look-up table. A large set of chemical descriptors of molecular structures and fragments has been reported in the literature.[21,26]

Parameterization of chemical structures or substructures is not only of great interest to QSAR studies, but has much current interest in definitions of molecular similarity and diversity. This information may be used in molecular modeling studies or in combinatorial chemistry projects aimed at generating large molecular diversity in order to improve lead finding chances.

Classes of descriptors include

- Physicochemical ($\log P$, $\log D$, molar refraction, pK_a, solubility, etc.)
- Size/shape (molecular weight, moments of inertia, shadow descriptors, Verloop's STERIMOL substituent constants, etc.)
- Topological (connectivity indices, etc.)
- Hydrogen bonding (number of H-bond donors or acceptors, Abraham's α and β descriptors, PSA, etc.)
- Electronics/charge (formal charge, partial atom charges, dipole moment and vectors, HOMO and LUMO energy, F and R substituent constants, etc.)
- Fragment-based descriptors[27]

A selected subset of these properties is usually considered in the systematic variation of a selected substitution site during HTL and LO.

4. *Selection and scaling of descriptors*

The choice of descriptors to use in a QSAR analysis is dependent on several factors:

- What property data can be readily measured?
- What descriptors may be important to model the particular endpoint (biological activity or physicochemical property)?
- What can be readily calculated?
- What is particularly relevant to the therapeutic target?
- What variation is relevant to the compound series?

Ideally, a QSAR model is preferred which uses properties that can provide information on mechanism of action. Thus, properties which are readily understood, such as $\log P$ and pK_a, are ideal, while connectivity indices, for example, can be harder to interpret. Measured properties are generally more accurate than calculated properties but have the limitation that they are not available to predict the activity of novel analogs yet to be synthesized.

It is commonly the case that a wide variety of properties can be included in a QSAR analysis and a decision must be made on whether to include all possibilities or limit the number of descriptors. This decision depends on the size of the data set and the correlation matrix between the properties. Large sets of property data contain a lot of redundancy of information. For example, molecular weight, surface area and molar refraction are always highly correlated, therefore a decision to use only molecular weight could be made. Some multivariate statistical analysis methods are tolerant of data sets which contain more properties than compounds, for example, PLS, while others are not, for example, linear discriminant analysis (LDA). Ideally, a set of uncorrelated properties is desirable as this is most likely to give a robust, interpretable model.

In general, there is a trade-off between the predictivity and interpretability of a particular descriptor subset in combination with a particular data modeling tool. Thus often a linear model is built for interpretability and a non-linear model for predictivity. A prediction can be based on the consensus between the two (or more) models.

When using ensemble modeling, that is, calculating multiple models with the same technique, each model might use a different subset of descriptors. An alternative is to do the descriptor subset selection first.[28]

The choice of properties has a major influence on pattern recognition methods (unsupervised multivariate statistical or neural network methods) in particular and different property sets can result in quite different patterns of similarity between compounds.[29] Several methods are available to make selections of subsets of uncorrelated properties which can be used

FIGURE 23.4 The analysis of multivariate data to create predictive models. Two major types of studies can be defined: (1) determining the relationship between biological and physicochemical data using supervised methods such as MLR and PLS and (2) classification of compounds or descriptors using pattern recognition. This is shown schematically for a data set comprising *m* compounds in the training set and *n* compounds in the test set, characterized by *i* biological measurements and *j* physicochemical properties.

for QSAR studies. Some of these methods are discussed in Section IV.B.3.).

It is important to consider the different scales of various descriptors and whether or not scaling is performed during a data analysis. For example, molecular weight is commonly in the range 300–600, while absolute charge covers a much smaller data range and may vary from −2 to +2. Some analysis methods automatically center the descriptors to a mean of zero and scale them to give a standard deviation (SD) of 1, termed auto-scaling, so all descriptors are equally weighted. PLS does this, while MLR is generally performed on the raw data. Thus if molecular weight and absolute charge both contribute equally to activity, PLS would generate approximately equal coefficients for the descriptors, while MLR will give a much larger coefficient to the charge descriptor than the molecular weight. Unsupervised approaches commonly auto-scale the data. Thus, care needs to be taken in the interpretation of the size of the coefficients of the descriptors in a model.

B. Methods for building predictive models

1. *Modeling linear, non-linear and discontinuous data using supervised and unsupervised methods*

There are several ways of subdividing the various QSAR methods. One is to look at the nature of the activity data and its likely relationship to the descriptors: linear, non-linear, or discontinuous. Linear relationships for continuous activity data are commonly handled with MLR and PLS, while classified activity data can be analyzed with LDA. Non-linear and discontinuous relationships are also often found in biology and can best be treated with pattern recognition and machine learning methods. These include rule induction, recursive partitioning (RP), classification and regression tree (CART), Treenet, Random Forest (RF), rule discovery system (RDS), Bayesian methods, genetic algorithms (GA), *k*-nearest neighbor (*k*-NN), support vector machines (SVM), trend vectors, self-organizing maps (SOMs). Linear relationships can also be handled by most of these methods.[30]

The elements of QSAR or QSPR studies are depicted in Figure 23.4. The biological and physicochemical data relevant to a certain project may be represented as two tables and may be analyzed in various ways. On one hand, high quality and relevant biological data are required, whereas on the other hand relevant chemical descriptors should be defined. A further critical element is the proper choice of a model to investigate relationships within and between these tables. Taking the biological or physicochemical data separately, pattern recognition or classification studies may be useful to detect redundancy in the test systems or classify the compounds in a particular way which may be related to their specific mechanism of action. Clustering and classification of compounds based on their properties is central to molecular similarity studies. Finding correlation or other relationships between the biological and chemical data are of course useful to rationalize SAR. If the right prerequisites are met, relevant information may be extracted from the data, which can be used to get better understanding of the molecular structures and possibly the mode of action at the molecular level. This information may then be used to predict the properties and activities of new compounds. The statistical relevance of a model always needs to be evaluated (see Section III.B.4. and Box 23.1). Data sets that have the propensity to produce a chance correlation are typified by small sets, those skewed towards high and low values or those that have a low signal/noise ratio (e.g. HTS data, see also Section IV.D.).[31]

The following section looks at supervised methods for analyzing biological activity in terms of physicochemical properties and/or structural features. QSAR studies were initially focused on understanding activity in terms of variation in substituents using MLR and typified by the Hansch and FW approaches. Such analysis are now considered as traditional or classical QSAR methods. Recently, chemometrics has had a significant impact on QSAR analysis, and the methodology has been even further augmented by advances in artificial neural networks, GA and artificial intelligence. The various methods described below have different data requirements and different degrees of ease of

BOX 23.1 Key QSAR Statistics

In a QSAR equation, the activity is typically referred to as the Y variable and the descriptors as the X variables; the SD of each coefficient is given in parentheses; n is the number of compounds in the study; r is the correlation coefficient, where 1.0 indicates perfect correlation and 0 is totally uncorrelated and a good relationship should ideally be >0.8 and where r^2 (×100%) is the variance explained by the equation; s is the SD of the regression, which should have a value near to the experimental error in the biological dependent variable. It is good practice to report r^2 to see the explained variance. In more recent papers,[51] root mean square error (RMSE) is often reported instead of s. This is the more correct way of comparing the statistical quality of the model to the experimental data and error. The cross-validated correlation coefficient r^2_{cv} or q^2 is often included. A robust model has a $q^2 > 0.5$. Finally, the F-value is a measure for the statistical significance of the regression model and is calculated as the ratio between regression and residual variances. This value should be higher than a value which can be found in a Fisher F-statistics table, and is a function of the degrees of freedom and the significance level.[9] This is an important statistic to evaluate whether a model could have been found by chance. The traditional approach using the F-statistic has recently been challenged and adapted.[114] The error in the regression coefficients should not be larger than the coefficient itself. It is preferable to report 95% confidence intervals, rather than SDs which are smaller by a factor of 2 and may give a too optimistic figure.

Dependent (Y) data:

SD in experimental Y data: $$SD = \sqrt{\frac{\sum_{i=1}^{n}(Y_{observed} - Y_{mean})^2}{n-1}}$$

Standard error in the mean of the Y data: $$SE = \frac{SD}{\sqrt{n}}$$

Mean of Y data: $$Y_{mean} = \frac{\sum_{i=1}^{n} Y_{observed}}{n}$$

Correlation:

Pearson correlation coefficient:

$$r^2 = \frac{\sum_{i=1}^{n}(Y_{predicted} - Y_{mean})^2}{\sum_{i=1}^{n}(Y_{observed} - Y_{mean})^2}$$

Adjusted correlation coefficient: $$r^2_{adj} = 1 - (1 - r^2)\frac{n-1}{n-k-1}$$

where k is the number of descriptors/variables.

Standard error *of estimate of the regression equation*:

$$s = \sqrt{\frac{\sum_{i=1}^{n}(Y_{observed} - Y_{predicted})^2}{n-k-1}}$$

F-ratio: $$F = \frac{(n-k-1)\sum_{i=1}^{n}(Y_{predicted} - Y_{mean})^2}{k\sum_{i=1}^{n}(Y_{observed} - Y_{predicted})^2}$$

Root mean square error:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_{observed} - Y_{predicted})^2}{n}}$$

Bias or mean error: $$ME = \frac{\sum_{i=1}^{n}(Y_{observed} - Y_{predicted})}{n}$$

(also gives direction of bias)

Mean absolute error: $$MAE = \frac{\sum_{i=1}^{n}\left|Y_{observed} - Y_{predicted}\right|}{n}$$

Correlation to the line of unity: $$r_0^2 = 1 - \frac{n}{n-1}\left(\frac{RMSE^2}{SD^2}\right)$$

r_0^2 is calculated from the RMSE, the SD in the measured values and the number of compounds n. Hence, where the error in the model is larger than the SD in the training data (and thus RMSE/SD > 1), the r^2 around the line of unity (r_0^2) can be a negative value.

Cross-validation:

Using leave-one-out (LOO), models are built for n-1 compounds and a prediction is made for the omitted sample. Doing this for every compound leads to the following definitions:

Predictive residual sum of squares:

$$PRESS = \sum_{i=1}^{n}(Y_{observed} - Y_{predicted})^2$$

Sum of squares of Ys: $$SSY = \sum_{i=1}^{n}(Y_{observed} - Y_{mean})^2$$

Cross-validated correlation coefficient:

$$q^2 = r^2_{CV} = 1 - \frac{PRESS}{SSY}$$

interpretability. It is essential to select the appropriate analysis methods for the data set.

2. *Prediction of activity/property*

Supervised approaches leading to linear models are discussed first:

a. Multiple linear regression

This is the most common approach in QSAR to develop a quantitative linear correlation model, although not necessary always best for a given data set. The simplest means to obtain such a quantitative relationship is to use MLR available in any statistical software package. MLR generates predictive equations of the form:

$$\text{Activity} = \alpha(\text{descriptor_1}) + \beta(\text{descriptor_2})\cdots + \chi \tag{23.2}$$

Activity is called the "dependent" variable, and the descriptors are the "independent" variables. The equation generally contains a relatively small number of descriptors and so the model is easy to interpret. In order to avoid statistically insignificant relationships or chance correlations, one should always apply the following rules of thumb:

- The ratio of compounds to descriptors should be >5.
- The descriptors should not be inter-correlated (inter-descriptor correlation coefficient (r^2) should be < 0.5).

Also non-linear regression, that is, using quadratic terms such as $(\log P)^2$ and cross terms, may be used. However, as described in detail,[9] there are a number of pitfalls to this method. A statistically more robust method which could be used instead of MLR is the PLS regression method.

b. Partial least squares (PLS)

MLR is not suitable for analyzing data sets containing large numbers of properties or for properties which are inter-correlated. Both these data attributes are acceptable to PLS. PLS outputs an equation similar to MLR to describe activity, however, it is built up in a series of steps (components) and a decision needs to be taken on the optimal number of components to include based on the stability of the model as each new component is added. Stability is measured using CV (see section III.B.4. and Box 23.1). To improve the interpretation of a PLS model, so-called hierarchical modeling has been proposed.[27] This consists of building several sub PLS models and combining these to generate an overall prediction.

c. Discriminant analysis

In many biological experiments only discrete (or classification, or categorical) data are obtained, such as inactive/active, agonist/antagonist. In some data sets a clear separation may be found between such classes in multi-dimensional descriptor space. An appropriate method to describe separation between classes is LDA. LDA calculates hyper-planes which separate the different classes. For example, it was possible using LDA to distinguish 24 calmodulin inhibitors in three different activity groups (with different associated binding modes) using descriptors which described the positive potential surface area on the side chain, as well as the total and neutral surface areas on the ring in the inhibitor molecules.[32] This group assignment information was used to classify 29 additional inhibitors. A related method used to analyze categorical data is adaptive least squares (ALS).

Supervised approaches to model non-linear relationships are best explored with the following methods:

d. Soft independent modeling of class analogy

It often occurs that active compounds cannot be well separated from inactive ones using linear models such as PLS or LDA. This may be because the active compounds cluster together in a specific area of property space and they are surrounded by inactive compounds. Such data are called embedded or asymmetric data. Several methods have been developed to treat such data sets; the best known is the soft independent modeling of class analogy (SIMCA) algorithm. The SIMCA method is a tool for recognizing patterns in a data set.[14] The basic idea is to build local models for each class using disjoint principal component analysis (PCA) (see section "Principal component analysis") from a training data set. Predictions can then be made for test compounds as to which activity class the new compound belongs. A related method for analyzing embedded data is single-class discrimination.[33]

Figure 23.5 depicts the differences between data sets which can be analyzed by LDA and those best modeled using SIMCA.

e. Neural networks

A number of techniques related to artificial intelligence and natural computing have been investigated in QSAR. The methods use so-called natural algorithms, which are based on principles in nature, such as natural selection in evolution. Examples are artificial neural networks,[34] learning machines,[35] rule induction[36] and GA (see below). Different types of neural networks have been conceived, but for QSAR studies back-propagation (BP)[34] and Bayesian neural networks (BNN)[37,38] appear to be the most suitable approaches. A BP network generally consists of three layers; input nodes (one for each property), hidden nodes and output nodes (the predicted activity), linked by weights. The network is trained, that is, the values of the weights are optimized, using the property values of the compounds to accurately reproduce the activity of the compounds in the

FIGURE 23.5 Schematic diagram showing the difference between linearly separable activity classes (a) and an embedded (non-linear) structure (b). For example, the active compounds (yellow) in (a) tend to have higher values of PC1 and lower values of PC2 than the inactives (blue). While in (b), activity only occurs within a limited range of values of both PC1 and PC2 and compounds outside this region are inactive. Data in (a) could be classified by LDA, while the data in (b) could be analyzed using SIMCA for example.

FIGURE 23.6 Schematic representation of a 5-2-1 BP network. The network consists of an input layer containing five nodes into which values for descriptors D1–D5 are input. This is connected by weighted non-linear transfer functions to a hidden layer of two nodes, which is connected by weighted non-linear transfer functions, the final output layer of one node which is the activity value. The network is trained in an iterative fashion by adjusting the weights until the predicted activity values best match the measured activity values.

training set. An example of the architecture of a 5-2-1 BP network is given in Figure 23.6.

The advantage of neural networks is that few statistical assumptions are made *a priori*. Among the disadvantages are that no real statistical validation method has been developed. There is a danger of over-fitting the data, resulting in poor predictions outside the training set. The current best practice is to divide the data set into three: one training set, one test set and one validation set. The network is trained using the training set and training is terminated once the predictions for the test set start to deteriorate due to over-fitting of the training set. The ability of the network to make accurate predictions is then tested using the validation set. The results are not always easy to interpret in terms of chemistry due to the complexity and non-linearity of the neural network model. However, the ability to deal with non-linear relationships is an attractive feature of neural networks.

A special case of neural networks are so-called regularized BNN.[37,38] They have far less potential for overfitting as compared to more traditional backpropagation neural networks.

f. Decision trees

The term decision trees[36] is used here to cover a variety of related methods including rule induction and RP. Decision trees classify compounds into local activity groups using sets of rules. The data set is sequentially divided using rules which maximize the separation of active and inactive compounds at each step. For example, a simple rule which states $\log P > 2$ may serve to distinguish the majority of active compounds from more hydrophilic inactive ones. The data set is sequentially split in this manner, resulting in a tree diagram where the final leaves (nodes) contain, ideally, a single activity class.

Figure 23.7 shows a simple example of a decision tree where the majority of active compounds can be defined by a single set of rules. In more complex trees, one active group of compounds may be defined by a set of rules which state

$$\log P > 2,\ \mathrm{MW} < 450,\ \text{dipole moment} < 4$$

while another group of actives may be defined by a very different set of rules which state

$$\log P < 2,\ \mathrm{p}K_{\mathrm{a}} > 7.4,\ \mathrm{MW} < 300$$

This may indicate that the two sets of active compounds are acting via different mechanisms.

Decision trees come in different flavors such as CART, RF,[39] and Treenet, among others, but are nearly equivalent in practice in terms of predictive results.

g. Genetic algorithms

The statistically "best" set of descriptors is selected automatically in MLR, although there may be several good solutions within a larger set of descriptors, some of which may be more interpretable than the "best" solution. This problem occurs in particular when large descriptor sets are being considered. The identification of a selection of appropriate descriptor set solutions in a QSAR study can be done by several techniques, including genetic algorithms. These algorithms are inspired by population genetics and are believed to produce higher-quality predictive models.[40] Genetic algorithms are being widely applied in drug discovery software for QSAR and molecular modeling for optimization applications.[41]

FIGURE 23.7 Schematic diagram of a decision tree. Nine of the eleven active compounds (yellow) are characterized by MW < 450, charge = 1 and $\log D < 3.8$. Two active compounds are misclassified.

h. *k*-nearest neighbors

k-NN bases its prediction of activity of a test compound on its vicinity to a certain number (*k*) nearest neighbor compounds in multivariate descriptor space. For example, if *k* is set to three, then if the three closest neighbors in descriptor space to an unknown compound are all active, then the unknown is predicted to be active. If two of the three are active, then the unknown has a 66% probability of being active. Varying *k* may alter the prediction. This approach has, for example, been used to predict melting points.[115]

i. Support vector machines

SVMs were originally designed as a classification method using advanced mathematics to position a hyperplane to define and separate two or more classes. In later versions, it can also be used to predict continuous data. They are becoming increasingly popular in QSAR studies.[116]

j. Gaussian processes (GP)

These originate from the machine learning field and compare well and often exceed artificial neural networks for data modelling. The method prevents overtraining and does not require CV.[117]

3. *Exploiting molecular similarity*

a. Overview

Pattern recognition or unsupervised multivariate methods are used in QSAR to visualize similarity and clustering in a data set. The methods can be used to look for potential clustering of variables or compounds, by considering chemical or biological data separately or together. Another application is to look at clustering behavior of substituent descriptors which is useful to make a good experimental design.[42] When compounds cluster together in a multivariate descriptor space, this means that they are "similar" with respect to the variables considered. This forms the basis of the science of molecular diversity studies. Alternatively, if we perform a pattern recognition using biological data, clustered compounds have a similar activity profile. This section describes the methods of PCA, cluster analysis (CA), non-linear mapping (NLM) and Kohonen mapping (KM). All these methods can be used to reduce the dimensionality of a data set and give a visual display of similarity. PCA is the most widely used method in QSAR and has therefore been described in most detail in this section. A brief overview of the other methods has been provided. A comparison of PCA, CA, NLM and KM is given in.[43]

b. Principal component analysis

Large data tables may hide information which is not easily detected by simple inspection of the various columns. PCA and some closely related techniques such as factor analysis (FA), correspondence factor analysis (CFA) and NLM, reduce a data matrix to a smaller number of new super variables retaining a maximum of information (or variance) from the original data matrix. These new variables are called latent variables or principal components (PCs), and are orthogonal vectors composed of linear combinations of the original variables. This concept is shown schematically in Figure 23.8. If for example three variables are highly correlated, then they can be replaced by a single PC without loosing much information. The first component (PC1) is the axis through the data that explains the greatest amount of variance in the data set, the "best fit" line. The second component (PC2) is the axis explaining the maximum amount of variance remaining in the data set subject to the constraint that it is orthogonal to PC1. Thus, PCs 1 and 2 together define the best fit plane to the data. The PCs summarize the information in the data set in a smaller number of new, orthogonal variables and can be used to display similarity in 2D or 3D plots.

Two types of plot can be obtained from a PCA. Compound (row) similarity can be displayed on a 2D or 3D "scores" plot

FIGURE 23.8 Diagram showing the relationship between a set of descriptors (D1–D10) and their first three PCs (PC1–PC3). PC1 explains the greatest percentage of variation (depicted above by a longer vector), PC2 is orthogonal to PC1 and smaller, PC3 is the smallest PC and is orthogonal to PC1 and PC2. PC1 is highly correlated with descriptors D1–D4, PC2 is correlated with descriptors D5–D7, PC3 is correlated with descriptors D8 and D9. Descriptor 10 is not well represented by any of the PCs.

or the similarity of descriptors (columns) to each other and to the PCs can be displaid on a "loadings" plot.

The usual objective of PCA is to reduce the dimensionality of a data matrix, or determine its intrinsic dimensionality. The PCs can also be used in other QSAR methods including linear regression models (termed principal component regression, PCR). However, PLS gives similar results and is generally preferred to PCR.

c. Cluster analysis

CA is commonly used to investigate and display compound similarity, however, it can also be used for descriptor selection from a larger set. CA relies on the fact that similarity and dissimilarity among two points in multi-dimensional space can be quantified by calculating their inter-point distance. The most common measure being the Euclidean distance. Both hierarchical and non-hierarchical approaches exist. CA is often used complementary to PCA.

A number of different hierarchical clustering algorithms are available, for example, single, average or complete linkage. Ward's method is especially favored in QSAR and modeling studies. The difference between the methods lies in the definition of the spatial distance among pairs of data points and/or the cluster center. The results are presented as a dendrogram in hierarchical methods, that is, a tree-like figure, where very similar compounds or descriptors are close together and linked by short branches. Clustering may still be partially due to chance and unrelated to the underlying chemical or biological meaning. Sampling compound diversity may be achieved by selecting a compound from each cluster. Compounds in a single cluster would be expected to show similar properties and/or activities. The biggest drawback to hierarchical clustering algorithms is that they are compute-intensive on large data sets, so cannot be used to cluster corporate databases for example.

Non-hierarchical clustering methods are based on partitioning the data set and include, for example, the k-Means procedure and Jarvis-Patrick methods. k-Means is an iterative relocation clustering method. It is necessary to predefine the number of clusters required. Then, starting from a selected set of seed molecules which represent the initial cluster centroids, compounds or samples are assigned and re-assigned to clusters in an iterative manner until optimization is achieved. Jarvis-Patrick clustering is commonly encountered in QSAR for grouping compounds based on similarity, especially when clustering is undertaken using Daylight fingerprints or similar as the molecular descriptor.

Cluster significance analysis (CSA) is a related, supervised method that can be used to determine subsets of properties that cause active compounds to cluster together.[44,45]

d. Non-linear mapping

Like PCA, NLM or multi-dimensional scaling, is a method for visualizing relationships between objects, which in medicinal chemistry context often are compounds, but could equally be a number of measured activities.[46] It is an iterative minimization procedure which attempts to preserve interpoint distances in multi-dimensional space in a 2D or 3D representation. Unlike PCA, however, the axes are not orthogonal and are not clearly interpretable with respect to the original variables. However, it can be valuable in cases where the first two or three PCs are influenced by outliers (extreme data points) or only explain a small percentage of the original data. NLM has been used to cluster aromatic and aliphatic substituents,[47,48] for example.

e. Kohonen mapping

The KM, also known as a SOM, is a technique to map multi-dimensional properties on to a 2D grid. This is another unsupervised learning process which can be used to detect patterns in large data sets, and to predict whether a compound belongs to a particular activity class. A KM consists of a 2D grid of nodes and compounds with similar properties cluster in the same or adjacent nodes following training of the map. Other more complex representations of KMs have been reported such as 3D torus (doughnut-like) maps. KM has also been used to display the surface properties of a molecule on a 2D grid.[49]

A comparison of the results obtained by applying PCA, hierarchical CA, NLM and KM to a data set of 15 substituents characterized by five substituent constants (π, F, R, MR and Verloop's L) is given in Figure 23.9. It can be seen that the PCA and NLM plots are quite similar. This is to be expected when the first two PCs explain a high percentage

FIGURE 23.9 Comparison of (a) PCA, (b) non-linear mapping, (c) KM and (d) Ward's CA to display similarity of 15 substituents characterized by five *para* substituent constants (π, F, R, MR and L).

of the variation in the data (73% in this case). The main difference between the plots is on the left-hand side where the H, F, OH and NH_2 are less closely clustered in the NLM than in the PCA plot. Similar clustering is apparent in the CA and KM. In fact all four approaches give similar, but not identical, representations of similarity in this small example.

4. *Building and testing of predictive models*

It is good practice to build a model using a split of the data and use upto *ca.* 75–80% as the training set and the remainder as a test set. The selection can be based on a random choice, sorted biological activity, clustering,[50] or simply based on time, that is, early compounds as training set, later compounds as test set. Alternatively, statistical molecular design, such as coupling onion and D-optimal design, can be used to select a balanced training set[27] (see also Section III.E.).

The Pearson correlation coefficient (r^2) between measured and predicted data is often used as a measure for model quality. However, this can be misleading, since r^2 depends heavily on the variance (range) in the dependent (Y) variable. The RMSE for the unity line is the primary statistic of choice and also relates directly to the experimental variability[51] (see Box 23.1).

Modern internal validation techniques used in PLS and other multivariate statistical methods include CV, bootstrapping and Y-randomization. CV evaluates a model not by how well it fits the data, but by how well it predicts data. The data set consisting of n compounds is divided into groups. The MLR or PLS model is recalculated omitting one of the groups and the predicted values for that group are calculated. This is repeated for each group until every compound has been left out once and only once and has a predicted and measured value. This approach is referred to as the LOO method when each group only contains one compound. Many authors use the LOO method, although it has been shown that the leave-several-out (LSO) approach is preferable.[52] A recommendation is to divide the data set into seven groups. Using the predicted values the predictive residual sum of squares (PRESS) and SD values, a cross-validated r^2, called q^2, can be obtained (see Box 23.1). q^2 will always be smaller than r^2. When $q^2 > 0.5$ a model is considered significant.

Although CV may seem a robust validation technique, some difficulties should not be overlooked. Variables that do not contribute to prediction, that is, cause noise in the

model, may have detrimental effects on CV. This may particularly play a role when many variables have to be considered, such as in a 3D-QSAR CoMFA analysis (see Chapter 30). To overcome this problem, a procedure for variable selection in the case of many variables has been developed and is named generating optimal linear PLS estimations (GOLPE).[53] CV may also not work well when compounds are strongly grouped. Examples have shown that CV is misleading when it is applied after variable selection in stepwise MLR.[53] Thus, although CV is considered a state-of-the-art statistical validation technique, its results are only relevant when correctly applied.[54,55] Therefore, q^2 could be viewed as a measure of a model's self-consistency rather than its predictive power.

The related technique of boot-strapping is less commonly applied than CV. Subsets of samples are omitted in a random fashion in boot-strapping and confidence intervals for coefficients are generated rather than a PRESS value.

Another popular internal validation consists of scrambling the Y (dependent) variables and recalculating r^2. If the new r^2 is low then the model is assumed to have predictive value. This is also known as Y-randomization. In more rigorous approaches, Y is randomized about 500–1,000 times and r^2 values calculated, resulting in a measure of the probability that the equation is significant.[56]

The best way to detect the true predictive power of a QSAR model is external validation, using an independent test set. It is furthermore good to demonstrate the value of a model to predict compounds being made in the future and to look at a test set (called a temporal test set) for several months before implementing a new model for routine use.

It is important to understand the "applicability domain" (AD) of any model. For example, care needs to be taken when applying models over time as new compounds can be quite different from the compounds used to build the original predictive model. Such shifts in chemotype typically result in less accurate prediction. A measure on how well a model is expected to perform is to calculate a "distance-to-model" statistic.[27] This uses the model descriptors to measure how similar a new compound is to the compounds used to create the model. For example, it can be based on the Euclidean or Mahalanobis distance from the compound to the model space.[37] Apart from regularly rebuilding the model to take new chemical space into account, an elegant method to get the best possible predictions is to use associative or correction libraries.[37,57,58,64] These use the most recent experimental data to make adjustments in the errors in prediction.

A so-called confusion or contingency matrix is generally calculated for categorical classifications, where one models presence or absence of activity (see Box 23.2). This provides a measure of the prediction accuracy for known active compounds (sensitivity), non-active compounds (specificity), and overall predictivity (concordance or accuracy). To compare performances of a number of classification models, the receiver operating characteristic (ROC) or enrichment curve can be used. Good models show a high true positive rate and low false positive rate.

C. Global and local models, and consensus modeling

Predictive models can be built using all available experimental data measured under the same protocol and covering a wide range of chemotypes. Such models are called global models. Models for smaller subsets are called local models. The definition is nebulous, since local can mean various things. In some cases, local is site-specific (typically for larger pharmas), in others project-specific or series-specific. There is no *a priori* reason for a local or a global model to be more predictive. A strategy has been proposed to construct local models on the fly using the query compound and its nearest neighbors for which experimental data are available. This is called local lazy regression.[59,60] Another approach is to do similarity clustering first and then build a set of local models.[61] Such models combine the simplicity of local linear models with the complexity of global non-linear models.

Since the nature of the data set is not always clear, a good approach in prediction is to build a consensus model using a combination of several methods.[62] There are different ways of achieving this. One is to run the same approach, for example, a neural network, many times and build an ensemble or committee model.[63] The other is to use several approaches, for example, PLS, RF, NN, and calculate the weighted or unweighted average in a consensus model. The advantage is that you cover various aspects of the nature of the data and each individual model uses a potentially different descriptor subset.

D. Time-series behavior and autoQSAR

It is to be expected that the predictive performance of global models will detoriate over time as projects move on and produce new chemotypes not used in the generation of the model.[27] Literature on time-dependent QSAR model behavior is still sparse. In a recent investigation this effect was shown for a human plasma protein binding (hPPB) model.[17] A solution to this behavior is the use of correction libraries (see Section III.B.4.) and/or to update (rebuild) the model regularly.[64] An automated approach to this problem has been published[65] and it is expected that this technology (autoQSAR) will become more widespread in the future.[59,115]

It is important to realize that with modern technology (such as pipelining tools) allowing regular model updating predictions will change. These changes are due, for example, to updates in descriptors and use of new modeling

BOX 23.2 Confusion or Contingency Matrix for Classification Models

	Measured		
Predicted	Positive (good)	Negative (bad)	Total
Positive	a	b	$a + b$
Negative	c	d	$c + d$
Total	$a + c$	$b + d$	n
a: true positive	b: false positive	c: false negative	d: true negative

Sensitivity (= recall)	$a/(a + c)$ Fraction of positives correctly predicted positive; true positive rate
Specificity	$d/(b + d)$ Fraction of negatives correctly predicted negative; true negative rate
Positive predictive value (= precision)	$a/(a + b)$ Fraction predicted positives that are actually positive
Negative predictive value	$d/(c + d)$ Fraction predicted negatives that are actually negative
Prevalence	$(a + c)/n$ Fraction of positives
Concordance (= accuracy)	$(a + d)/n$ Fraction correctly classified
False positive rate	$b/(b + d)$ Fraction of negatives that are predicted positive = 1 − specificity
False negative rate	$c/(a + c)$ Fraction of positives that are predicted negative = 1− sensitivity
Overall agreement, Kappa(κ)	$\text{overall agreement} = \dfrac{\text{observed agreement} - \text{chance agreement}}{1 - \text{chance agreement}}$ $\kappa = \dfrac{(a + d) - [(a + c)(a + b) + (b + d)(c + d)]/n}{n - [(a + c)(a + b) + (b + d)(c + d)]/n}$ $\kappa < 0.4$: poor agreement κ 0.4 − 0.74: good agreement $\kappa > 0.75$: excellent agreement

tools. The overall intention is to get more robust models with predictions closer to the experimental values and with a smaller error. Obviously, the error in the predictions will never be better than those in the experimental data.

E. Experimental design

Optimization covers many aspects in medicinal chemistry. The optimization of the affinity and often selectivity for the biological target and the pharmacokinetic properties of a lead compound are the primary goals of most preclinical research projects. Secondly, optimization strategies may be applied to synthesis procedures to minimize the cost of goods.[66] In both cases, a number of variables have to be taken into account simultaneously. Strategies changing only one variable at a time take much time and many experiments are needed. In contrast to the sequential approach, the full variable space can be covered by far less experiments by using a proper selection of a limited number of experiments. Experimental design schemes are therefore of great help to focus on the most informative experiments. These techniques have been applied in many types of synthetic programs, including peptide design and substituent variation for example. Below we will describe how physicochemical descriptors for aromatic and aliphatic substituents may be used for substituent

selection. In a similar way, relevant descriptors for amino acids may be used. It is widely acknowledged that a series of compounds based on some experimental design plan are most likely to produce significant QSAR equations. Therefore, careful selection of appropriate substituents and variables is an important matter. More recently, design strategies have been applied to the creation of virtual and combinatorial libraries and selection of sets of compounds for purchase or screening. Specifically, attention has been focused on the relevance of such compounds for drug discovery purposes.

Various strategies have been advocated in order to cover the physicochemical parameter space of a series of new compounds as well as possible. Familiar strategies go back to proposals by Topliss[67] and Craig.[68] Both are schemes used for substituent variation at a selected site. The Topliss substitution scheme can be used to optimize aromatic and aliphatic substituents using a fixed set of substituents and rules. A Craig plot is a 2D plot of selected descriptors, for example, Hammett σ (electronic properties) and Hansch π values (lipophilicity). Substituents can be selected from each quadrant of this plot such that they vary widely in their properties, for example, lipophilic and hydrophilic, electron-donor and electron-acceptor, and to ensure the two properties are not correlated in the selected set which is preferable for the generation of stable QSAR models.

A further extension would be to consider a 3D Craig plot using three descriptors, for example, reflecting steric, lipophilic and electronic properties of the substituents. In that case, substituents may be chosen from the eight octants. If one wants to consider even more descriptors, this approach becomes impractical. In that case, more advanced experimental design techniques may be applied. One approach taken by Hansch and Leo was to use CA to define sets of aliphatic and aromatic substituents useful in the design of compounds for synthesis, such that various aspects of the substituents are taken into account in a balanced way.[69]

Factorial designs (FD) are valuable when more than three properties are to be systematically varied. In order to limit the number of combinations, each variable may be considered at two levels, for example, lipophilic versus hydrophilic. A two-level FD with k variables requires 2^k experiments. A Craig plot is an example of a 2^2 FD. Or in other words, the minimum number of compounds to synthesize using two descriptors is four (one from each quadrant). As stated above, with many variables this number rapidly becomes impractical and fractional factorial designs (FFD) are preferred. Using a reduction factor, r, the number of experiments then becomes 2^{k-r}. This reduction factor is, in practice, chosen rather pragmatically, such that one has to consider 8 or 16 compounds. Further design schemes are known as central composite and D-optimal designs. In this latter method, the determinant of the variance–covariance matrix of considered properties is calculated. This determinant has a maximum value for those combinations of substituents which have a maximum variance and minimum covariance in their physicochemical descriptors. The variance–covariance matrix is an important mathematical cornerstone in matrix operations used in MLR and PCA, and to obtain a correlation matrix among a set of selected variables.

A large number of substituent descriptors have been reported[21,69] in the literature. In order to use this information for substituent selection, appropriate statistical methods may be used. Pattern recognition or data reduction techniques, such as PCA or CA are good choices. As explained in Section III.B.3. in more detail, PCA consists of condensing the information in a data table into a few new descriptors made of linear combinations of the original ones. These new descriptors are called PCs or latent variables. This technique has been applied to define new descriptors for amino acids, as well as for aromatic or aliphatic substituents, which are called principal properties (PPs). These PPs can be used in FD methods or as variables in QSAR analysis.[70]

F. Inverse QSAR and multi-objective optimization

QSAR normally tries to describe molecules from properties. The inverse has also been looked at, that is, how can properties help to engineer better molecules.[71,72] The challenge is to develop an automated iterative design taking QSAR models to suggest improved structures with better properties and using appropriate filters to test the AD and synthetic feasibility. Current approaches to such knowledge-based design include transformation vectors[73] and matched-pairs analysis.[74,75] It is expected that in the near future we will see more use of inverse QSAR/QSPR workflows using multi-objective optimization schemes.[72,76]

IV. PRACTICAL APPLICATIONS

A. Limitations and appropriate use

Although molecular modeling studies may be more appealing to most medicinal chemists, QSAR studies appear to be of considerable interest to many projects. In many cases, the two approaches are complementary. Nevertheless, one should be aware that in QSAR (for endpoints involving receptors or enzymes) as well as in modeling studies, most models are being developed and used under the assumption of a single binding mode. However, X-ray studies have shown very elegantly that different binding modes may occur, even within a series of closely related structures such as thrombin inhibitors.

Genomics, bioinformatics and chemoinformatics, combinatorial chemistry, and high-throughput screening are key contributors to modern medicinal chemistry, particularly for finding new targets and lead compounds. Medicinal chemists

apply parallel and traditional synthesis methods to optimize these leads to drug candidates with high target affinity and selectivity. In parallel pharmacokinetic and physicochemical properties need to be optimized, which taken together produce the desired therapeutic profile. Quantitative approaches to structure–activity, structure–permeability, structure–pharmacokinetic and structure–metabolism relationships will continue to play a role in this optimization process.[77] The wide recognition of the importance of the latter three types of relationships to molecular structure, as well as the desire for good library design and HTS analysis has renewed a keen interest in various QSAR approaches.[78,79]

Indeed, it is often valuable to use QSAR *in silico* predictions in combination with simple *in vitro* filters during the optimization phase. A recent example of such an *in combo* approach[80] is the prediction of skin permeation.[81]

A wide variety of chemometric statistical tools may be used to investigate QSAR or more general structure–property correlations. Some of these techniques require expert support. However, the bench chemist may successfully use a number of techniques, when the basic guidelines discussed in this chapter are followed. The most important methods are:

- Supervised predictive models:
  - Multiple linear regression (MLR)
  - Partial least squares (PLS) regression
  - BP neural networks (BPNN)
  - Bayesian neural networks (BNN)
  - Recursive partitioning (RP), classification trees, decision trees
  - Support vector machines (SVM)
- Pattern recognition models:
  - Principal component analysis (PCA)
  - Non-linear mapping (NLM)
  - Cluster analysis (CA)
  - Kohonen map neural networks (KM or SOM)

The first question to consider is "which physicochemical or structural descriptors should be used"? The answer may be rather pragmatic: all those available to me. First of all, any experimental physicochemical property can be used, such as lipophilicity data,[82] aqueous solubility, membrane transport properties and ionization constants. A number of descriptors can be readily calculated, such as the molecular weight, octanol/water partition coefficients, molar refractivity, molecular volume, surface area, PSA, number of H-bond donating and accepting groups. Since drugs and their targets are 3D objects, it is of course appropriate to consider 3D molecular properties. However, 3D molecular properties and quantum mechanical electronic properties, such as partial atomic charges, will probably require expert assistance from the modeling group. The number of descriptors may be too large to generate stable models and variable selection tools will need to be applied and again, this will probably be best left to the experts. Once the descriptor set has been generated, the medicinal chemist can apply QSAR models to analyze the entire data set or subsets of related analogs.

There are now approaches in which one builds a family of QSAR models where each model represents a different compromise between complexity and accuracy.[76] Either the best or the full range of models can be presented to the user depending on the user's objectives with the model.

It is important in any prediction to know whether the query compounds fall within the AD of the model. Ideally, the AD should cover the structural, physicochemical and response space (e.g. mode of action) of the model.[83] The distance-to-model measure can be used as a measure for the AD (see Section III.B.4.).[27,64]

B. Examples

1. *Biological activity*

An example of a Hansch analysis (see section III. B. 2.a) using MLR is a study on substituted tetrahydroisoquinolines with affinity for both phenylethanolamine *N*-methyltransferase (PNMT) and the α_2-adreno-receptor[84] (see Figure 23.10). The multiple regression equations obtained were:

$$\begin{aligned} \mathrm{p}K_{\mathrm{i}}(\mathrm{PNMT}) &= 0.60(\pm 0.17)\pi - 0.073(\pm 0.027)\mathrm{MR} \\ &\quad + 1.55(\pm 0.92)\sigma_{\mathrm{m}} + 5.80(\pm 0.48) \\ n = 27 \quad r^2 &= 0.78 \quad s = 0.57 \quad F = 27.6 \end{aligned} \tag{23.3}$$

$$\begin{aligned} \mathrm{p}K_{\mathrm{i}}(\alpha_2) &= 0.60(\pm 0.13)\pi - 0.054(\pm 0.019)\mathrm{MR} \\ &\quad - 0.95(\pm 0.62)\sigma_{\mathrm{m}} + 6.45(\pm 0.34) \\ n = 27 \quad r^2 &= 0.84 \quad s = 0.40 \quad F = 40.5 \end{aligned} \tag{23.4}$$

Two types of information can be obtained from these equations:

1. statistical quality and relevance (see Box 23.1)
2. chemical implications for maximizing activity and selectivity

In QSAR studies it is common to transform biological activities to their negative logarithmic form, for example, $\log 1/IC_{50} = -\log IC_{50} = \mathrm{p}IC_{50}$ and similarly for binding affinities, $\mathrm{p}K_{\mathrm{i}}$ is used. Thus, the most active compounds have largest values. Among all the descriptors evaluated in

FIGURE 23.10 Optimization of activity and specificity of 7-substituted tetrahydroquinolines by MLR.

this study, only those which are relevant appear in the final equation. π is the lipophilic constant for the substituent, MR is its molar refraction and σ_m is the electron-donating effect.

This is an interesting example as the equations show that the lipophilicity (π) and bulk (MR) effects are similar, but that the sign for the coefficient in σ_m is opposite for the two targets. Thus, selectivity can be achieved by exploiting differences in the electron-donating properties of the substituents.

2. *Predictive* ADMET

The rational design of new chemical entities intended for use as drugs can be based on several methods. Structure-based drug design has been very successful for the optimization of binding to the biological target. However, a good drug not only needs high and selective affinity for its target, it should also have appropriate pharmacokinetic[77,85] and biopharmaceutic properties.[86] Taking these properties into account early on has been called property-based design.[87] *In silico* predictions of ADMET properties has received much attention in recent years.[88,89] High oral absorption and bioavailability are important properties in many projects. There is therefore considerable interest in developing predictive models for oral absorption.[90] Equation (23.5) is the result of a careful selection of good data on human oral absorption ($A\%$).[91]

$$A\% = 2.94E + 4.10S + 10.6V - 21.7A - 21.1B + 92$$
$$n = 169 \qquad r^2 = 0.74 \qquad s = 14 \qquad F = 93 \tag{23.5}$$

This equation is based on Abraham's solvation equation[91] which uses five molecular descriptors: excess molar refraction (E), solute polarity/polarizability (S), McGowan characteristic volume (V), solute overall acidity (A) and basicity (B). The steric (size/shape) descriptors E, S and V have a positive effect on oral absorption, while the descriptors related to H-bonding, A and B, have a negative effect. The model accounts for 74% of the variance (r^2) in the data and the predictions have a 14% standard error (s). This is nearly as good as it gets, since the experimental biological variance is *ca*. 15%.

There are two important other points to mention here. The original paper[91] did not report a SD or confidence interval for the regression coefficients of the equation, so it is difficult to judge its stability. The physicochemical theory behind the Abraham approach means these five descriptors are always forced into the equation, which is against good statistical practice. It is only acceptable, if there is low cross-correlation ($r^2 < 0.5$) among the descriptor pairs. This may lead to an over-estimation of the statistical relevance of the equation and therefore its applicability in predictions.

The literature reports many other human absorption models, using more complex methods such as SVM, ANN, PLS, etc. However, the predictivity for all approaches is very similar, so simplicity and interpretability should prevail over complexity and black box in these situations.

RP has successfully been used to derive predictive ADME models. Good examples include the prediction of CYP1A2 and CYP2D6 inhibition using MOE descriptors[92] and blood-brain barrier penetration using 19 simple molecular descriptors and various fragmentation schemes.[93] These studies show on one hand the importance of the data set, both in quality and chemical diversity,[92] and on the other hand the beauty of having simple intuitive descriptors in the models.[93]

Drug-induced QT prolongation is related to the blockage of the human ether-a-go-go-related gene (hERG) ion channel and has led to drug withdrawals from the market. Early prediction of hERG binding is therefore topical. An example is a literature study comprising 90 compounds, in which SVM, PLS and RF are compared using fragment fingerprints as descriptors.[94] The best results in this case were achieved using SVM with an $r^2 = 0.85$ and RMSE = 0.60 for a test set. In another study by a pharmaceutical company,[27] a hierarchical PLS was used to derive a predictive hERG model involving 1,312 compounds in the training set and 436 as test set. PLS was used because it can deal with many descriptors; 606 were used here. These examples show the difference in scale between an academic and an industrial model. Both examples show that there is no best statistical method, but the combination of data, descriptors and analysis tool needs to be evaluated each time.

3. *Data pre-processing and pattern recognition*

A typical use of PCA is illustrated by an example from antibacterial research. The antibacterial effects of sulfones and sulfonamides in whole-cell and cell-free systems has been analyzed. In this example, some missing activity data (19%) have been estimated by an iterative process using PCA.[95] However, estimation of missing values should be done with care and preferably avoided.

Analyzing the minimal inhibition concentration (MIC) data for nine different strains, resulted in two significant PCs being obtained, accounting for 77.1% and 16.1%, respectively, of the data variance. Thus, the intrinsic dimensionality of the data matrix of nine assays was only two. The loading plot, that is, a plot of the calculated PCs with respect to the descriptors, shows that the first component is mainly related to the seven cell-free test systems, while the second component represents the two whole-cell test results. Thus, there was much correlation between the results of the seven cell-free assays and the two whole-cell assays were correlated with each other but not with the cell-free assays. In other words, much redundant information was obtained

from the measurement results of nine test systems; two would have been sufficient. This separation means that the potency in both test systems is governed by different physicochemical properties.

The PCs derived from the activities can be correlated to physicochemical properties using MLR. Thus, it was found that component one (PC1) appears to be dominated by electronic factors (equation 23.6), while in component two (PC2) transport (lipophilicity) properties (equation 23.7) play a role. The following parameters are used: $\Delta\text{ppm}(NH_2)$ is the nuclear magnetic resonance (NMR) chemical shift of the amino protons relative to the unsubstituted congener, f_i is the fraction ionized at pH 7.4, and $\log k'$ is the lipophilicity measured by HPLC. Equation (23.7) shows an example of a non-significant constant term, since the SD is larger than the term itself. In such cases, the equation should be forced through the origin.

$$\begin{aligned} \text{PC1} &= -7.02(\pm 1.25)\,\Delta\text{ppm}(NH_2) \\ &\quad + 1.81(\pm 0.42)\,f_i - 0.93(\pm 0.19) \\ n &= 17 \quad r^2 = 0.94 \quad s = 0.26 \quad F = 107.9 \end{aligned} \tag{23.6}$$

$$\begin{aligned} \text{PC2} &= 1.40(\pm 0.52)\log k' - 3.49(\pm 1.32)\log(0.098k' + 1) \\ &\quad + 0.51(\pm 0.73) \\ n &= 17 \quad r^2 = 0.87 \quad s = 0.40 \quad F = 22.3 \end{aligned} \tag{23.7}$$

Thus, PCA may be used to filter out the most relevant information in a data set. Variations of PCA, such as correspondence factorial analysis and NLM may sometimes have small advantages with particular data sets, but require expert support.

Another example of a descriptor pre-processing is a multivariate analysis of HIV-1 protease inhibitors.[96] The aim of the study is to develop a method to predict on the basis of simple descriptors whether compounds are likely to trigger resistance or are effective against mutant HIV strains. PCA is used to reduce the 12 original descriptors to 4 uncorrelated descriptors. These four are then used to build a LDA model.

C. Library design, compound acquisition and profiling

The design of new compound libraries may be based around a lead compound or X-ray structural information about the target or a combination of both. Cheeseright *et al.*[97] have described using the molecular field of an active molecule as a design template, when no X-ray data of the target is available.

In the absence of a clear model for activity, an ideally designed set of compounds should consist of a variety of structural shapes and molecular properties, while avoiding redundancies.[98] The above principle has been extended for use in the design of combinatorial libraries as well as for the selection of compounds for high-throughput screening (HTS). Basically, for this purpose a large set of 1D, 2D and 3D molecular representations can be used together with appropriate statistical tools as discussed in this chapter.

The choice of appropriate substituents and building blocks for library synthesis depends on synthetic feasibility, availability and costs, but should also be based on an understanding of the physicochemical properties of the substituents, as well as the predicted properties of the targeted compound. Therefore, we discuss below how these choices may be made as rational as possible by considering design techniques. Indeed, there has been much debate as to whether combinatorial library design should be based on reagent (monomer) diversity or on the diversity of the final product.[99] Some relative merits of each approach are outlined below:

- Advantages of using final product designs
  - Can calculate the physicochemical property profile of the final array.
  - Suitable for pharmacophore-based designs.
  - More descriptors are relevant to final products than reagents.
  - "Greater diversity" can be obtained using a final product design.[99]
- Advantages of using reagent designs
  - Easy substitution of reagent if it is found to be poorly reactive.
  - Reagent diversity is quicker to calculate than fully enumerated final product libraries.
  - Clustering studies on reagent sets can be re-used for new arrays.
  - More sympathetic to parallel synthesis logistics.

An interesting hybrid approach, termed reagent-biased product design, selects reagent sets based on final product diversity. This approach was initially proposed by Good and Lewis.[100] The method uses a genetic algorithm (see Section III.B.1.) to optimize the selection of sets of monomers in an iterative fashion to maximize diversity in the final all-combinations array. Reagent-based product design is now quite widely used and commonly the method of choice.[99,101–103] See[104] for a review on library design strategies.

The descriptors most commonly applied to measure diversity in large compounds sets (combinatorial libraries or screening collections) include 2D fingerprints (for substructural diversity) and 3D three-point pharmacophores[100] (to measure the range of pharmacophoric groups that are present within the compound set).

Virtual screens are applied to assist in library (array) design, decisions on commercial library acquisition and the

selection of screening sets of compounds. They are used to measure and focus both diversity and relevance to drug discovery as part of the overall experimental design process. One approach is to test for "drug-likeness" properties. Other virtual screens or filters may include, for example, screening compounds against pharmacophores derived for a specific target,[102] screening for molecular field similarity to an active ligand,[97] high-throughput docking studies,[105] screening for unwanted reactive groups, screening for suitable physicochemical properties,[101,103] screening for diversity of 3D pharmacophores, etc.

Various analyses of historical data on existing drugs have led to a better understanding of what constitute *drug-like properties*.[23,106] Several groups have trained neural networks to distinguish between "drug-like" and "non-drug-like" compounds.[106,107] The training sets for the network consist of one large database of drug molecules (e.g. from Derwent's World Drug Index, WDI) and one database on non-drugs (e.g. MDL's Available Chemicals Directory, ACD). Both databases are pre-processed to remove unwanted substances. The trained network can then be used to predict the drug-likeness of combinatorial arrays and screening sets.

An analysis of compounds in the WDI has resulted in the *rule-of-fives*,[23] which states that poor oral absorption is expected for compounds with a molecular weight above 500, a calculated log P >5, the number of hydrogen bond donors >5, and the number of hydrogen bond acceptors >10. Exceptions to this rule may be compounds with an active uptake mechanism. Such simple rules are now widely used in the design of combinatorial libraries. More detailed knowledge on relevant properties can be used to optimize oral absorption and PK,[77,85] an approach which has been called property-based design.[87]

D. HTS analysis

The analysis of HTS data is another area where predictive models and filters have proved useful to identify the most promising compounds and series.[108–111] The challenge with HTS data is their rather noisy nature. To address this, a technique that increases the significance of active compounds has been presented.[112] In another study, the corporate database was analyzed using rigorous statistical methods in an attempt to identify compound scaffolds that appear to give technology-related screening artifacts or demonstrate target family specific activities.[113]

E. Software

Most of the QSAR software listed in Box 23.3 is readily available in PC packages (e.g. JMP), or as client–server applications such as SAS and Pipeline Pilot. Some is freeware such as R or Orange. Unfortunately, much of the

BOX 23.3 QSAR Software and Vendors

Some of the most commonly used software in QSAR studies is listed below. Many vendors sell a variety of software for QSAR. For a more comprehensive list visit the individual vendors web site or see www.ndsu.edu/qsar_soc/resource/software.htm.

QSAR/data mining

Vendor	URL	*Software*
Accelrys	www.accelrys.com	TSAR, Cerius2, Pipeline Pilot
Applied InSilico	www.appliedinsilico.com	Evolutionary Learning Environment
BioWisdom	www.omniviz.com	Omniviz
Chemical Computing Group	www.chemcomp.com	MOE
Coalesix	www.coalesix.com	Mobius
Compumine	www.compumine.com	Rule Discovery System (RDS)
Cresset BMD	www.cresset-bmd.com	FieldAlign, FieldTemplater
IDBS	www.idbs.com	PredictionBase
Salford Systems	www.salford-systems.com	TreeNet, CART, MARS, RandomForests
SAS Institute	http://www.sas.comwww.sas.com	SAS
	www.jmp.com	JMP
StatSoft	www.statsoft.com	Statistica
The R Foundation	www.r-project.org	R
TIBCO	www.spotfire.com	Spotfire
Tripos	www.tripos.com	CoMFA, HQSAR
Umetrics	www.umetrics.com	SIMCA
University of Konstanz	www.knime.org	Knime
University of Ljubljana	www.ailab.si/orange	Orange
University of Waikato	www.cs.waikato.ac.nz/~ml/weka/index.html	Weka

(Continued)

BOX 23.3 (Continued)

Physical chemistry

Company	URL	*Software*
ACD/labs	www.acdlabs.com	ACD Log D Sol Suite
Biobyte	www.biobyte.com	Bio-Loom, CQSAR
ChemSilico	www.chemsilico.com	Various
CompuDrug	www.compudrug.com	Various

Descriptors

Company	URL	*Software*
ChemAxon	www.chemaxon.com	Calculator plugins
Molecular Networks	www.molecular-networks.com	Adriana Code
Molinspiration	www.molinspiration.com	Property calculator
Pharma Algorithms	www.ap-algorithms.com	ABSolv
Talete	www.talete.mi.it/	Dragon
Timtec	www.timtec.net/software/	HYBOT

ADMET

Company	URL	*Software*
BioFocus DPI	www.biofocusdpi.com/ADME_optimization	Admensa
Bio-Rad Laboratories	www.biorad.com	KnowItAll
Fujitsu	http://www.fqs.pl/life_science	ADMEWORKS
Molecular Discovery	www.moldiscovery.com	VolSurf
Simulations Plus	www.simulations-plus.com	ADMET Predictor
Strand Genomics	www.trupk.strandgenomics.com	TruPK

Societies and other links

Organization	URL
Cheminformatics and QSAR Society	www.qsar.org
UK QSAR and Chemoinformatics Group	www.ukqsar.co.uk
QSARWorld	www.qsarworld.com

software is targeted at expert users, but there are exceptions. Visualization of the data and its information content can be done easily with packages such as Spotfire. JMP is a relatively simple data analysis package. Mobius is a library design tool for chemists. FieldAlign and FieldTemplater allow chemists to explore molecular field SAR in a qualitative way. Many companies rely at least in part on in-house developments for chemists.

The most difficult step in building QSAR models is getting access to relevant descriptors and the selection of adequate descriptors. TSAR provides one of the best source of descriptors for chemists. Often the statistical tool is of less importance than the descriptors in the predictivity of a model. In practice, consensus models using a variety of properties and multivariate methods can often maximize usefulness.

REFERENCES

1. Engel, Th. Basic overview of chemometrics. *J. Chem. Inf. Model.* 2006, *46*, 2267–2277.
2. Chen, W. L. Chemoinformatics: past, present and future. *J. Chem. Inf. Model* 2006, *46*, 2230–2255.
3. Stouch, T. R., Kenyon, J. R., Johnson, S. R., Chen, X. Q., Doweyko, A., Li, Y. In silico ADME/Tox: why models fail. *J. Comput.-Aided Mol. Des.* 2003, *17*, 83–92.
4. Martin, Y. C. What works and what does not: lessons from experience in a pharmaceutical company. *QSAR Comb. Sci.* 2006, *25*, 1192–1200.
5. Hansch, C., Selassie, C. Quantitative structure–activity relationship – a historical perspective and the future. In *Comprehensive Medicinal Chemistry* 2nd Ed., vol. 4, *Computer-Assisted Drug Design,* (Taylor, J. B., Triggle, D. J., Eds), Elsevier: Oxford, 2007, pp. 43–63.
6. Hansch, C. The physicochemical approach to drug design and discovery (QSAR). *Drug Dev. Res.* 1981, *1*, 267–309.
7. Hansch, C. On the state of QSAR. *Drug Inf. J.* 1984, *18*, 115–122.
8. Craig, P. N. QSAR – origins and present status: a historical perspective. *Drug Inf. J.* 1984, *18*, 123–130.
9. Kubinyi, H. Methods and principles in medicinal chemistry, Vol. 1, *QSAR: Hansch Analysis and Related Approaches*, VCH: Weinheim, 1993.
10. Hansch, C. Quantitative structure–activity relationships and the unnamed science. *Acc. Chem. Res.* 1993, *26*, 147–153.
11. Hyde, R. M., Livingstone, D. J. Perspectives in QSAR: computer chemistry and pattern recognition. *J. Comput.-Aided Mol. Des.* 1988, *2*, 145–155.
12. Kubinyi, H. The Free-Wilson method and its relationship to the extrathermodynamic approach. In *Comprehensive Medicinal*

Chemistry, Vol. 4, *Quantitative Drug Design* (Hansch, C., Sammes, P. G., Taylor, J. B., Eds), Pergamon Press: New York, 1990, pp. 589–643.
13. Cramer, R. D. Partial least squares (PLS): its strength and limitations. *Perspect. Drug Discov. Des.* 1993, *1*, 269–278.
14. Eriksson, L., Johansson, E. Multivariate design and modeling in QSAR. *Chemom. Intell. Lab. Syst.* 1996, *34*, 1–19.
15. Fleischer, R., Frohberg, P., Büge, A., Nuhn, P., Wiese, M. QSAR analysis of substituted 2-phenylhydrazonoacetamides acting as inhibitors of 15-lipoxygenase. *Quant. Struct.-Act. Relat.* 2000, *19*, 162–172.
16. Gedeck, P., Rohde, B., Bartels, C. QSAR – How good is it in practice? Comparison of descriptor sets on an unbiased cross section of corporate data sets. *J. Chem. Inf. Model.* 2006, *46*, 1924–1936.
17. Rodgers, S. L., Davis, A. M., van de Waterbeemd, H. Time-series QSAR analysis of human plasma protein binding data. *QSAR Comb. Sci.* 2007, *26*, 511–521.
18. Young, R. C., Mitchell, R. C., Brown, T. H., Ganellin, C. R., Griffiths, R., Jones, M., Rana, K. K., Saunders, D., Smith, I. R., Sore, N. E., Wilks, T. J. Development of a new physicochemical model for brain penetration and its application to the design of centrally acting H_2 receptor histamine antagonists. *J. Med. Chem.* 1988, *31*, 656–671.
19. van de Waterbeemd, H., Kansy, M. Hydrogen-bonding capacity and brain penetration. *Chimia* 1992, *46*, 299–303.
20. Lombardo, F., Shalaeva, M. Y., Tupper, K. A., Gao, F., Abraham, M. H. ElogPoct: a tool for lipophilicity determination in drug discovery. *J. Med. Chem.* 2000, *43*, 2922–2928.
21. van de Waterbeemd, H., Testa, B. The parametrization of lipophilicity and other structural properties in drug design. *Adv. Drug Res.* 1987, *16*, 85–225.
22. El Tayar, N., Testa, B., Carrupt, P. A. Polar intermolecular interactions encoded in partition coefficients: an indirect estimation of hydrogen-bond parameters of polyfunctional solutes. *J. Phys. Chem.* 1992, *96*, 1455–1459.
23. Lipinski, C. A., Lombardo, F., Dominy, B. W., Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug. Deliv. Rev.* 1997, *23*, 3–25.
24. Bevan, C. D., Lloyd, R. S. A high-throughput screening method for the determination of aqueous drug solubility using laser nephelometry in microtiter plates. *Anal. Chem.* 2000, *72*, 1781–1787.
25. Testa, B., Kier, L. B. The concept of molecular structure in structure–activity relationship studies and drug design. *Med. Res. Rev.* 1991, *11*, 35–48.
26. Todeschini, R., Consonni, V. Methods and principles in medicinal chemistry, Vol. 11. In *Handbook of Molecular Descriptors*, VCH: Weinheim, 2000.
27. Gavaghan, C. L., Hasselgren-Arnby, C., Blomberg, N., Strandlund, G., Boyer, S. Development, interpretation and temporal evaluation of a global QSAR of hERG electrophysiology screening data. *J. Comput.-Aided Mol. Des.* 2007, *21*, 189–206.
28. Dutta, D., Guha, R., Wild, D., Chen, T. Ensemble feature selection: consistent descriptor subsets for multiple QSAR models. *J. Chem. Inf. Model.* 2007, *47*, 989–997.
29. Rose, V. S., Rahr, E., Hudson, B. D. The use of Procrustes analysis to compare different property sets for the characterization of a diverse set of compounds. *Quant. Struct.-Act. Relat.* 1994, *13*, 152–158.
30. Dudek, A. Z., Arodz, T., Galvez, J. Computational methods in developing quantitative structure–activity relationships (QSAR): a review. *Comb. Chem. High Through. Scr.* 2006, *9*, 213–228.
31. Tong, W., Hong, H., Xie, Q., Shi, L., Fang, H., Perkins, R. Assessing QSAR limitations – a regulatory perspective. *Curr. Comput.-Aided Drug Des.* 2007, *1*, 195–205.
32. Liu, Q., Hirono, S., Moriguchi, I. Quantitative structure–activity relationships for calmodulin inhibitors. *Chem. Pharm. Bull.* 1990, *38*, 2184–2189.
33. Rose, V. S., Wood, J., MacFie, H. J. H. Single class discrimination using principal component analysis (SCD-PCA). *Quant. Struct.-Act. Relat.* 1991, *10*, 359–368.
34. Manallack, D. T., Livingstone, D. J. Neural networks in drug discovery: have they lived up to their promise?. *Eur. J. Med. Chem.* 1999, *34*, 195–208.
35. King, R. D., Hirst, J. D., Sternberg, M. J. E. New approaches to QSAR: neural networks and machine learning. *Perspect. Drug Discov. Des.* 1993, *1*, 279–290.
36. A-Razzak, M., Glen, R. C. Rule induction applied to the derivation of quantitative structure–activity relationships. In *Methods and Principles in Medicinal Chemistry,* Vol. 3. *Advanced Computer-Assisted Techniques in Drug Discovery* (van de Waterbeemd, H., Ed.), VCH: Weinheim, 1995, pp. 319–331.
37. Bruneau, P., McElroy, N. R. Log D7.4 modeling using Bayesian regularized neural networks. Assessment and correction of the errors of prediction. *J. Chem. Inf. Model.* 2006, *46*, 1379–1387.
38. Winkler, D., Burden, F. R. Bayesian neural nets for modeling in drug discovery. *Drug Discov. Today* 2004, *2*, 104–111.
39. Palmer, D. S., O'Boyle, N. M., Glen, R. C., Mitchell, J. B. O. Random forest models to predict aqueous solubility. *J. Chem. Inf. Model.* 2007, *47*, 150–158.
40. Rogers, D., Hopfinger, A. J. Application of genetic function approximation to quantitative structure–activity relationships and quantitative structure–property relationships. *J. Chem. Inf. Comput. Sci.* 1994, *34*, 854–866.
41. Clarke, D. E. *Evolutionary Algorithms in Molecular Design*. Wiley-VCH: Weinheim, 2000.
42. van de Waterbeemd, H., El Tayar, N., Carrupt, P. A., Testa, B. Pattern recognition study of QSAR substituent descriptors. *J. Comput.-Aided. Mol. Des.* 1989, *3*, 111–132.
43. Rose, V. S., Croall, I. F., MacFie, H. J. H. An application of unsupervised neural network methodology (Kohonen topology-preserving mapping) to QSAR analysis. *Quant. Struct.-Act. Relat.* 1991, *10*, 6–15.
44. McFarland, J. W., Gans, D. J. On the significance of clusters in the graphical display of structure–activity data. *J. Med. Chem.* 1986, *29*, 505–514.
45. Rose, V. S., Wood, J. Generalized cluster significance analysis and stepwise cluster significance analysis with conditional probabilities. *Quant. Struct.-Act. Relat.* 1998, *17*, 348–356.
46. Kowalski, B. R., Bender, C. F. Pattern recognition. II Linear and non-linear methods for displaying chemical data. *J. Am. Chem. Soc.* 1973, *95*, 686–693.
47. Domine, D., Devillers, J., Chastrette, M. A nonlinear map of substituent constants for selecting test series and deriving structure–activity relationships. I. Aromatic series. *J. Med. Chem.* 1994, *37*, 973–980.
48. Domine, D., Devillers, J., Chastrette, M. A nonlinear map of substituent constants for selecting test series and deriving structure–activity relationships. II. Aliphatic series. *J. Med. Chem.* 1994, *37*, 981–987.
49. Anzali, S., Barnickel, G., Krug, M., Sadowski, J., Wagener, M., Gasteiger, J., Polanski, J. The comparison of geometric and electronic properties of molecular surfaces by neural networks: application to the analysis of corticosteroid-binding globulin activity of steroids. *J. Comput.-Aided Mol. Des.* 1996, *10*, 521–534.
50. Leonard, J. T., Roy, K. On selection of training and test sets for the development of predictive QSAR models. *QSAR Comb. Sci.* 2006, *25*, 235–249.
51. Gleeson, M. P. Plasma protein binding affinity and its relationship to molecular structure: an *in silico* analysis. *J. Med. Chem.* 2007, *50*, 101–112.
52. Wold, S., Eriksson, L. (van de Waterbeemd, H., Ed.) *Methods and Principles in Medicinal Chemistry,* Vol. 2. *Chemometric Methods in Molecular Design*, VCH: Weinheim, 1995, pp. 309–318.
53. Baroni, M., Costantino, G., Cruciani, G., Riganelli, D., Valigi, R., Clementi, S. Generating optimal linear PLS estimations (GOLPE): an advanced chemometric tool for handling 3D-QSAR problems. *Quant. Struct.-Act. Relat.* 1993, *12*, 9–20.

54. Golbraikh, A., Tropsha, A. Beware of q^2!. *J. Mol. Graph. Model.* 2002, *20*, 269–276.
55. Aptula, A. O., Jeliazkova, N. G., Schultz, T. W., Cronin, M. T. D. The better predictive model: high q^2 for the training set or low root mean square error of prediction for the test set?. *QSAR Comb. Sci.* 2005, *24*, 385–396.
56. Wakeling, I. N., Morris, J. J. A test of significance for partial least squares regression. *J. Chemom.* 1993, *7*, 291–304.
57. Tetko, I. V., Tanchuk, V. Y. Application of associative neural networks for prediction of lipophilicity in ALOGPS2.1 program. *J. Chem. Inf. Comput. Sci.* 2002, *42*, 1136–1145.
58. Tetko, I. V., Bruneau, P., Mewes, H. W., Rohrer, D. C., Poda, G. L. Can we estimate the accuracy of ADME-Tox predictions?. *Drug Discov. Today.* 2006, *11*, 700–707.
59. Zhang, S., Golbraikh, A., Oloff, S., Kohn, H., Tropsha, A. A novel automated lazy learning QSAR (ALL-QSAR) approach: method development, applications, and virtual screening of chemical databases using validated ALL-QSAR models. *J. Chem. Inf. Model.* 2006, *46*, 1984–1995.
60. Guha, R., Dutta, D., Jurs, P. C., Chen, T. Local lazy regression: making use of the neighbourhood to improve QSAR predictions. *J. Chem. Inf. Model.* 2006, *46*, 1836–1847.
61. Yuan, H., Wang, Y., Cheng, Y. Local and global quantitative structure–activity relationship modeling and prediction for the baseline toxicity. *J. Chem. Inf. Model.* 2007, *47*, 159–169.
62. De Cerqueira Lima, P., Golbraikh, A., Oloff, S., Xiao, Y., Tropsha, A. Combinatorial QSAR modeling of P-glycoprotein substrates. *J. Chem. Inf. Model.* 2006, *46*, 1245–1254.
63. Arodz, T., Yuen, D. A., Dudek, A. Z. Ensemble of linear models for predicting drug properties. *J. Chem. Inf. Model.* 2006, *46*, 416–423.
64. Rodgers, S., Davis, A. M., Tomkinson, N., Van de Waterbeemd, H. QSAR modeling using automatically updating correction libraries: application to a human plasma protein binding model. *J. Chem. Inf. Model.* 2007, *47*, 2401–2407.
65. Cartmell, J., Enoch, S., Krstajic, D., Leahy, D. E. Automated QSPR through competitive workflow. *J. Comput.-Aided. Mol. Des.* 2005, *19*, 821–833.
66. Carlson, R., Nordahl, A. Exploring organic synthetic experimental procedures. *Top. Curr. Chem.* 1993, *166*, 1–64.
67. Topliss, J. G. Utilization of operational schemes for analog synthesis in drug design. *J. Med. Chem.* 1972, *15*, 1006–1011.
68. Craig, P. N. Interdependence between physical parameters and selection of substituent groups for correlation studies. *J. Med. Chem.* 1971, *14*, 680–684.
69. Hansch, C., Leo, A. *Substituent Constants for Correlation Analysis in Chemistry and Biology*. John Wiley & Sons: New York, 1979.
70. Norinder, U., Högberg, T. PLS-based quantitative structure–activity relationship for substituted benzamides of clebopride type. Application of experimental design in drug design. *Acta Chem. Scand.* 1992, *46*, 363–366.
71. Lewis, R. A. A general method for exploiting QSAR models in lead optimization. *J. Med. Chem.* 2005, *48*, 1638–1648.
72. Brown, N., McKay, B., Gasteiger, J. A novel workflow for the inverse QSPR problem using multiobjective optimization. *J. Comput.-Aided Mol. Des.* 2006, *20*, 333–341.
73. Sheridan, P. P., Hunt, P., Culberson, J. C. Molecular transformations as a way of finding and exploiting consistent local QSAR. *J. Chem. Inform. Model.* 2006, *46*, 180–192.
74. Haubertin, D. Y., Bruneau, P. A database of historically-observed chemical replacements. *J. Chem. Inf. Model.* 2007, *47*, 1294–1302.
75. Leach, A. G., Jones, H. D., Cosgrove, D. A., Kenny, P. W., Ruston, L., MacFaul, P., Wood, J. M., Colclough, N., Law, B. Matched molecular pairs as a guide in the optimization of pharmaceutical properties; a study of aqueous solubility, plasma protein binding and oral exposure. *J. Med. Chem.* 2006, *49*, 6672–6682.
76. Nicolotti, O., Gillet, V., Fleming, P., Green, D. Multi-objective optimisation in quantitative structure–activity relationships: deriving accurate and interpretable QSARs. *J. Med. Chem.* 2002, *43*, 5069–5080.
77. Testa, B., Van de Waterbeemd, H., Folkers, G., Guy, R. (Eds.). *Pharmacokinetic Optimization in Drug Research: Biological, Physicochemical and Computational Strategies*. Wiley-VCH: Weinheim and Zurich, 2001.
78. Livingstone, D. *Data Analysis for Chemists*. Oxford University Press: Oxford, 1995.
79. Hansch, C., Leo, A. *Exploring QSAR. Fundamentals and Applications in Chemistry and Biology*. American Chemical Society: Washington, DC, 1995.
80. Dickins, M., Van de Waterbeemd, H. Simulation models for drug disposition and drug interactions. *Drug Discov. Today: BioSilico* 2004, *2*, 38–45.
81. Ottaviani, G., Martel, S., Carrupt, P. A. *In silico* and *in vitro* filters for the fast estimation of skin permeation and distribution of new chemical entities. *J. Med. Chem.* 2007, *50*, 742–748.
82. Pliska, B., Testa, B., Van de Waterbeemd, H. (Eds). *Lipophilicity in Drug Action and Toxicology*. Wiley-VCH: Weinheim, 1996.
83. Schultz, T. W., Hewitt, M., Netzeva, T. I., Cronin, M. T. D. Assessing applicability domains of toxicological QSARs: definition, confidence in predicted values, and the role of mechanisms of action. *QSAR Comb. Sci.* 2007, *26*, 238–254.
84. Grunewald, G. L., Dahanukar, V. H., Jalluri, R. K., Criscione, K. R. Synthesis, biochemical evaluation and classical and three-dimensional quantitative structure–activity relationship studies of 7-substituted-1,2,3,4-tetrahydroisoquinolines and their relative affinities toward phenylethanolamine *N*-methyltransferase and the α2-adrenoreceptor. *J. Med. Chem.* 1999, *42*, 118–134.
85. Smith, D. A., Jones, B. C., Walker, D. K. Design of drugs involving the concepts and theories of drug metabolism and pharmacokinetics. *Med. Res. Rev.* 1996, *16*, 243–266.
86. Chan, O. H., Stewart, B. H. Physicochemical and drug-delivery considerations for oral drug bioavailability. *Drug Discov. Today* 1996, *1*, 461–473.
87. van de Waterbeemd, H., Smith, D. A., Beaumont, K., Walker, D. K. Property-based design: optimisation of drug absorption and pharmacokinetics. *J. Med. Chem.* 2001, *44*, 1313–1333.
88. Yamashita, F., Hashida, M. *In silico* approaches for predicting ADME properties of drugs. *Drug Metab. Pharmacokin.* 2004, *19*, 327–338.
89. van de Waterbeemd, H., Gifford, E. ADMET *in silico* modeling: towards prediction paradise?. *Nat. Rev. Drug Discov.* 2003, *2*, 192–204.
90. van de Waterbeemd, H. *In silico* models to predict oral absorption. *In ADME/Tox Approaches, Comprehensive Medicinal Chemistry* 2nd Ed (Taylor, J. B., Triggle, D. J., Testa, B., van de Waterbeemd, H., Eds.) Vol. V., Elsevier: Oxford, 2007, pp. 669–697.
91. Zhao, Y. H., Le, J., Abraham, M. H., Hersey, A., Eddershaw, P. J., Luscombe, C. N., Boutina, D., Beck, G., Sherborne, B., Cooper, I., Platts, J. A. Evaluation of human intestinal absorption data and subsequent derivation of a quantitative structure–activity relationship (QSAR) with the Abraham descriptors. *J. Pharm. Sci.* 2001, *90*, 749–784.
92. Burton, J., Ijjaali, I., Barberan, O., Petitet, F., Vercauteren, D. P., Michel., A. Recursive partitioning for the prediction of cytochromes P450 2D6 and 1A2 inhibition: importance of the quality of the dataset. *J. Med. Chem.* 2006, *49*, 6231–6240.
93. Zhao, Y. H., Abraham, M. H., Ibrahim, A., Fish, P. V., Cole, S., Lewis, M. L., De Groot, M., Reynolds, D. P. Predicting penetration across the blood-brain barrier from simple descriptors and fragmentation schemes. *J. Chem. Inf. Model.* 2007, *47*, 170–175.
94. Song, M., Clark, M. Development and evaluation of an *in silico* model for hERG binding. *J. Chem. Inf. Model.* 2006, *46*, 392–400.
95. Coats, E. A., Cordes, H.-P., Kulkarni, V. M., Richter, M., Schaper, K.-J., Wiese, M., Seydel, J. K. Multiple regression and principal component analysis of antibacterial activities of sulfones and sulfonamides in whole cell and cell-free systems of various DDS

sensitive and resistant bacterial strains. *Quant. Struct.-Act. Relat.* 1985, *4*, 99–109.
96. Almerico, A. M., Tutone, M., Lauria, A., Diana, P., Barraja, P., Montalbano, A., Cirrincione, G., Dattolo, G. A multivariate analysis of HIV-1 protease inhibitors and resistance induced by mutation. *J. Chem. Inf. Model.* 2006, *46*, 168–179.
97. Cheeseright, T., Mackey, M., Rose, S., Vinter, A. Molecular field technology applied to virtual screening and finding the bioactive conformation. *Expert Opin. Drug Discov.* 2007, *2*, 131–144.
98. Gorse, D., Lahana, R. Functional diversity of compound libraries. *Curr. Opin. Chem. Biol.* 2000, *4*, 287–294.
99. Gillet, V. J., Nicolotti, O. Evaluation of reactant-based and product-based approaches to the design of combinatorial libraries. *Perspect. Drug Discov. Dev.* 2000, *20*, 265–287.
100. Good, A. C., Lewis, R. A. A new methodology for profiling combinatorial libraries and screening sets: cleaning up the design process with HARPick. *J. Med. Chem.* 1997, *40*, 3926–3936.
101. Martin, E., Wong, A. Sensitivity analysis and other improvements to tailored combinatorial library design. *J. Chem. Inf. Comput. Sci.* 2000, *40*, 215–220.
102. Sheridan, R. P., SanFeliciano, S. G., Kearsley, S. K. Designing targeted libraries with genetic algorithms. *J. Mol. Graph. Model.* 2000, *18*, 320–334.
103. Brown, R. D., Hassan, M., Waldman, M. Combinatorial library design for diversity, cost efficiency and drug-like character. *J. Mol. Graph. Model.* 2000, *18*, 427–437.
104. Rose, S., Stevens, A. Computational design strategies for combinatorial libraries. *Curr. Opin. Chem. Biol.* 2003, *7*, 331–339.
105. Baxter, C. A., Murray, C. W., Waszkowycz, B., Li, J., Sykes, R. A., Bone, R. G. A., Perkins, T. D. J., Wylie, W. New approach to molecular docking and its application to virtual screening of chemical databases. *J. Chem. Inf. Comput. Sci.* 2000, *40*, 254–262.
106. Sadowski, J., Kubinyi, H. A scoring scheme for discriminating between drugs and nondrugs. *J. Med. Chem.* 1998, *41*, 3325–3329.
107. Ajay, A., Walter, W. P., Murko, M. A. Can we learn to distinguish between 'drug-like' and 'nondrug-like' molecules?. *J. Med. Chem.* 1998, *41*, 3314–3324.
108. Nell, P. G., Mundt, S. M. Chemoinformatic tools in high-throughput screening data analyses. In *Methods and Principles in Medicinal Chemistry,* Vol. 35. *High Throughput-Screening in Drug Discovery* (Hüser, J., Ed.), VCH: Weinheim, 2006, pp. 207–257.
109. Lobell, M., Hendrix, M., Hinzen, B., Keldenich, J., Meier, H., Schmeck, C., Schohe-Loop, R., Wunberg, T., Hillisch, A. *In silico* ADMET traffic lights as a tool for the prioritization of HTS hits. *ChemMedChem* 2006, *1*, 1229–1236.
110. Yan, S. F., Asatryan, H., Li, J., Zhou, Y. Novel statistical approach for primary high-throughput screening hit selection. *J. Chem. Inf. Model.* 2005, *45*, 1784–1790.
111. Fischer, H. P., Heyse, S. From targets to leads: the importance of advanced data analysis for decision support in drug discovery. *Curr. Opin. Drug Discov. Dev.* 2005, *8*, 334–346.
112. Kolossov, E., Lemon, A. Medicinal chemistry tools: making sense of HTS data. *Eur. J. Med. Chem.* 2006, *41*, 166–175.
113. Yan, S. F., King, F. J., He, Y., Caldwell, J. S., Zhou, Y. Learning from the data: mining of large high-throughput screening databases. *J. Chem. Inf. Model.* 2006, *46*, 2381–2395.
114. Salt, D. W., Ajmani, S., Chrichton, R., Livingstone, D. J. An improved approximation to the estimation of the critical F values in best subset regression. *J. Chem. Inf. Model.* 2007, *47*, 143–149.
115. Nigsch, F., Bender, A., Van Buuren, B., Tissen, J., Nigsch, E., Mitchell, J. B. O. Melting point prediction employing k-nearest neighbor algorithms and genetic parameter optimization. *J. Chem. Inf. Model.* 2006, *46*, 2412–2422.
116. Huang, J., Ma, G., Muhammad, I., Cheng, Y. Identifying P-glycoprotein substrates using a support vector machine optimized by a particle swarm. *J. Chem. Inf. Model.* 2007, *47*, 1638–1647.
117. Obrezanova, O., Csanyi, G., Gola, J. M. R., Segall, M. S. Gaussian processes: a method for automatic QSAR modeling of ADME properties. *J. Chem. Inf. Model.* 2007, *47*, 1847–1857.

Part V

Spatial Organization, Receptor Mapping and Molecular Modeling

David J. Triggle

Section Editor

Chapter 24

Overview: The Search for Biologically Useful Chemical Space

David J. Triggle

I. INTRODUCTION
II. HOW BIG IS CHEMICAL SPACE?
III. BIOLOGICAL SPACE IS EXTREMELY SMALL
IV. LIMITED BIOLOGICAL SPACE AS AN EFFECTIVE BIOLOGICAL STRATEGY
REFERENCES

Two roads diverged in a yellow wood
And sorry I could not travel both

Then took the other, as just as fair
And having perhaps the better claim.

Robert Frost (1874–1963)

I. INTRODUCTION

The search for and use of biologically active chemicals by man dates at least to the dawn of *Homo sapiens*, and doubtless to his ancestors. One can speculate about why such early human societies used such agents. Likely, anything that could provide relief from the undoubted harshness and danger of life – described by the 17th century English philosopher Thomas Hobbes as, "poor, nasty, brutish, and short" – would be considered desirable. This search for biologically active chemicals continues today, albeit at a more sophisticated and rational level.

The contemporary process of drug discovery has its roots in the late 19th and early 20th centuries with the work of Crum Brown and Fraser, Langley, Ehrlich, Pasteur, Cushny, Fisher, Clark and others who introduced the concepts of structure–activity relationships (SARs), specific binding sites at which drugs interact, molecular complementarity, stereoselectivity of molecular interaction, the quantification of dose–response relationships and the concept of drug selectivity of action. Although the paradigm of drug discovery has changed significantly in the past several decades through the introduction of genomics, combinatorial chemistry and high-throughput screening.

(Figures 24.1 and 24.2) the fundamental challenges remain unchanged – these are the identification and validation of targets for drug action and the discovery of biologically active molecules. These are not, of course, the only issues facing the would-be drug discoverer: pharmacokinetic and toxicological issues remain critical – insoluble and/or toxic agents are scarcely useful therapeutic agents – and these aspects of the drug discovery process are now increasingly considered simultaneously with the manipulation of chemical structure that generates activity and selectivity. However, it remains true that absent a biologically active molecule and a validated target little else matters. Thus, in some respects the medicinal chemist may be thought of as a chemical astronaut – "looking for life in the chemical universe."

This section of "Principles of Medicinal Chemistry" concerns itself primarily with the issues surrounding the exploration of chemical space, including the relationship between chemical and biological space, stereoselectivity of drug interaction, privileged structures, structure–function relationships, selectivity and non-selectivity of drug action

Copyright © 2008, Elsevier Ltd
All rights reserved.

FIGURE 24.1 The two pathways of drug discovery. The classic pathway typically starts with a lead – a natural product, a clinical observation, a phenotypic consequence of a chemical action, etc., – that is screened in a variety of tissue and animal preparations, followed by more detailed study in a pharmacological model of a disease state and ultimately (if successful) to a single clinical candidate. The new pathway identifies targets from gene interrogation and the application of the technologies of combinatorial chemistry and high-throughput screening yields, in principle, multiple candidates.

FIGURE 24.2 The contemporary organization of drug discovery.

and the elucidation of biological space through X-ray crystallography.

II. HOW BIG IS CHEMICAL SPACE?

The theoretically available chemical space is remarkably large, so large in fact that an entirely trivial portion only has been explored or utilized by either nature or man.[1,2] Using, for example, the standard 20-amino acid repertoire and assuming an average protein of 300 residues then there exists more than 10^{390} possible proteins and their collective mass at the combined single molecule level would exceed the mass of the known universe.[1] A similar calculation for small molecule ligands of molecular weight of <500 (approximating small molecule drugs) indicates $>10^{60}$ possible structures.[3] These numbers represent unattainable challenges even for Nature, the world's most competent synthetic chemist. Indeed, it is quite clear that life as we know it on planet Earth has evolved to function within a very small area (or areas) of potential chemical space. Thus, it has been estimated that that the total number of small molecules employed biologically is but a few thousand and the number of proteins between 1,000 and 100,000 only, thus mapping trivial fractions of available chemical space.[4,5] Such conclusions are quite in accord with the (initially) surprisingly small number of genes identified in the human genome and with observations that a minimal genome for a living single cell organism can be as small as approximately 200 genes or less.[6,7]

These arguments do not, of course, mean that life on other planets or on other universes has been subject to the same constraints.[8] Indeed, it is entirely possible that in another universe life employs a different biological code, is not necessarily carbon-based and may be a mirror image of ours in terms of stereoselectivity of molecular interaction. Such a small genome size makes eminently feasible the construction of entirely novel genomes and organisms[9] and with an expanded genetic code[10] to explore currently biologically uncharted areas of chemical space for their significance.

III. BIOLOGICAL SPACE IS EXTREMELY SMALL

These considerations indicate that the amount of chemical space mapped by biologically active compounds is extremely small. Confirmation of this for small molecule ligands that represent the dominant fraction of therapeutically available drugs is provided by an analysis of the structural properties of known drugs.[11,12] In a set of over 5,000 known drugs it was found that approximately 1,200 individual frameworks could be detected, but that 50% of these known drugs could be described by only 32 of these frameworks. Similarly, an analysis of the side chains of these drugs revealed that of some 1,300 different side chains only 20 described approximately 70% of known drugs.

Thus, biological space may be viewed as a very small island in an almost infinitely large universe of chemical space. This view parallels that of our own physical universe where life is likely extremely rare. A further parallel is that just as life may be found in scattered and divergent areas of the universe so may biologically useful space be scattered in the chemical universe.[13] There will be a continuum of space that determines the biological activity of molecules,

FIGURE 24.3 The 1,4-dihydropyridine nucleus as a privileged structure. In addition to classically established antagonist and activator interactions at the L-type of voltage-gated calcium channel appropriately substituted 1,4-dihydropyridines can interact, often with high selectivity, at a variety of other ion channels and receptors as depicted.

but also discrete regions of space that control various ADME (absorption, distribution, metabolism, excretion) properties as well as regions that may control toxicological properties. Thus, chemical space is, like our physical universe, "lumpy" in its arrangement of biological utility. An analysis of the mapping of pharmacological space has been attempted by Hopkins and his colleagues[14] through the integration of SAR data from diverse sources.

Since biologically useful chemical space is but a small component of total chemical space it is not surprising, as already noted[11,12] that a discrete number of basic molecular features are regularly found in biologically active molecules. This is seen very clearly with so-called privileged structures where certain molecular scaffolds such as benzodiazepines[15] and 1,4-dihydropyridines[16] (Figure 24.3) generate, when appropriately decorated with different side-chains, high biological activity at discrete targets.[17] That biological activity resides in a limited number of molecular features also makes possible the design of "multiple ligands" – drugs that exert activity at a number of targets rather than at a single molecular entity – the "magic shotgun" versus the "magic bullet".[18,19] Such existing "dirty" drugs appear to be already useful in the treatment of mood disorders and schizophrenia and would likely also be useful in other disorders such as ischemic stroke where a variety of activated pathways contribute to the pathologic condition.[20]

IV. LIMITED BIOLOGICAL SPACE AS AN EFFECTIVE BIOLOGICAL STRATEGY

At first consideration it might appear that the use of an extremely limited amount of the chemical universe for biological purposes is an inefficient strategy. This is not so, for several reasons. First, the focus on a limited number of bases, amino acids, proteins and small molecules prevents the wasteful scattering of cellular energy across the metabolic landscape. Second, the restriction to limited areas of chemical space makes possible the efficient use of common pathways and molecular arrangements to achieve diverse functions and these pathways and arrangements can be reused if the original function is rendered obsolete

or unnecessary. Third, the selectivity of molecular complementarity, including stereoselectivity, that characterizes biological systems is adequately provided for since as little as one residue change in a protein can render all-or-none changes in biological activity.[21,22]

REFERENCES

1. Dobson, C. M. Chemical space and biology. *Nature* 2004, *432*, 824–828.
2. Weiss, G. A. Exploring the milky way of molecular diversity. Combinatorial chemistry and molecular diversity. *Curr. Opin. Biotechnol.* 2007, *11*, 241–243.
3. Bohacek, R. S., McMartin, C., Guida, W. C. The art and practice of structure-based drug design. Med. Res. Rev. 1996, *16*, 3–50.
4. Goto, S., Okuno, Y., Hattori, M., Nishioka, T., Kanehisa, M. LIGAND: database of chemical compounds and reactions in biological pathways. *Nucleic Acids Res.* 2002, *30*, 402–404.
5. Lander, E. S. *et al*. Initial sequencing and analysis of the human genome. *Nature* 2001, *409*, 806–921.
6. Gil, R., Silva, F. J., Pereto, J., Moya, A. Determination of the core of a minimal bacterial gene set. *Microbiol. Mol. Biol. Rev.* 2004, *68*, 518–537.
7. Forster, A. C., Church, G. M. Towards synthesis of a minimal cell. *Mol. Syst. Biol.* 2006, 1–10.
8. The Limits of Organic Life in Planetary Systems, Committee on the Limits of Organic Life in Planetary Systems, Committee on the Origins and Evolution of Life. National Research Council, Washington, DC, 2007. www.nap.edu/catalog/11919.html
9. Lartigue, C. *et al*. Genome transplantation in bacteria: changing one species to another. Science Express, 2007, June. doi 10.1126/1144622.
10. Wang, L., Xie, J., Schultz, P. G. Expanding the genetic code. *Annu. Rev. Biophys. Biomol. Struct.* 2006, *35*, 225–249.
11. Bemis, G. W., Murcko, M. The properties of known drugs. 1. Molecular frameworks. *J. Med. Chem.* 1996, *39*, 2887–2893.
12. Bemis, G. W., Murcko, M. The properties of known drugs. 2. Side chains. *J. Med. Chem.* 1999, *42*, 5095–5099.
13. Lipinski, C., Hopkins, A. Navigating chemical space for biology and medicine. *Nature* 2004, *432*, 856–861.
14. Paolini, G. V. *et al.* Global mapping of pharmacological space. *Nat. Biotechnol.* 2006, *24*, 805–815.
15. Evans, B. *et al.* Methods for drug discovery: development of potent, selective, orally effective cholecystokinin antagonists. *J. Med. Chem.* 1988, *31*, 2235–2246.
16. Triggle, D. J. The 1,4-dihydropyridine nucleus: a pharmacophoric template. *Mini-Rev. Med. Chem.* 2003, *3*, 215–223.
17. Muller, G. Medicinal chemistry of target family-directed masterkeys. *Drug Discov. Today* 2003, *8*, 5–15.
18. Roth, B. L., Sheffler, D. J., Kroeze, W. K. Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. *Nat. Rev. Drug Discov.* 2004, *3*, 353–359.
19. Morphy, R., Rankovic, Z. Designed multiple ligands: an emerging drug discovery paradigm. *J. Med. Chem.* 2005, *48*, 6523–6543.
20. O'Collins, V. E. *et al*. 1,026 Experimental treatments in acute stroke. *Ann. Neurol.* 2006, *59*, 467–477.
21. Hirst, W. D. *et al.* Differences in the central nervous system distribution and pharmacology of the mouse 5-HT-6 receptor compared with rat and human receptors investigated by radioligand binding, site-directed mutagenesis, and molecular modeling. *Mol. Pharmacol.* 2003, *64*, 1295–1308.
22. Souza, S. C. *et al*. A single arginine residue determines species specificity of the human growth hormone receptor. *Proc. Natl. Acad. Sci. USA* 1995, *92*, 959–963.

Chapter 25

Pharmacological Space

Andrew L. Hopkins

I. WHAT IS PHARMACOLOGICAL SPACE?
II. CHEMICAL SPACE
A. Drug-like space
III. TARGET SPACE
A. Druggability
B. Structure-based druggability
C. Degrees of druggability
D. Druggable genome
IV. CONCLUSIONS
ACKNOWLEDGMENTS
REFERENCES

"When you know you've come up with something completely novel, the medical imperative is to come up with a good use for it. That takes imagination. And the final step is to persuade the patents people that something has actually been discovered."

Paul Janssen

I. WHAT IS PHARMACOLOGICAL SPACE?

The medicinal chemist is faced with an apparent endless series of choices in undertaking a drug discovery project, such as which disease to attempt to design a treatment for, which proteins should be targeted, which assays should be employed, and which compounds should be synthesized and in what order. Sir James Black elegantly outlined six critical requirements for decision-making in a drug discovery project (Black's Rules, Box 25.1). To aid the drug discovery in decision-making, the concept of "pharmacological space" provide a theoretical framework for navigating the apparently infinite number of choices in drug discovery. Pharmacological space attempts to chart the limits of chemical space, targets space and disease space in order to reduce and systematize the search for new drugs in these spaces. Thus charting pharmacological space is an attempt to outline the areas where opportunities for new drugs may lie based on an extrapolation of the knowledge we have gained from our experiments to date. Here in this chapter we shall outline some of the recent theoretical arguments and empirical evidence for navigating target space and chemical space for drug discovery.

BOX 25.1 Black's Rules

Sir James Black's (Nobel Laureate in Physiology and Medicine, 1988) requirements for a drug discovery project:[46]

1. Is the project purged of wishful thinking?
2. Is a chemical starting point identified?
3. Are relevant bioassays available?
4. Will it be possible to confirm laboratory-specificity in humans?
5. Is a clinical condition relevant to the specificity mentioned in point four?
6. Does the project have a champion – someone with the necessary passion, conviction and energy?

II. CHEMICAL SPACE

At the invention of modern drug discovery, Paul Ehrlich screened just over 600 synthetic compounds to discover arsphenamine (Salvarsan),[1] a novel treatment for syphilis. With advances in screening technology, researchers can now routinely test millions of compounds in protein-based

Copyright © 2008, Elsevier Ltd
All rights reserved.

bioassays. Yet even the compound files of the largest pharmaceutical companies – which typically contain $\sim 10^6$ compounds – offer only a cursory examination of all the possible organic compounds that comprise 'chemical space", which even up to a limit of 500 Da molecular weight (MW) per molecules is, for all practical purposes, infinite and limited only by the chemist's imagination. Yet the medicinal chemist's goal is not to explore chemical space for its own sake, but to identify the small, discrete islands of compounds that interact with biological systems from the vast ocean of possible chemical structures. Measured in terms of physicochemical properties and topological descriptors, therapeutically useful compounds – that is, drugs – appear to form clusters in chemical space rather than be evenly distributed. The primary explanation of the distribution of drugs is chemical space as clusters is because historically about eight out of ten approved drugs are incremental improvements on existing classes of drugs and therefore often possess similar chemical structures. On average of the 30 drugs approved each year over the past two decades only five new proteins are targeted by new drugs each year of which less than two targets belong to previously undrugged target classes or protein families.[2] A similar discrete phenomenon is also observed in the chemical scaffolds of known drugs. Whilst chemical space may be apparently limitless the repertoire of molecular framework – the scaffolds of atom connectivity in a molecule – is extremely low. Only 32 molecular frameworks represent the chemical structures of half of all drugs,[3] whilst 73% of all side chains in drugs attached to the molecular frameworks are represented by only 20 side chain groups.[4]

Our experience to date, from a century of medicinal chemistry and thousands of high-throughput screening (HTS) programs suggests that compounds that bind to certain "target classes" – that is, proteins from the same protein family, such as G-protein-coupled receptors (GPCRs) – are clustered together in discrete regions of a chemical space that can be defined by particular chemical descriptors (see Figure 25.1). The figure depicts a cartoon representation of the relationship between the continuum of chemical space (light blue) and the discrete areas of chemical space that are occupied by compounds with specific affinity for biological molecules, such as those from major gene families (shown in yellow, with specific gene families color-coded as proteases (purple), lipophilic GPCRs (blue) and kinases (red), in terms of molecular property descriptors. The independent intersection of compounds with drug-like properties, or (absorption, distribution, metabolism and excretion) "ADME space", is shown in green.

FIGURE 25.1 Graphical representation of property space have been observed for compounds active against individual genes families[5,52,53] and for bio-availability.[6]

A. Drug-like space

A further restriction on chemical space is that a compound not only has to be biologically active, but also contain the desired physicochemical properties to be administered as a drug. A biologically active compound may be too lipophilic to be orally absorbed, too polar to cross the gastrointestinal wall or may have too much vulnerable chemistry functionality that is attacked by liver metabolizing systems and so does not exist intact long enough to generate a useful *in vivo* biological effect. Observations of the characteristics of compounds that are far more likely to yield safe, orally bioavailable medicines has led to the concept of "drug-likeness" to describe compounds that have the potential to be developed into orally administered drugs, which are generally favored owing to their ease of use by patients. Essential to the design of a drug are the physicochemical characteristics of the lead compound. A balance of solubility and polar/hydrophobic properties is crucial for specific routes of absorption and membrane permeabilities and other biological barriers that a drug needs to penetrate to reach the desired site of action, in order to affect the biological equilibrium of a whole organism. The presence of such biological barriers limits the range of molecular properties, and thus the chemical space the medicinal chemists can design within.

The distribution of molecular properties of small molecule, launched drugs, has changed little in the past 20 years, despite changes in the range of indications and targets[5] and despite changes in the underlying properties of investigational compounds (see Box 25.2). Lipinski's seminal analysis of the Derwent World Drug Index introduced the concept of *drug-likeness*: orally administered drugs are far more likely to reside in areas of chemical space defined by a limited range of molecular properties, which have been encapsulated in Lipinski's "Rule–of-five".[6] The analysis by Lipinski *et al.* shows that, historically, 90% of orally absorbed drugs had fewer than five hydrogen-bond donors, less than 10 hydrogen-bond acceptors, MW of less than 500 Da and $\log P$ values (a measure of lipophilicity) of less than five.[6] Since this work, various definitions of, and methods to predict, drug-likeness have been proposed

BOX 25.2 Changes in Medicinal Compound Properties

Over the past 25 years, there has been a steady, inexorable rise in the median molecular weight of reported medicinal chemistry compounds[5,22] (Figure 25.2a). Comparing 5 year averages from 1986–1990 to those of 1999–2003, the median MW of all reported medicinal chemistry compounds in the literature rose to 68 Da (20%), from 354 to 422 Da, respectively. Interestingly, this growth is also reflected in the increase of the median MW of disclosed ligands for several gene families. For example, compounds binding to aminergic GPCRs have increased in MW by around 56 Da, from 337 to 393 Da between the two 5-year periods. No significant increase in mean or median potency is observed in the data to explain the increase in MW although it is suspected that with the rise of the availability of recombinant proteins assay, there is been an increase in selectivity driving this increase. The changing target portfolio of the industry is also a significant driver in the increase in MW of new lead compounds and investigational drugs. Even so, this rise in MW contrasts with the much slower rise in the mean MW of approved drugs of only 2.2 Da per launch year over the same period. In contrast, there is a steady decline in MW through each subsequent stage of clinical development and increase in the proportion of compounds that are rule-of-five compliant [5,47–51] (Figure 25.2b).

FIGURE 25.2 (a) Increase in MW over time of published medicinal compounds. (b) Mean MW through clinical development (*Source*: Investigational Drugs Database). Similar results found by Blake [50] and Wenlock *et al.*[51]. The increase in MW between discovery compounds (leads) and preclinical compounds (candidates) has been studied in detail by Oprea *et al.*[54].

in the literature,[7–18] but the consensus is that drug-likeness is defined by a range of molecular properties and descriptors that can discriminate between drugs and non-drugs for such characteristics as oral absorption, aqueous solubility and permeability – what could be considered as ADME space. Computational property filters can be used to rapidly assess the drug-likeness of chemical libraries *in silico* before purchase or synthesis.

In recent times, toxicity has replaced poor drug metabolism properties as the major cause of failure in the early clinical phase of drug development. Toxicity may derive from modulation of the biological target or from the compound's structure or physicochemical properties. Extrapolating from large-scale gene knock-out studies in model organisms, approximately a fifth of human proteins are thought to be essential, at least in embryo development. Despite this, only a handful of targets are considered routinely as general "anti-targets", such as the HERG ion channel, the binding of which should be avoided. In some circumstance inhibition of a vital metabolic process may be beneficial in specific tissue types (such as statins predominately targeting liver HMG-CoA reductase) or in specific circumstances where disease-specific differentiation is possible (such as bortezomib inhibition the proteasome). The most common tactic medicinal chemists apply to decrease the chance of isosyncratic toxicity is to avoid compounds that contain specific chemical groups (toxicophores) that have been associated with toxic effects, or compounds that interact covalently with protein targets, which suffer from problems such as lack of specificity and unsuitability for optimization by medicinal chemistry techniques.[19–21] Although several well-known drugs, such as omeprazole and β-lactamase inhibitors, are known to act *via* irreversible mechanisms, medicinal chemists and toxicologists are becoming more wary of incorporating reactive groups within tools or drugs that can form covalent bonds to the target and/or other proteins. Recent studies have also highlighted the relationship between high lipophilicity and the increased chance of toxicity *via* the increased target promiscuity of lipophilic molecules.[22]. To avoid promiscuous off-targets effects Leeson *et al.*[22] recommend medicinal chemists should be wary synthesizing compounds with $c\,\mathrm{Log}\,P > 3.5$.[22]

III. TARGET SPACE

Drugs act by binding to and modifying the function of biological macromolecules, predominately proteins, but also DNA, RNA, carbohydrates and phospholipids membranes. Drug target space, unlike chemical space is limited, discrete and with the availability of genome sequences is capable of being fully defined.[2,23–28] Of the 1,357 unique drugs approved by the US Food and Drug Administration (FDA), 1,204 are classed as "small-molecule drugs" and 166 are classed as "biological" drugs. Of the 1,204 small-molecule drugs, 803 can be administered orally, 421 can be dosed parenterally and 275 can be used as topical agents (including buccal, rectal, inhalational and other such routes of administration for "topical" agents).[2] At least 192 (16%) of small-molecule drugs are administered as prodrugs. Of the 1,204 small-molecule drugs 885 pass the rule-of-five test; of these, 619 (70%) are dosed orally, whereas 159 (20%) of orally dosed drugs fail at least one of the rule-of-five parameters.[2] For all the current drugs for which we can determine the mode of action, 324 distinct molecular drug targets have been identified, of these, 266 are human proteins, and the remainder are present in bacterial, viral, fungal or other pathogenic organism targets. Small-molecule drugs modulate 248 proteins, of which 207 are targets encoded by the human genome. Oral small-molecule drugs target 227 molecular targets, of which 186 are human targets[2] (Table 25.1).

The pharmacopoeia of molecular drug targets represents approximately only 1% of the total number of predicted genes in the human genome. Medicinal chemistry has, however, explored a greater number of targets for which chemical tools, lead compounds and investigational drugs have been discovered. Paolini *et al*. have attempted by the large-scale integration of proprietary and published screening data to identify the number of unique molecular targets for which chemical tools, leads or drugs have been discovered.[5] The global survey of the data from Pfizer, Warner-Lambert and Pharmacia integrated with a large body of medicinal

TABLE 25.1 Molecular Targets of FDA Approved Drugs (2006).

Class of drug target	Species	Number of molecular targets
Targets of approved drugs	Pathogen and human	324
Human genome targets of approved drugs	Human	266
Targets of approved small-molecule drugs	Pathogen and human	248
Targets of approved small-molecule drugs	Human	207
Targets of approved oral small-molecule drugs	Pathogen and human	227
Targets of approved oral small-molecule drugs	Human	186
Targets of approved therapeutic antibodies	Human	15
Targets of approved biologicals	Pathogen and human	76

chemistry SAR results published in the literature identified for over 1,300 proteins from 55 organisms, with biologically active chemical matter. These include a non-redundant list of 836 genes in the human genome for which small-molecule chemical tools have been discovered, of which 727 human targets have at least one compound with binding affinity below 10 μM compliant with Lipinski's "rule-of-five" criteria for oral drug absorption[6] and 529 human targets have at least one "rule-of-five" compound below 100 nM (Table 25.2).

A. Druggability

Development of the ideas of drug-likeness of compounds has lead to the parallel concept of druggability of proteins. The druggability hypothesis proposes that the likelihood of discovering a "drug-like" ligand for a protein can be assessed *a priori* derives from biophysical basis of molecular recognition.[25,29–31] In short, the binding sites on biological molecules must be complementary in terms of volume, topology and physicochemical properties with their ligands, then only certain binding sites on putative drug targets are

TABLE 25.2 Human Proteins with Identified Chemical Tools, by Gene Family, Lipinski rule-of-five Properties and Potency[5].

Gene taxonomy	All targets at <10 μM*	Human targets at <10 μM	Human targets at <10 μM	Human targets at <10 μM Ro5† $n > 1$	Human targets at <10 μM	Human targets at <100 nM Ro5 $n > 1$
Protein kinases	131	105	99	98	86	83
Peptide GPCRs	10	63	59	59	55	42
Transferases	75	49	42	36	33	24
Aminergic GPCRs	72	35	35	35	35	35
GPCRs class A – others	68	44	44	40	38	32
Oxidoreductases	68	40	36	38	29	25
Metalloproteases	63	44	41	41	36	35
Hydrolases	56	36	29	30	25	21
Ion channels – ligand gated	55	29	28	24	25	22
Nuclear hormone receptors	47	24	24	22	23	19
Serine proteases	37	30	30	28	29	21
Ion channels – others	24	18	16	16	13	11
PDEs	23	19	19	19	18	18
Cysteine proteases	20	16	16	14	14	13
GPCRs class C	20	10	10	10	6	6
Kinases – others	16	12	9	11	6	5
GPCRs class B	14	7	7	4	7	3
Aspartyl proteases	10	7	7	4	6	4
Others	241	139	119	108	83	63
Enzymes – others	156	109	97	90	69	47
Total	**1,306**	**836**	**767**	**727**	**639**	**529**

*Molecular targets for 1357 unique FDA approved drugs (including New Chemical Entities and New Biological Entities) as derived by Overington *et al.* (Nature Reviews Drug Discovery, (2006) 5(12):993-6) from a normalised database analysis of the FDA Orange Book. Following a comprehensive analysis of the literature unique molecular targets could be assigned to a1065 FDA drugs as the mechanisms of action.

†Human proteins with known small molecule chemical tools, lead or drugs as determined by Paolini *et al.* (Nature) survey of the medicinal chemistry literature and corporate databases. Biologically active compounds are diffined as this with a binding affinity below 10mM against the molecular target. Lipinski's 'rule-of-five' criteria15 of fewer than 5 H-bond donors, fewer than 10 H-bond acceptors, MW below 0.5 kDa and clog P below 5. Compounds that fail Lipinski's criteria are more likely to show poor absorption or permeation because such compounds are unlikely to show good oral bioavailability. Over 761 proteins have more than one compound reported active.

compatible with the high affinity binding of compounds with "drug-like" properties.[25] The binding energy (ΔG) of a ligand to a molecular target such as a protein, RNA, DNA or carbohydrate is defined as $\Delta G = -RT \ln K_d$. The van der Waals and entropy components are considered to be the predominate contributors to the binding energy by the burying of hydrophobic surfaces and the liberation of ordered waters. A low affinity ligand of $K_i = 1\,\mu M$ affinity equates to a binding energy of $\Delta G = -8.4$ kcal/mol. A high affinity drug molecule binding with an affinity of $K_i = 10$ nM requires a binding energy (ΔG) of -11 kcal/mol. Thus, 1.36 kcal/mol of binding energy is equivalent to a 10-fold increase in potency. The binding energy potential of a ligand is approximately proportional to the available surface area and its properties, assuming there are no strong covalent or ionic interactions between the ligand and the protein. For small-molecule (less than 500 Da) drug-like molecules a near linear correlation between molecular surface area and MW has been observed. The van der Waals attractions between atoms, and the hydrophobic effect from the displacement of water contributes approximately 0.03 kcal/mol/Å^2.[32] Thus assuming there are no strong ionic interactions between the protein and the ligand, a ligand with a 10 nM dissociation constant would be required to bury 370Å^2 of hydrophobic surface area. The contribution of the hydrophobic surface to binding energy is demonstrated by the medicinal chemistry phenomenon of the "magic methyl," where a single methyl group placed in the correct position, can increase ligand affinity by 10-fold. The accessible hydrophobic surface area of a methyl group is approximately 46 Å^2 (if one assumes all of the hydrophobic surface area is encapsulated by the protein binding site and thus forms full contact with the protein) with a hydrophobic effect of 0.03 kcal/mol/Å^2 approximately equal to 1.36 kcal/mol, equivalent to the observed 10-fold affinity increase: approximately the maximal affinity per non-hydrogen atom.[33] In addition to the predominantly hydrophobic contribution to the binding of many drugs, ionic interactions, such as those found in zinc proteases (such as ACE inhibitors) contribute to the binding energy. The attraction of complementary polar groups contributes up to 0.1 kcal/mol/Å^2, with ionic salt bridge approximately 3 times greater, allowing low MW compounds to bind strongly. Unlike hydrophobic interactions complementary polar interactions are dependent on the correct geometry.

B. Structure-based druggability

The physicochemical and energetic constraints of molecular recognition leads to the conclusion that a drug target needs a "pocket," whether the pocket is predefined by the protein's architecture or formed on binding by allosteric mechanisms. Druggable cavities on proteins that are complementary with the high affinity binding of non-covalent, small molecule, "rule-of-five" compliant ligands (whose binding energy is predominantly driven the entropic, hydrophobic and van der Waals contributions) are predominately apolar cavities of 400–1,000 Å^3, where over 65% of the pocket is buried or encapsulated, with an accessible hydrophobic surface area of at least 350 Å^2.[29] Encapsulated cavities maximize the ratio of the surface area to the volume and are thus capable of binding low MW compounds with high affinities.

The hypothesis that the physicochemical properties of cavities on protein structures can be analyzed *a priori* to predict the druggability of a protein has been developed further into automatic algorithms to assess the protein structures in the Protein Data Bank (PDB) and the stream of novel structures determined by the structural genomics initiative.[3130] Empirical druggability predictions have been explored experimentally using heteronuclear NMR to identify and characterize the binding surfaces on protein at Abbott by screening ~10,000 low MW molecules (average MW 220, average c Log P 1.5).[30] In a small sample of 33 proteins, the screening results reveal that about 90% of the ligands bind to sites known to be small-molecule ligand binding sites. Only 3 out of the 23 proteins were distinct, uncompetitive new binding sites discovered. In the relatively small sample of proteins studied, Hajduk *et al.* note a high correlation between experimental NMR hit rates and the ability to find high affinity ligands. From the experimental screening hit rates, Hajduk *et al.* constructed a simple model that included physicochemical property descriptors such as cavity dimensions, surface complexity, polar and apolar surface area that accurately predicts the experimental screening hit rates with an R^2 of 0.72, an adjusted R^2 of 0.65.

A decision tree approach to assessing the druggability of protein structure has been developed by Al-Lazikani and Overington.[34] A range of physicochemical properties of the identified binding sites and cavities were calculated from the protein structures including volume, depth, curvature, accessibility, hydrophobic surface area and polar surface area. The algorithm was trained set against a test set of 400 protein complexes binding small molecule, rule-of-five compliant ligands. From this analysis a decision tree was derived to predict the druggability of a binding site or cavity from calculated physicochemical properties. The decision tree predicts whether a cavity is druggable within the statistical confidence of the tree. A success rate of 91% when predicting druggability on the protein drug targets has been claimed for this approach.[34] The method requires either an experimentally derived structure or a high-quality homology model. Ideally, because of the inherent flexibility of many protein–ligand binding sites, a sample of multiple conformations is preferred. The decision tree method was applied to the entire PDB (December 2004 release). Following a cleanup process, 27,409 files were suitable for analysis, further classified into 76,322 structural domains using SCOP[35] of which 28% (21,522) were found to have at least one site predicated to have some degree of druggability. From this

analysis a non-redundant set of 427 human proteins were predicted to contain a druggable binding site of which 281 of these proteins having no prior known compounds or drugs developed against them. In a similar analysis Hudjuk *et al.* calculated the druggability of 1,000 non-redundant human proteins derived from the PDB, of which 35% of entries contain at least one site predicted to be highly druggable, sightly higher, but comparable with Al-Lazikani's prediction. A third approach to structure-based druggability has been proposed by Cheng *et al.*[31] who developed the topical and surface area arguments of druggability, as outlined above, into a automated algorithm for analyzing binding sites on protein structures. Assuming a hydrophobic binding energy of de-solvated flat surface inside a binding site where a drug-like compound of 500 Da MW ligand is assumed from observation to correlate 300 Å^2 solvent accessible surface area. A limitation of this method is that drugs that exploit strong ionic interactions such as those targeting angiotensin converting enzyme (ACE) or HIV-1 integrase are predicted to have a low druggability due to the lack of hydrophobic surface area in the binding site.

C. Degrees of druggability

Distinct differences in the distribution of molecular properties between sets of compounds active against different gene families are observed (Figure 25.3). The relationship between target class and the physicochemical properties of ligands has been explored by calculating a set of physicochemical descriptors of hundreds of thousands of biologically active compounds, across over a thousand proteins where the protein sequences assigned to each of the pharmacological targets were classified into gene families. For example, ligands for the nuclear hormone receptors are significantly most lipophilic, as measured by *c* Log *P*, mirroring the properties of steroids. In comparison the mean MW of ligands binding to aminergic GPCRs is close to the mean MW of approved drugs (383 Da, St. Dev. = 155 Da), while the mean MW of peptide GPCR ligands is greater at close to Lipinski's "rule-of-five" limited of 500 Da.

By linking predicted druggable targets to orthologs and homologs with known chemical matter in a structure–activity database, the likely physicochemical properties of potential ligands can be assessed. This premise is based on the assumption that protein targets that are closely related in sequence space are closely related in chemical space. Analysis of the diversity of the physicochemical properties of ligands for a protein family supports this general assumption. Prediction of the likely physicochemical properties of ligands for a novel drug targets can be used to assess their drug-like properties and attractiveness for drug discovery programs. Lipinski's rule-of-five is commonly used as a metric to assess drug-like properties, however Lipinski's parameters ($N + O < 10$, $\log P < 5$, MW <500, H-bond donors <5) do not allow for a continuum of probabilities of compound properties to be assessed.

The ligand efficiency, or binding energy per atom (Δg) of a compound can be calculated by converting the K_d into the free energy of binding (equation (25.1)) at 300 K and dividing by the number of "heavy" (i.e. non-hydrogen atoms) atoms (equation (25.2)):

Free energy of ligand binding:

$$\Delta G = -RT \ln K_d \tag{25.1}$$

Binding energy per atom (ligand efficiency):

$$\Delta g = \frac{\Delta G}{N_{\text{non-hydrogen atoms}}} \tag{25.2}$$

The logarithmic relationship between free energy of binding and dissociation constant potency means that every ΔG change of -1.36 kcal/mol results in a 10-fold change in potency. Kuntz *et al.* surveyed the dissociation or IC_{50} values of ~150 ligand complexes and concluded that the maximum affinity per atom for organic compounds is –1.5 kcal/mol/non-hydrogen atom. The medicinal chemistry phenomenon of "magic methyls", the addition of a single methyl group increasing potency by 10-fold is explained in terms of the maximum achievable from burying the surface area of a single "heavy" atom. The vast majority of medicinal chemistry compounds have efficiencies far below the observed maximal affinity per atom. A simple calculation can define the lowest limit of acceptable ligand efficiency in a typical pharmaceutical project where we wish to obtain a compound with a potency of 10 nM and an upper MW of 500:

- 500 MW compounds contain on average 38 non-hydrogen atoms.
- 10 nM binding constant = 10.99 kcal/mol.
- Therefore a 500 MW compound with a binding constant of 10 nM compound possesses a ligand efficiency of 0.29 kcal/mol/non-H atom.

Small differences in ligand efficiency (Δg) may have large consequences for the type of compounds that may be possible in a chemical series or against a particular target. For example:

- Compound with a $\Delta g = -0.27$ kcal/mol/non-H atom requires 41 atoms (541 MW) to bind with $K_d = 10$ nM.
- Compound with a $\Delta g = -0.36$ kcal/mol/non-H atom requires only 30 atoms (405 MW) to bind with $K_d = 10$ nM.

Potency is an important criteria for assessing leads (or "hits" discovered in HTS), however potency alone is often a false prophet. Indeed the screening parameters, reagent concentrations and false positive filters used make the detection of weak, low MW leads unlikely in many HTS. The bias of the HTS toward high MW compounds has often confounded further optimization as increase in potency often track increase in MW and result in compounds falling

MW	Am	Pep	Lip	Kin	PDE	Prot	Nuc
Mean	368	455	427	381	412	484	407
S.dev.	87	98	76	104	101	150	103

c log *P*	Am	Pep	Lip	Kin	PDE	Prot	Nuc
Mean	3.8	4.6	4.3	3.8	3.4	3.5	5.2
S.dev.	1.7	1.8	1.8	1.9	1.6	2	1.8

PSA	Am	Pep	Lip	Kin	PDE	Prot	Nuc
Mean	53.3	71.4	89.8	90.2	92.3	126.7	67.6
S.dev.	28.7	40.8	28	38.8	38.1	44.6	28.4

% Rot	Am	Pep	Lip	Kin	PDE	Prot	Nuc
Mean	17.8	23.2	27.7	19.4	20.5	30.1	19.4
S.dev.	7.7	7.3	11.8	9	7.2	12	10.1

# Acceptors	Am	Pep	Lip	Kin	PDE	Prot	Nuc
Mean	3	3.7	3.8	3.8	4.4	5	3.2
S.dev.	1.5	1.6	1.3	1.6	1.8	2	1.4

# Donors	Am	Pep	Lip	Kin	PDE	Prot	Nuc
Mean	1.2	1.5	1.3	1.9	1.4	3.2	1.4
S.dev.	1.1	1.7	0.8	1.3	1.2	1.8	0.9

FIGURE 25.3 Physicochemical property distributions of approximately 75,000 biologically active ligands found in Pfizer's screening data as a function of target type. *Source*: Data courtesy of Barker, Snarey, Groom and Hopkins.

outside of the "rule-of-five" profile for acceptable absorption and permeability properties.[3,10] Scaffold and lead series selection could be aided by considering a parameter that "normalizes" the potency of a lead, with respect to MW, to allow comparisons between different series and scaffolds. Indeed small compounds with low molecular complexity are predicted to have an improved probability of binding to the target of interests.[36]

Some chemical series may exploit a specific binding site more successfully than others, yet the fundamental physiochemical nature of the target site defines the upper limits of the binder energy available. Thus a quantifiable value of the "druggability" of a particular target or gene family, can be derived from the mean ligand efficiency of active compounds. An example of the use of ligand efficiency as a measure of the "degree of druggability" of a drug target

TABLE 25.3 Distribution of Ligand Efficiencies against GPCR Sub-families.

Gene family (ligand sub-family)	Mean ligand efficiency[a]	Mean efficiency of top 10%[b]	Mean MW of top 10%	Mean IC_{50} (nM) of top 10%
GPCR (metabotropic class)	0.40	0.48	281	1,384
GPCR (aminergic)	0.36	0.44	364	191
GPCR (proteins)	0.29	0.34	449	643
GPCR (small peptides)	0.28	0.33	486	99
GPCR (peptide)	0.28	0.32	471	424
GPCR (secretin class)	0.28	0.32	468	279
GPCR (amidated peptides)	0.28	0.31	490	291
GPCR (cyclic peptides)	0.28	0.29	467	365
GPCR (lipophilics)	0.27	0.33	437	323
GPCR (nucleotide)	0.25	0.32	536	10

Note: This data is derived from over 31,000 compounds. Each sub-family is represented by between 3 and 29 targets.
[a]kcal/mol/non-hydrogen atom.
[b]Mean of upper decile of compounds ranked by potency.

is illustrated by analyzing ligands of the various sub-families of GPCRs. Over 31,000 compounds with recorded activity, measurements against members of the GPCR class of targets were analyzed in Pfizer's chemogenomics database.[5] Within this set of compounds, we observe significant differences in the MW and ligand efficiencies against the distinct GPCR sub-families (Table 25.3). Aminergic GPCR compounds tend to be significantly smaller and have higher ligand efficiencies than compounds that are active against peptide-binding GPCRs. GPCR sub-families differ in terms of medicinal chemistry tractability. Peptide-binding GPCR targets, we note with maximum ligand efficiencies, are around 0.29 kcal/mol/non-H atom. This implies that one would need a compound with a MW of around 500 to achieve potency in the nanomolar range. Analysis across a range of targets suggests ligand efficiency is a useful metric for assessing target druggability[5] as well as attractive, efficient lead compounds.

D. Druggable genome

The knowledge of which proteins medicinal chemistry has developed drugs, leads and tool compound against can be used to infer the subset of the proteins expressed by the human genome that have a high probability of being "druggable", that is capable of binding drug-like small molecules with high affinity. The first systematic estimate of the number of druggable proteins – the "druggable genome" – following the publication of the draft human genome[37,38] was based on a search for membership of an extensive list of druggable gene families.[25] Gene family-based analysis assumes that the sequence and functional similarities underlies a conservation of binding site architecture between protein family members. Thus, the explicit assumption being that if one member of a gene family is modulated by a drug molecule, other members of the druggable protein domain family are likely to also be able to bind a compound with similar physicochemical properties. Thus, analysis based on druggable protein families or domains are likely to over estimate the number of druggable targets. Following the construction of a drug target sequence database of 399 targets of approved and experimental drugs and leads, 376 sequences could be assigned to 130 drug-binding domains as captured by their InterPro domain annotation. Of these, 130 InterPro domains have orthologs present in the human proteome. At the time of the initial draft of the human genome[37,38] 3,051 genes were identified as belonging to the 130 druggable protein domains and thus predicted to encoded proteins that are inferred to bind a drug-like molecules. Further refinements of the initial druggable genome analysis have been published[39,40] reflecting how the number of predicted protein expressing genes in the human genome has been modified since the initial draft. Orth *et al.* estimate that there are 3,080 genes belonging the druggable genome with over 2,950 druggable gene sequences in public databases in 2004 based on an estimate of the InterPro domain assignments of druggable gene families.[40] An analysis conducted by Russ and Lampel of the 130 druggable protein domains using InterPro and PFAM on the final assembly of the human genome.[3941] Overall the PFAM protein domain annotation predicted fewer false positives than the InterPro classification used. When corrected for the overestimate of olfactory and taste GPCRs the author identify, again, 3,050 druggable genes from the previously defined set of druggable protein domains,[25] but with some significant changes within individual gene families. Using more stringent predictions

for enzyme proteases and other sub-families a conservative estimate of around approximately 2,200 druggable genes are identified.[39]

In order to expand the homology analysis methodology for identifying which targets the survey include that most recent catalogs of known biological targets of drugs and lead compounds. Al-Lazikani and Overington have conducted the most extensive analysis, to date, on identifying the druggable genome, based on the homology to chemically tractable drug targets.[34] Using the BLAST sequence alignment algorithm to search each of the sequences against the human genome, over 900 distinct genes are identified that show close homology to 200 human proteins of approved small molecular drugs,[2] at a cut off of 30% sequence identity and *E*-value less than or equal to 10^{25}. Expanding the analysis to include human proteins from the small molecule chemical leads as published in the medicinal chemistry journal data (i.e. *J. Med. Chem.* 1980–2004, *Bioorg. Med. Chem. Letts.* 1990–2004), a set of 1,155 protein targets known with at least one drug or lead compound with a binding affinity below 10 μM, 707 of which are human molecular targets. BLAST sequence analysis of this database of medicinal chemistry literature[5,34] identified 2,921 protein sequences, within the same sequence identify cut-offs, which are predicted to be druggable proteins expressed by the human genome.

The distribution of the population size of gene families in the druggable genome follows a power law (Figure 25.4) with the top five target classes of GPCRs, protein kinases, proteases, transporters and ion channels accounting for 54% of identified sequences in the druggable genome. The remaining 46% consists of a long tail of over 130 small gene families and singleton enzymes. Thus, many of the new leads that appear in the literature are targeting new members of existing large gene families such as protein kinases and GPCRs do not necessarily increase the size of the predicted druggabe genome. At present novel, potent chemical tools and lead compounds are reported in the

FIGURE 25.4 (a) Gene family distributions of human druggable genome[34]. (b) Population distribution of gene families size in the human genome as represented by a plot of top 500 InterPro families (http://www.ensembl.org/Homo_sapiens/interpro/IPtop500.html). (c) Distribution of gene families size in the human genome as represented as a logarithm scale to illustrate the power of law distribution of gene family population size.

peer-reviewed literature for 80–100 new molecular targets each year.[5] No doubt, many more are only disclosed only in patents. The increase in the rate of discovery of chemical tools for new targets doubled from an average of 30 new targets with leads being disclosed in the 1980s to an average of 60 new targets per year in the 1990s. However, in comparison, an average of four new targets, for first-in-class drugs, have reached the market each year during the 1990s.[5] As those new leads that target new families and proteins only increase the size of the druggable genome in small incremental fashion as most gene families are small. For example, the discovery that imiquimod acts *via* its agonist binding to toll-like receptor 7[42] suggests that the other members of the TLR gene family may be druggable with small molecule compounds, yet this new, potential druggable gene family introduced only 10 new members to the druggable genome.

IV. CONCLUSIONS

Sir James Black once remarked that, "the most fruitful basis for the discovery of a new drug is to start with an old drug." Indeed, approximately 80% of new drugs approved for marketing over the past decade by the FDA have been incremental improvements upon existing chemical designs.[2] Whilst incremental innovations have been challenged by critics of the pharmaceutical industry, it must be remembered that no two drugs have exactly the same biological profile. A range of drugs in the same class expanded the treatment options available to clinicians, as patients often respond differently to different drugs. Incremental innovations, however, illustrated starkly how important *a priori* knowledge of drug targets and chemical structure are the medicinal chemist in design new drugs. However, *a priori* information can be used in the hunt for innovative new medicines and new classes which can be used in all stages of medicinal chemistry design, from selection of targets to identifying lead chemical matter and in the choice of chemical design-modification and substituent replacement.[43,44] Analysis of the properties of drugs reveals, despite changes in chemical structures and the targets of drugs, the overall distribution of their physical properties, has changed little over the past 25 years, leading to concepts such as "drug-likeness" and the rule of five.[5,16,22] Knowledge of the shared experience of compound library screening and the elucidation of X-ray structures of protein–ligand complexes lead to the hypothesis of the druggability of a protein is a function of the physicochemical properties and topology of a binding site. Whilst concepts like drug like-ness and druggability are often criticized for limiting options of the drug hunter, these concepts have been derived from empirical observations and fundamental physical principles of molecular recognition. Given the fact that chemistry space is practically infinite, drug discovery resources are finite, and cost and time to market equate prolonged patient suffering; medicinal chemistry is wise to learn from its previous endeavors to attempt to chart areas of chemical space which may be rich in pharmacological compounds. We started this chapter with a quote from the most successful drug discoverers, Dr. Paul Janssen. In many ways Janssen's drug discovery strategy demonstrated the power of using knowledge of pharmacological space in the search from drugs. Janssen's produced a large number of new medicines by the careful exploration of the pharmacological space and medicinal applicability around a narrow seam of chemical space centred on piperidine derivatives.[45] Within this relatively narrow area of chemical space Janssen *et al.* explored the wider range of molecular targets and disease indications. This successful strategy of exploring pharmacological space in its widest sense is in contrast to drug discovery, as it usual conducted today, which focuses on the narrow spectrum of a single disease or molecular target, but screens a wide sample of chemical space.

ACKNOWLEDGMENTS

I would like to thank Gaia Paolini, Colin Groom, Carolyn Barker, Mike Snarey, Bissan Al-Lazikani and John Overington for all their efforts to understand pharmacological space.

REFERENCES

1. Sneader, W. *Drug Prototypes and their Exploitation*. John Wiley & Sons: London, 1996.
2. Overington, J. P., Al-Lazikani, B., Hopkins, A. L. How many drug targets are there?. *Nat. Rev. Drug Discov.* 2006, *5*(12), 993–996.
3. Bemis, G. W., Murcko, M. A. The properties of known drugs. 1. Molecular frameworks. *J. Med. Chem.* 1996, *39*(15), 2887–2893.
4. Bemis, G. W., Murcko, M. A. Properties of known drugs. 2. Side chains. *J. Med. Chem.* 1999, *42*(25), 5095–5099.
5. Paolini, G. V., Shapland, R. H. B., van Hoorn, W. P., Mason, J. S., Hopkins, A. L. Global mapping of pharmacological space. *Nat. Biotechnol.* 2006, *24*(7), 806–815.
6. Lipinski, C. A., Lombardo, F., Dominy, B. W., Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 1997, *23*, 3–25.
7. Ajay, A., Walters, W. P., Murcko, M. A. Can we learn to distinguish between "drug-like" and "nondrug-like" molecules? *J. Med. Chem.* 1998, *41*(18), 3314–3324.
8. Wang, J., Ramnarayan, K. Towards designing drug-like libraries: a novel computational approach for prediction of drug feasibility of compounds. *J. Comb. Chem.* 1999, *1*(6), 524–533.
9. Walters Ajay, W. P., Murcko, M. A. Recognizing molecules with drug-like properties. *Curr. Opin. Chem. Biol.* 1999, *3*(4), 384–387.
10. Walters, W. P., Murcko, M. A. Prediction of 'drug-likeness'. *Adv. Drug Deliv. Rev.* 2002, *54*(3), 255–271.
11. Lipinski, C. A. Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharmacol. Toxicol. Meth.* 2000, *44*(1), 3–25.
12. Podlogar, B. L., Muegge, I., Brice, L. J. Computational methods to estimate drug development parameters. *Curr. Opin. Drug Discov. Dev.* 2001, *4*(1), 102–109.

13. Muegge, I. Selection criteria for drug-like compounds. *Med. Res. Rev.* 2003, *23*(3), 302–321.
14. Muegge, I., Heald, S. L., Brittelli, D. Simple selection criteria for drug-like chemical matter. *J. Med. Chem.* 2001, *44*(12), 1841–1846.
15. Veber, D. F., Johnson, S. R., Cheng, H. Y., Smith, B. R., Ward, K. W., Kopple, K. Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* 2002, *45*(12), 2615–2623.
16. Proudfoot, J. R. Drugs, leads, and drug-likeness: an analysis of some recently launched drugs. *Bioorg. Med. Chem. Lett.* 2002, *12*(12), 1647–1650.
17. Egan, W. J., Walters, W. P., Murcko, M. A. Guiding molecules towards drug-likeness. *Curr. Opin. Drug Discov. Dev.* 2002, *5*(4), 540–549.
18. Lajiness, M. S., Vieth, M., Erickson, J. Molecular properties that influence oral drug-like behavior. *Curr. Opin. Drug Discov. Dev.* 2004, *7*(4), 470–477.
19. Williams, D. P., Naisbitt, D. J. Toxicophores: groups and metabolic routes associated with increased safety risk. *Curr. Opin. Drug Discov. Dev.* 2002, *5*, 104–115.
20. Willams, D. P., Park, K. Idiosyncratic toxicity: the role of toxicophores and bioactivation. *Drug Discov. Today* 2003, *8*(22), 1044–1050.
21. Hakimelahi, G. H., Khodarahmi, G. A. The identifications of toxicophores for the prediction of mutagenicity, hepatotoxicity and cardiotoxicity. *J. Iran. Chem. Soc.* 2005, *2*(4), 244–267.
22. Leeson, P. D., Springthorpe, B. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat. Rev. Drug Discov.* 2007, *6*(11), 881–890.
23. Drews, J., Ryser, S. Classic drug targets. *Nat. Biotechnol.* 1997, *15*, 1318–1319.
24. Overington, J. Prioritizing the proteome: identifying pharmaceutically relevant targets. *Drug Discov. Today* 2002, *7*(9), 516–521.
25. Hopkins, A. L., Groom, C. R. The druggable genome. *Nat. Rev. Drug Discov.* 2002, *1*, 727–730.
26. Golden, J. Towards a tractable genome: knowledge management in drug discovery. *Curr. Drug Discov.* 2003(February), 17–20.
27. Golden, J. B. Prioritizing the human genome: knowledge management for drug discovery. *Curr. Opin. Drug Discov. Dev.* 2003, *6*(3), 310–316.
28. Wishart, D. S., Knox, C., Guo, A. C., Shrivastava, S., Hassanali, M., Stothard, P., Chang, Z., Woolsey, J. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res.* 2006, *34*(1), D668–D672.
29. Hopkins, A. L., Groom, C. R. Target analysis: a priori assessment of druggability. *Ernst Schering Res. Found. Workshop.* 2003, *42*(11–17).
30. Hajduk, P. J., Huth, J. R., Fesik, S. W. Druggability indices for protein targets derived from NMR-based screening data. *J. Med. Chem.* 2005, *48*, 2518–2525.
31. Cheng, A. C., Coleman, R. G., Smyth, K. T., Cao, Q., Soulard, P., Caffrey, D. R., Salzberg, A. C., Huang, E. S. Structure-based maximal affinity model predicts small-molecule druggability. *Nat. Biotechnol.* 2007, *25*(1), 71–75.
32. Chothia, C. Hydrophobic bonding and accessible surface area in proteins. *Nature* 1974, *248*, 338–339.
33. Kuntz, I. D., Chen, K., Sharp, K. A., Kollman, P. A. The maximal affinity of ligands. *Proc. Natl. Acad. Sci. USA* 1999, *96*(18), 9997–10002.
34. Al-Lazikani, B., Gaulton, A., Paolini, G., Lanfear, J., Overington, J., Hopkins, A. The molecular basis of predicting druggability. In *Chemical Biology* (Wess, G., Schreiber, S., Eds). John Wiley & Sons, 2007.
35. Hubbard, T. J., Ailey, B., Brenner, S. E., Murzin, A. G., Chothia, C. SCOP: a structural classification of proteins database. *Nucleic Acids Res.* 1999, *27*(1), 254–256.
36. Hann, M. M., Leach, A. R., Harper, G. Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inf. Comput. Sci.* 2001, *41*(3), 856–864.
37. Lander, E. *et al.* Initial sequencing and analysis of the human genome. *Nature* 2001, *409*, 860–921.
38. Venter, J. *et al.* The sequence of the human genome. *Science* 2001, 1304–1351.
39. Russ, A. P., Lampel, S. The Druggable Genome. 2005 (in press).
40. Orth, A. P., Batalov, S., Perrone, M., Chanda, S. K. The promise of genomics to identify novel therapeutic targets. *Expert Opin. Ther. Targets.* 2004, *8*(6), 587–596.
41. Consortium, article. Finishing the euchromatic sequence of the human genome. *Nature* 2004, *431*(7001), 931–945.
42. Hemmi, H., Kaisho, T., Takeuchi, O., Sato, S., Sanjo, H., Hoshino, K., Horshino, T., Tomizawa, H., Takeda, K. Small anti-viral compounds activate immune cells via the TLR7 MyD88-dependent signalling pathway. *Nat. Immunol.* 2002, *3*(2), 196–200.
43. Hopkins, A. L., Polinsky, A. Knowledge and intelligence in drug design. *Annu. Rep. Med. Chem.* 2006, *41*, 425–437.
44. Stewart, K. D., Shiroda, M., James, C. A. Drug Guru: a computer software program for drug design using medicinal chemistry rules. *Bioorg. Med. Chem.* 2006, *14*(20), 7011–7022.
45. van Gestel, S., Schuermans, V. Thirty-three years of drug discovery and research with Dr. Paul Janssen. *Drug Dev. Res.* 1986, *8*(1–4), 1–13.
46. Black, J. W. Receptors as pharmaceutical targets. In *Textbook of Receptor Pharmacology* (Johanson, J. F. T., Ed.). CRC Press: Boca Raton, FL, 2003, pp. 271–279.
47. Blake, J. F. Identification and evaluation of molecular properties related to preclinical optimization and clinical fate. *Med. Chem.* 2005, *1*(6), 649–655.
48. Proudfoot, J. R. The evolution of synthetic oral drug properties. *Bioorg. Med. Chem. Lett.* 2005, *15*(4), 1087–1090.
49. Leeson, P. D., Davis, A. M. Time-related differences in the physical property profiles of oral drugs. *J. Med. Chem.* 2004, *47*(25), 6338–6348.
50. Blake, J. F. Examination of the computed molecular properties of compounds selected for clinical development. *BioTechniques* 2003, *34*, S16–S34.
51. Wenlock, M., Austin, R. P., Barton, P., Davis, A. M., Leeson, P. D. A comparison of physiochemical property profiles of development and marketed oral drugs. *J. Med. Chem.* 2003, *46*(7), 1250–1256.
52. Morphy, R. The influence of target family and functional activity on the physicochemical properties of pre-clinical compounds. *J. Med. Chem.* 2006, *49*(10), 2969–2978.
53. Vieth, M., Sutherland, J. J. Dependence of molecular properties on proteomic family for marketed oral drugs. *J. Med. Chem.* 2006, *49*(12), 3451–3453.
54. Oprea, T. I., Davis, A. M., Teague, S. J., Leeson, P. D. Is there a difference between leads and drugs? A historical perspective. *J. Chem. Inf. Comput. Sci.* 2001, *41*(5), 1308–1315.

Chapter 26

Optical Isomerism in Drugs

Camille G. Wermuth

I. INTRODUCTION
II. EXPERIMENTAL FACTS AND THEIR INTERPRETATION
A. Stereoselectivity in biologically active compounds
B. The three-point contact model
C. Diastereoisomers
D. Stereoselectivity ratios
E. Pfeiffer's rule
III. OPTICAL ISOMERISM AND PHARMACODYNAMIC ASPECTS
A. Differences in potency and antagonism between two enantiomers
B. Differences in the pharmacological profile of two enantiomers
IV. OPTICAL ISOMERISM AND PHARMACOKINETIC ASPECTS
A. Isomer effects on absorption and distribution
B. Isomer effects on metabolism
C. Isomer effects on uptake
D. Isomer effects on excretion
V. PRACTICAL CONSIDERATIONS
A. Racemates or enantiomers?
B. The distomer counteracts the eutomer
C. Racemic switches
D. The distomer is metabolized to unwanted or toxic products
E. Deletion of the chiral center
F. Usefulness of racemic mixtures
REFERENCES

Most natural organic compounds, the essential products of life, are asymmetric and possess such asymmetry that they are not superposable on their images… This establishes perhaps the only well marked line of demarcation that at present can be drawn between the chemistry of dead matter and the chemistry of living matter.

(Pasteur, Van t'Hoff, Le Bel and Wislicenus, Memoirs, 1901)

I. INTRODUCTION

This chapter is concerned with bioactive compounds bearing on their skeleton one or more asymmetric carbon atom(s). For such compounds the term *configuration* defines the implantation mode of the four covalent linkages on the central, asymmetric, carbon atom. The terms *optical isomers*, *optical antipodes*, *enantiomorphs* or *enantiomers*, are synonyms and relate to molecules which are mirror images one to another and are not, therefore, superimposable. Owing to their non-identical 3D structure, enantiomers may elicit differentiated biological responses and thus provide useful information on drug–receptor interactions and on receptor characteristics. A great number of books deal with chirality and drug design.[1–8]

II. EXPERIMENTAL FACTS AND THEIR INTERPRETATION

A. Stereoselectivity in biologically active compounds

Toward a biological target, the potency of two enantiomers can sometimes differ considerably, sometimes be very similar (Table 26.1). Often the activity is concentrated in only one enantiomer. When such a high stereoselectivity arises, it is admitted that the mechanism of action at the molecular level involves a highly specific interaction between the ligand, a chiral molecule and the recognition site, a chiral environment. It is to be expected that the most active isomer, in

Copyright © 2008, Elsevier Ltd
All rights reserved.

TABLE 26.1 Differences in Activity or in Affinity (Eudismic Index) Between Couples of Enantiomers

SUBSTANCE		EUDISMIC INDEX
R-Thiorphan	K_i values for inhibition of neutral endopeptidase 24.11 (former enkephalinase)	S/R = 1,2[9]
R(-)-Noradrenaline [R(-)-Norepinephrine] natural enantiomer	Adrenergic activity at rat aorta α_1 sites	R(−)/S(+) = 33[10]
S(-)-Nicotine (Natural nicotine)	Affinity constants for rat brain thalamus sites labeled by [3H]-(−)-nicotine	S(−)/R(+) = 35[11]
S(+)-Chlorpheniramine (Polaramine)	K_i values for human brain frontal lobe sites labeled by [3H]-mepyramine	S(+)/R(−) = 83[12]
S(+)-Oxaprotiline	IC_{50} values for noradrenaline uptake into rat brain synaptosomes	S(+)/ R(−) = 1000[13]

(Continued)

TABLE 26.1 (Continued)

SUBSTANCE		EUDISMIC INDEX
S(+)-Dexetimide	[^{3}H]-dexetimide binding to brain muscarinic receptors	S(+)/ R(−) = 2000[14]
(+)-Lysergide (LSD)	(+)-[3H] lysergic acid diethylamide binding on rat forebrain suspension	R(+)/ S(−) = 24000[15]

terms of affinity, achieves a better steric complementarity to the receptor than the less active one.

When considering *in vivo* activities, the difference in activity observed for the two enantiomers is neither always and nor exclusively the result of the quality of the ligand–receptor fit. It must be kept in mind that *in vivo* the pharmacokinetic processes (ADME) may account for the observed difference in activity. The interpretation of pharmacological data obtained from *in vivo* assays should thus be questioned and does not allow anticipating the quality of the ligand–receptor interaction.

B. The three-point contact model

When, in a compound exhibiting stereoselectivity, only one asymmetric center is present in the molecule, it is thought that the substituents on the chiral carbon atom make a three-point contact with the receptor. Such a fit ensures a very specific molecular orientation which can only be obtained for one of the two isomers. A three-point fit of this type was first suggested by Easson and Stedman,[16] and the corresponding model proposed by Beckett[17] in the case of R(−)-adrenaline (= R(−)-epinephrine). The more active natural R(−)-adrenaline establishes contacts with its receptor through the following three interactions (Figure 26.1):

1. Acceptor–donor or hydrophobic interaction between the aromatic ring of adrenaline and an aromatic ring of the receptor protein.
2. A hydrogen bond at the alcoholic hydroxyl.
3. An ionic bond between the protonated amino group and an aspartic or glutamic carboxylic group of the receptor.

The combination of these interactions can generate binding energies in the order of 12–17 kcal/M, corresponding to binding constants in the order of 10^{-9} to 10^{-12} M.[18] The biologically weak optical isomer, S(+)-adrenaline, can make contact through only two groups. According to this hypothesis, it would be anticipated that deoxyadrenaline (epinine) should have much the same activity as S(+)-adrenaline. This has been found to be basically true.[16,19,20]

Computer-generated receptor models for protein-G linked receptors are now available,[21,22] and Figure 26.2 illustrates the fit of R(−)-adrenaline into the active site of the β_2-adrenergic receptor. It appears clearly that the docking involves more interactions than only the above-mentioned three points:

1. The two phenolic hydroxyl groups exchange hydrogen bonds with Ser^{505} and Ser^{508} respectively
2. The aromatic ring of adrenaline is stabilized by means of π-π interactions with Phe^{509} and Phe^{617}
3. The cationic head exerts a coulombic interaction with the Asp^{311} carboxylate and is located in a hydrophobic pocket made of Trp^{307}, Phe^{616}, and Trp^{613}
4. Finally the secondary benzylic hydroxyl exchanges a hydrogen bond with Ser^{410}.

FIGURE 26.1 Interaction capacities of the natural R(−)-epinephrine and its S(+) antipode. In simply assuming that the natural R(−)epinephrine establishes a three-point interaction with its receptor (a) the combination of the donor–acceptor interaction, the hydrogen bond and the ionic interaction will be able to generate energies in the order 12 to 17 kcal/mol, that corresponds to binding constants of 10^{-9} to 10^{-12} M[18] The less active isomer, S(+) epinephrine, may establish only a two-point contact (b). The loss of the hydrogen bond interaction equals to approximately 3 kcal/mol, this isomer should therefore possess an approximately 100-fold lesser affinity. The experience confirms this estimate. If we consider less abstract models it becomes apparent that the less potent enantiomer also is able to develop three intermolecular bonds to the receptor, provided that it approaches the receptor in a different manner. However, the probability of this alternate binding mode to trigger the same biological response is close to null.

FIGURE 26.2 Interaction capacities of the natural R(−) epinephrine with a model of its receptor (after references[21,22]).

Even in taking into account these newer findings, it can be speculated that the Easson–Stedman hypothesis still holds. The non-natural S(+)-adrenaline, having the wrong orientation of its benzylic hydroxyl, is unable to exchange a hydrogen bond with Ser[410], and achieves therefore a weaker interaction with the receptor.

An alternative model of the adrenergic receptor-active site shows that natural R-(−) epinephrine (adrenaline) can establish a hydrogen bond with the Ser[410] alcoholic group whereas this interaction is not possible with the non-natural S-(+) epinephrine (Figure 26.3).

C. Diastereoisomers

When more than one asymmetric center is involved the complexity of the problem increases rapidly. For the four isomers of ephedrine, which represent a set of diastereoisomers, the R configuration of the β-carbon (as found for adrenaline, noradrenaline, nordefrin, phenylephrin, and octopamine, see Patil *et al.*[23]), is not automatically associated with the highest alpha-agonistic activity. Both (−)-ephedrine and (−)-pseudo-ephedrine possess the β-(R) configuration, yet only (−)- ephedrine acts as an agonist. This anomaly can be explained if one takes into account the preferred conformations of these two compounds, calculated by using the ETH (Extended Hückel Theory).[24,25] In the (−)-ephedrine molecule, the methyl group attached to the carbon in alpha-position to the amino function is projected above the plane of the phenylethyl-amino group, whereas in (−)-pseudo-ephedrine the methyl group is oriented below the plane and thus prevents an efficient interaction of the drug with the receptor (Figure 26.4).

The antibacterial activity of chloramphenicol isomers represents a similar example. Significant activity is only found for the (−)-threo-chloramphenicol.[26] The clinical formulation of the adrenergic receptor-blocking agent labetalol consists of a mixture of equal proportions of the four optical isomers (RR, SS, RS, and SR). Each possesses different pharmacological properties. The most active RR enantiomer was developed some years ago as dilevalol,[27] but had to be withdrawn after some months due to a slightly higher than average degree of hepatic toxicity.[28]

The antihistaminic drug clemastine (Tavegyl), despite the fact that it contains two chiral centers, provides one of the few examples of chiral antihistamines employed clinically in the form of a single isomer. Data on the antihistaminic activity of clemastine and its isomers[29] are summarized on Table 26.2.

The stereoisomers of some oxotremorine analogs containing two chiral centers and acting as oxotremorine antagonists show *in vivo* (tremorolytic activity) stereoselectivity ratios as high as 1 to 200.[30]

FIGURE 26.3 In an adrenergic receptor model the interaction of the natural R-(−) epinephrine implies.

FIGURE 26.4 Preferred conformations of D-(−)-ephedrine and of D-(−)-pseudo-ephedrine.

D. Stereoselectivity ratios

Stereoselectivity was defined by Rauws[31] as follows: "Stereoselectivity is the extent to which an enzyme or other macromolecule, or macromolecular structure (antibody or receptor) exhibits affinity towards one molecule of a pair of isomers in comparison with and in contrast to the other isomer.". Lehmann[32,33] has expressed this in a mathematical form: the ratio of activity of the better fitting enantiomer (eutomer; greek, "eu" = good), to that of the less fitting enantiomer (distomer; greek, "dys" = bad) is defined eudismic ratio. From this an eudismic affinity quotient can be derived (Table 26.3).

In a series of agonists or antagonists one can write the equation (26.1):

$$\text{EI} = a + b \text{ Log Affin.}Eu \qquad (26.1)$$

in which: a is a constant b is the quotient of eudismic affinity (QEA) which precisely accounts for the stereoselectivity.

When the activity of the eutomer "Eu" is compared to that of the racemic mixture "Rac," four possibilities can arise:[34,35]

1. The activity ratio is equal to 2: Eu/Rac = 2/1. In this case the activity is only concentrated in the eutomer and the distomer does not contribute significantly to the observed activity. The chiral compound shows stereoselectivity.
2. The activity ration is higher than 2: Eu/Rac > 2 (e.g. Eu/Rac = 2/0.3). This means that the distomer represents a competitive antagonist of the eutomer. In the practice such a situation is rather exceptionally encountered (see section "V Practical Considerations; B. The distomer counteracts the eutomer."

TABLE 26.2 Antihistamine Activities of Clemastine and its Isomers

R, R-Clemastine

Antihistamine activities of clemastine and its isomers[29]

Isomer	Prevention of histamine toxicity ED_{50} (mg/kg s)	Prevention of histamine spasm	pA_2
RR (clemastine)	0.04	~ +7	9.45
SS	5.1	~−1.5	7.99
SR	11.0	~−6	8.57
RS	0.28	~+5	9.40

TABLE 26.3 Nomenclature and Definitions in Drug Stereoselectivity

Eutomer (Eu): Enantiomer presenting the *highest* affinity (or activity)
Distomer (Dis): Enantiomer presenting the *lowest* affinity (or activity)
Eudismic quotient: $\frac{\text{Affin.}Eu}{\text{Affin.}Dis}$
Eudismic index (EI): Log Affin. $_{Eu}$-Log Affin.$_{Dis}$

3. If the activity ratio is lower than 2: Eu/Rac < 2 (e.g. Eu/Rac = 2/1.6), we are in the presence of two active isomers. The distomer reinforces the activity of the eutomer. Such a situation indicates a decrease of the receptor selectivity.
4. The activity ratio is Eu/Rac = 1; in this case both isomers are equipotent and no stereoselectivity is observed. This can be explained by the assumption: (a) that the compounds act through a non-specific mechanism, (b) that the active compound and the receptor make only a two-point contact with the chiral center, (c) that the chiral center is not involved in the contact (is located in a "silent region").

E. Pfeiffer's rule

One usually admits that the discriminative effect between the two enantiomers increases with the proximity of the chiral center to the site of interaction with the receptor. An empirical rule published by Pfeiffer in 1956[36] states that the isomeric activity ratio (eudismic quotient) of a highly active couple of isomers is always superior to that of a less active couple. In other words: *"The greater the difference between the pharmacological activity of the R and the S isomers, the greater is the potency of the active isomer."* However, there are some exceptions to Pfeiffer's rule. Some of the reasons are conformational flexibility of the ligands[37] others reside in an improper selection of "homologous" sets of compounds as illustrated with muscarinic agonists and antagonists.[38] Quantitative analyses of the correlations between biological activity and the structure of stereoisomeric compounds are difficult.[39,40]

III. OPTICAL ISOMERISM AND PHARMACODYNAMIC ASPECTS

The biological response induced by a pair of enantiomers can differ in potency (quantitative difference) or in nature (qualitative difference). In the latter case, it is assumed that one enantiomer acts at one receptor site, whereas its antipode is recognized by other sites and possesses a different activity and toxicity profile.

A. Differences in potency and antagonism between two enantiomers

Two optical isomers *are never antagonists*, at least at comparable dosages. This comes from the space relationship required for the interaction with the receptor site which is only slightly altered by passing from S to R forms, or vice-versa. If one of the enantiomers achieves the optimal

TABLE 26.4 Differences in pharmacological profile of couples of enantiomers

Racemate	Levorotatory enantiomer	Dextrorotatory enantiomer	References
Quinine/quinidine (racemate not in use)	Quinine: antipyretic, antimalarial	Quinidine: antiarrythmic antimalarial	White *et al.*,[43] White *et al.*,[44] Alexander *et al.*[45]
Sotalol	(−)-Sotalol β-adrenoceptor blocker	(+)-Sotalol antiarrythmic agent	Drayer[27]
Racemorphane	(−)-N-Methyl-3-methoxy-morphinane antitussive	(+)-N-Methyl-3-methoxy-morphinane analgesic	Benson *et al.*[42]
Indacrinone	R(−)-Indacrinone diuretic	S(+)-Indacrinone uricosuric	Drayer[27]
Propoxyfene	α-Levopropoxyfene (Novrad) antitussive	α-Dextropropoxyfene (Darvon) analgesic	Drayer[27]
Tetramisole	S(−)-Levamisole: nematocidal, immunostimulant	R(+)-Dexamisole: antidepressant	Bullock *et al.*,[46]Schnieden[47]
3-Amino-1-hydroxypyrrolid-2-one (HA-966)	3R-(+)-HA-966 partial agonist at the glycine site of the NMDA receptor	3S-(−)-HA-966 γ-butyrolactone-like sedative	Singh *et al.*[48]
3-Methoxy-cyproheptadine	(−)-3-Methoxy-cyproheptadine anticholinergic activity	(+)-3-Methoxy-cyproheptadine antiserotonin activity	Remy *et al.*[49]

fit to the receptor site in exchanging the highest number of non-covalent linkages, its antipode can only give rise to a weaker interaction, even in the most favorable conditions (Figure 26.1). From a practical point of view, this absence of *stoechiometric* antagonism entails two consequences:

1. If a racemic mixture does not show any activity, it is useless to carry out the separation of the two antipodes.
2. A racemic mixture usually has the average potency of both constituents, thus, the maximal benefit one can achieve in resolving racemic mixtures is to increase the activity of one of the antipodes to twice of that of the racemate.

B. Differences in the pharmacological profile of two enantiomers

Besides the difference in potency, it often happens that two enantiomers show differences in their pharmacological profile. In such a case, resolving the racemic mixture can generate two pharmacologically different and useful compounds and also separate the more active compound from its less well tolerated or more toxic isomer. In the quinine–quinidine couple (Table 26.4), both isomers share antimalarial, antipyretic, oxytoxic, as well as skeletal and cardiac muscle depressant activities. However, whereas antipyresis and treatment of malaria represent the main use of quinine, quinidine is more effective on the cardiac muscle and is used in the therapy of atrial fibrillation and in certain other arrhythmias.[41] In the *N*-methyl-3-methoxy morphinane racemate (racemorphane) most of the analgesic and addictive properties are concentrated in the (+)-isomer. The corresponding (−)-isomer is non-addictive and retains only antitussive properties.[42] The same kind of discrimination is found for the antitussive levopropoxyphene and its well-known analgesic enantiomer dextropropoxyphene.[27]

The substituted imidazo-thiazole, dexamisole, has antidepressant properties and its isomer, levamisole, possesses anthelmintic and immunostimulant properties.[46,47] Enantiomers of HA-966 (3-amino-1-hydroxypyrrolid-2-one) exhibit distinct central nervous system effects: (+)-HA-966 is a selective glycine/*N*-methyl-D-aspartate receptor antagonist, but (−)-HA-966 is a potent γ-butyrolactone-like sedative.[48] A comparison of (+) and (−)-3-methoxycyproheptadine shows that all of the anticholinergic activity of the (±)-3-methoxycyproheptadine resides solely in the dextrorotatory enantiomer, while the antiserotonin activity resides in the levorotatory enantiomer.[49]

Table 26.5 shows some experimental data for the active isomers of the fluoro analogs of the tricyclic neuroleptic clotepin as compared to the corresponding racemate.[50] In this example, it appears clearly that the neuroleptic activity is concentrated in the dextrorotatory compound (+)**2** whereas the toxicity resides in the (−)**3** antipode.

In the present case, the 1:10 therapeutic index of the racemate – unsatisfactory for a clinical outlook – was risen to the much more acceptable 1:50 ratio for the isolated S(+) antipode.[50]

IV. OPTICAL ISOMERISM AND PHARMACOKINETIC EFFECTS

After administration and before it arrives in the vicinity of its receptor site, a drug is subjected to a variety of physiological processes : absorption, distribution metabolism,

TABLE 26.5 Comparison of the Racemate and the Two Enantiomers of a Fluoro Analogs of the Tricyclic Neuroleptic Clotepin[50]

Test	Measurements	RS(±) 1	S(+) 2	R(−)3
Increase in brain homovanillic acid	At 100 mg/Kg	256%	316%	128%
Adenylate-cyclase inhibition	$c = 10^{-6}$ M	48%	72%	27%
Inhibition of conditioned flight reflexes in rats	ED_{50} (mg/Kg) per os	14	10	>100
Inhibition of apomorphine-induced emesis in dogs	ED_{50} (mg/Kg) per os	20	12	>30
Acute toxicity in mice	LD_{50} (mg/Kg) per os	200	515	68

uptake at storage sites, and excretion. Many of these processes are stereoselective, for reviews see Jamali *et al.*[51] and Kroemer *et al.*[52]

A. Isomer effects on absorption and distribution

The higher narcotic potency *in vivo* of the S(+)-isomer of hexobarbital was shown to be related to higher central nervous system (CNS) levels than for the R(−) form, this seems to be due to a better crossing of the blood–brain barrier (BBB).[53] In a distribution study of [^{14}C] (+) and (−)-alpha-methyl-DOPA in the rat after intravenous injection,[54] the (−)-isomer attained higher concentrations than the (+)-form in most organs, in accordance with the fact that of the two isomers, only the (−) isomer has hypotensive activity.[55]

B. Isomer effects on metabolism

Since all enzymes are chiral in nature, and therefore probably possess some degree of asymmetry at the reactive center, it is not surprising that most metabolic reactions of isomers lead to qualitative and quantitative differences in the metabolites formed. For review articles see Testa,[1,2] Vermeulen,[56] and Kroemer *et al.*[52]

1. *Differential metabolism of two antipodes*

The levo isomers of 3-hydroxy-*N*-methyl-morphinan and of methadone are demethylated by rat liver 2–3 times more rapidly than the corresponding dextro antipodes.[57,58] The S(+)-enantiomer of hexobarbital (Figure 26.4) is metabolized almost twice as rapidly as the R(−)-enantiomer by allylic hydroxylation[59] and, in the dog, the dextrorotatory isomer of 5-ethyl-5-phenyl-hydantoin affords 10 times more para-hydroxy-metabolite than the levorotatory isomer.[60] Hydroxylation takes place alpha to a carbonyl in the dextrorotary enantiomer of glutethimide whereas the levorotamer is hydroxylated on the methylene group of the ethyl side chain.[61] Numerous other examples are found in the literature.[62,63]

2. *Enzymatic inversion*

The energy requirements necessary for the conversion of a given sp^3 configuration into its optical antipode imply the formation of an intermediary carbenium ion, carbanion, or free radical and are unlikely to arise in biological systems. Thus, racemization or epimerization involving non-oxygenated sp^3 carbon atoms are generally not encountered in mammals. They are usually restricted to microorganisms (e.g. alanine-racemase). One case of this unusual phenomenon is described in mammals for arylpropionic acids. More precisely, for the non-steroidal antiinflammatory agent ibuprofen (R,S-para-isobutyl-hydratropic acid), it has been

(a) (b)

FIGURE 26.5 An artistic illustration of the R,R/S,S relationship in diastereoisomers. Auguste Rodin's famous sculpture "The Cathedural (image of (a)) represents two right hands. Its mirror image (image of the (b)) represents two left hands. The images symbolize the R,R versus S.S relationship as found in diastereoisomers.

demonstrated that only the S(+)-isomer is active *in vitro* as inhibitor of the prostaglandin-synthesizing enzyme cyclo-oxygenase. Surprisingly no significant differences could be observed *in vivo* between the S(+) or the R(−)-enantiomers and the racemate (= ibuprofen).[64] It was therefore concluded that *in vivo* there must be an almost complete *inversion* of the poorly active R(−) form to the much more active S(+)-isomer. In man, the main metabolites isolated after administration of (racemic) ibuprofen were dextrorotatory[65] and also the pure R(−)-enantiomer is converted to the S(+) –isomer.[66] A biochemical investigation, using deuterium-labeled R(−)-isomer, led to the hypothesis of the existence of an R-arylpropionic acid isomerase ("R-APAI") enzyme system proceeding via the enzymes of lipid catabolism and anabolism as outlined in Figure 26.5.

It is assumed that the coenzyme A ester (CoA-ester) of the R(−)-enantiomer acts as a substrate for the fatty acid deshydrogenase, thus eliminating the chiral center. The next step may, or may not, take place, depending whether or not the CoA-ester must be transferred to an acyl-carrier protein or another site in the fatty acid synthetase system, so that a stereoselective reduction by an enoylreductase can take place. Thus the nature of X is unknown.[67] Similar epimerization reactions were also described for some other arylpropionic acids such as benoxaprofen,[68] carprofen,[69] and isopropyl-indanyl-propionic acid.[70] It was demonstrated that the configural inversion does not take place in the liver, and that the responsible enzyme, R-(−)-arylpropionic acid isomerase, is located in the gut wall.[67,71]

C. Isomer effects on uptake

As drugs are usually absorbed by passive diffusion and, since enantiomers do not differ in their aqueous and lipid solubilities, absorption is not usually considered to be a stereoselective process. However, stereoselectivity has been described for drugs that are transported by a carrier-mediated process. Typical uptake selectivity is observed for neurotransmitter reuptake inhibitors such as nipecotic acid, oxaprotiline, fluoxetine, and venlafaxine. Uptake of drugs by various organs can also be enantioselective, for example the liver/plasma concentration ratios of S(−) and R(+)-phenprocoumon in the rat were found to be different (6.9 and 5.2 respectively), indicating a preferential uptake of the more potent isomer.[72]

D. Isomer effects on excretion

The kinetics of excretion are a direct consequence of the kinetics of metabolic transformations. The faster a drug is metabolized, the faster its elimination can be expected. In accordance with this assertion, rats given R,S(±), S(+), and R(−)-amphetamine, were found to excrete less (+)-*p*-hydroxy-amphetamine than its (−)-isomer; this may be the basic explanation of the more pronounced pharmacological properties of the dextro-, compared to the levoamphetamine.[73] For the hypnotic agent hexobarbital, the elimination half-life in man is about three times longer for the more active (+)-isomer then for the less active (−)-isomer. This was attributed to a difference in hepatic metabolic clearance and not in volumes of distribution or plasma binding between the enantiomers.[74]

V. PRACTICAL CONSIDERATIONS

A. Racemates or enantiomers?

Many drugs having a center of asymmetry are still used in clinical practice as racemates, and racemic mixtures were estimated to represent 10–15% of all the marketed drugs.[3] For certain types of therapeutics, such as the β-adrenergic agents, β-adrenergic blockers, antiepileptics, and oral anticoagulants, up to 90% of the compounds are, according to Ariëns,[3] in fact racemic mixtures. For antihistaminics and local anesthetics this holds true for about 50% of the drugs currently used.[3] Often racemic drugs were introduced in clinical practice because the animal and the clinical pharmacology, the toxicology, and the teratology were performed with the racemates. The reasons for that is, that at the time of the discovery of the drug, the resolution (or the chiral synthesis) appeared to be too difficult or too costly, or even impossible.

The question now arises to decide when and why to use rather racemic mixtures or pure enantiomers. Although it seems good sense to use pure eutomers and to consider the distomer as an unwanted load of xenobiotic (a kind of pollution, or even an impurity) there are however instances where it is recommended to use racemates rather then eutomers.

Thus racemates may be more stable, more active or less toxic or present a favorable combination of the properties of each separate isomer (see below). Finally one can ask if it would not be wise to design effective drugs without centers of asymmetry.

B. The distomer counteracts the eutomer

Contrariwise to a well-established belief, there are no examples of inactive racemates in which the distomer antagonizes, in a stoichiometric manner, the activity of the eutomer.

Thus, a dihydropyridine-derived calcium inhibitor, the R(−) enantiomer of compound Sandoz 202-791 inhibits the uptake of [$^{45}Ca^{++}$] with an IC_{50} of 4.3×10^{-8}, whereas its S(+) enantiomer increases the uptake with an IC_{50} of approximately 10^{-6} to 10^{-7} M.[75] The corresponding racemic mixture inhibits the uptake with an IC_{50} of 1.7×10^{-7} M. Some other examples are reported in Table 26.6.

As shown in Table 26.6, a more or less important residual activity is always present in the racemate but resolution would generally be beneficial. Picenadol (LY 150720) seems to be an exception to the rule that pure eutomers should be used when the distomer shows antagonistic properties. For clinical trials as narcotic analgesic the racemate was the preferred preparation owing to its partial agonist profile.

C. Racemic switches

Presently, a general trend in the pharmaceutical industry is to switch from racemates to single enantiomers. Examples are given by (R)-(−)-verapamil, (S)-fluoxetin, (S)-ketoprofen, (R)-albuterol, levofloxacin, esoprazole, cetirizine, cisapride, and many others.[81,82] In addition to the quality improvement of the drug, this switch represents also a way to prolong its life insofar that the isolated eutomer is legally considered as a new drug entity. As a consequence, drug companies are increasingly adopting racemic switches as a management strategy. The company first develops a chiral drug as a racemate, and, later on, patents and develops the single isomer.[82] This strategy does not always work successfully. This is illustrated by the S(−)-eutomer of propranolol. This compound shows reduced β-blocking activity when administered as single isomer compared with its bioavailability when administered as a racemate, suggesting that the presence of R(+)-propranolol had a beneficial effect on the availability of S(−)-propranolol.[83] The same phenomenon happened when the racemate of fluoxetin was compared with its eutomer. The consequence was that the management at Eli Lilly decided not to practice the racemic switch for this compound.

D. The distomer is metabolized to unwanted or toxic products

Racemic deprenyl, a monoamine–oxidase inhibitor used in the treatment of depression, is metabolized to (+)- and (−)-metamphetamine,[84] the former being much more active than its (−)-isomer as central stimulant leading to drug abuse (Table 26.7).

On the other hand the (−)-isomer of deprenyl is a much more potent MAO-B inhibitor than the (+)-isomer. For these reasons racemic deprenyl has been replaced by (−)-deprenyl in clinical practice. In the racemic local anesthetic prilocaïne (Figure 26.6) only the R-(−)-isomer is metabolized to an aniline derivative (ortho-toluidine) and to the corresponding para- and ortho-aminophenols that are highly toxic and responsible for met hemoglobinemia.[85]

The S-(+) enantiomer is not a substrate for the metabolizing enzyme and would probably be chemically safe.

TABLE 26.6 Antagonism in Couples of Enantiomers

Compound	Eutomer	Distomer	Racemate	Reference
N-Isopropyl-norepinephrine	(−) α-Adrenergic agonist	(+) Inactive competitive antagonist	(±) Partial agonist	Page 15 of ref.[3]
5-Ethyl-5(1,3-dimethylbutyl)barbituric acid	(+) Convulsant	(−) Depressant	(±) Convulsant	Hof and Adron-Harris.[76]
Ozolinone (metabolite of etazoline)	(−) Diuretic	(+) Inhibits low doses of (−) or of furoxemide	(±) Diuretic	Greven *et al.*[77]
Picenadol	(+) Morphinomimetic	(−) Narcotic antagonist	(±) Partial agonist	Zimmerman and Gesellchen[78]
Alpha-(2,4,5)-trichloro-phenoxy-propionic acid	(+) Auxin-like plant growth regulator	(−) Decreases activity of (+)	(±) Auxin-like plant growth regulator	Smith *et al.*[79]
6-Ethyl-9-oxaergoline (EOE)	(−) Dopamine agonist	(+) Dopamine antagonist	(±) Dopamine agonist	Lotti and Taylor[80]

Many of the side effects (e.g. granulocytopenia) encountered with racemic DOPA were not seen with levo-DOPA and therefore can be attributed to the (+)-enantiomer.[86] For this reason the racemate is no longer given. Post-anesthesia reactions to the anesthetic and analgesic agent ketamine are overwhelmingly associated with the R(−) antipode.[87]

In vitro studies suggest that the beneficial antiarrhythmic properties of disopyramide are concentrated in the S(+)-isomer whereas the negative inotropic effect predominates in the R(−)-isomer.[88] In addition, the pharmacokinetics (clearance and protein binding) differ.[89,90] For these reasons, selection of the S(+) isomer may have led to the development of a very effective drug with significantly fewer therapeutic problems.

E. Deletion of the chiral center

Nowadays it is well accepted that racemates and both enantiomers are usually three different pharmacological entities and that it requires extensive pharmacological, toxicological, and clinical pharmacological research before it can be decided whether it is advantageous to use racemates or enantiomers in clinical practice. According to Soudijn,[91] these research efforts could be reduced to about one third when drugs without centers or planes of asymmetry could be developed with the same or higher affinity.

Effectively *asymmetry is far from being an absolute requisite for activity!* The alkaloid morphine possesses five chiral centers, on the other hand its synthetic derivative fentanyl is devoid of any asymmetry center and nonetheless belongs to the most potent analgesics known. In some instances the chiral centers can at least partially be eliminated. This is the case for the synthetic analogs of the HMG-CoA reductase inhibitor mevinolin. Mevinolin itself (Figure 26.8) has eight asymmetric centers but structure-activity relationship (SAR) studies rapidly revealed that the six chiral centers contained in the hexahydronaphtalene unit are unnecessary for HMG-CoA inhibition. The second generation of mevinolin analogs, illustrated in Figure 26.7 by the compound HR 780, retains only two of the initial eight chiral centers.[92]

Usually chiral centers are eliminated in creating symmetry. Thus, in a series of muscarinic agonists derived from 3-aminopyridazines, one of the most favorable side-chains, was the racemic 2-*N*-ethylpyrrolidinyl-methyl chain, that is, the side-chain of sulpiride (Figure 26.8). The 5-methyl-6-phenylpyridazine bearing this basic chain at its 3-amino function presented a 0.26 micromolar affinity for M1 muscarinic receptor preparations.[93]

After resolution of the racemate, the corresponding enantiomers show only a six-fold difference in M1 affinity. It was therefore decided to eliminate the chiral center by introducing symmetry either by ring opening, or by ring closure or even by replacing the 2-*N*-ethylpyrrolidinyl-methyl unit by the non-chiral tropane ring. The modified structures show affinities similar to that of the corresponding chiral molecule.[93]

TABLE 26.7 Activities of Deprenyl Enantiomers and Their Metabolites

Formula	Activity	Ratio
Deprenyl	MAO-B inhibition amphetamine effects	(−) >> (+) (−) < (±)
Metamphetamine	Amphetamine effects	(−) << (+)
Amphetamine		

F. Usefulness of racemic mixtures

In practice, if both optical isomers are of similar potency and do have similar pharmacokinetic profiles, it may be

FIGURE 26.6 The structures of hexobarbital, 5-ethyl-5-phenyl-hydantoin, and glutethimide.

FIGURE 26.7 Mechanism of the enzymatic inversion of R(−)-ibuprofen.[67]

FIGURE 26.8 Stereoselective metabolic attack yielding toxic metabolites.[85]

useless to proceed to the resolution of the racemic mixture. Such situations are infrequent but may occur. An example is given by the antithrombotic acids 21-X and 21-Y (Figure 26.9).[94] The corresponding pure enantiomers were first compared to the corresponding racemates for their *in vitro* activities.

In both series almost equipotent activities were observed for thromboxane receptor antagonism and thromboxane synthase inhibition (IC_{50} = 2–30 nM). Upon oral administration to guinea pigs the enantiomers inhibited the *ex vivo* U-46619-induced platelet aggregation with potencies similar to that of the corresponding racemates. This indicates that the enantiomers have pharmacologic profile and bioavailability similar to that of the corresponding racemic compound (Figure 26.10 and Figure 26.11).

The racemates can even be more potent than either of the enantiomers used separately, this is observed with the antihistaminic drug isothipendyl.[95] In other cases it may be of interest to racemize a natural optically active molecule. Thus, to warrant a constant pharmacological activity of ergotamine, which is racemized in solution, producing inactive ergotaminine, the commercial solution is produced as an equilibrium mixture of the two antipodes.[96] Another example of the utility of a racemic mixture is given by the lysine salts of aspirine. The acetylsalicylate prepared from (R,S)-lysine is a stable, crystalline white powder which is freely soluble in water giving a tasteless, odorless and colorless solution, suitable for parenteral injections. Surprisingly, the corresponding salts of pure (R)-lysine or pure (S)-lysine do not crystallize.[97] Finally, when the distomer is converted to the eutomer *in vivo*, as seen above for ibuprofen and its analogs it becomes also preferable to commercialize the racemate.

The recommendations of the European Community Working Party on drug quality, safety and efficacy, take into account two situations.[98] For already *well-established racemates* the clinical use can continue as such, no specific study of the isolated enantiomers is required. For *newly introduced chiral drugs* both enantiomers have to be prepared and studied separately with regard to their activity as well as their disposition *in vivo*. However the final decision to introduce the drug on the market as enantiomer or as racemate belongs to the producer.

FIGURE 26.9 Deletion of six out of eight chiral centers yield still highly potent mevinolin analogs.[92]

FIGURE 26.10 Introducing symmetry and abolishing thus a chiral center (affinity values for M_1 receptor preparations expressed as micromoles).

FIGURE 26.11 Isoactive antithrombotic enantiomers.[94]

BOX 26.1 The water-soluble D, L-lysine salt of aspirine: a success story based on luck and serendipity

Before the World War II, in the little city of Chef-Boutonne in Frances's Poitou-Charente region, Gaston Baetz prepared, in the back shop of his pharmacy, some in-house remedies able to be sold over the counter. His start-up business created in 1933 and named ATP (Association Technique Pharmaceutique) flourished reasonably well.

At the same period, about 160 miles eastward, in the Auvergne city of Commentry, a chemical plant named Alimentation Equilibrée Commentry (AEC), produced some synthetic food additives such as vitamins, methionine and lysine for the intensive farming of chicken and porks. A spin-off company of AEC named l'Equilibre Biologique was then created with the objective to develop some amino-acid derived drugs for human use. The two companies, ATP and l'Equilibre Biologique, merged in 1945 under the control of l'Alimentation Equilibrée.

In 1953, Gaston Baetz, his oldest son Jacques and some friends had the opportunity to take over the control of l'Equilibre Biologique and to develop their pharmaceutical company under the name Egic. Egic was specialized in hospital formulations of sterile injectable nutriments such as lipid emulsions, amino-acid mixtures, and glucose solutions.

With the objective of preparing water-soluble salts of aspirin (which is known to be badly soluble in water and anyway rapidly hydrolyzed), the chemists of l'Equilibre Biologique prepared and patented some aspirin salts of the basic amino-acids lysine and arginine which were provided by AEC (French patent 1 295 304, May 7th, 1956). The salts proved to be water-soluble, but the laboratory procedure used for their preparation yielded slightly hygroscopic salts which partly decomposed to acetic acid and free salicylic acid. The industrial development was therefore precluded for some years.

In 1967, under the scientific direction of Dr. Pierre Baetz (the younger brother of Jacques), the question arose if an adequate pharmaceutical formulation could not rescue the compound. However, a preliminary physicochemical re-investigation had to be undertaken first. Indeed, the original batches of lysine aspirinate were prepared from the available AEC lysine. As this lysine was produced by synthesis, it was racemic, [(D,L)-lysine]. For many scientists of the company, notably Pierre Baetz, Francis Rosé and Abkar Vartanian, it appeared that the natural L-lysine salt would be preferable. As a consequence, attempts were made to prepare the aspirin salt of the natural L-lysine. However, despite good will and obstinate efforts the salt did not crystallize, nor did the aspirin salt of D-lysine. The decision was then taken to develop the D,L-salt corresponding to the synthetic racemic lysine. A *posteriori* it is interesting to note that fortunately, at the time of the preparation of the first batches of lysine aspirinate, the AEC chemists were probably the only ones in the world using racemic lysine …

The combination of a modified preparation procedure yielding a very anhydrous lyophilized salt (French Demande 2 115 060, August 11, 1972) on one hand, and working under controlled atmosphere on the other hand, allowed the industrial production of the racemic lysine salt. This compound is a white crystalline powder which is freely soluble in water, odorless and tasteless. It allowed for the first time to prepare injectable preparations of aspirin.

A possible explanation of the higher propensity of the racemate to crystallize can reside in a closer fitting of the molecules in the crystal grid. As some textbooks mention:

"It is easier to put a pair of shoes into a box than two left shoes (or two right ones)…"

REFERENCES

1. Testa, B. Chiral aspects of drug metabolism. *Tr. Pharmacol. Sci.* 1986, *7*, 60–64.
2. Testa, B., Mayer, J. M. Stereoselective drug metabolism and its significance in drug research. In *Progress in Drug Research* (Jucker, E., Ed.). Birkhäuser: Basel, 1988, pp. 249–303.
3. Ariëns, E. J. Stereoselectivity of bioactive agents: general aspects. In *Stereochemistry and Biological Activity of Drugs* (Ariëns, E. J., Soudjin, W., Timmermans, P. B. M. W. M., Eds). Blackwell Scientific Publications: Oxford, 1983, pp. 11–32.
4. Ariëns, E. J. Stereospecificity of bioactive agents. In *Stereoselectivity of Pesticides* (Ariëns, E. J., van Rensen, J. J. S., Welling, W., Eds). Elsevier: Amsterdam, 1988.
5. Simonyi, M. *Problems and Wonders of Chiral Molecules*. Akadémia Kiado: Budapest, 1990, p. 400.
6. Brown, C. *Chirality in Drug Design and Synthesis*. Academic Press: London, 1990, p. 243.
7. Casy, A. F., Dewar, G. H. *The Steric Factor in Medicinal Chemistry Dissymmetric Probes of Pharmacological Receptors*. Plenum Press: New York, 1993.
8. Federsel, H.-J., Chiral drug discovery and development-from concept stage to market launch. In *Comprehensive Medicinal Chemistry II*, (Taylor, J. B., Triggle, D. J., Eds.), Vol. 2, Elsevier: Amsterdam, 2007, pp. 713–734.
9. Fournié-Zaluski, M. C., Lucas-Soroca, E., Devin, J., Roques, B. P. ^{1}H NMR configural correlation for retro-inverso dipeptides: application to the determination of the absolute configuration of "enkephalinase" inhibitors. *J. Med. Chem.* 1986, *29*, 751–757.
10. Jordan, R., Midgeley, J. M., Thonoor, C. M., Williams, C. M. Beta-adrenergic activities of octopamine and synephrine stereoisomers on guinea-pig atria and trachea. *J. Pharm. Pharmacol.* 1987, *39*, 752–754.
11. Martino-Barrows, A. M., Kellar, K. J. [3H]Acetylcholine and [3H](−)nicotine label the same region in brain. *Mol. Pharmacol.* 1987, *31*, 169–174.
12. Chang, R. S. L., Tran, V. T., Snyder, S. H. Heterogenicity of histamine H1-receptors: species variations in [3H]mepyramine binding of brain membranes. *J. Neurochem.* 1979, *32*, 1653–1663.
13. Waldmeier, P. C., Baumann, P. A., Hauser, K., Maitre, L., Storni, A. Oxaprotiline, a noradrenaline uptake inhibitor with an active and an inactive enantiomer. *Biochem. Pharmacol.* 1982, *31*, 2169–2176.
14. Laduron, P. M., Vervimp, M., Leysen, J. E. Stereospecific *in vitro* binding of [^{3}H]-dexetimide to brain muscarinic receptors. *J. Neurochem.* 1979, *32*, 421–427.
15. Lovell, R. A., Freedman, D. X. Stereospecific receptor sites for *d*-lysergic acid diethylamide in rat brain: effects of neurotransmitters, amine antagonists, and other psychotopic drugs. *Mol. Pharmacol.* 1976, *12*, 620–630.

16. Easson, L. H., Stedman, E. Studies on the relationship between chemical constitution and physiological action. V. Molecular dissymmetry and physiological activity. *Biochem. J.* 1933, *27*, 1257–1266.
17. Beckett, A. H. Stereochemical factors in biological activity. In *Fortschritte der Arzneimittel Forschung* (Jucker, E., Ed.), Vol. I. Birkhäuser Verlag: Basel, 1959, pp. 455–530.
18. Farmer, P. S., Ariëns, E. J. Speculations on the design of nonpeptide peptidomimetics. *Tr. Pharmacol. Sci.* 1982, *3*, 362–365.
19. Patil, P. N., LaPidus, J. B., Campbell, D., Tye, A. Steric aspects of adrenergic drugs. II. Effects of DL isomers and deoxy derivatives on the reserpine-pretreated vas deferens. *J. Pharmacol. Exp. Ther.* 1967, *155*, 13–23.
20. Ruffolo, R. R., Jr Stereoselectivity in adrenergic agonists and adrenergic blocking agents. In *Stereochemistry and Biological Activity of Drugs* (Ariëns, E. J., Soudjin, W., Timmermans, P. B. M. W. M., Eds). Blackwell Scientific Publications: Oxford, 1983, pp. 103–125.
21. Hibert, M. F., Trumpp-Kallmeyer, S., Bruinvels, A., Hoflack, J. Three-dimensional models of neurotransmitter G protein coupled receptors. *Mol. Pharmacol.* 1992, *40*, 8–15.
22. Trumpp-Kallmeyer, S., Hoflack, J., Bruinvels, A., Hibert, M. Modelling of G protein-coupled receptors. Application to dopamine, adrenaline, serotonin, acetylcholine and mammalian opsin receptors. *J. Med. Chem.* 1992, *35*, 3448–3462.
23. Patil, P. N. Steric aspects of adrenergic drugs.8. Optical isomers of beta adrenergic receptor antagonists. *J. Pharmacol. Exp. Ter.* 1968, *160*, 308–314.
24. Kier, L. B. The preferred conformations of ephedrine isomers and the nature of the *alpha* adrenergic receptor. *J. Pharmacol. Exp. Ther.* 1968, *164*, 75–81.
25. Portoghese, P. S. Stereochemical studies on medicinal agents. IV Conformational analysis of ephedrine isomers and related compounds. *J. Med. Chem.* 1967, *10*, 1057–1063.
26. Maxwell, R. E., Nickel, V. S. The antibacterial activity of the isomers of chloramphenicol. *Antibiot. Chemother.* 1954, *4*, 289–295.
27. Drayer, D. E. Pharmacodynamic and pharmacokinetic differences between drug enantiomers in humans: an overview. *Clin. Pharmacol. Ther.* 1986, *40(2)*, 125–133.
28. Fell, A. F. Current perspectives on chiral drug development. *EUFEBS News Lett.* 1998, *7*, 1–2.
29. Ebnöther, A., Weber, H.-P. Synthesis and absolute configuration of clemastine and its isomers. *Helv. Chim. Acta.* 1976, *59*, 2462–2468.
30. Ringdahl, B., Resul, B., Dahlbom, R. Stereoselectivity of some oxotremorine antagonists containing two chiral centers. *J. Pharm. Pharmacol.* 1979, *31*, 837–839.
31. Rauws, A. G. Origin and basis of stereoselectivity in biology. In *Stereochemistry and Biological Activity of Drugs* (Ariëns, E. J., Soudijn, W., Timmermans, P. B. M. W. M., Eds). Blackwell Scientific Publications: Oxford, 1983, pp. 1–10.
32. Lehman, F. P. A., Ariëns, E. J., Rodrigues de Miranda, J. F. Stereoselectivity and affinity in molecular pharmacology. In *Progress in Drug Research* (Jucker, E., Ed.), Vol. 20. Birkhäuser Verlag: Basel, 1976, pp. 101–142.
33. Lehman, P. A. Stereoisomerism and drug action. *Tr. Pharmacol. Sci.* 1986, *7*, 281–285.
34. Casy, A. F. Stereochemistry and biological activity. In *Medicinal Chemistry* (Burger, A., Ed.), Vol. 1. Wiley-Interscience: New York, 1970, pp. 81–107.
35. Schröder, E., Rufer, C., Schmiechen, R.*Arzneimittelchemie*, Vol. 1. Georg Thieme Verlag: Stuttgart, 1976. p. 48.
36. Pfeiffer, C. C. Optical isomerism and pharmacological action, a generalization. *Science* 1956, *124*, 29–31.
37. Barlow, R. B. Enantiomers: how valid is Pfeiffer's rule?. *Trends Pharmacol. Sci.* 1990, *11*, 148–150.
38. Gualtieri, F. Pfeiffer's rule OK?. *Trends Pharmacol. Sci.* 1990, *II*, 315–316.
39. Lien, E. J., Rodrigues de Miranda, J. F., Ariëns, E. J. Quantitative structure-activity correlation of optical isomers: a molecular basis for Pfeiffer's rule. *Mol. Pharmacol.* 1976, *12*, 598–604.
40. Portoghese, P. S., Williams, D. A. Stereochemical studies on medicinal agents. VIII. Absolute stereochemistries of isomethadol isomers. *J. Med. Chem.* 1970, *13*, 626–630.
41. Roden, D. M. Antiarrhythmic drugs. In *The Pharmacological Basis of Therapeutics* (Hardman, J. G., Limbird, L. E., Molinoff, P. B., Ruddon, R. W., Goodman Gilman, A., Eds), 9th ed. Mc Graw-Hill: New York, 1995, p. 1905.
42. Benson, W. M., Stefko, P. L., Randall, L. O. Comparative pharmacology of levorphan, racemorphan and dextrorphan and related methyl ethers. *J. Pharmacol. Exptl. Therap.* 1953, *109*, 189–200.
43. White, N. J., Looareeswan, S., Warrel, D. A. Quinidine in falciparum malaria. *Lancet* 1981, *2*, 1069–1071.
44. White, N. J., Looareeswan, S., Warrel, D. A. Quinine and quinidine: a comparison of EKG effects during the treatment of malaria. *J. Cardiovasc. Pharmacol.* 1983, *5*, 173–175.
45. Alexander, F., Gold, H., Katz, L. N. The relative value of synthetic quinidine, dihydroquinidine, commercial quinidine and quinine in the control of cardiac arrhythmias. *J. Pharmacol. Exptl. Ther.* 1947, *90*, 191–201.
46. Bullock, M. W., Hand, J. J., Waletzky, E. Resolution and racemization of dl-tetramisole, dl–6-phenyl-2,3,5,6-tetrahydroimidazo-[2,1-b] thiazole. *J. Med. Chem.* 1968, *11*, 169–171.
47. Schnieden, H. Levamisole – a general pharmacological perspective. *Int. J. Immunopharmacol.* 1981, *3*, 9–13.
48. Singh, L., Donald, A. E., Foster, A. C. Enantiomers of HA-966 (3-amino-1-hydroxypyrrolid-2-one) exhibit distinct central nervous system effects: (+)-HA-966 is a selective glycine/*N*-methyl-D-aspartate receptor antagonist, but (−)-HA-966 is a potent γ-butyrolactone-like sedative. *Proc. Natl. Acad. Sci. USA* 1990, *87*, 347–351.
49. Remy, D. C., Rittle, K. E., Hunt, C. A., Anderson, P. S., Engelhardt, E. L., Clineschmidt, B. V., Scriabine, A. (+) and (−)-3-Methoxycyproheptadine a comparative evaluation of the antiserotonin, antihistaminic, anticholinergic, and orexigenic properties with cyproheptadine. *J. Med. Chem.* 1977, *20*, 1681–1684.
50. Aschwanden, W., Kyburz, E., Schönholzer, P. Stereospezifizität der neuroleptischen Wirkung und Chiralität von (+)-3-{2-[Fluor-2-methyl-10,11-dihydrodibenzo[*b,f*]thiepin-10-yl)-1-piperazinyl]-äthyl}-2-oxazolidinon (16). *Helv. Chim. Acta.* 1976, *59*, 1245–1252.
51. Jamali, F., Mehvar, R., Pasutto, F. M. Enantioselective aspects of drug action and disposition: therapeutic pitfalls. *J. Pharm. Sci.* 1989, *78*, 695–715.
52. Kroemer, H. K., Gross, A. S., Eichelbaum, M. Enantioselectivity in drug action and drug metabolism: influence on dynamics. In *Pharmacodynamics and Drug Development* (Cutler, N. R., Sramek, J. J., Narang, P. K., Eds). John Wiley & Sons: Chichester, 1994, pp. 103–114.
53. Buch, H., Rummel, W., Brandenburger, V. Versuche zur Aufklärung der Ursachen der unterschiedlichen narkotischen Wirksamkeit von (+)- und (−)-Evipan. *Arch. Pharmakol. Exptl. Pathol.* 1967, *257*, 270–271.
54. Duhm, B., Maul, W., Medenwald, H., Platzschke, K., Wegner, L. A. Experimental animal studies with alpha-methyldopa-14C, with special attention to the optical isomers. II. Organ distribution. *Z. Naturforsch.* 1967, *22b*, 70–84.
55. Sjoerdsma, J., Udenfriend, S. Pharmacology and biochemistry of α-methyl-dopa in man and experimental animals. *Biochem. Pharmacol.* 1961, *8*, 164.
56. Vermeulen, N. P. E. Stereoselective biotransformation: its role in drug disposition and drug action. In *Innovative Approaches in Drug Research* (Harms, A. F., Ed.). Elsevier Science Publishers B.V.: Amsterdam, 1986, pp. 393–416.
57. Axelrod, J. The enzymatic *N*-demethylation of narcotic drugs. *J. Pharmacol. Expt. Therap.* 1956, *117*, 322–330.

58. Elison, C., Elliott, H. W., Look, M., Rapoport, H. Some aspects of the fate and relationship of the *N*-methyl group of morphine to its pharmacological activity. *J. Med. Chem.* 1963, *6*, 237–246.
59. Degwitz, E., Ullrich, V., Staudinger, H., Rummel, W. Metabolism and cytochrome P-450 binding spectra of (+) and (−) hexobarbital in rat liver microsomes. *Hoppe-Seylers Z. Physiol. Chem.* 1969, *350*, 547–553.
60. Kupfer, A., Bircher, J., Preisig, R. Stereoselective metabolism, pharmacokinetics and biliary elimination of phenylethylhydantoin (nirvanol) in the dog. *J. Pharmacol. Exptl. Therap.* 1977, *203*, 493–499.
61. Keberle, H., Riess, W., Hoffman, K. The stereospecific metabolism of the optical antipodes of β-phenyl-β-ethylglutarimide (Doriden). *Arch. Int. Pharmacodyn.* 1963, *142*, 117–124.
62. Jenner, P., Testa, B. The influence of stereochemical factors on drug metabolism. *Drug Metab. Rev.* 1973, *2*, 117–184.
63. Testa, B., Jenner, P. *Concepts in Drug Metabolism. Part A*. Marcel Dekker: New York, 1980.
64. Adams, S. S., Bresloff, P., Mason, C. G. Pharmacological differences between the optical isomers of ibuprofen: evidence for metabolic inversion of the (−)-isomer. *J. Pharm. Pharmacol.* 1976, *28*, 256.
65. Adams, S. S., Cliffe, E. E., Lessel, B., Nicholson, J. S. Some biological properties of 2-(4-isobutylphenyl)-propionic acid. *J. Pharm. Sci.* 1967, *56*, 1686.
66. Vane, J. R. Inhibition of prostaglandin synthesis as a mechanism of action for aspirin-like drugs. *Nature* 1971, *231*, 232–235.
67. Wechter, W. J., Loughhead, D. G., Reisher, R. J., van Geissen, G. J., Kaiser, D. G. Enzymatic inversion at saturated carbon: nature and mechanism of the inversion of R(−) *p-iso*-butyl hydratropic acid. *Biochem. Biophys. Res. Commun.* 1974, *61*, 833–837.
68. Bopp, R. J., Nash, J. F., Ridolfo, A. S., Shepard, E. R. Stereoslective inversion of (R)-(−)-benoxaprofen to the (S)-(+)-enantiomer in humans. *Drug Metab. Dispos.* 1979, *7*, 356–359.
69. Kemmerer, J. M., Rubio, F. A., Mc Clain, R. M., Koechlin, B. A. Stereospecific assay and stereospecific disposition of carprofen in rats. *J. Pharm. Sci.* 1979, *68*, 11274–11280.
70. Tanaka, Y., Hayashi, R. Stereospecific inversion of configuration of 2-(2-isopropylindan-5-yl)-propionic acid in rats. *Chem. Pharm. Bull.* 1980, *28*, 2542–2545.
71. Simmonds, R. G., Woodage, T. J., Duff, S. M., Green, J. N. Stereospecific inversion of (R)-(−)-benoxafen in rat and in man. *Eur. J. Drug Metab. Pharmacokin.* 1980, *5*, 169–172.
72. Schmidt, W., Jahnchen, E. Stereoselective drug distribution and anticoagulant potency of phenprocoumon in rats. *J. Pharm. Pharmacol.* 1977, *29*, 266–271.
73. Gunne, L. M., Galland, L. Stereoselective metabolism of amphetamine. *Biochem. Pharmacol.* 1967, *16*, 1374–1377.
74. Vermeulen, N. P. E., Breimer, D. D. Stereoselectivity in drug and xenobiotic metabolism. In *Stereochemistry and Biological Activity of Drugs* (Ariëns, E. J., Soudjin, W., Timmermans, P. B. M. W. M., Eds). Blackwell Scientific Publications: Oxford, 1983.
75. Hof, R. P., Rüegg, U. T., Hof, A., Vogel, A. Stereoselectivity at the calcium channel: opposite action of the enantiomers of a 1,4-dihydropyridine. *J. Cardiovasc. Pharmacol.* 1985, *7*, 689–693.
76. Ho, I. K., Harris, R. A. Mechanism of action of barbiturates. *Ann. Rev. Pharmacol. Toxicol.* 1981, *21*, 93–111.
77. Greven, J., Defrain, W., Glaser, G., Meywald, K., Heidenreich, O. Studies with the optically active isomers of the new diuretic drug ozolinone. *Pflüger's Arch* 1980, *384*, 57–60.
78. Zimmerman, D. M., Gesellchen, P. D. Analgesics (peripheral and central), endogenous opioids and their receptors. *Ann. Rep. Med. Chem.* 1982, *17*, 21–30.
79. Smith, S. M., Wain, R. L., Wightman, F. Studies on plant growth-regulating substances.V. Steric factors in relation to mode of action of certain aryloxyalkylcarboxylic acids. *Ann. Appl. Biol.* 1952, *39*, 295–307.
80. Lotti, V. J., Taylor, D. A. α2-Adrenergic agonist and antagonist activity of the respective (−) and (+)-enantiomers of 6-ethyl-9-oxaergoline. *Eur. J. Pharmacol.* 1982, *85*, 211–215.
81. Stinson, S. T. Chiral drugs. *Chem. Eng. News.* 1995, 44–74.
82. Stinson, S. C. Chiral drug interactions. *Chem. Eng. News.* 1999, 101–120.
83. Lindner, W., Rath, M., Stochitzky, K., Semmelrock, H. J. Pharmacokinetic data of propranolol enantiomers in a comparative human study with (S)- and (R,S)-propranolol. *Chirality* 1989, *1*, 10–13.
84. Reynolds, G. P., Elsworth, J. D., Blau, K., Sandler, M., Lees, A. J., Stern, G. M. Deprenyl is metabolized to metamphetamine and amphetamine in man. *Br. J. Clin. Pharmacol.* 1978, *6*, 542–544.
85. Akerman, B., Astrom, A., Ross, S., Telc, A. Studies on the absoption, distribution and metabolism of labeled prilocaine and lidocaine in some animal species. *Acta Pharmacol. Toxicol.* 1966, *24*, 389–403.
86. Cotzias, G. C., Papavasiliow, P. S., Gellene, R. Modification of parkinsonism: chronic treatment with L-dopa. *New Engl. J. Med.* 1969, *280*, 337–345.
87. White, P. F., Ham, J., Way, W. L., Trevor, A. J. Pharmacology of ketamine isomers in surgical patients. *Anesthesiology* 1980, *52*, 231–239.
88. Kidwell, G. A., Lima, J. J., Schaal, S. F., Muir, W. M. Hemodynamic and electrophysiologic effects of disopyramide enantiomers in a canine blood superfusion model. *J. Cardiovasc. Pharmacol.* 1989, *13*, 644–655.
89. Lima, J. J., Boudoulas, H., Shields, B. J. Stereoselective phamacokinetics of disopyramide enantiomers in man. *Drug Metab. Dispos.* 1985, *13*, 572–577.
90. Giacomini, K. M., Nelson, W. L., Pershe, R. A., Valdevieso, L., Turner-Tamayasu, K., Blaschke, T. F. *In vivo* interactions of the enantiomers of disopyramide in human subjects. *J. Pharmacokin. Biopharmacol.* 1986, *14*, 335–356.
91. Soudijn, W. Advantages and disadvantages in the application of bioactive racemates or specific isomers as drugs. In *Stereochemistry and Biological Activity of Drugs* (Ariëns, E. J., Soudijn, W., Timmermans, P. B. M. W. M., Eds). Blackwell Scientific Publications: Oxford, 1983, pp. 89–102.
92. Baader, E., Bartmann, W., Beck, G., Bergmann, A., Granzer, E., Jendrella, H., Kerekjarto, B.v., Kesseler, K., Krause, R., Paulus, E., Schubert, W., Wess, G. Rational approaches to enzyme inhibitors: new HMG-CoA reductase inhibitors. In *Trends in Drug Research* (Claassen, V., Ed.). Elsevier: Amsterdam, 1990, pp. 49–71.
93. Wermuth, C. G. Aminopyridazines – an alternative route to potent muscarinic agonists with no cholinergic syndrome. *Il Farmaco* 1993, *48*, 253–274.
94. Bhagwat, S. S., Gude, C., Cohen, D. S., Dotson, R., Mathis, J., Lee, W., Furness, P. Thromboxane receptor antagonism combined with thromboxane synthase inhibition. 5. Synthesis and evaluation of enantiomers of 8-{[(4-chlorophenyl)sulfonyl]amino}-4-(3-pyridinylalkyl)octanoic acid. *J. Med. Chem.* 1993, *36*, 205–210.
95. Yamamura, S., Oda, K., Mizoguchi, T., Saito, S., Iwasawa, Y., Kiyomoto, A. Study on the structure-activity relationships of adrenergic, beta-mimetic benzylamine derivatives. V. 9-aryl-1*H*-2,3,7,8,9,10-hexahydro-benzo[d,e]quinolines. *Chem. Pharm. Bull.* 1979, *27*, 858–869.
96. Stoll, A. Über Ergotamin. *Helv. Chim. Acta.* 1945, *28*, 1283–1308.
97. Baetz, J. *Scientific Director, Egic Laboratories, Paris*. In Personal Commmunication: Paris, 1982.
98. Knabe, J. Synthetische Enantiomere als Arzneistoffe. *Pharmazie in unserer Zeit* 1995, *24*, 324–330.

Chapter 27

Multi-target Drugs: Strategies and Challenges for Medicinal Chemists

Richard Morphy and Zoran Rankovic

I. INTRODUCTION
II. STRATEGIES FOR LEAD GENERATION
III. MAIN AREAS OF FOCUS IN DML DISCOVERY (1990–2005)
A. SERT-Plus DMLs for Depression
B. Dopamine D_2-Plus DMLs for Schizophrenia
C. DMLs Targeting the Angiotensin System for Hypertension
D. Histamine H_1-Plus DMLs for Allergies
E. AChE-Based DMLs for Alzheimer's Disease
F. PPAR-Based DMLs for Metabolic Disease
G. DMLs that Inhibit Multiple Kinases for Treating Cancer
H. DMLs Targeting the Arachidonic Acid Cascade
I. Mu-Opioid-Plus DMLs for Treating Pain
IV. OPTIMIZATION OF THE ACTIVITY PROFILE AND WIDER SELECTIVITY
V. THE PHYSICOCHEMICAL CHALLENGE
VI. SUMMARY
REFERENCES

When the whole is greater than the sum of the parts

I. INTRODUCTION

Historically, the compounds produced by medicinal chemists were screened by *in vivo* pharmacologists in whole animal models of disease. This approach provided a means to identify, in a single test, compounds that exhibited a rare combination of desirable pharmacokinetic (PK) and pharmacodynamic (PD) properties. The downside of this approach was that the animal model was essentially a "black box," so when compounds were inactive, it was unclear whether this was because they no longer interacted with the PD target(s), or whether they had failed to reach the required site of action due to poor pharmacokinetics. Often, the molecular targets driving both the desired PD effect and any detrimental side effects were unclear so a rational, reductionist approach to drug discovery was impossible. In the latter decades of the 20th century, the drug discovery paradigm became cemented in a "one-target-one-disease" philosophy, increasingly dominated by *in vitro* high throughput screening (HTS) technologies. Many successful drugs, that are selective for a single target, have emerged from this strategy, but despite the best efforts of drug discoverers, many diseases remain inadequately treated by such an approach. Recent evidence suggests that the main causes of failure of compounds in the clinic are now lack of efficacy and poor safety.[1] Since agents that modulate multiple targets simultaneously (polypharmacology) have the potential to enhance efficacy or improve safety relative to drugs that address only a single target, it is not surprising that this area is attracting the attention of increasing number of drug discoverers.[2,3,4]

There are three distinctly different approaches to multi-target therapy (Figure 27.1). Traditionally, clinicians have treated unresponsive patients by combining therapeutic mechanisms with cocktails of drugs. Most frequently the cocktail is administered in the form of two (or more) individual tablets (scenario A).[5,6] However, the benefits of this

Copyright © 2008, Elsevier Ltd
All rights reserved.

FIGURE 27.1 The three different approaches to multi-target therapy (polypharmacology).

approach are often compromised by poor patient compliance, particularly for treating asymptomatic diseases, such as hypertension.[7] Recently, there has been a move toward fixed dose combination (FDC) drugs, whereby the two (or more) agents are co-formulated in a single tablet to make dosing regimes simpler and thereby improve patient compliance (scenario B).[8,9] An alternative strategy is to develop a multiple ligand which is a single chemical entity that is able to modulate multiple targets simultaneously (scenario C).[2]

Across the pharmaceutical industry, the FDC approach is increasingly providing an attractive opportunity for enhancing R&D output.[10] Several FDCs are very successful commercially. Vytorin, combining the cholesterol absorption inhibitor, ezetimibe with the statin, simvastatin, for treating hypercholesterolemia had sales of 2 billion dollars in 2006. Advair, combining a glucocorticoid, fluticasone, with a long-acting bronchodilator, salmeterol, for treating asthma had sales of 6.5 billion dollars in 2006. However, there can be significant risks involved in the development of FDCs. There is the commercial uncertainty arising from the risk that clinicians might still prefer prescribing combinations of existing monotherapies that may offer greater dose flexibility and lower cost treatment in the case of generic drugs. This is illustrated by the sales of the hypertension/hyperlipidemia FDC, Caduet, being 370 million dollars in 2006 compared with multi-billion dollar sales for the individual drugs, atorvastatin and amlodipine.[10] Differences in the relative rates of metabolism between patients can produce highly complex PK/PD relationships for FDCs, leading to unpredictable variability between patients and necessitating extensive and expensive clinical studies.

Compared to FDCs, the multiple ligand approach has a profoundly different risk–benefit profile (Table 27.1). A downside is that it is significantly more difficult to adjust the ratio of activities at the different targets. However, this increased complexity in the design and optimization of such ligands, occurs at an earlier and therefore less expensive stage of the drug discovery process. The risks and costs of developing multiple ligands is in principle no different to the development of any other single entity. Another advantage of a single chemical entity is a lower risk of drug–drug interactions compared to cocktails or FDCs.[11]

TABLE 27.1 Risk–Benefit Profile of FDCs and Multiple Ligands

Risks/benefits	FDCs	Multiple ligands
Patient compliance	Improved when compared to drug cocktails	Improved when compared to drug cocktails
PK/PD relationship	Often highly complex PK/PD correlation that requires sophisticated formulation solutions	Single chemical entity – generally no issues
Drug–drug interactions	Increased risk of drug–drug interactions	Risk similar to any other single compound entity
Titration of activities	Possible, but may be difficult and costly to develop – requires full clinical development, production and marketing of a series of dose combinations	Not possible
R&D challenges	Potentially fast progress toward proof-of-concept. However, clinical development can be complicated by the requirement to demonstrate the superiority of combination versus individual agents, as well as potentially increased risk of drug–drug interactions and formulation issues	Can be challenging to design a multiple ligand with the required ratio of activities and adequate selectivity at the discovery stage. However, the development program and regulatory approval process is the same as for a standard NCE
Intellectual property	Patent life of old drugs can be prolonged when combined with a new drug	Standard NCE position

Several drugs currently on the market have been found to have activity at more than one target. In some cases, this non-selective activity serendipitously increases efficacy whereas in others it is associated with side effects. Whereas for these historical drugs, the multiple activity was not designed, a recent trend has been to deliberately and rationally design ligands that act selectivity on multiple targets (*selectively non-selective* drugs). Numerous terms have been used to describe such ligands, dual ligand, heterodimer, promiscuous drug, pan-agonist and triple blocker

being just a few of many examples. The complexity and inconsistency of this nomenclature has partly obscured overall developments in this field so to improve communication and awareness, a common term, designed multiple ligands (DMLs), has recently been introduced.[2]

In a number of disease areas, drug discoverers have followed a three-stage evolutionary journey, from a non-selective drug with undesirable side effects, to a target-selective ligand with a safer profile, and onward toward a *selectively non-selective* DML which attempts to provide a more optimal balance of efficacy and safety. An example of a non-selective ligand is the atypical anti-psychotic drug, Clozapine, which shows antagonist activity at multiple aminergic G-protein-coupled receptors (GPCRs). To circumvent the side effects of Clozapine, a number of ligands that are selective for single receptors targeted by Clozapine were developed, such as dopamine D_4 and serotonin 5-HT_{2a} antagonists, but these lacked sufficient efficacy in the clinic.[12] Research then shifted toward DMLs, such as the dual D_2/5-HT_{2a} antagonists.[13,14] Non-selective tricyclic antidepressants such as Amitryptyline were superseded by selective serotonin (5-HT) transporter inhibitors (SSRIs) which increased safety, but had a slow onset of action and lacked efficacy in some patients. Dual serotonin and norepinephrine (NA) re-uptake inhibitors (SNRIs) were subsequently developed with the hope of addressing these deficiencies.[15] The same trend is observed in the area of nonsteroidal anti-inflammatory drugs (NSAIDs), starting from non-selective agents such as aspirin, to selective cyclooxygenase-2 (COX-2) inhibitors and then to dual COX-2/5-lipoxygenase (5-LOX) inhibitors.[16] Similarly, for the treatment of asthma, non-selective adrenergic agonists (e.g. epinephrine) have been replaced by selective β_2-adrenoceptor agonists such as salbutamol, with a significantly improved therapeutic window. Most recently, dual M3 antagonist/β_2 agonist and D_2/β_2 agonist have been developed.[17,18]

II. STRATEGIES FOR LEAD GENERATION

As with single target projects, medicinal chemists have access to a number of different ways of generating the chemical matter with which to commence a DML project. Conceptually, there are two quite different methods of generating lead compounds, *screening* approaches that rely largely upon serendipity and *knowledge-based* approaches that exploit information either from the general literature or proprietary information from within an organization (Figure 27.2).

The screening of either diverse or focused compound libraries can deliver a single molecule that has at least minimal activity at each of the targets of interest. To date, there have not been many reported examples of DMLs derived via the HTS approach. This could be due to the fact that HTS has only become the *de rigeur* method of lead generation in the last decade or so and there is an inevitable time lag to publication. Other factors could be the logistical complications of screening against multiple targets in parallel or to an inherently low probability of detecting a compound with a multiple profile of therapeutic interest from screening compounds at random. Due to the large number of compounds typically involved in diversity-based screening, they will usually be screened first at one target of interest and any actives will then be filtered on the basis of activity at the other target(s). Even if activity is observed for the second target, usually the balance of affinities is non-optimal so the activity ratio must be adjusted during optimization.

FIGURE 27.2 The screening of diverse or focused libraries can deliver a compound that has at least minimal activity at each target of interest. However, it is unlikely that the hit compound has the optimal affinity for all targets so the profile must be balanced during optimization. Alternatively, screening might deliver a compound that in addition to the desired activities has undesired activities and these must be designed out during optimization.

Compared to HTS, there are many more examples in the literature of the screening of focused libraries of compounds selected from single target projects or using prior knowledge of the targets. In focused screening, compound classes that are already known to be active against one of the targets of interest are screened against another target. For example, DMLs for kinase targets are usually discovered serendipitously through the cross-screening of ligands from selective kinase programes against other kinases. In addition to the desired activities, screening frequently provides hit compounds that bind to other targets. To minimize the risk of side effects the medicinal chemist will need to "design out" these undesired activities.

The second most common lead generation strategy reported in the literature is a knowledge-based approach known as *framework combination*. It starts with two compounds, one of which binds with high selectively to one of the targets and the other with high selectively to the other target. In this case, the first goal is to "design in" both activities

FIGURE 27.3 Framework combination is a knowledge-based approach to generating DMLs. There is a continuum in the degree of merger of the frameworks of the target-selective starting ligands. In linked DMLs, the frameworks are connected via a definable linker, which in some cases is designed to be cleaved *in vivo* to release 2 independently acting drugs. In fused DMLs, the frameworks are directly attached and in the commonest form of DML, the frameworks are merged together.

TABLE 27.2 Features of the Screening and Framework Combination Approaches to DMLs

Screening approach	Framework combination approach
Can provide novel chemotypes not found in selective ligands	Can be employed where selective ligands are known for each target
Can provide ligands with complex profiles unavailable by framework combination	Can employ existing SAR knowledge from selective ligand projects to assess feasibility and facilitate optimization
Can provide ligands with improved physicochemical and pharmacokinetic properties compared to framework combination	Can readily provide linked DMLs for use as i.v. drugs or biochemical tools, even for two targets with incompatible pharmacophores
Chance of success is low for unrelated targets	Can be difficult to incorporate a 2nd activity whilst retaining the 1st activity and good physicochemical properties

into a single lead molecule by combining the frameworks (and the underlying pharmacophores) of the two selective molecules (Figure 27.3). The intellectual elegance of the framework combination stems from the fact that often a wealth of structure–activity relationships (SAR) knowledge is on hand from previous selective ligand projects which can be used to guide the optimization process.

DMLs arising from framework combination can be viewed as *linked*, *fused* or *merged* depending on the degree to which the frameworks have been integrated (Figure 27.3). In linked DMLs (conjugates), the molecular frameworks are not at all integrated and there is a distinct linker group between the two components that is not found in either of the selective ligands. This linker is usually intended to be metabolically stable so that the single compound is capable of interacting with both targets, albeit different ends of the molecule may be responsible for the activity at the different targets. Some linked DMLs contain a cleavable linker that is designed to be metabolized to release two ligands that interact independently with each target. This scenario represents a half way point between a true DML and a FDC.

If the frameworks are essentially touching, so there is neither a discernable linker nor any framework overlap, the DML can be viewed as fused. In the most common and most sought after type of DML, the frameworks are merged together by taking advantage of commonalities in the structures of the starting compounds. Medicinal chemists will normally aspire to maximize the degree of overlap in order to produce smaller and simpler molecules. The degree of framework combination for the examples reported in the literature forms a continuum, with high molecular weight (MW) DMLs with lengthy linker groups at one extreme, and small DMLs with highly merged frameworks at the other.

The screening and framework combination approaches to lead generation have various advantages and disadvantages that influence which one is best applied to a particular project (Table 27.2). Indeed, given the added challenges of multiple ligand projects in general, it would make sense to employ both strategies if feasible to increase the overall chance of success. A major advantage of the screening approach is that you start from a compound that already has multiple activities built in, albeit these may be quite weak. Screening can add particular value if there is a lack of selective ligands for the targets of interest or little of the SAR information required for a knowledge-driven approach. Screening can deliver novel and unexpected chemotypes, sometimes providing hits for unusual target combinations that span unrelated receptor families. Since the framework combination strategy almost invariably produces dual ligands, discovering ligands that bind to more than two targets, usually demands that a screening approach is followed. Screening can also provide ligands with improved physiochemical and PK properties compared to framework combination (see Section V on physicochemical properties).

In the case of framework combination, incorporating a second activity into a compound that has no measurable affinity for that target, whilst retaining affinity for the original target, is by no means an easy task. However, many literature examples testify to the fact that it can often be achieved by effectively leveraging SAR knowledge from historical selective ligand projects. Compared to screening, framework combination can provide rapid entry to conjugate molecules that can be used as intravenously (i.v.) administered drugs or biochemical tools, even for targets that are very different at the pharmacophore level

FIGURE 27.4 Secondary activities that have been added to a clinically validated primary target in an effort to enhance efficacy and reduce side effects. An example of a drug selective for the primary target is shown in red. Abbreviations: 5-HT_{1A}, 5-HT_{1A} receptor; 5-HT_{1D}, 5-HT_{1D} receptors; 5-HT_{2A}, 5-HT_{2A} receptor; 5-LOX, 5-lipoxygenase; alpha2, alpha2 adrenergic receptor; ACE, angiotensin-converting enzyme; AChE, acetylcholinesterase; AT_1, angiotensin-1 receptor; AT_2, angiotensin-2 receptor; B2, bradykinin-2 receptor; COX-2, cyclooxygenase-2; delta, delta-opioid receptor; D_2, dopamine-2 receptor; D_4, dopamine-4 receptor; DAT, dopamine transporter; DHFR, dihydrofolate reductase; ECE, endothelin-converting enzyme; EGFR, epidermal growth factor receptor; ET_A, endothelin-A receptor; FGFR1, fibroblast growth factor receptor 1; GR, glucocorticoid receptor; H_1, histamine-1 receptor; H_3, histamine-3 receptor; kappa, kappa-opioid receptor; I2, Imidazoline-2 receptor; LTD_4, leukotriene D4 receptor; mu, mu-opioid receptor; MAO, monoamine oxidase; NEP, neutral endopeptidase; NET, norepinephrine transporter; NK1, neurokinin-1 receptor; NO, nitric oxide; PAF, platelet-activating factor receptor; PDGFRb, Platelet-derived growth factor receptor beta; PPAR, peroxisome proliferator-activated receptor; SERT, serotonin transporter; TxA2, thromboxane-A2 receptor; TxS, thromboxane-A2 synthase; VEGFR-1, vascular endothelial growth factor receptor-1; VEGFR-2, vascular endothelial growth factor receptor-2.

(see Section V on physicochemical properties). The chance of success with a random screening approach would be expected to rapidly diminish as the targets in a combination become more dissimilar.

III. MAIN AREAS OF FOCUS IN DML DISCOVERY (1990–2005)

Historically, the most common disease areas for DML projects have been psychiatry, neurodegeneration, oncology, as well as metabolic, cardiovascular and allergic disease. A common theme is to focus on a primary target, which has previously been well validated in the clinic for a given disease, and then add one or more secondary activities in an effort to enhance efficacy and reduce side effects. For example, there are a large number of reported combinations containing the 5-HT transporter (serotonin transporter (SERT)) for depression, the histamine H_1 receptor for allergy and the vascular endothelial growth factor receptor-2 (VEGFR-2) kinase for cancer (Figure 27.4). A relatively small number of target combinations have predominated in terms of their percentage share of the total number of publications in the literature between 1990 and 2005. The six most commonly reported combinations were as follows:

1. Angiotensin-converting enzyme/neutral endopeptidase (ACE/NEP) for hypertension.
2. Cyclooxygenase-2/5-lipoxygenase (COX-2/5LOX) for inflammatory pain.

3. Thromboxane-A2 receptor/thromboxane-A2 synthase (TxA2/TxS) as anti-platelet agents.
4. Serotonin transporter/5-HT1A receptor (SERT : 5-HT1A) for depression.
5. Neurokinin-1 receptors (NK1/NK2) for asthma.
6. Peroxisome proliferator-activated receptors (PPARalpha/PPARgamma) for diabetes.

A. SERT-Plus DMLs for Depression

Depression is associated with reduced levels of serotonin (5-HT) in the brain. Drugs that inhibit the re-uptake of 5-HT, such as fluoxetine, have been used clinically for many years. In an attempt to address the deficiencies of SSRIs as anti-depressants, in terms of efficacy or time of onset, SERT inhibition has been supplemented with activity at a secondary monoamine target, such as the 5-HT_{1A}, 5-HT_{1D}, alpha2, NET or DAT (Figure 27.4). The delayed onset time for SSRIs has been attributed to the need for 5-HT_{1A} autoreceptors to become desensitized by sustained SERT blockade. By mimicking this desensitization with a 5-HT_{1A} antagonist, the onset time might be accelerated. The following three examples of dual 5-HT_{1A}/SERT blockers illustrate how the various lead generation methods of screening and framework combination have been employed in this area.

Van Niel *et al.* designed a focused screening library based on the 3-aryloxy-2-propanolamine scaffold found in the 5-HT_{1A} antagonist, pinadol **1** (Figure 27.5).[19] The variations at the amine and phenol positions included privileged structures, as well as fragments reported to have affinity for either 5-HT_{1A} or SERT. The SAR around the indole region was reasonably tolerant for both targets but the only amine group that provided reasonable SERT inhibition was a spiro-piperidine **2**. This compound provided balanced inhibition as well as good oral exposure (F = 65%) and brain penetration in the rat.

Using a framework combination approach, compounds with dual 5-HT_{1A}/SERT activity were designed by Mewshaw *et al.*, starting from a template known to possess

FIGURE 27.5 SERT-plus DMLs for depression.

robust SERT activity **3** and adding 5-HT_{1A} features in the form of the aryloxyethyl group found in **4** (Figure 27.5).[20] The presence of a basic nitrogen was the common pharmacophoric feature that allowed the two frameworks to be merged to give **5**. The degree of framework overlap in this example (shown in magenta) is quite extensive helping to produce a DML with a relatively low MW of 358 Da.

A high throughput screen provided a multiple ligand **6** with a surprising combination of activities at a peptide GPCR, the neurokinin NK_1 receptor, and a monoamine transporter, SERT.[21] The two targets in this combination have individually generated much interest for treating depression. Whilst a NK_1-selective ligand, MK-869, gave disappointing results in clinical trials for depression, it is interesting that the target is now being pursued in combination with a clinically validated target, SERT. Although the hit **6** had only modest activity, systematic optimization of each aromatic moiety in turn provided a more potent compound with a balanced activity at both targets **7** (Figure 27.5). An aryl ether moiety was introduced to reduce lipophilicity.

B. Dopamine D_2-Plus DMLs for schizophrenia

D_2-selective antagonists, such as haloperidol, are efficacious against the positive symptoms of schizophrenia, for example, hallucinations and delusions. However, unlike the atypical anti-psychotic drugs such as Clozapine, they do not address the negative symptoms (such as social withdrawal) and cause extrapyramidal side effects (EPS) such as Parkinsonism. The aim of combining D_2 antagonism with activity at other targets (Figure 27.4) is to mimic the advantages of Clozapine in terms of efficacy without producing the disadvantages such as weight gain. One of a number of possible explanations for Clozapine's atypical profile is its higher antagonist affinity for the 5-HT_2 receptor than for the D_2 receptor. This observation lead to the so-called D_2/5-HT_2 ratio hypothesis whereby agents with $>$10-fold selectivity for 5-HT_2 over D_2 were sought.

Using a framework combination approach, the structure of the endogenous agonist for the D_2 receptor, dopamine **8**, was fused with a large lipophilic group from the 5-HT ligand **9** (Figure 27.6).[22] This transformed the D_2-agonist activity of the endogenous ligand into an antagonist. This DML is of the fused type since there is only one nitrogen atom overlap between the frameworks of the starting compounds. Fused DMLs can have an undesirably high MW if the starting compounds are already quite large, but since the starting ligands are much smaller in this case, the resulting DML has a relatively low MW of 371 Da. Various heterocyclic groups were selected containing hydrogen bonding groups that might mimic the phenolic interaction, such as the oxindole found in **10**. Further optimization involved replacing the naphthyl group by a 1,2-benzisothiazole group **11**, which provided D_2 blockade comparable in potency to the typical anti-psychotic haloperidol, together with a desirable D_2/5-HT_2 ratio of 11, comparable to the atypical agent, clozapine.[23] The D_2/alpha1 ratio of 0.44 for **11** is substantially lower than that for clozapine, suggesting the former should have a lower propensity to cause orthostatic hypotension. The ratio hypothesis was validated by clinical studies and **11** (Ziprasidone) was launched in 2001 by Pfizer for the treatment of schizophrenia.

It has also been hypothesized that the unique profile of Clozapine in treating psychosis might be due to a precise ratio of D_2 and D_4 receptor affinities, with higher affinity required at D_4 than D_2. Zhao *et al.* tried to reproduce this exact ratio with the goal of obtaining D_4 affinity of less than 10 nM and D_2 affinity of less than 200 nM.[24] They started from a non-selective D_2/D_4 compound **12,** with undesired alpha-1 affinity, discovered via a screening approach (Figure 27.6). Introduction of a methyl group in the 2-position of the indoline ring gave an improvement in D_2 activity, **13** and also good selectivity against a diverse range of other targets including alpha-1.[25] It displayed activity in an *in vivo* test of psychosis, the inhibition of amphetamine-induced locomotor activity and showed low activity in a catalepsy test, suggesting a low propensity to cause EPS. The behavioral data for this dual antagonist provided support for the "D_2/D_4 ratio" hypothesis, although the approach still needs clinical validation.

To maximize the efficacy and safety profile of an anti-psychotic drug, much evidence now suggests that it is necessary to address more than 2 receptors. Using a screening approach, Garzya *et al.* discovered a molecule **14** that had five activities regarded as being critical for an anti-psychotic drug, blocking the D_2, D_3, 5-HT_{2A}, 5-HT_{2C} and 5-HT_6 receptors.[26] Careful optimization produced a DML **15** with the optimal balance of affinities.

C. DMLs Targeting the Angiotensin System for Hypertension

The vasoconstricting peptide, angiotensin II, is a principal component of the renin-angiotensin system (RAS), a hormone system that helps to regulate blood pressure and extracellular volume in the body. ACE inhibitors and angiotensin-1 receptor antagonists (AT_1), such as Captopril and Losartan, respectively, have gained widespread acceptance for the treatment of hypertension and congestive heart failure (Figure 27.7). It has been postulated that DMLs, such as dual ACE/NEP inhibitors or dual AT_1/endothelin receptor 1 (ET_A) antagonists may produce a beneficial synergistic effect in the management of hypertension and congestive heart failure (Figure 27.7).

A rational approach toward one of the earliest ACE/NEP dual inhibitors, dipeptide **18**, demonstrates how a good understanding of the pharmacophore requirements for

FIGURE 27.6 Dopamine D_2-plus DMLs for schizophrenia.

the targets is highly desirable when designing multiple ligands. For example, knowledge that NEP favors a hydrophobic substituent in the S1′ pocket, preferably a benzyl group such as the one present in the NEP selective inhibitor **16**, whereas ACE is more tolerant in this region but strongly favors a proline residue at P_2', as in ACE selective inhibitor captopril **17**, was instrumental in the design of **18** (Figure 27.7).[27] In order to further improve the *in vitro* and *in vivo* potency of **18**, a range of diverse constrained analogs were designed, drawing extensively from the SAR generated around selective inhibitors. A particular challenge for this approach was a relatively tight SAR for NEP, which

FIGURE 27.7 DMLs targeting the Angiotensin system for hypertension.

fortuitously was counterbalanced by a remarkably flexible SAR for ACE. The optimization efforts led to the discovery of a 7,6-bicyclic oxazepinone series which produced omapatrilat **19**, a potent ACE/NEP inhibitor displaying a good PK profile and efficacy *in vivo*.[28]

In another example of a merged DML, Murugesan *et al.* were interested in simultaneous blockers of AT_1 and ET_A receptors since a combination of the AT_1 selective antagonist, Losartan, and the ET_A/ET_B selective antagonist, SB-290670, produced an additive reduction in blood pressure compared to either drug alone. Fortuitously, the selective AT_1 and ET_A antagonists, **20** and **21** respectively, both contained a biaryl core (Figure 27.7) and the heterocycle in the 4′-position of the biaryl, required for AT_1 activity, was

tolerated by ET_A, albeit with reduced affinity.[29] The acylsulphonamide moiety was found to be a carboxylic acid bioisostere that was suitable for both receptors. By introducing a new substituent in the C2′-position of the biaryl **22**, a balanced dual activity at AT_1 and ET_A receptors was obtained. Compound **22** has a high MW of 660, reflecting the size of the starting compounds used for framework combination. Nonetheless, good oral bioavailability was observed for **22** in rats (F = 38%).

D. Histamine H_1-Plus DMLs for Allergies

Histamine is a primary mediator of the systemic inflammatory response to allergens in humans. H_1-antagonists, such as loratidine, have found widespread utility in the treatment of hay fever and other allergic reactions. However, selective H_1-antagonists have been largely ineffective for the treatment of asthma.[30] Almost all the H_1-antagonists that show some efficacy against asthma are reported to possess additional activities, suggesting that other chemical mediators are also involved in its pathogenesis. As a result, various groups have aimed to produce DMLs combining H_1 antagonism with a range of additional activities (Figure 27.8).

For example, the TxA_2 receptor has also been linked to allergic disease. Although both the H_1 and TxA_2 receptors are GPCRs, they might be expected to possess very different binding sites given that the endogenous ligand for the former is a small polar amine **23** and for the latter a lipophilic acid **24** (Figure 27.8). It might be anticipated that rationally "designing in" activity for targets with highly dissimilar endogenous ligands might be particularly difficult. However, a large number of recent examples show that this need not necessarily be a barrier to the discovery of a DML. It was observed by Ohshima *et al.* that molecules with a common benzoxepine scaffold, the selective H_1-antagonist, **25**, and the TxA_2R antagonist **26,** bound to both targets.[31] The tertiary amine group in **27** successfully mimicked the benzimidazole moiety that was known to be crucial for the TxA_2 activity of **26**. Compound **27** was active at both GPCRs, albeit with rather different binding affinities, as well as being selective over related GPCRs.

Perhaps an even more striking example is provided by the dual H_1-antagonist/5-lipoxygenase (5-LOX) inhibitor **31** that inhibits an enzyme that oxidizes highly lipophilic arachidonic acid **28**, whilst also antagonizing a GPCR that binds highly polar histamine **23**. The starting points for framework combination were the selective H_1-antagonist

FIGURE 27.8 Histamine H_1-plus DMLs for allergies.

29 and the 5-LOX inhibitor **30** (Figure 27.8).[32] The strategy took advantage of the flat SAR around the basic nitrogen of the anti-histamine to introduce a butynyl-hydroxyurea group into **31** required for 5-LOX inhibition.

E. AChE-Based DMLs for Alzheimer's Disease

Alzheimer's disease is associated with a progressive loss of cholinergic neurons in the brain that results in memory disturbances and cognitive dysfunction. One strategy for the treatment of Alzheimer's patients has been the use of acetylcholinesterase (AChE) inhibitors such as rivastigmine to enhance cholinergic activity in the central nervous system, although the result of such treatment is far from satisfactory in many patients. In an attempt to increase efficacy, AChE inhibition has been combined with SERT and monoamine oxidase (MAO) activity.

Kogen *et al.* described work toward a dual AChE and 5-HT transporter (SERT) inhibitor, **35**.[33] This work also represents another example of a DML that crosses different proteomic families. A notable feature of this work is the elegant use of biostructural information to guide the combination of the frameworks of the starting compounds. A model of the active site of AChE showed that the AChE-selective inhibitor, rivastigmine **32**, possessed only three elements of the proposed AChE pharmacophore, lacking a fourth hydrophobic binding site (Figure 27.9). If the phenoxyethyl motif from the SERT blocker, fluoxetine **33**,

FIGURE 27.9 AChE-based DMLs for Alzheimer's disease.

could provide this hydrophobic interaction, potency should be improved relative to rivastigmine. Thus, hybridization of the two inhibitors, followed by optimization of the carbamate and phenoxy substituents, provided a dual inhibitor **34**. Conformational constraint using a seven-membered ring then gave a compound **35,** with potent and balanced inhibition at the two diverse targets.

Youdim *et al.* described dual AChE/MAO inhibitors as another approach to the treatment of Alzheimer's disease.[34] The structural framework of rivastigmine **32** was in this case combined with a selective MAO-B inhibitor rasagiline **36**, yielding a dual inhibitor ladostigil **37** (Figure 27.9). Reported SAR around this compound indicates that carbamate and propargylamine groups are key pharmacophoric elements responsible for the AChE and brain MAO inhibition, respectively. Ladostigil has shown efficacy in rhesus monkey cognition and neuroprotection models.[35]

F. PPAR-Based DMLs for Metabolic Disease

The realization that the fibrate and glitazone classes of drugs, used to treat dyslipidemia and type-2 diabetes, respectively, exert their effects through activation of PPARalpha and PPARgamma, respectively, lead to the development of selective ligands for each of the PPAR receptor subtypes. However, findings suggesting that insulin resistance, dyslipidemia and obesity can be seen as components of a complex mixture of abnormalities known as "metabolic syndrome" have stimulated interest in developing dual PPARalpha and PPARgamma agonists.[36]

An interesting combination of screening and structure-based approaches was reported by Xu *et al*.[37] Their screening efforts resulted in identification of carboxylic acid **38** containing a bulky lipophilic group in the α-position as a moderate dual PPARα/γ agonist (Figure 27.10). The fact that **38** exhibited activity for both targets despite of lacking the lipophilic "tail" characteristic of PPAR ligands suggested that the α-benzyl group might improve the binding affinity of **39**, a well balanced but weak dual agonist. The α-benzyl derivative **40** indeed showed improved activity at both PPARα and PPARγ. Shifting the oxygen adjacent to the quaternary stereogenic center in **40** to the alternative benzylic position provided a significantly more potent dual agonist **41**.

G. DMLs that Inhibit Multiple Kinases for Treating Cancer

Systematically targeting multiple kinases is currently of great interest in the fight against various forms of cancer. Whilst there are a number of literature examples of the framework combination approach being applied to GPCRs, transporters, nuclear receptors, proteases and oxidases, we have so far identified no such examples for kinases. The absence of the framework combination approach is probably due to the fact that obtaining selective ligands for kinases is still a major challenge and this step precedes the rational "designing in" of multiple activities, driven by knowledge of the selective ligand SARs. The most feasible strategy for designing multi-kinase inhibitors is focused screening to identify a non-selective inhibitor and then attempting to "design out" undesired kinase activities.

The first kinase inhibitor to be developed for clinical use was imatinib **42**, first marketed in 2001 for chronic myelogenous leukemia (CML). The clinical effectiveness of imatinib for the treatment of CML is now thought to be due to its multi-kinase activity, inhibiting PDGFR and c-KIT, in addition to its well known activity as a Bcr-Abl kinase inhibitor. Resistance to imatinib can become a problem due to mutations in the Abl gene.[38,39] Dual Src/Abl inhibitors are currently of interest for the treatment of CML

FIGURE 27.10 Dual PPARα/γ agonist for treating metabolic disease.

in patients who are resistant to imatinib. Whereas imatinib itself has no measurable activity against Src, Boschelli *et al.* use a focused screening approach to identify an inhibitor **43** with dual Src/Abl activity (Figure 27.11).[40] They found a very close correlation between the Src and Abl SARs, reflecting the close homology of these kinases.

Soon after the first generation of kinase inhibitors appeared on the market, the first multi-kinase inhibitors that were intentionally designed to have a particular profile have been introduced for cancer treatment. For example, building on the success of the VEGFR-1/VEGFR-2 blocking antibody, Avastin, a small molecule VEGFR-2 and PDGFRβ inhibitor, sunitinib **44** was introduced in 2006.[41] Similarly, the selective HER2 blocking antibody, herceptin, was followed to the market by the dual epidermal growth factor receptor (EGFR)/erbB2 inhibitor, lapatinib **45** in 2007.

To enhance efficacy, various other kinase targets with a potential role in angiogenesis and tumor growth have been combined with VEGFR-2. Becknell *et al.* developed a dual TIE-2/VEGFR-2 inhibitor **46** by cross-screening molecules from an earlier selective VEGFR-2 project (Figure 10.11).[42] By inhibiting angiogenesis, such multi-kinase agents are showing promise for the treatment of solid tumors, for example, in the breast and kidney that were previously highly resistant to therapy.

H. DMLs Targeting the Arachidonic Acid Cascade

NSAIDs exert their anti-inflammatory effect by inhibiting cyclooxygenases-1 and -2 (COX-1 and COX-2), key enzymes in prostaglandin (PG) biosynthesis from arachidonic acid.[43] Side effects often limit their use, in particular gastrointestinal ulcerogenic activity and renal toxicity.[44] A "single target" strategy resulted in the development

FIGURE 27.11 DMLs that inhibit multiple kinases for treating cancer.

of COX-2 inhibitors like celecoxib **47** (Figure 27.12). So-called selective drugs have often been found to possess unexpected polypharmacological profiles and can therefore provide attractive starting points for a DML project. For example, the COX-2 inhibitor, celecoxib **47,** was reported to potently inhibit carbonic anhydrases, hCA II and IX.[45]

A number of DML approaches targeting multiple key proteins involved in the arachidonic acid biosynthesis, have been reported including COX/5-lipoxygenase (5-LOX), 5-LOX/TxA_2 and TxA_2/TxA_2 synthase (TxS). In particular, the combination of 5-LOX with COX-2 inhibitory activity has attracted much attention in recent years.[46] Henichart *et al.* reported a dual COX-2/5-LOX inhibitor designed by fusing the tricyclic moiety present in celecoxib **47** with an aryltetrahydropyran moiety from the 5-LOX inhibitor, ZD-23138 **48** (Figure 27.12).[47] Both starting compounds were completely inactive at the second target, but the resulting DML **49** possessed nanomolar potencies for both enzymes.

Nitric oxide-releasing NSAIDs, such as NO-aspirin **50** (NCX-4016) and the ibuprofen derivative **51,** contain a cleavable ester linker to a nitric oxide-releasing moiety (Figure 27.12).[48,49] It was hoped that this dual activity would translate into a superior anti-inflammatory and anti-thrombotic profile in patients with cardiovascular diseases, whilst sparing the gastrointestinal tract.

It might seem that bridging two diverse proteomic superfamilies, the GPCRs and the oxidases, to design dual TxA_2/TxS inhibitors would be very challenging. However, a good understanding of target pharmacophore requirements proved to be very helpful in one example.[50] The essential structural features of TxS inhibitors like isbogrel **52** are a pyridine nitrogen and carboxylic group separated by between 8.5 and 10 Å (Figure 27.12). Since TxS is a cytochrome P-450 enzyme, it was postulated that the pyridine moiety forms a complex with the heme group of the enzyme catalytic site. A key feature of TxA_2 receptor

FIGURE 27.12 DMLs targeting the arachidonic acid cascade.

antagonists like daltroban **53** is a carboxylic acid separated by a non-specific spacer from a benzenesulphonamide group. Integration of the TxS and TxA_2 features produced compounds such as Samixogrel **54**, which showed low nanomolar activity at both targets.[51]

I. Mu-Opioid-Plus DMLs for Treating Pain

Mu-opioid receptor agonists such as morphine and fentanyl remain the gold standard treatment for severe pain. The use of these agents is limited by mechanism-dependent side effects such as euphoria, respiratory depression, sedation, tolerance and dependence. In an attempt to maintain efficacy whilst reducing these side effect liabilities, mu-opioid activity has been combined with agonist activity at delta- and kappa-opioid receptors **55**.[52] One unusual example of a DML **56**, in terms of contrasting functional activity, combines mu-opioid agonism with delta-opioid antagonism again with the aim of circumventing mu-based side effects.[53]

Montero *et al.* combined agonism at the mu-opioid and I_2-imidazoline receptors in a single molecule (Figure 27.13).[54,55] A guanidinium group from the I_2 ligand, agmatine **57**, was incorporated into the opioid, fentanyl **58**. The lead compound, **59**, possessed activity at both receptors but the activity was unbalanced, having significantly higher affinity for the opioid receptor. In this example, the frameworks of the starting compounds are slightly merged with the agmatine-derived alkyl chain replacing the aniline system in fentanyl. The identification of such a "tolerant region" for both receptors is a first key step in any DML program. The compound with an 8-carbon spacer, **60** possessed the best balance of activities.

IV. OPTIMIZATION OF THE ACTIVITY PROFILE AND WIDER SELECTIVITY

No matter whether the lead compound is obtained by a screening and framework combination approach, the compound will usually lack the optimal ratio of activities. Thus, a medicinal chemist working on a DML lead optimization project is faced with the twin challenges of balancing the desired activities at an appropriate level whilst, if necessary, removing any undesired side activities. Establishing what the desired level of modulation for each target should be for optimal efficacy and safety is not a straightforward task. Moreover, understanding the relationship between *in vivo* target modulation and activity in a simple *in vitro* test, such as receptor affinity in a recombinant cell assay, is difficult. Factors such as the distribution of the compound, whether the targets are located in different tissues, the receptor/enzyme densities and receptor reserve in different tissues, will influence the optimal balance of *in vitro* activities. Ideally, knowledge from clinical studies will guide researchers toward the optimal ratio, though for novel mechanisms of action, this clearly will not be

FIGURE 27.13 Mu-opioid-plus DMLs for treating pain.

available. In the absence of this knowledge, the aim of most historical DML projects has been to obtain the same degree of *in vitro* activity for each target, with the assumption that this will also lead to similar levels of enzyme modulation or receptor occupancy *in vivo*. Assuming a validated animal model is available, the testing of a lead candidate *in vivo* may help to clarify the required ratio of *in vitro* activities.

In the anti-depressant field, there has been a historical trend toward developing agents with both potent and balanced activity at both the serotonin (SERT) and norepinephrine (NET) transporters, starting with fluoxetine **33**, moving to venlafaxine **61** and most recently duloxetine **62** (Figure 27.14). Although classified as a dual SERT/NET blocker (SNRI), venlafaxine has a 30-fold difference in *in vitro* potency at the two transporters meaning that it behaves as a multiple ligand *in vivo* only at high doses.[56] The newer drug, duloxetine, has a more potent and balanced *in vitro* profile.[57]

A difference in the *in vitro* activities may sometimes be desirable where a different level of receptor occupancy for each target is associated with a desired pharmacological effect. The "atypical" profile of the anti-psychotic drug, clozapine, has been variously associated with its lower activity at the D_4 or 5-HT_{2A} receptors compared to the D_2 receptor. Neuroimaging studies have shown that an optimal D_2 receptor occupancy of 60–70% is sufficient to produce an atypical anti-psychotic effect, and if D_2 receptor occupancy is too high, the atypical profile can be lost even in the presence of high 5-HT_2 occupancy.[58] Several atypical anti-psychotics with low D_2/5-HT_2 binding ratios have now been introduced onto the market such as ziprasidone **11**.

As the number of targets to be balanced increases, the complexity of the task for a medicinal chemist can increase exponentially. It is therefore not surprising that the vast majority of reported DMLs are dual ligands. However, for targets which are closely related such as combinations of monoamine transporters, monoamine GPCRs, proteases or kinases, triple blockers are known. By concurrently blocking the re-uptake of dopamine as well as serotonin and NA, super mixed uptake blockers (SMUBs), such as **63**, may possess mood-elevating properties and deliver better control of depression than either SSRIs or SNRIs. One of potential limitations of the ACE/NEP dual inhibition approach for the management of hypertension and congestive heart failure is an increase in plasma levels of another vasoconstricting peptide, endothelin-I (ET-1). This might be overcome by additionally inhibiting a closely related zinc metallopeptidase, endothelin-converting enzyme-1 (ECE-1).

FIGURE 27.14 Optimization of a DML profile to enhance efficacy and safety.

Triple blockers of ACE, NEP and ECE-1, such as **64,** may produce a beneficial synergistic effect.

In addition to adjusting the ratio of activities, optimizing wider selectivity against a broad panel of targets is often required. Many publications in the DML area do not even discuss the key issue of global selectivity, so it is frequently difficult to judge whether real selectivity for the disease-relevant targets has been achieved. Animal models and subsequent clinical studies can provide essential feedback on the level of cross-reactivity that can be tolerated. In cases where a large number of closely related receptor subtypes or isozymes exist and the compound possesses more than one undesired activity, the task of achieving wider selectivity will be particularly intricate. Two therapeutic areas where this is currently a critical issue are psychiatric drugs, which frequently hit multiple monoamine GPCRs and transporters, and oncology drugs, which often hit multiple kinases. In both these areas, it can be difficult, if not impossible, to obtain absolute selectivity for the desired targets with no affinity for any off-target. This current reality has lead to a pragmatic approach whereby DMLs are developed which are deemed to be "selective enough" to be progressed into toxicity testing in animal studies.

DMLs for monoamine GPCRs and transporters frequently hit adrenergic GPCRs that are associated with cardiovascular side effects. Bonnert *et al.* successfully "designed out" adrenergic α_1 receptor activity from a dual dopamine D_2/adrenergic beta-2 (β_2) agonist[59] and Atkinson *et al.* removed adrenergic receptor β_2 activity from a 5-HT_{1A}/SERT ligand.[60] Atypical anti-psychotics often have a complex multi-receptor profile and off-target activities can be associated with undesirable side effects. In particular, cross-reactivity at the histamine H_1 receptor has been highlighted as the main cause of the weight gain caused by agents such as clozapine and olanzepine.[61] During the optimization of the anti-psychotic agent **15**, Garzya *et al.* had to balance the five desired activities (D_2, D_3, 5-HT_{2A}, 5-HT_{2C} and 5-HT_6), whilst avoiding undesired activity at the H_1 receptor, as well as at other monoamine targets, α_{1B}, M_{1-4} and β_{1-3}.[26]

Similarly, discovering multi-kinase inhibitors is complicated by the risk of inhibiting kinases that are critical to normal cellular function. At the present time, it is difficult to design an inhibitor that inhibits 2 or 3 kinases specifically whilst being inactive at all others. Unanticipated activities for even well studied inhibitors such as imatinib **42** are still being found via panel screening.[62] This information can reveal which kinases are "safe" to inhibit and which are critical to normal cellular function and should be avoided. Although the recently launched drug, Sunitinib, was designed as a dual VEGFR-2 and PDGFRβ inhibitor,[41] it was later reported to inhibit no less than eight kinases with IC_{50} values of less than 100 nM and yet has an acceptable side effect profile in man (Figure 27.11).[63] It remains to be seen whether such a pragmatic approach to kinase selectivity profiles can be extended beyond oncology to non-life threatening disease areas such as inflammation, where side effect liabilities will be particularly critical. Cardiotoxicity associated with multi-kinase inhibition is one area of possible concern.[64] Screening for multi-kinase inhibitors sometimes provides compounds with undesired off-target activity at non-kinase targets. In a recent example, activity at the hERG ion channel was successfully designed out of a multi-kinase inhibitor.[65]

There are several examples that give encouragement to the medicinal chemist that surprising activity and selectivity profiles can sometimes be achieved. The dual AChE/SERT blocker **35** possesses high selectivity over several closely related targets, including butyrylcholinesterase and the NET/DAT.[33] Similarly the COX-2/5-LOX inhibitor **49** possesses surprising selectivity for COX-2 over COX-1 and the AT_1/ET_A antagonist **22** is inactive at AT_2 and ET_B receptors.[47]

V. THE PHYSICOCHEMICAL CHALLENGE

Compared to optimizing the balance of affinities and the wider selectivity, an even greater challenge for medicinal chemists when confronted with the challenge of designing multiple ligands is to obtain physicochemical and PK properties consistent with developing an oral drug.[2] The influence of physicochemical properties on the PK behavior of orally administered drugs has been the subject of intense interest over the past few years since the publication of Lipinski's seminal work on the "Rule-of-5" (RO5) in 1997.[66]

On average, the current generation of DMLs have been found to be larger and more lipophilic than marketed drugs[67] or pre-clinical compounds in general (Figure 27.15).[68,69]Larger and more lipophilic molecules are often associated with poorer oral absorption profiles, and yet this route of administration is required for most DMLs.[70,71] So optimizing the pharmacokinetics, in addition to attaining a balanced profile, can easily become

FIGURE 27.15 The median MW and *c*Log*P* values for DMLs are higher than those for oral drugs[67] or a general set of preclinical compounds from Organon's SCOPE database[68].

FIGURE 27.16 A example of a "fused" DML with a high MW and low oral absorption.

the most challenging aspect of working with DMLs. One explanation for this has been the popularity of the knowledge-based framework combination strategy whereby the molecular frameworks from two selective ligands are combined. Given that the selective ligands used as the starting points are already drug-like in size and the extent to which the frameworks can be integrated is often low, this process can result in large property increases which compromise oral bioavailability.

This Achilles' heel of the framework combination strategy is illustrated by the example in Figure 27.16 wherein the framework of a selective gastrin receptor antagonist **65** was combined with that of a histamine H_2 ligand **66**.[72] Compound **67** is a classic example of a "fused" DML since the degree of overlap that was possible was just a single carbon atom. The incompatibility of the hydrophobic gastrin pharmacophore with the hydrophilic H_2 pharmacophore produces "tolerated regions" that are only relevant for binding at one of the targets, having the effect of increasing the size of the resulting molecule (MW 744) and compromising oral absorption.

Nonetheless, the framework combination approach is a conceptually elegant knowledge-driven strategy that effectively uses SAR knowledge derived from selective ligand projects. Furthermore, there are successful examples of oral drugs having been discovered by this strategy reaching the market such as ziprasidone **11**.[23] To achieve an orally active DML, it is important that the degree of framework overlap is maximized and the size and complexity of the selective ligands is minimized. These goals will typically be more feasible for targets with simple endogenous ligands and conserved binding sites, such as monoamine GPCRs and transporters.

The MW for screening-derived DMLs is frequently lower than for the framework combination strategy, suggesting that this approach may provide a route to smaller and less complex leads (Figure 27.17). A starting compound

FIGURE 27.17 Median MW of DMLs derived via framework combination and screening compared to a general set of preclinical compounds.

obtained via screening already possesses multi-target activity to some extent. During optimization, the activities are balanced usually by adding modestly sized groups or modifying the existing functionality. This typically had less of an effect on the overall size and physicochemical properties of the molecule than the combination of two frameworks.

Over recent years, there has been an increasing amount of evidence that physicochemical properties are less favorable for the ligands from some proteomic target families of interest in drug discovery than for others which makes the discovery of orally active drugs for those targets more challenging.[68] Similar trends amongst the target families have also been reported for DMLs.[69] The target family that has consistently given the highest property values for both pre-clinical compounds in general and DMLs is the peptide GPCRs (Figure 27.18). For example, DMLs for peptide GPCRs had a median MW of 636 and a median $c\,\mathrm{Log}\,P$ of 5.1, figures in excess of those defined in the "rule-of-5" for druglikeness.[66] At the other end of the spectrum, the

FIGURE 27.18 Median MW of DMLs classified according to proteomic target family.

ligands for transporters, monoamine GPCRs and oxidases generally possess favorable physicochemical properties and the feasibility of such targets for DML projects using a variety of lead discovery strategies will be relatively high.

The analysis indicates that designing DMLs for peptide GPCRs will be a more difficult endeavor than for other types of GPCR or indeed for selective ligands for individual peptide GPCRs. However, with perseverence and skill, even difficult families such as peptide GPCRs can sometimes be addressed with a framework combination approach such as the dual AT_1/ET_A antagonist program that delivered compounds, such as **22**, with good oral bioavailability.[29] In such cases, a strong emphasis is often required during lead optimization on simplifying the structure of the lead compound (Figure 27.7).

In a number of other literature examples, the combination of a desirable *in vitro* profile with the PK profile required for the development of an oral drug was not achievable. Where the pharmacophores are fundamentally different, it may not be possible to integrate the requirements of both binding sites into a small, compact molecule and a higher MW compound may be unavoidable. Inevitably, this will mean that some combinations of targets will be more difficult, if not impossible, to address with a drug-like molecule, illustrated by example 67 in Figure 27.16.

While the framework combination strategy tends to produce large molecules, this is less of an issue when the goal is the discovery of pharmacological tools for validating novel target combinations or the production of injectable drugs. An important goal for future research in this field, particularly in academic institutions, will be to develop high quality pharmacological tools to explore the potential therapeutic value of novel target combinations. Here, less attention can be paid to oral exposure and overall developability criteria. More important will be the wider selectivity profile of these pharmacological tools. Portoghese *et al.* reported a range of homo- and hetero-dimeric conjugates with varying linker length designed to investigate pharmacodynamic and organizational features of opioid receptors.[73] For example, recently reported heterodimeric conjugates containing delta-antagonist (naltrindole) and kappa-agonist (ICI-199,441) pharmacophores tethered by variable length oligoglycyl-based linkers **68** (Figure 27.19) were demonstrated to possess significantly greater potency and selectivity compared to their monomer congeners providing further evidence for the opioid receptor hetero-oligomerization phenomenon.[74]

The use of alternative routes of administration, such as intravenous and transdermal, is applicable for some DML applications. High MW DMLs (conjugates), containing a linker group separating the frameworks of the two selective ligands, have been successfully employed as i.v. administered drugs. Van Boeckel *et al.* designed compound **69** with a metabolically stable linker (Figure 27.19) as a dual inhibitor of thrombin, via NAPAP, and anti-thrombin III (ATIII)-mediated factor Xa, via a heparin-derived pentasaccharide fragment.[75] The polyethylene glycol linker in this anti-thrombotic compound confers good aqueous solubility, making it suitable for intravenous administration. Since the pentasaccharide demonstrated a much longer half-life (13–15 h in man) than NAPAP (18 min), the authors postulated that a conjugate with NAPAP might possess improved duration of action. *In vivo* studies confirmed that **69** provided a stronger and longer-lasting anti-thrombotic effect than a mixture of free pentasaccharide and NAPAP.

FIGURE 27.19 Linked DMLs derived via a framework combination approach can make useful biochemical tools or i.v. administered drugs.

VI. SUMMARY

Compounds that act at multiple targets (DMLs) can deliver superior efficacy against complex diseases compared to compounds with high specificity for a single target. The medicinal chemistry literature contains many elegant, and increasingly rational, approaches to the discovery of DMLs, a small cross-section of which has been described above. To address the "physicochemical challenge," new design strategies will certainly be needed and some new approaches, such as fragment-based approaches, have been proposed.[76] Inevitably, medicinal chemists will face target combinations that are particularly compelling in terms of biological rationale, but problematical from the perspective of combining appropriately balanced *in vitro* and *in vivo* activities with acceptable oral bioavailability, duration of action and safety. In many cases, alternative formulations and routes of administration will need to be investigated. Without doubt, the field of multiple ligands will present future generations of medicinal chemists with many challenges, but also numerous opportunities to discover a range of new and superior medicines.

REFERENCES

1. Kola, I., Landis, J. Can the pharmaceutical industry reduce attrition rates? *Nat. Rev. Drug Discov.* 2004, *3*, 711–716.
2. Morphy, R., Kay, C., Rankovic, Z. From magic bullets to designed multiple ligands. *Drug Discov. Today* 2004, *9*, 641–651.
3. Millan, M. J. Multi-target strategies for the improved treatment of depressive states: conceptual foundations and neuronal substrates, drug discovery and therapeutic application. *Pharmacol. Ther.* 2006, *110*, 135–370.
4. Keith, C. T., Borisy, A. A., Stockwell, B. R. Multicomponent therapeutics for networked systems. *Nat. Rev. Drug. Discov.* 2005, *4*, 1–7.
5. Law, M. R., Wald, J., Morris, J. K., Jordan, R. E. Value of low dose combination treatment with blood pressure lowering drugs: analysis of 354 randomised trials. *Br. Med. J.* 2003, *326*, 1427–1431.
6. Larder, B. A., Kemp, S. D., Harrigan, P. R. Potential mechanism for sustained antiretroviral efficacy of AZT-3TC combination therapy. *Science* 1995, *269*, 696–699.
7. Eisen, S. A., Miller, D. K., Woodward, R. S., Spitznagel, E., Przybeck, T. R. The effect of prescribed daily dose frequency on patient medical compliance. *Arch. Intern. Med.* 1990, *150*, 1881–1884.
8. Skolnik, N. S., Beck, J. D., Clark, M. Combination antihypertensive drugs: recommendations for use. *Am. Fam. Physician* 2000, *61*, 3049–3056.
9. Glass, G. Cardiovascular combinations. *Nat. Rev. Drug. Discov.* 2004, *3*, 731–732.
10. Herrick, T. M., Million, R. P. From the analyst's couch: tapping the potential of fixed-dose combinations. *Nat. Rev. Drug. Discov.* 2007, *6*, 513–514.
11. Edwards, I. R., Aronson, J. K. Adverse drug reactions: definitions diagnosis and management. *Lancet* 2000, *356*, 1255–1259.
12. Kramer, M. S., Last, B., Getson, A., Reines, S. A. The effects of selective D_4 dopamine receptor antagonist (L-745, 870) in acutely psychotic inpatients with schizophrenia. *Arch. Gen. Psychiatry* 1997, *54*, 567–572.
13. Campiani, G., Butini, S., Fattorusso, C., Catalanotti, B., Gemma, S., Nacci, V., Morelli, E., Cagnotto, A., Mereghetti, I., Mennini, T., Carli, M., Minetti, P., Di Cesare, M. A., Mastroianni, D., Scafetta, N., Galletti, B., Stasi, M. A., Castorina, M., Pacifici, L., Vertechy, M., Di Serio, S., Ghirardi, O., Tinti, O., Carminati, P. Pyrrolo[1,3]benzothiazepine-based serotonin and dopamine receptor antagonists. Molecular modeling, further structure–activity relationship studies, and identification of novel atypical antipsychotic agents. *J. Med. Chem.* 2004, *47*, 143–157.
14. Jones, H. M., Pilowsky, L. S. New targets for antipsychotics. *Expert Rev. Neurother.* 2002, *2*, 61–68.
15. Stahl, S. M., Entsuah, R., Rudolph, R. L. Comparative efficacy between venlafaxine and SSRIs: a pooled analysis of patients with depression. *Biol. Psychiatry* 2002, *202*, 1166–1174.
16. Charlier, C., Michaux, C. Dual inhibition of cyclooxygenase-2 (COX-2) and 5-lipoxygenase (5-LOX) as a new strategy to provide safer non-steroidal anti-inflammatory drugs. *Eur. J. Med. Chem.* 2003, *38*, 645–659.
17. Chao, R., Rapta, M., Colson, P. J., Lee, J. *US Patent 0182092*, 2005.
18. Investigational Drugs Database, 590916.
19. Van Niel, M., Beer, M., Castro, J., Cheng, S., Evans, D., Heald, A., Hitzel, L., Hunt, P., Mortishire-Smith, R., O'Connor, D., Watt, A., MacLeod, A. Parallel synthesis of 3-aryloxy-2-propanolamines and evaluation as dual affinity 5-HT_{1A} and 5-HT re-uptake ligands. *Bioorg. Med. Chem. Lett.* 1999, *9*, 3243–3248.
20. Mewshaw, R., Meagher, K., Zhou, P., Zhou, D., Shi, X., Scerni, R., Smith, D., Schechter, L., Andree, T. Studies toward the discovery of the next generation of antidepressants. Part 2: incorporating a 5-HT1A antagonist component into a class of serotonin reuptake inhibitors. *Bioorg. Med. Chem. Lett.* 2002, *12*, 307–310.
21. Ryckmans, T., Balançon, L., Berton, O., Genicot, C., Lamberty, Y., Lallemand, B., Pasau, P., Pirlot, N., Quéré, L., Talaga, P. First dual NK_1 antagonists–serotonin reuptake inhibitors: synthesis and SAR of a new class of potential antidepressants. *Bioorg. Med. Chem. Lett.* 2002, *12*, 261–264.
22. Lowe, J., Seeger, T., Nagel, A., Howard, H., Seymour, P., Heym, J., Ewing, F., Newman, M., Schmidt, A. 1-Naphthylpiperazine derivatives as potential atypical antipsychotic agents. *J. Med. Chem.* 1991, *34*, 1860–1866.
23. Howard, H., Lowe, J., Seeger, T., Seymour, P., Zorn, S., Maloney, P., Ewing, F., Newman, M., Schmidt, A., Furman, J., Robinson, G., Jackson, E., Johnson, C., Morrone, J. 3-Benzisothiazolylpiperazine derivatives as potential atypical antipsychotic agents. *J. Med. Chem.* 1996, *39*, 143–148.
24. Zhao, H., Thurkauf, A., He, X., Hodgetts, K., Zhang, X., Rachwal, S., Kover, R., Hutchison, A., Peterson, J., Kieltyka, A., Brodbeck, R., Primus, R., Wasley, J. Indoline and piperazine containing derivatives as a novel class of mixed D_2/D_4 receptor antagonists. Part 1: identification and structure–activity relationships. *Bioorg. Med. Chem. Lett.* 2002, *12*, 3105–3109.
25. Zhao, H., He, X., Thurkauf, A., Hoffman, D., Kieltyka, A., Brodbeck, R., Primus, R., Wasley, J. Indoline and piperazine containing derivatives as a novel class of mixed D_2/D_4 receptor antagonists. Part 2: asymmetric synthesis and biological evaluation. *Bioorg. Med. Chem. Lett.* 2002, *12*, 3111–3115.
26. Vincenzo Garzya Ian, T. F., Gribble, A. D., Hadley, M. S., Lightfoot, A. P., Payne, A. H., Smith, A. B., Douglas, S. E., Cooper, D. G., Stansfield, I. G., Meeson, M., Dodds, E. E., Jones, D. N. C., Wood, M., Reavill, C., Scorer, C. A., Worby, A., Riley, G., Eddershaw, P., Ioannou, C., Donati, D., Hagan, J. J., Ratti, E. A. Studies towards the identification of a new generation of atypical antipsychotic agents. *Bioorg. Med. Chem. Lett.* 2007, *17*, 400–405.
27. Robl, J., Sieber-McMaster, E., Asaad, M. M., Bird, J. E., Delaney, N. G., Barrish, J. C., Neubeck, R., Natarajan, S., Cohen, M., Rovnyak, G. C., Huber, G., Murugesan, N., Girotra, R., Cheung, H. S., Waldron, T., Petrillo, E. W. Mercaptoacyl dipeptides as dual inhibitors of angiotensin converting enzyme and neutral peptidase. Preliminary structure–activity studies. *Bioorg. Med. Chem.* 1994, *4*, 1783–1789.
28. Robl, J. A., Sun, C.-Q., Stevenson, J., Ryono, D. E., Simpkins, L. M., Cimarusti, M. P., Dejneka, T., Slusarchuk, W. A., Chao, S., Stratton, L., Misra, R. N., Bednarz, M. S., Asaad, M. M., Cheung, H. S., Abboa-Offei, B. E., Smith, P. L., Mathers, P. D., Fox, M., Schaeffer, T. R., Seymour, A. A., Trippodo, N. C. Dual metalloprotease inhibitors: mercaptoacetyl-based fused heterocyclic dipeptide mimetics as inhibitors of angiotensin-converting enzyme and neutral endopeptidase. *J. Med. Chem.* 1997, *40*, 1570–1577.
29. Murugesan, N., Tellew, J., Gu, Z., Kunst, B., Fadnis, L., Cornelius, L., Baska, R., Yang, Y., Beyer, S., Monshizadegan, H., Dickinson, K., Panchal, B., Valentine, M., Chong, S., Morrison, R., Carlson, K., Powell, J., Moreland, S., Barrish, J., Kowala, M., Macor, J. Discovery of *N*-isoxazolyl biphenylsulfonamides as potent dual angiotensin II and endothelin A receptor antagonists. *J. Med. Chem.* 2002, *45*, 3829–3835.
30. Eiser, N. Histamine antagonists and asthma. *Pharm. Ther.* 1982, *17*, 239–250.
31. Ohshima, E., Takami, H., Harakawa, H., Sato, H., Obase, H., Miki, I., Ishii, A., Ishii, H., Sasaki, Y. Dibenz[*b*,*e*]oxepin derivatives: novel antiallergic agents possessing thromboxane A_2 and histamine H_1 dual antagonizing activity, 1. *J. Med. Chem.* 1993, *36*, 417–420.
32. Lewis, T., Bayless, L., Eckman, J., Ellis, J., Grewal, G., Libertine, L., Nicolas, J., Scannell, R., Wels, B., Wenberg, K., Wypij, D. 5-Lipoxygenase inhibitors with histamine H1 receptor antagonist activity. *Bioorg. Med. Chem. Lett.* 2004, *14*, 2265–2268.
33. Toda, N., Tago, K., Marumoto, S., Takami, K., Ori, M., Yamada, N., Koyama, K., Naruto, S., Abe, K., Yamazaki, R., Hara, T., Aoyagi, A., Abe, Y., Kaneko, T., Kogen, H. A conformational restriction approach

to the development of dual inhibitors of acetylcholinesterase and serotonin transporter as potential agents for Alzheimer's disease. *Bioorg. Med. Chem.* 2003, *11*, 4389–4415.

34. Sterling, J., Herzig, Y., Goren, T., Finkelstein, N., Lerner, D., Goldenberg, W., Miskolczi, I., Molnar, S., Rantal, F., Tamas, T., Toth, G., Zagyva, A., Zekany, A., Lavian, G., Gross, A., Friedman, R., Razin, M., Huang, W., Krais, B., Chorev, M., Youdim, M. B., Weinstock, M. Novel dual inhibitors of AchE and MAO derived from hydroxy aminoindan and phenethylamine as potential treatment for Alzheimer's disease. *J. Med. Chem.* 2002, *45*, 5260–5270.
35. Sagi, Y., Drigues, N., Youdim, M. B. The neurochemical and behavioral effects of the novel cholinesterase-monoamine oxidase inhibitor, ladostigil, in response to l-dopa and l-tryptophan, in rats. *Br. J. Pharmacol.* 2005, *146*, 543–552.
36. Henke, B. Peroxisome proliferator-activated receptor α/γ dual agonists for the treatment of type 2 diabetes. *J. Med. Chem.* 2004, *47*, 4118–4127.
37. Xu, Y., Rito, C. J., Etgen, G. J., Ardecky, R. J., Bean, S. J., Bensch, W. R., Bosley, J. R., Broderick, C. L., Brooks, D. A., Dominianni, S. J., Hahn, P. J., Dale, S. L., Mais, D. E., Montrose-Rafizadeh, C., Ogilvie, K. M., Oldham, B. A., Peters, M., Rungta, D. K., Shuker, A. J., Stephenson, G. A., Tripp, A. E., Wilson, S. B., Winneroski, L. L., Zink, R., Kauffman, R. K., McCarthy, J. R. Design and Synthesis of α-aryloxy-α-methylhydrocinnamic acids: a novel class of dual peroxisome proliferator-activated receptor α/γ agonists. *J. Med. Chem.* 2004, *47*, 2422–2425.
38. Lackey, K. E. Lessons from the drug discovery of lapatinib, a dual ErbB1/2 tyrosine kinase inhibitor. *Curr. Top. Med. Chem.* 2006, *6*, 435–460.
39. Daub, H., Specht, K., Ullrich, A. Strategies to overcome resistance to targeted protein kinase inhibitors. *Nat. Rev. Drug Discov.* 2004, *12*, 1001–1010.
40. Boschelli, D., Wang, Y., Johnson, S., Wu, B., Ye, F., Sosa, A., Golas, J., Boschelli, F. 7-Alkoxy-4-phenylamino-3-quinolinecarbonitriles as dual inhibitors of Src and Abl kinases. *J. Med. Chem.* 2004, *47*, 1599–1601.
41. Sun, L., Liang, C., Shirazian, S., Zhou, Y., Miller, T., Cui, J., Fukuda, J. Y., Chu, J.-Y., Nematalla, A., Wang, X., Chen, H., Sistla, A., Luu, T. C., Tang, F., Wei, J., Tang, C. Discovery of 5-[5-fluoro-2-oxo-1,2-dihydroindol-(3*Z*)-ylidenemethyl]-2,4-dimethyl-1*H*-pyrrole-3-carboxylic acid (2-diethylaminoethyl)amide, a novel tyrosine kinase inhibitor targeting vascular endothelial and platelet-derived growth factor receptor tyrosine kinase. *J. Med. Chem.* 2003, *46*, 1116–1119.
42. Becknell, N. B., Zulli, A. L., Angeles, T. S., Yang, S., Albom, M. S., Aimone, L. D., Robinson, C., Chang, H., Hudkin, R. L. Novel C-3 N-urea, amide, and carbamate dihydroindazolo[5,4-*a*]pyrrolo[3,4-*c*] carbazole analogs as potent TIE-2 and VEGF-R2 dual inhibitors. *Bioorg. Med. Chem. Lett.* 2006, *16*, 5368–5372.
43. Vane, J. R. *Nat. New Biol.* 1971, 232–235. Dannhard, G., Kiefer, W. Cyclooxigenase inhibitors-current status and future prospects. *Eur. J. Med. Chem.* 2001, *36*, 109–126.
44. Fosslien, E. Adverse effects of nonsteroidal antiinflammatory drugs on the gastrointestinal system. *Ann. Clin. Lab. Sci.* 1998, *28*, 67–81.
45. Weber, A., Casini, A., Heine, A., Kuhn, D., Supuran, C. T., Scozzafava, A., Klebe, G. Unexpected nanomolar inhibition of carbonic anhydrase by COX-2-selective celecoxib: new pharmacological opportunities due to related binding site recognition. *J. Med. Chem.* 2004, *46*, 550–557.
46. Goossens, L., Pommery, N. Henichart, COX-2/5-LOX. Dual acting anti-inflammatory drugs in cancer chemotherapy. *Curr. Top. Med. Chem.* 2007, *7*, 283–296.
47. Barbey, S., Goossens, L., Taverne, T., Cornet, J., Choesmel, V., Rouaud, C., Gimeno, G., Yannic-Arnoult, S., Michaux, C., Charlier, C., Houssin, R., Henichart, J. P. Synthesis and activity of a new methoxytetrahydropyran derivative as dual cyclooxygenase-2/5-lipoxygenase inhibitor. *Bioorg. Med. Chem. Lett.* 2002, *12*, 779–782.
48. Di Napoli, M. F., Papa, F. NCX-4016 NicOX. *Curr. Opin. Invest. Drugs*. 2003, *4*, 1126–1139.
49. Lolli, M. L., Cena, C., Medana, C., Lazzarato, L., Morini, G., Coruzzi, G., Manarini, S., Fruttero, R., Gasco, A. A new class of Ibuprofen derivatives with reduced gastrotoxicity. *J. Med. Chem.* 2001, *44*, 3463–3468.
50. Kato, K., Ohkawa, S., Terao, S., Terashita, Z. I., Nishikava, K. Thromboxane synthetase inhibitors (TXSI). Design, synthesis, and evaluation of a novel series of omega-pyridylalkenoic acids. *J. Med. Chem.* 1985, *28*, 287–294.
51. Soyka, R., Heckel, A., Nickl, J., Eisert, W., Muller, T. H., Weisenberger, H. 6,6-Disubstituted hex-5-enoic acid derivatives as combined thromboxane A2 receptor and synthetase inhibitors. *J. Med. Chem.* 1994, *37*, 26–39.
52. Gengo, P. J., Pettit, H. O., O'Neill, S. J., Wei, K., McNutt, R., Bishop, M. J., Chang, K.-J. DPI-3290 [(+)-3-((alpha-R)-alpha-((2*S*,5*R*)-4-Allyl-2,5-dimethyl-1-piperazinyl)-3-hydroxybenzyl)-*N*-(3-fluorophenyl)-*N*-methylbenzamide]. Mixed opioid agonist with potent antinociceptive activity. *J. Pharmacol. Exp. Ther.* 2003, *307*, 1221–1226.
53. Ananthan, S., KezarIII, H. S., Carter, R. L., Saini, S. K., Rice, K. C., Wells, J. L., Davis, P., Xu, H., Dersch, C. M., Bilsky, E. J., Porreca, F., Rothman, R. B. Synthesis, opioid receptor binding, and biological activities of naltrexone-derived pyrido-and pyrimidomorphinans. *J. Med. Chem.* 1999, *42*, 3527–3538.
54. Montero, A., Goya, P., Jagerovic, N., Luis, F., Callado, L., Meana, J., Girón, R., Goicoechea, C., Martín, I. Guanidinium and aminoimidazolinium derivatives of *N*-(4-piperidyl)propanamides as potential ligands for opioid and I_2-imidazoline receptors: synthesis and pharmacological screening. *Bioorg. Med. Chem.* 2002, *10*, 1009–1018.
55. Dardonville, C., Jagerovic, N., Callado, L., Meana, J. Fentanyl derivatives bearing aliphatic alkaneguanidinium moieties: a new series of hybrid molecules with significant binding affinity for μ-opioid receptors and I_2-imidazoline binding sites. *Bioorg. Med. Chem. Lett.* 2004, *14*, 491–493.
56. Stahl, S. M. Comparative efficacy between venlafaxine and SSRIs: a pooled analysis of patients with depression. *Biol. Psychiatry* 2002, *52*, 1166–1174.
57. Bymaster, F. P., Beedle, E. E., Findlay, J., Gallagher, P. T., Krushinski, J. H., Mitchell, S., Robertson, D. W., Thompson, D. C., Wallace, L., Wong, D. T. Duloxetine (Cymbalta TM), as dual inhibitor of serotonin and norepinephrine reuptake. *Bioorg. Med. Chem. Lett.* 2003, *13*, 4477–4480.
58. Nyberg, S., Eriksson, B., Oxenstierna, G., Halldin, C., Farde, L. Suggested minimal effective dose of risperidone based on PET-measured D2 and 5-HT(2A) receptor occupancy in schizophrenic patients. *Am. J. Psychiatry* 1999, *156*, 869–875.
59. Bonnert, R. V., Brown, R. C., Chapman, D., Cheshire, D. R., Dixon, J., Ince, F., Kinchin, E. C., Lyons, A. J., Davis, A. M., Hallam, C., Harper, S. T., Unitt, J. F., Dougall, I. G., Jackson, D. M., McKechnie, K., Young, A., Simpson, W. T. Dual D_2-receptor and b_2-adrenoreceptor agonists for the treatment of airway diseases. 1. Discovery and biological evaluation of some 7-(2-aminoethyl)-4-hydroxybenzothiazol-2(3*H*)-one analogues. *J. Med. Chem.* 1998, *41*, 4915–4917.
60. Atkinson, P., Bromidge, S., Duxon, M., Laramie, M., Gaster, L., Hadley, M., Hammond, B., Johnson, C., Middlemiss, D., North, S., Price, G., Rami, H., Riley, J., Scott, C., Shaw, T., Starr, K., Stemp, G., Thewlis, K., Thomas, D., Thompson, M., Vong, A., Watson, J. 3,4-Dihydro-2*H*-benzoxazinones are 5-HT1A receptor antagonists with potent 5-HT reuptake inhibitory activity. *Bioorg. Med. Chem. Lett.* 2005, *15*, 737–741.
61. Meltzer, H. Y. Illuminating the molecular basis for some antipsychotic drug-induced metabolic burden. *Proc. Natl. Acad. Sci. USA* 2007, *104*, 3019–3020.
62. Fabian, M. A., Biggs, W. H., Treiber, D. K., Atteridge, C. E., Azimioara, M. D., Benedetti, M. G., Carter, T. A., Ciceri, P.,

Edeen, P. T., Floyd, M., Ford, J. M., Galvin, M., Gerlach, J. L., Grotzfeld, R. M., Herrgard, S., Insko, D. E., Insko, M. A., Lai, A. G., Lélias, J. M., Mehta, S. A., Milanov, Z. V., Velasco, A. M., Wodicka, L. M., Patel, H. K., Zarrinkar, P. P., Lockhart, D. J. A small molecule-kinase interaction map for clinical kinase inhibitors. *Nat. Biotechnol.* 2005, *23*, 329–336.

63. Roskoski, R. Sunitinib: a VEGF and PDGF receptor protein kinase and angiogenesis inhibitor. *Biochem. Biophys. Res. Commun.* 2007, *356*, 323–328.
64. Force, T., Krause, D. S., van Etten, R. A. Molecular mechanisms of cardiotoxicity of tyrosine kinase inhibition. *Nat. Rev. Cancer* 2007, *7*, 332–344.
65. Dinges, J., Albert, D. H., Arnold, L. D., Ashworth, K. L., Akritopoulou-Zanze, I., Bousquet, P. F., Bouska, J. J., Cunha, G. A., Davidsen, S. K., Diaz, G. J., Djuric, S. W., Gasiecki, A. F., Gintant, G. A., Gracias, V. J., Harris, C. M., Houseman, K. A., Hutchins, C. W., Johnson, E. F., Li, H., Marcotte, P. A., Martin, R. L., Michaelides, M. R., Nyein, M., Sowin, T. J., Su, Z., Tapang, P. H., Xia, Z., Zhang, H. Q. 1,4-Dihydroindeno[1,2-*c*]pyrazoles with acetylenic side chains as novel and potent multitargeted receptor tyrosine kinase inhibitors with low affinity for the hERG Ion channel. *J. Med. Chem.* 2007, *50*, 2011–2029.
66. Lipinski, C. A., Lombardo, F., Dominy, B. W., Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 1997, *23*, 3–25.
67. Vieth, M., Siegel, M. G., Higgs, R. E., Watson, I. A., Robertson, D. H., Savin, K. A., Durst, G. L., Hipskind, P. A. Characteristic physical properties and structural fragments of marketed oral drugs. *J. Med. Chem.* 2004, *47*, 224–232.
68. Morphy, J. R. The influence of target family and functional activity on the physicochemical properties of pre-clinical compounds. *J. Med. Chem.* 2006, *49*, 2969–2978.
69. Morphy, J. R., Rankovic, Z. The physicochemical challenges of designing multiple ligands. *J. Med. Chem.* 2006, *49*, 4961–4970.
70. Lipinski, C. A., Lombardo, F., Dominy, B. W., Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* 1997, *23*, 3–25.
71. Veber, D. F., Johnson, S. R., Cheng, H. Y., Smith, B. R., Ward, K. W., Kopple, K. D. Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* 2002, *45*, 2615–2623.
72. Kawanishi, Y., Ishihara, S., Tsushim, T., Seno, K., Miyagoshi, M., Hagishita, S., Ishikawa, M., Shima, N., Shimamura, M., Ishihar, Y. Synthesis and pharmacological evaluation of highly potent dual histamine H_2 and gastrin receptor antagonists. *Bioorg. Med. Chem. Lett.* 1996, *6*, 1427–1430.
73. Portoghese, P. S. From models to molecules: opioid receptor dimers, bivalent ligands, and selective opioid receptor probes. *J. Med. Chem.* 2001, *44*, 2259–2269.
74. Daniels, D. J., Kulkarni, A., Xie, Z., Bhushan, R. G., Portoghese, P. S. A bivalent ligand (KDAN-18) containing δ-antagonist and κ-agonist pharmacophores bridges δ2 and κ1 opioid receptor phenotypes. *J. Med. Chem.* 2005, *48*, 1713–1716.
75. Buijsman, R. C., Basten, J. E., van Dinther, T. G., van der Marel, G. A., van Boeckel, C. A., van Boom, J. H. Design and synthesis of a novel synthetic NAPAP-pentasaccharide conjugate displaying a dual antithrombotic action. *Bioorg. Med. Chem. Lett.* 1999, *9*, 2013–2018.
76. Morphy, R., Rankovic, Z. Fragments, network biology and designing multiple ligands. *Drug Discov. Today* 2007, *12*, 156–160.

Chapter 28

Pharmacophore Identification and Pseudo-Receptor Modeling

Wolfgang Sippl

I. INTRODUCTION
A. Historical background
B. Definitions
C. Importance of the pharmacophore concept
D. Application of pharmacophores
II. METHODOLOGY
A. Pharmacophore modeling
III. ADVANCED APPROACHES
A. Structure-based pharmacophores
B. Pseudo-receptor models
IV. APPLICATION STUDY
A. Pharmacophore-based screening for novel histamine H_3-receptor antagonists
B. Pharmacophore determination process
C. Pharmacophore-based screening of compound libraries
V. CONCLUSIONS
REFERENCES

In theory, theory and practice are the same. In practice, they are not.

Lawrence Peter Berra

I. INTRODUCTION

In the large majority of cases the basis for a pharmacodynamic effect is the interaction of a certain substance with a biomacromolecule of physiological importance. Above all, proteins like enzymes, receptors, and ionchannels, but also nucleic acids, serve as physiological binding partners. In all cases, there must exist a highly specific 3D binding epitope which serves as complementary binding site for a drug molecule. Compounds exerting similar activities at the same enzyme or receptor therefore possess in most cases closely related binding properties. That is, these molecules present to the macromolecule structural elements of identical chemical features in sterically consistent locations. The highest common denominator of a group of ligands exhibiting a similar biological effect recognized by the same binding site is named pharmacophore.[1] Thus, in other words, a pharmacophore is an abstraction of the crucial molecular features responsible for the binding of a set of ligands to a macromolecular target.

As a practical matter, computer-aided molecular design is frequently split into disciplines that focus on either structure-based or ligand-based methods. When the 3D structure of a target protein and the binding site is available, then it is possible to invoke structure-based approaches. New candidate ligands may be docked into a particular binding site in order to study if they can interact with the protein in an optimal way. If, however, knowledge about the structure of the macromolecular target is limited, but a sufficient number of active analogs have already been discovered, then pharmacophore-based methods are applied to design novel active molecules. It may seem straightforward to develop new ligands for known proteins by applying structure-based approaches, but there are significant problems invoked. Induced fit mechanism, multiple binding modes, solvation, or entropic effects are some of the problems that

Copyright © 2008, Elsevier Ltd
All rights reserved.

must be overcome to end up with reliable models. Beside these problems, many target proteins of high pharmaceutical interest are membrane-bound receptors (e.g. G-protein-coupled receptors) and attempts to crystallize them have been not successful as yet. In the absence of the 3D structure of a protein of interest, ligand design may be performed by the use of a pharmacophore-based method.

A. Historical background

The idea that bioactive substances interact with receptors began with Langley in 1878, who introduced the term "receptive substance."[2] However, the term "receptor" was introduced several years later by Paul Ehrlich.[3] He also introduced the term "pharmacophore" to describe those parts of a molecule which are responsible for its activity. Together with the "lock and key" concept of Emil Fischer it became clear that not all parts of a molecule, the "key," are equally important for exerting its biological effect on the "lock."[4] Thus sometimes, small variations of distinct parts of a molecule can dramatically influence the activity, whereas variations of other parts only cause minor changes in the biological activity. The concepts of Langley, Ehrlich, and Fischer have constituted the cornerstones of modern drug discovery and development up to this day. Half a decade later their concepts were confirmed in an impressive manner by the first solved crystal structures of protein–ligand complexes.[5]

Even before the advent of computers and modeling software, simple pharmacophores were described in the literature and considered as tools for the discovery of novel molecules. Based on initial structure–activity relationship (SAR) considerations simple 2D models were introduced in the 1940s. With the advent of computers and modeling programs, the idea of displaying and manipulating 3D structures became true.[6] Kier and Marshall have pioneered the development of the pharmacophore concept and the application in SARs.[7,8] Several years later Peter Gund implemented the first *in silico* screening methodology with a program to screen a substance library for pharmacophoric patterns in the 1970s.[9] The active analog approach developed by Garland Marshall's group was one of the first automated tools for pharmacophore generation. Marshall's approach was the basis for many following pharmacophore modeling programs in that area. Since these early days a variety of automated pharmacophore discovery programs have been developed in academia and software developing companies.[10]

B. Definitions

The term pharmacophore is used by different group of scientists not always in accordance with the official definition elaborated by the IUPAC working party which stated:[1] "*A pharmacophore is the ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interaction with a specific biological target structure and to trigger (or block) its biological response.*" Many scientists use the term "pharmacophore" or "pharmacophoric group" to define distinct functional groups or substance classes possessing biological activity, for example, sulfonamides or dihydropyridines. In this context, the term pharmacophore is mixed with another concept of structure and activity, namely "privileged structures." The retrospective analysis of the chemical structures and scaffolds of drug molecules led to the detection of some structural motifs that are often associated with biological activity. Such motifs were called "privileged structures" by Evans *et al.*, to represent substructures that confer activity toward two or more different targets.[11] The idea behind is that the privileged structure provides the scaffold and the substitution the reason for specificity. However, in the terms of the IUPAC definition, the pharmacophore represents the common molecular interaction features of a set of molecules toward their receptor.[12]

A pharmacophoric element (also called feature) is generally defined as an atom or a group of atoms (e.g. a hydrogen bond donor atom or an aromatic ring system) common to active compounds with respect to a target protein and essential for the activity. Thus, a pharmacophore model can also be regarded as the representation of a collection of pharmacophore features.

The above described definition of a pharmacophore is based on a 3D point of view of molecules. It reflects the way medicinal chemists characterize the binding ability of molecules for a given target protein. However, depending on the different research areas, scientists have different views. Computational chemists often use the term pharmacophore in a more abstract way. Influenced by the structural representation of molecules, a set of topological connections is used to define the properties and dimension of a molecule in 2D. Here, the spatial and topological distribution of pharmacophoric features are converted to a lower dimensional representation, for example, vectors. Such vectors, which represent pharmacophore descriptors, are referred to as "fingerprints," "keys," "bitstrings," or "correlation vectors," depending on the type of information stored. The pharmacophore descriptors or fingerprints can be regarded as a transformed molecular representation instead of an explicit 3D structure. These fingerprints are often used to rapidly screen large compound libraries. In this chapter we will only focus on 3D-pharmacophore concepts.

Starting from a preliminary pharmacophore model a hypothetical receptor consisting of individual amino acid residues can be constructed surrounding a set of superimposed ligands. Guided by permanent correlation of biological data and model-derived calculated free energies of binding, a complex system is generated, mimicking reasonably well the interaction pattern of a real binding site.

FIGURE 28.1 Comparison of atom-based alignment and experimentally derived position of methotrexate and dihydrofolate in dihydrofolatreductase.

The resulting hypothetical receptor model is named mini-receptor or pseudo-receptor and can be used to derive three-dimensional quantitative structure–activity relationships (3D-QSAR). The concept was originally developed in the 1980s by several groups.[13–16]

C. Importance of the pharmacophore concept

A pharmacophore captures the concept of bioisosterism by not only comparing topological similarities but structural groups at similar locations with the same chemical functionality. It is important to concentrate on the pharmacophoric features since topological molecule characteristics are often misleading in the superpositioning of two molecules with respect to their binding mode. Figure 28.1 shows the well-known example of dihydrofolate reductase ligands.[17] For the two shown ligands a topological overlay would result in a wrong prediction of the binding mode. If the pharmacophoric features (the hydrogen bonding pattern in this example) are taken into account for the superimposition the correct overlay mode can be deduced. The pharmacophore-based superposition is similar to the binding mode observed in the crystal structures of methotrexate and dihydrofolate with dihydrofolatreductase (Figure 28.2, 1rx2.pdb, 1rb3.pdb).

The increasing number of accessible compounds that can be used nowadays as starting point for a biological

FIGURE 28.2 (a) Comparison of the pharmacophore-based alignment of dihydrofolate reductase ligands and (b) the experimentally derived protein-based alignment.

target screening makes it necessary to have fast and reliable *in silico* screening tools. Structure-based methods are often too slow to virtually screen compound databases with millions of molecules. Besides the speed, there are other problems in structure-based design and docking programs that need to be addressed. For instance, most of the current docking programs do not take into account protein flexibility. Only very recently, programs were developed (e.g. AutoDock4,[18] GOLD3.0,[19] Glide,[20] or FlexE[21]) which consider protein side-chain flexibility for docking. Other problems which often occur in ligand docking, are the correct placement of water molecules within the binding site (which represent putative ligand binding partners), the treatment of solvation effects (on the ligand and protein site) and consideration of the internal strain of a docked ligand. Structure-based approaches are able to provide important information about the interaction between a ligand and a macromolecule, but the accurate prediction of the binding affinity is still an unsolved problem. A detailed discussion about the limitations of docking and scoring programs can be found in recently published reviews.[22,23]

Another reason why pharmacophore-based approaches are often used in drug design is the missing 3D structure of many interesting macromolecules. Many current drug targets are membrane-bound and so far only very few membrane proteins have been successfully crystallized. In the absence of an experimentally determined 3D protein structure, the use of indirect ligand-based approaches, including pharmacophores, is the only way to rationally design novel bioactive molecules.[24]

D. Application of pharmacophores

Pharmacophore modeling in computer-aided drug design is generally applied in three domains. The first is the definition of relevant pharmacophoric features in a drug molecule necessary to achieve a certain biological effect and to establish clear SARs. A well-developed pharmacophore

model, preferentially including information about the dimension of the receptor-binding cavity, may be employed to design novel and more active molecules which fit the model. Often, such pharmacophore models are the starting point for 3D-QSAR analysis, for example, CoMFA,[25] by which quantitative predictions may be made (see chapter 30 in this book). The second is scaffold hopping, that is detecting molecules with different scaffolds (novel chemotypes) by virtually screening large compound libraries.[26] The third one is the use of parallel pharmacophore-based screening in order to predict pharmacological profiles for lead structures *in silico*. With the use of pharmacophore fingerprints it is hoped to predict unwanted side effect in very early stages of the drug discovery process and therefore to reduce the risk of late failure of drug candidates.[27]

II. METHODOLOGY

A. Pharmacophore modeling

To end up with a predictive pharmacophore model, it is necessary to start with reliable structural and biological data. First of all, it is important to have correct 3D structures of all compounds under study. Thus, atomic valences, bond orders, protonation state and stereochemistry have to be checked carefully. Also the consideration of different possible tautomers is necessary when the bioactive form is not exactly known. Another prerequisite is the existence of a similar binding mode of all ligands under study. Experimental data, from competition experiments or protein–ligand crystal structures, can clearly point out that the ligands interact with the same binding epitope in a similar way and not on distinct binding sites.

The four steps in the development of a pharmacophore model are: (1) selection of a set of active ligands known to bind to the same target (same binding site), (2) conformational analysis for all ligands, (3) assignment of pharmacophoric features, and (4) molecular superimposition of the ligand conformations to develop a common 3D-pharmacophore. The majority of automated pharmacophore generation programs use qualitative pharmacophore models that do not consider the activity of the ligands. The ultimate goal of all these programs is to search for a unique conformation of all congeners, where most if not all assigned pharmacophoric features of the ligands are presented in a superimposed manner. Most of the programs are based on minimizing the root-mean-square (RMS) superposition error between conformations of the ligands under study while trying to increase the fit of the pharmacophoric features. To compare the different conformations for a data set of given active molecules, a superpositioning procedure is needed. The assignment of the pharmacophoric features and the generation of the ligand alignment is carried out in an automated way by most of the current pharmacophore modeling programs (e.g. Catalyst,[28] DISCO,[29] Galahad,[30] LigandScout,[31] Phase,[32] MOE PHP[33]). The scope of this chapter is not to describe all available software packages in detail, but to illustrate the different steps of the pharmacophore development process. For a recently published overview of current pharmacophore modeling programs the reader is referred to the literature.[10,34]

1. *Conformational analysis of ligand molecules and bioactive conformation*

Since molecules are flexible and not static, a conformational analysis has to be carried out first to generate an ensemble of low-energy conformations. This is probably one of the most critical steps in the pharmacophore discovery process, since the goal is not only to consider the global minima of a molecule, but also to include the bioactive conformation as part of an ensemble of low-energy conformations.

In order to bind to a receptor with high affinity, a ligand must match the binding pocket. The steric match will thereby depend primarily on the ligand conformation. Within a binding pocket, the ligand will not necessarily be present in its lowest energy conformation, as the gain in interaction energy with the receptor can compensate for a conformation with higher energy.[35] Still, it can be expected that for a high-affinity ligand, the bioactive conformation is at least energetically favorable, as otherwise the conformational energy cost would reduce binding affinity. The relation between a high energetic binding conformation and the loss of free energy of binding ΔG is given by equation (28.1):

$$\Delta G = -2.303\, RT \log K_{\mathrm{i}} \tag{28.1}$$

Under physiological conditions ($T = 310\,\mathrm{K}$), the free energy (in kcal/mol) and the binding affinity are related by:

$$\Delta G = -1.42 \log K_{\mathrm{i}} \tag{28.2}$$

Thus, if a compound binds in a conformation that deviates 1.42 kcal/mol from the global minimum structure, its affinity will be decreased by one order of magnitude. High-affinity compounds can thus be expected to bind in an energetically favorable conformation. To analyze the conformational space of molecules experimental and theoretical approaches are applied. Experimental techniques like nuclear magnetic resonance (NMR) only provide information on one or a few conformations of a molecule. A complete overview about the conformational space of molecules can be gained only by theoretical techniques.[36] Correspondingly, a variety of theoretical methods for conformational analysis has been developed. The most general conformational analysis methods are those that are able to identify all minima on the potential energy surface. However, as the number of minima dramatically increases with the number of rotatable bonds, an exhaustive

detection of all minima becomes a difficult and time-consuming task. Commonly used methods for this purpose are listed below (described in depth in[37]):

a. Systematic search

Each bond is rotated incrementally and the resulting structures are minimized. Systematic search algorithms have the advantage that they sample the conformational space very well, yet, in case of a high number of rotatable bonds this method may be computationally impracticable.[38]

b. Random search

In a random search, one can move from one region of the energy-surface to a completely unconnected region in a single step. A commonly applied method is the Metropolis Monte Carlo scheme that starts with a minimized conformation A of a molecule. Then a random move on the energy-landscape is carried out (e.g. torsion angles are rotated by a random amount) and the structure is minimized. The potential energy of the output structure B is evaluated. If $E_{pot}(B) < E_{pot}$ (A), the new conformation is accepted. If $E_{pot}(B) > E_{pot}(A)$, the move may still be accepted depending on the transition probability that in turn depends on the temperature. Monte Carlo methods efficiently sample the conformational space; however, there is no guarantee – as with all random search tools – that the entire energetic landscape will be sampled. Another sampling technique applied to the problem of improved conformational searching is known as Poling.[39] Poling has been applied within the Catalyst[26] program in order to enable the generation of large multi-conformer databases in reasonable time, which are needed for pharmacophore-based screening.

c. Simulated annealing or molecular dynamics (MD) simulations

The aim of MD simulations is to reproduce the time-dependent motional behavior of a molecule. MD is based on molecular mechanics. It is assumed that the atoms in the molecule interact with each other according to the rules of an employed force field. MD simulations generate an ensemble of structures that does however not only represent minimum structures. In a simulated annealing MD protocol, the system temperature is periodically increased resulting in a significant rise of kinetic energy which makes it easier to overcome barriers of potential energy. Subsequently, the system is cooled down, thereby trapping the molecule in an energetically favorable conformation. MD simulation techniques for sampling the conformational space are quite time-consuming and are therefore used only for smaller ligand data sets. Again, there is no guarantee of sampling the entire potential energy-surface.[40]

There is an ongoing discussion in the literature which ligand conformations (i.e. within which energy range) have to be considered in a pharmacophore generation process. Several recently carried out studies on protein–ligand X-ray structures have shown that many conformational search tools yield ensembles including the experimentally observed bioactive conformation.[41] The energy difference between the co-crystallized conformation of a ligand and its global minimum calculated with molecular mechanic programs is depending on the force field employed, and therefore a general energy range to be considered cannot be defined.[42]

Which conformational analysis performs best? A clear-cut answer cannot be given as it depends on the individual data set to be studied and the problems to address. If only a limited number of ligands are considered, more computationally intensive methods such as the systematic search can be applied. If a compound library with hundreds of thousand of entries has to be converted into a multi-conformer database, fast and simplified approaches have to be used (e.g. in Catalyst[26] or Omega[43]).[44]

2. *Ligand superposition techniques*

Several algorithms for ligand superposition exist, including techniques that superimpose molecules by mapping and comparing shape and field properties of the structures (e.g. DISCO,[27] Catalyst,[26] or LigandScout[29]) or by incrementally building up a test molecule upon a rigid reference molecule (e.g. FlexS[45]).[46] When applying the FlexS algorithm, the flexible compound is first partitioned into fragments. Subsequently, an anchor fragment is selected and placed onto the reference compound in a way that similar interactions can be established by both compounds. Then, the remaining fragments of the flexible molecule are incrementally added. Flexibility is considered by allowing each fragment to adopt a discrete set of energetically favorable conformations. Each superposition is then assigned a score that will be higher the better the match between the reference and target molecule is.

3. *Assignment of pharmacophoric elements*

The assignment of pharmacophoric features shall be described using as an example the histamine H_3 receptor antagonist shown in Figure 28.3.[47] Table 28.1 lists the pharmacophoric features assigned in the ligand structure by comparison with other known active antagonists. Thus, the protonated nitrogen atom of the piperidyl moiety can be translated into a center of a sphere with coordinates corresponding to the location of the nitrogen atom and a radius defining a volume around this atom. If a molecule is compared to this pharmacophore model and its protonated nitrogen atom will lie within the sphere, this pharmacophoric feature will be said to be matched. The bigger the sphere, the easier it will be for a ligand conformation to match

FIGURE 28.3 (a) Pharmacophoric features observed in the ligand (by comparison with other known active ligands). (b) The resulting pharmacophore is based on features and shape. (c) The molecule's shape can serve as an additional constraint in pharmacophore searches.

TABLE 28.1 Pharmacophoric features observed in the ligand shown in Figure 28.3

Feature	Colour	Representation
Positive charge	Red	Sphere
H-bond donor	Magenta	Sphere-vector-sphere
H-bond acceptor	Green	Sphere-vector-sphere
Hydrophobic aliphatic	Blue	Sphere
Aromatic ring	Orange	Plane, center of plane, vector
Hydrophobic	Light blue	Sphere

the pharmacophoric features. Similarly, an aromatic or a hydrophobic aliphatic moiety can be defined by a center of sphere and radius. Hydrogen-bond acceptors and donors are represented by vectors in order to account for the directionality of H-bonds while aromatic rings can be either defined by spheres or the combination of center, plane and vector. If defined this way, the orientation of the aromatic plane in respect to the rest of the molecule is considered, too.

Again, the shape of the molecule can be incorporated into a pharmacophore by translating the van der Waals volume into an additional feature. Especially, if the ligand is known to fill the binding pocket well, the available volume can be taken into account. The abstract definition of a molecule in form of a pharmacophore as defined in Figure 28.3 facilitates comparison with other molecules. In the given example most features of the antagonist were considered for the generation of the pharmacophore model resulting in an almost unique fingerprint of the molecule. Depending on the number of features included in the model and the tolerances defined, it will be more or less difficult for other molecules to match the pharmacophore model. Matching a pharmacophore additionally depends strongly on the conformation that is adopted by the molecule that is compared to the pharmacophore model. Even a different conformation of the ligand might not match the pharmacophore model defined above. Thus, in order to increase the likelihood of a specific molecule fitting a pharmacophore model, each molecule of interest is associated with a conformational ensemble. When searching for similarities with the pharmacophore model, all conformations of a molecule are tested on the pharmacophore before the best fit is evaluated. The difficulty in defining a "useful" pharmacophore model lies in the restriction to only essential pharmacophoric features observed in the active ligands.

If a pharmacophore is used for database screening in order to retrieve new compounds based on the similarity of pharmacophoric features, a model is useful when it is able to identify known actives among a number of inactive molecules. In order to screen commercial compound databases with a pharmacophore model, firstly, a so called multi-conformer database must be generated. This means a set of conformations must be generated for all compounds deposited in the compound databases. Since such databases can number in the hundreds of thousand of compounds, fast algorithms are paramount. In addition, the conformation database should not lead to an explosion in storage requirements for the millions of conformers. Finally, the database program should be able to handle the pharmacophore search within reasonable time. The most widely used programs for building large multi-conformation databases are Catalyst,[26] UNITY,[48] Omega,[40] and MOE.[31] Whereas Catalyst, UNITY and MOE are also used to carry out the pharmacophore generation and the pharmacophore search, Omega can only be used to generate multi-conformer databases. A recent comparison of the performance of the different programs can be found in the literature.[38]

III. ADVANCED APPROACHES

A. Structure-based pharmacophores

If the 3D structure of a protein–ligand complex is known from either X-ray crystallography or protein NMR, the most

obvious way of deriving a picture of the relevant ligand interactions is to analyze the molecule's complementarity within the corresponding protein binding site. A commonly used structure-based design approach is the previously mentioned molecular docking of ligands into a target binding pocket assuming that the binding site is more or less rigid whereas the ligand is flexible. Molecular docking is still the most popular method for structure-based drug design. However, pharmacophore-based approaches have shown clear advantages regarding the computational demand and accuracy for virtual screening.[49] Especially, in regard to the number of false positives, which are often observed in classical docking-based virtual screening, the idea to combine structural information derived from a protein–ligand complex and the use of a rapid pharmacophore-based screening technique is obvious. There is no competition between ligand-based and structure-based pharmacophore modeling, and both approaches can be used fruitfully in a complementary manner.[50] In fact, a variety of pharmacophore modeling programs allow one to take advantage of additional information provided by a protein or protein–ligand complex structure to help improve the reliability of the generated model. The development of several novel programs for deriving structure-based pharmacophores in the last few years has clearly shown that pharmacophore-based virtual screening is very successful in identifying novel bioactive molecules.[51–55] On the other side, it was also recognized that the consideration of pharmacophores in docking programs can increase the reliability and accuracy. Several docking programs are now available which apply the pharmacophore concept to better discriminate between false and real binding modes (e.g. Glide,[19] FlexX-Pharm,[20] GOLD[18]).

As an illustration, the generation of a structure-based pharmacophore and its application for virtual screening of ABL tyrosine kinase inhibitors is given.[56] STI-571 (Gleevec®) has been approved for the treatment of chronic myelogenous leukemia (CML) and was the first antitumor drug from the family of tyrosine kinase inhibitors.[57] Several crystal structures of STI-571 in complex with different tyrosine kinases (ABL, c-KIT, SYK) have been obtained in the last few years showing that the compound can bind in varying conformations (open and closed conformation) to different forms of tyrosine kinases. In the case of ABL tyrosine kinase STI-571 binds to the inactive enzyme form and prevents in the activation.[58] Wolber *et al.* have generated several pharmacophores on the available X-ray structures of ABL in complex with STI-571 and analogs (1iep.pdb, 1fpu.pdb, 1opj.pdb).[59] In a straightforward approach, the different pharmacophore models were merged using the program LigandScout. The merged pharmacophore contained four lipophilic aromatic areas, two acceptor features and eight excluded volume spheres. As an example the structure-based pharmacophore extracted from the X-ray structure 1iep is shown in Figure 28.4. Subsequently,

FIGURE 28.4 3D structure of Gleevec® (capped sticks) bound to ABL kinase (top). The structure-based pharmacophore generated with the program LigandScout is shown in the middle and the extracted pharmacophore together with the excluded volumes (gray spheres) is shown at the bottom. (Yellow spheres = hydrophobic features, hydrogen bond donors = green arrows, hydrogen bond acceptors = red arrows).

a virtual screening was carried out using two different ligand databases. The first one was a collection of 2,765 drug-like ligands from the complexes in the Protein Data Bank (PDB), the second one was the Maybridge compound library (containing ~59,000 molecules). The pharmacophore model was able to identify all STI-571 entries from the PDB database and did not result in false positives. In addition seven compounds were identified from the Maybridge database (MDB), which might represent potential lead structures for the development of novel ABL tyrosine kinase inhibitors.

Several successful applications of the LigandScout program have been reported recently and have supported the

feasibility of structure-based pharmacophores to identify novel active molecules.[47–50]

B. Pseudo-Receptor Models

Starting in the 1980s, a combination of pharmacophore modeling and structure-based design was introduced and referred as pseudo-receptor modeling or receptor mapping.[60] Based on a preliminary pharmacophore model a hypothetical receptor consisting of individual amino acid residues is constructed surrounding a set of superimposed ligands. The placement of the individual amino acid residues is guided by experimental data (e.g. from site-directed mutagenesis data). Höltje successfully applied the receptor-mapping technique for several target proteins, for which no 3D structure was available.[61–63] Using a data set of 20 5-HT_{2A} receptor antagonists from different chemical families, a pharmacophore was generated, which was able to explain the SAR of the ligand.[64] The receptor-mapping, i.e. the placement of the individual amino acids was based on a homology model of the 5-HT_{2A} receptor generated on the basis of the low-resolution 3D structure of bacteriorhodopsin (a related membrane protein). Using the derived pseudo-receptor, a predictive QSAR model could be obtained which was subsequently applied to design novel potent antagonists.[64]

1. *Yak, PrGen, Flo*

Whereas the first pseudo-receptor models were generated more or less intuitively "by hand", which sometimes resulted in irreproducible results, a broader distribution of this concept was achieved by the commercial software packages Yak and PrGen.[14] Both programs allow the generation of a pseudo-receptor in a more or less automated way. In addition, guided by extensive correlation of experimental and model-derived free energies of binding, a host–guest system is created, mimicking reasonably well the interaction at a real binding site. The fundamental basis of a pseudo-receptor is the placement of the individual amino acid residues. In Yak and PrGen, ligand-specific interaction vectors (the pharmacophoric elements) are calculated and saturated with individual residues from a database of pre-calculated conformations of amino acids. Subsequently, a receptor minimization is carried out by relaxing all residues keeping the position, orientation and conformation of the ligands unchanged. To achieve a correlation between the experimentally derived binding affinities (or other biological data) and the calculated interaction energies, a coupling constant is introduced and the system is minimized (correlation-coupled minimization). In a next step, the ligand alignment is allowed to relax within the fixed pseudo-receptor (ligand relaxation). This process, i.e. thus correlation-coupled minimization followed by unconstrained ligand relaxation, is repeated several times until a highly correlated pseudo-receptor is obtained. To validate the generated pseudo-receptor, its possibility to predict the binding affinities of novel ligands must be examined.[59] Therefore, classical QSAR methods such as cross-validation via leave-one-out and/or prediction of external test set compounds are applied. In case of test set or novel ligands, the molecules have to be placed equally to the training set molecules in the pseudo-receptor and have to be minimized applying the same protocol as for the training set ligands.

Another pseudo-receptor modeling approach has been developed by Bohacec et al.[65] Their program Flo generates an ensemble of low-energy conformers of each compound of a training set. The conformations are then optimized to simultaneously minimize the internal energy and maximize the match of chemically similar moieties. Then a pseudo-receptor is composed of functional groups that will mimic the binding cavity. For example, a guanidinium group is selected to form hydrogen bonds with an acidic group of the ligands. The selected residues are positioned around the aligned training set ligands and anchored to the chemically complementary ligand atoms applying a distance constraint. The remaining volume of the pseudo-receptor is filled with propane molecules to mimic a binding site's hydrophobic surface. In the last step the pseudo-receptor is equilibrated, comparable to the PrGen approach, by applying several round of dynamics. While a pseudo-binding site is quite artificial, the method has the advantage that the binding site can be visualized and used for ligand docking and structure-based design.

2. *Quasar and Raptor*

A further development of Vedani et al. was the simplification of the atomistic pseudo-receptor concept (Yak and PrGen) to a quasi-atomistic receptor approach (named Quasar).[66] Similar to the approach of Walters et al., who developed the program GERM,[67] Quasar uses a 3D binding-site surrogate surrounding the ligands instead of a shell of amino acid residues. Each of the virtual particles bears relevant atomistic properties (e.g. H-bond donor, hydrophobic particle). Quasar does not only take into account one conformer per ligand but represents each ligand by an ensemble of low energy conformations (called fourth dimension), thereby reducing the bias associated with the selection of a putative bioactive conformation. Binding of ligand molecules to a macromolecular binding pocket is often facilitated by an induced fit, that is, the adaptation of a protein to the ligand topology. This effect which is not considered in most of the pharmacophore and 3D-QSAR approaches, is considered by Quasar and Raptor[68] (the so-called fifth dimension). Quantitative models generated with these programs have therefore been named 4D- or 5D-QSARs.[69]

FIGURE 28.5 Molecules used to generate the pharmacophore for inducers of fetal hemoglobin.

FIGURE 28.6 Molecular structures of the two most active inducers identified by the pseudo-receptor.

3. *Application of Pseudo-Receptor Models*

The pseudo-receptor concept has been applied in recent years to analyze crucial ligand–receptor interaction sites and to establish 3D-QSARs for the prediction of biological activities of ligands.[70] A variety of application studies have shown that the pseudo-receptor concept is a versatile tool to establish 3D-QSAR models, often better in their predictive behavior compared to results obtained from classical 3D-QSAR approaches (e.g. CoMFA).[71] Several application studies have been published which have shown the value but also the limitations of this approach.[56,72]

In a recently published study by Bohacec et al. the pseudo-receptor concept was successfully applied to identify novel small molecule inducers of fetal hemoglobin.[73] Four available active compounds (Figure 28.5) were selected, based on activity and diversity, for the construction of an initial pharmacophore.

The initial pharmacophore was constructed using the Flo molecular modeling software.[61] The derived pharmacophore was then successfully tested on a larger ligand data set to see if it could distinguish between active and inactive compounds. Satisfied with the preliminary evaluation of the pharmacophore template, the authors used the model to design novel compounds. The model was sufficiently well defined to allow docking of 630 compounds and the selection of 30 compounds for testing. Of the 26 compounds acquired and tested, four displayed significantly greater activity than previously identified ligands, showing the feasibility of using pseudo-receptor and docking to identify novel bioactive molecules. The structures of the two most potent molecules are shown in Figure 28.6.

When working with pseudo-receptors, and in general with quantitative structure–activity relationships (QSAR) of any dimension, a word of caution is necessary with respect to the biological data that is used. These should preferably constitute binding affinities from a single laboratory, a prerequisite which is also true for all QSAR studies. Since the receptor models simulate interaction events (ΔH) in a highly simplified manner, the experimental data which are combined with them in a correlation analysis must be as close to the molecular level as possible. It is therefore nonsense to correlate the calculated interaction energies with biological *in vivo* data because the receptor interaction can be blurred or even completely hidden by transport and other pharmacokinetic processes. Sometimes even the use of *in vitro* data is dangerous if a reaction cascade separates the measured event from the receptor-binding interaction. Also the combination of biological data (e.g. IC_{50} values) from different laboratories or assays is extremely dangerous. The reliability and meaning of any QSAR model (3D-QSAR, pseudo-receptor, 4D-QSAR, 5D-QSAR, 6D-QSAR[74]) should always be assessed by the ultimate test of usefulness, by prediction of new compounds.[75] Very often, QSAR models are only internally validated but are never tested whether they are useful in designing novel, more potent compound.[76]

IV. APPLICATION STUDY

A. Pharmacophore-based screening for novel histamine H_3-receptor antagonists

An example from the author's laboratory shall give the reader an informative picture of the pharmacophore generation process and its application to develop novel bioactive compounds.[44] The example deals with antagonists of the human histamine H_3 receptor (hH_3R). hH_3R is a G-protein-coupled receptor (GPCR) for which no exact 3D structures is available, as for any other GPCR. The H_3 receptor modulates the release of various neurotransmitters in the central and peripheral nervous system and therefore is a potential target in the therapy of numerous diseases.[77] Although ligands addressing this receptor are already known, the discovery of alternative lead structures represents a challenging goal in drug design.[78] Experimental structure–activity data for the hH_3R antagonists can be summarized as follows. The pharmacological results suggest that a protonatable nitrogen atom (either in an aromatic imidazole or in a saturated ring system) and an aromatic system separated by a certain distance seem to constitute a potent hH_3R antagonist. Additional polar moieties in the spacer can enhance the antagonistic activity (Figure 28.7).

FIGURE 28.7 Molecular structures of hH_3R antagonists used for the pharmacophore development.

FIGURE 28.8 Interaction of compound **1** with the $hH3_R$ binding site as obtained from the docking study. Only the two important acidic amino acid residues of the binding site are shown for clarity. H-bonds are shown in orange.

FIGURE 28.9 Pharmacophoric features defined based on compound **1**. (Red sphere = any positively charged element, orange sphere = aromatic or hydrophobic group, cyan sphere = aromatic ring).

B. Pharmacophore determination process

Due to their high flexibility and huge structural diversity, hH_3R antagonists also provide difficulties in the generation of pharmacophore models by standard means which normally include the identification of common features required for binding from a ligand set. A dataset of 418 ligands for which hH_3R binding affinities were determined in a $[^3H]N^{\alpha}$-methylhistamine assay was available (pK_i from 5 to 10). A pharmacophore able to discriminate between active and inactive antagonists should be developed on the basis of the known antagonists and be used to virtually screen compound libraries for novel structurally diverse hH_3R antagonists.

For the available ligand dataset a multi-conformer database was generated using the Catalyst software. An energy cut-off of 20 kcal/mol from each energetic minimum structure was set in order to avoid high-energy structures. In a first step three individual pharmacophore models were generated based on the potent antagonists **1**, **2** and **3**. The bioactive conformation of the ligands was deduced from a conformational analysis of semi-rigid hH_3R antagonists and an extensive docking study carried out on a homology model of the hH_3R (Figure 28.8, for details see[79]). The docking study showed that the homology model is able to explain the interaction of the ligands which is in accordance with the known biochemical data (e.g. site-directed mutagenesis data). However, a carried out receptor-based virtual screening was not very successful in discriminating active from inactive antagonists. Therefore, the idea was to carry out a pharmacophore-based virtual screening.

Defining a pharmacophore model upon a ligand has the advantage that this way the individual features are already correctly aligned in space. In order to account for the great structural variability of hH_3R antagonists, the pharmacophores were defined as loose as possible in order to still retrieve most of the validated hH_3R ligands as hits.

Once a pharmacophore capable of retrieving known hH_3R antagonists had been defined, it could be used in a subsequent screening procedure of commercial compound libraries. As an example the generated pharmacophore on the basis of compound **1** is shown in Figure 28.9.

The choice of chemical features was based on functionalities observed in validated hH_3R antagonists and inspection of the binding pocket of the homology model. The linker moiety and the adjacent hydrophobic/π-electron-rich system of the ligands lie in a cleft between *trans*-membrane region (TM) 3, 6 and 7 of the hH_3R. In this region, several aromatic residues border the binding site which are able to interact with the electron-rich system in the hH_3R antagonists. No pharmacophoric features were defined upon the 4-aminoquinoline moiety as a high degree of chemical diversity is observed in active ligands within this region. Any restriction of chemical features was thus avoided.

Apparently, the derived pharmacophore model is too loose-fitting for screening a compound database. Thus, the

FIGURE 28.10 Pharmacophore model based on compound **1** including the shape feature (van der Waals volume) and the two forbidden volumes (black spheres).

van der Waals volume of ligand **1** was included as an additional constraint into the pharmacophore model. Default parameters were used for the definition of the shape query. Finally, also forbidden volumes (black spheres) were defined in order to account for the fact that some ligands extending into these areas were inactive although resembling other active compounds. Figure 28.10 shows the ligand **1** fitted into the complete pharmacophore model.

Using this model, 316 compounds from the 418 ligand dataset were found as hits in a pharmacophore search using the Catalyst[26] program. 93% of the ligands with highest activity were retrieved by the pharmacophore model; less satisfactorily, also 54% of a set of inactive compounds could pass the pharmacophoric filter (Figure 28.11, top). Application of the pharmacophore filter for screening the Maybridge database (MDB) and the World Drug Index (WDI) resulted in 249 and 929 hits, respectively. Thus, 70% of the active and moderate active hH_3R ligands (with a $pK_i > 7$) were retrieved by the pharmacophore, meanwhile from the pool of MDB and WDI ligands (MDB: 59,000 compounds, WDI: 48 00 compounds) 98.9% could be excluded.

However, the filter was still quite loose so that a subsequent definition of further pharmacophoric features could result in a better separation of in/actives. In order to further increase the percentage of active hH_3R ligands found during the virtual screening, further pharmacophore models were defined in a similar way based on compounds **2** ($K_i = 0.33$ nM) and **3** ($K_i = 69$ nM). For the definition of the pharmacophore derived from compound **2**, again the three features described above were used in combination with a shape query and forbidden volumes. The third individual pharmacophore model was defined based on ligand **3** (see Figure 28.11) capable of retrieving 68% of ligands deposited in the hH_3R database. By combining the three pharmacophore models, 369 of 398 (93%) hH_3R ligands with a $pK_i > 7$ could be obtained, while only 2,668 (2.5%) compounds were obtained as hits when screening the MDB and WDI database with together 107,599 structures deposited.

The small percentage of structures from commercial databases matching the pharmacophores showed that the generated models were stringent enough for a reasonable screening.

C. Pharmacophore-based screening of compound libraries

For a more stringent screening, a leave-one-out (LOO) filter was defined on the pharmacophoric features of **1**. The Catalyst LOO model consisted of a combination of five individual pharmacophore models each lacking one pharmacophoric feature found in compound **1** at a time, with the exception of the positive ionizable group and the spacer moiety which were required in all models. The screening of the 2,668 WDI and MDB compounds with the LOO filter reduced the number of hits to 320. In order to ensure that compounds selected by the pharmacophore-based screening could be accommodated into the hH_3R binding site, the 320 hits were docked into the hH_3R binding site using the GOLD[18] program and ranked according to their docking scores. From the top-ranked complexes, seven MDB compounds were selected for experimental testing. The selection of the seven compounds was guided by a cluster analysis in order to select the most structurally diverse compounds among the top-ranked molecules. All compounds showed affinity for the hH_3R with binding affinities ranging from 79 nM to 6.3 μM, thereby showing that the pharmacophore and hH_3R binding site model can be used to identify novel active antagonists. Two compounds, BTB-08079 and RJC-03033, were found to be active in the nanomolar range (Figure 28.12).[44]

In order to determine the structural similarity between the seven retrieved MDB compounds and the 418 hH_3R ligands we calculated similarity indices on the basis of different fingerprint systems (MACCS keys and graph-3-point pharmacophore fingerprints in MOE[31]). Using the different fingerprint systems low similarities were observed between the seven MDB compounds and the original hH_3R antagonists. Interestingly, for the most potent hit BTB-08079 (79 nM), the lowest similarity to the original hH_3R ligand structures was observed. The dimethylaminofuran fragment, which is already known from the potent histamine H_2 receptor antagonist Ranitidine, was not reported before as structural element of potent H_3R antagonists.

Compared to the receptor-based virtual screening, application of the pharmacophore-based search resulted in significantly improved results. While in the docking approach 66.6% of the hH_3R ligands were retrieved, though limiting the number of WDI and MDB compounds to approximately 1,720 structures, application of a pharmacophore-based search allowed retrieval of 93% of active compounds, while reducing the number of WDI and MDB structures to 2,668 compounds (2.5%). The ideal strategy for the flexible hH_3R ligand data set appeared to be, however, a combined approach comprising a pre-screening of commercial

FIGURE 28.11 Enrichment of hH_3R ligands by pharmacophore search based on compound **1**, **2**, and **3**. The percentage of hH_3R ligands retrieved by the individual pharmacophore model within each pK_i-cluster is depicted. The percentage of ligands found in each cluster (dark columns) is written in red numbers and compared to the population of pK_i clusters of all hH_3R compounds in the 418 ligand data set (light gray columns).

FIGURE 28.12 Selected hits from the pharmacophore-based virtual screening.

databases with relatively loose pharmacophore models that mainly reflect the available volume in the binding site (e.g. by considering shape queries of sterically demanding ligands and forbidden volumes derived from ligand superposition) and some essential requirements for binding such as the protonated head group. In order to ensure that compounds selected by the pharmacophore-based screening fit into the binding site, docking of this subset of ligands resulted in a selection of candidates for biological testing.

V. CONCLUSIONS

In spite of the recent success and popularity of pharmacophore-based drug design, one should not forget the limitations of pharmacophore modeling. As with any other model we should be aware with the abstraction that is applied to generate these models. All pharmacophore approaches are based on molecular mechanical abstractions. Thus, properties associated with the interaction of electrons, e.g. polarization effects, are not considered. Another limitation in many pharmacophore-based approaches, is the neglect of the dynamic nature of protein–ligand interaction. Although novel pharmacophore generation programs allow the parallel consideration of multiple/alternative pharmacophores (e.g. in Catalyst[28] or LigandScout[29]), modeling different binding modes is still a problem. It is becoming increasingly clear, that for some protein binding sites, one has to be prepared to consider different binding modes and therefore different pharmacophores.[80–82]

Whereas, in the past, pharmacophore models have been mainly generated using ligand-based strategies, novel programs have been developed and applied successfully in the last few years, by combining structure-based and pharmacophore-based approaches. This is mainly influenced by the rapidly growing number of protein–ligand 3D structures that are the basis for such combined approaches. Closely related to this, one can observe a general merging of different techniques in molecular modeling studies – pharmacophore modeling, 3D-QSAR, *de novo* design and docking[83,84] – which might be helpful for future drug design studies.

REFERENCES

1. Wermuth, C. G., Ganellin, C. R., Lindberg, P., Mitschler, L. A. Glossary of terms used in medicinal chemistry (IUPAC Recommendations 1997). *Annu. Rep. Med. Chem.* 1998, *33*, 385–395.
2. Langley, J. N. On the reaction of cells and nerve-endings to certain poissons, chiefly as regards the reaction of striated muscle to nictoine and to cuare. *J. Physiol.* 1905, *33*, 374–413.
3. Ehrlich, P., Morgenroth, J. Über Haemolysine. Dritte Mitteilung. *Berl. Klin. Wochenschr.* 1900, *37*, 453–457.
4. Fischer, E. Einfluss der Konfiguration auf die Wirkung der Enzyme. *Ber. Dtsch. Chem. Ges.* 1894, *27*, 2985–2993.
5. Perutz, M. F. X-ray analysis of haemoglobin. *Science* 1963, *140*, 863–869.
6. Marshall, G. R. Binding-site modelling of unknown receptors. In *3D QSAR in Drug Design, Theory Methods and Applications* (Kubinyi, H., Ed.), ESCOM: Leiden, 1993, pp. 80–116.
7. Barry, C. D., Ellis, R. A., Graesser, S., Marshall, G. R. Display and manipulation in three dimensions. In *Pertinent Concepts in Computer Graphics* (Faiman, M., Nievergelt, J., Eds), University of Illinois Press: Chicago, IL, 1969, pp. 104–153.
8. Kier, L. B., Aldrich, H. S. A theoretical study of receptor site models for trimethylammonium group interactions. *J. Theor. Biol.* 1974, *46*, 529–541.
9. Gund, P., Wipke, W. T., Langridge, R. Computer searching of a molecular structure file for pharmacophoric patterns. *Comput. Chem. Res. Educ. Technol.* 1974, *3*, 5–21.

10. Poptodorov, K., Liu, T., Hoffmann, R. D. Pharmacophore model generation software tools. In *Pharmacophores and Pharmacophore Searches* (Langer, T., Hoffmann, R. D., Eds), WILEY-VCH: Weinheim, Germany, 2006, pp. 17–47.
11. Evans, B. E. *et al*. Methods for drug discovery: development of potent selective, orally effective cholecystokinin antagonists. *J. Med. Chem.* 1988, *31*, 2235–2246.
12. Wermuth, C. G. Pharmacophores: historical perspective and viewpoint from a medicinal chemist. In *Pharmacophores and Pharmacophore Searches* (Langer, T., Hoffmann, R. D., Eds), WILEY-VCH: Weinheim, Germany, 2006, pp. 1–13.
13. Snyder, J. P., Rao, S. N., Koehler, A., Vedani, A. Minireceptors and pseudoreceptors. In *3D QSAR in Drug Design*, *Theory Methods and Applications* (Kubinyi, H. (Ed.), ESCOM, Leiden, 336–354.
14. Vedani, A., Zbinden, P., Snyder, J. P. Pseudo-receptor modelling: a new concept for the three-dimensional construction of receptor binding sites. *J. Recept. Res.* 1993, *13*, 163–177.
15. Höltje, H. D., Anzali, A. Molecular modelling studies on the digitalis binding site of the Na+/K(+)-ATPase. *Pharmazie* 1992, *47*, 691–698.
16. Frühbeis, H., Klein, R., Wallmeier, H. Computergestütztes Moleküldesign (CAMD) – ein Überblick. *Angew. Chem. Int. Ed. Engl.* 1987, *26*, 403–418.
17. Böhm, H. J., Klebe, G., Kubinyi, H. *Wirkstoffdesign*. Spektrum Akademischer Verlag: Heidelberg, Germany, 1996.
18. Osterberg, F., Morris, G. M., Sanner, M. F., Olson, A. J., Goodsell, D. S. Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in autodock. *Proteins* 2002, *46*, 34–40.
19. Gold, Cambridge Crystallographic Data Centre, Cambridge, UK. http://www.ccdc.cam.ac.uk
20. Glide, Schrödinger, Portland, OR. http://www.schroedinger.com
21. FlexE, BiosolveIT, St. Augustin, Germany. http://www.biolsolveit.de
22. Klebe, G. Virtual ligand screening. *Drug Discov. Today* 2006, *11*, 580–594.
23. Tame, J. R. Scoring functions–the first 100 years. *J. Comput.-Aided Mol. Des.* 2005, *19*, 445–451.
24. Hackling, A., Ghosh, R., Perachon, S., Mann, A., Höltje, H.-D., Wermuth, C. G., Schwartz, J.-G., Sippl, W., Sokoloff, P., Stark, H. *N*-(4-(2-Methoxyphenyl)piperazin-1-yl)alkylcarboxamides as dopamine D_2- and D_3-receptor ligands. *J. Med. Chem.* 2003, *46*, 3883–3899.
25. Cramer, R. D., III, Patterson, D. E., Bunce, J. D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* 1988, *110*, 5959–5967.
26. Ahlstrom, M. M., Ridderstrom, M., Luthman, K., Zamora, I. Virtual screening and scaffold hopping based on GRID molecular interaction fields. *J. Chem. Inform. Model.* 2005, *45*, 1313–1323.
27. Steindl, T. M., Schuster, D., Wolber, G., Laggner, C., Langer, T. High-throughput structure-based pharmacophore modelling as a basis for successful parallel virtual screening. *J. Comput.-Aided Mol. Des.* 2006, *20*, 703–715.
28. Catalyst, Accelrys Software, San Diego, CA. http://www.accelrys.com
29. DISCOtech, St. Louis, MO. http://www.tripos.com
30. Galahad, Tripos, St. Louis, MO. http://www.tripos.com
31. LigandScout, Inteligand, Wien, Austria. http://www.inteligand.com
32. Phase, Schrödinger, Portland, OR. http://www.schroedinger.com
33. MOE, Chemical Computing Group, Montreal, Canada. http://www.chemcomp.com.
34. Patel, Y., Gillet, V. J., Bravi, G., Leach, A. A comparison of the pharmacophore identification programs: Catalyst, DISCO and GASP. *J. Comput.-Aided Mol. Des.* 2002, *16*, 653–681.
35. Boström, J., Norrby, P.-O., Liljefors, T. Conformational energy penalties of protein-bound ligands. 1998, *12*, 383–396.
36. Agrafiotis, D. K., Gibbs, A. C., Zhu, F., Izrailev, S., Martin, E. Conformational sampling of bioactive molecules: a comparative study. *J. Chem. Inform. Model.* 2007, *47*, 1067–1086.
37. Höltje, H.-D., Sippl, W., Rognan, D., Folkers, G. *Molecular Modelling: Basic principles and applications*, 3rd Edition. Wiley-VCH: Weinheim, Germany, 2007.
38. Marshall, G. R., Barry, C. D., Bosshard, H. E., Dammkoehler, R. A., Dunn, D. A. The conformational parameter in drug design: the active analog approach. In *Computer-Assisted Drug Design, ACS Monograph 112* (Olsen, E. C., Christoffersen, R. E., Eds), American Chemical Society: Washington DC., 1979, pp. 205–226.
39. Smellie, A., Teig, S., Towbin, P. Poling: promoting conformational variation. *J. Comput. Chem.* 1994, *16*, 171–187.
40. Kearsley, S. K., Smith, G. M. An alternative method for the alignment of molecular structures: maximizing electrostatic and steric overlap. *Tetrahedron Comput. Methodol*. 1990, *3*, 615–633.
41. Kirchmair, J., Wolber, G., Laggner, C., Langer, T. Comparative performance assessment of the conformational model generators omega and catalyst: a large-scale survey on the retrieval of protein-bound ligand conformations. *J. Chem. Inform. Model.* 2006, *46*, 1848–1861.
42. Boström, J. Reproducing the conformations of protein-bound ligands: a critical evaluation of several popular conformational searching tools. *J. Comput.-Adided Mol. Des.* 2001, *15*, 1137–1152.
43. Omega, OpenEye Scientific Software, Santa Fee, NM. http://www.eyesopen.com
44. Boström, J., Greenwood, J. R., Gottfries, J. Assessing the performance of OMEGA with respect to retrieving bioactive conformations. *J. Mol. Graph. Model.* 2003, *21*, 449–462.
45. FlexS, BiosolveIT, St. Augustin, Germany. http://www.biosolveit.com
46. Lemmen, C., Lengauer, T. Computational methods for the structural alignment of molecules. *J. Comput.-Aided Mol. Des.* 2000, *14*, 215–232.
47. Schlegel, B., Meier, R., Laggner, C., Schnell, D., Langer, T., Seifert, R., Stark, H., Höltje, H.-D., Sippl, W. Generation of a homology model of the human histamine H3 receptor for ligand docking and pharmacophor based screening. *J. Comput.-Aided Mol. Des.* 2007, *21*, 437–453.
48. UNITY, Tripos, St. Louis, MO. http://www.tripos.com
49. Wolber, G., Langer, T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J. Chem. Inform. Model.* 2005, *45*, 160–169.
50. Hindle, S. A., Rarey, M., Buning, C., Lengauer, T. Flexible docking under pharmacophore type constraints. *J. Comput.-Aided Mol. Des.* 2002, *16*, 129–149.
51. Barreca, M. L., De Luca, L., Iraci, N., Rao, A., Ferro, S., Maga, G., Chimirri, A. Structure-based pharmacophore identification of new chemical scaffolds as non-nucleoside reverse transcriptase inhibitors. *J. Chem. Inform. Model.* 2007, *47*, 557–562.
52. Krovat, E. M., Fruhwirth, K. H., Langer, T. Pharmacophore identification, *in silico* screening, and virtual library design for inhibitors of the human factor Xa. *J. Chem. Inform. Model.* 2005, *45*, 146–159.
53. Maccari, R., Ottana, R., Ciurleo, R., Vigorita, M. G., Rakowitz, D., Steindl, T., Langer, T. Evaluation of *in vitro* aldose reductase inhibitory activity of 5-arylidene-2,4-thiazolidinediones. *Bioorg. Med. Chem. Lett.* 2007, *17*, 3886–3893.
54. Laggner, C., Schieferer, C., Fiechtner, B., Poles, G., Hoffmann, R. D., Glossmann, H., Langer, T., Moebius, F. F. Discovery of high-affinity ligands of sigma1 receptor, ERG2, and emopamil binding protein by pharmacophore modelling and virtual screening. *J. Med. Chem.* 2005, *48*, 4754–4764.
55. Spannhoff, A., Heinke, R., Bauer, I., Trojer, P., Metzger, E., Gust, R., Schüle, R., Brosch, G., Sippl, W., Jung, M. Target-based approach to inhibitors of histone arginine methyltransferases. *J. Med. Chem.* 2007, *50*, 2319–2325.
56. Wolber, G., Dornhöfer, A. A., Langer, T. Efficient overlay of small organic molecules using 3D pharmacophores. *J. Comput.-Aided Mol. Des.* 2006, *20*, 773–788.
57. Capdeville, R., Buchdunger, E., Zimmermann, J., Matter, A. Glivec (STI571, imatinib), a rationally developed, targeted anticancer drug. *Nat. Rev. Drug. Discov.* 2002, *7*, 493–502.

58. Mol, C. D., Fabbro, D., Hosfield, D. J. Structural insights into the conformational selectivity of STI-571 and related kinase inhibitors. *Curr. Opin. Drug Discov. Dev.* 2004, *5*, 639–648.
59. Wolber, G., Kosara, R. Pharmacophores from macromolecular complexes with LigandScout. In *Pharmacophores and Pharmacophore Searches* (Langer, T., Hoffmann, R. D., Eds), WILEY-VCH: Weinheim, Germany, 2006, pp. 131–150.
60. Schleifer, K. J. Concept and applications of pseudoreceptors. In *Pharmacophores and Pharmacophore Searches* (Langer, T., Hoffmann, R. D., Eds). WILEY-VCH: Weinheim, Germany, 2006, pp. 117–130.
61. Höltje, H.-D. Theoretical studies on drug-receptor model complexes. *Pharm. Acta Helv.* 1979, *54*, 125–134.
62. Marrer, S., Höltje, H.-D. Molecular modelling study of calcium channel activating and blocking 1,4-dihydropyridine derivatives: a hypothetical molecular receptor model. *J. Protein Chem.* 1989, *8*, 420–421.
63. Sippl, W., Stark, H., Höltje, H.-D. Development of a binding site model for histamine H_3-receptor agonists. *Pharmazie* 1998, *53*, 433–437.
64. Höltje, H. D., Jendretzki, U. K. Construction of a detailed serotoninergic 5-HT_{2a} receptor model. *Arch. Pharm. (Weinheim)* 1995, *328*, 577–584.
65. McMartin, C., Bohacek, R. S. QXP: powerful, rapid computer algorithms for structure-based drug design. *J. Comput.-Aided Mol. Des.* 1997, *11*, 333–344.
66. Vedani, A., Zbinden, P. Quasi-atomistic receptor modelling. A bridge between 3D QSAR and receptor fitting. *Pharm. Acta Helv.* 1998, *73*, 11–18.
67. Walters, D. E., Hinds, R. M. Genetically evolved receptor models: a computational approach to construction of receptor models. *J. Med. Chem.* 1994, *37*, 2527–2536.
68. Lill, M. A., Vedani, A., Dobler, M. Raptor: combining dual-shell representation, induced-fit simulation, and hydrophobicity scoring in receptor modelling: application toward the simulation of structurally diverse ligand sets. *J. Med. Chem.* 2004, *47*, 6174–6186.
69. Lill, M. A., Vedani, A. Combining 4D pharmacophore generation and multidimensional QSAR: modelling ligand binding to the bradykinin B2 receptor. *J. Chem. Inform. Model.* 2006, *46*, 2135–2145.
70. Hahn, M., Rogers, D. Receptor surface models. 2. Application to quantitative structure–activity studies. *J. Med. Chem.* 1995, *38*, 2091–2102.
71. Schleifer, K. J. Pseudoreceptor model for ryanodine derivatives at calcium release channels. *J. Comput.-Aided Mol. Des.* 2000, *14*, 467–475.
72. Vedani, A., Dobler, M., Dollinger, H., Hasselbach, K. M., Birke, F., Lill, M. A. Novel ligands for the chemokine receptor-3 (CCR3): a receptor-modeling study based on 5D-QSAR. *J. Med. Chem.* 2005, *48*, 1515–1527.
73. Bohacek, R., Boosalis, M. S., McMartin, C., Faller, D. V., Perrine, S. P. Identification of novel small-molecule inducers of fetal hemoglobin using pharmacophore and 'PSEUDO' receptor models. *Chem. Biol. Drug Des.* 2006, *67*, 318–328.
74. Vedani, A., Dobler, M., Lill, M. A. Combining protein modeling and 6D-QSAR. Simulating the binding of structurally diverse ligands to the estrogen receptor. *J. Med. Chem.* 2005, *48*, 3700–3703.
75. Sippl, W., Contreras, J. M., Parrot, I., Rival, Y. M., Wermuth, C. G. Receptor-based QSAR and design of novel acetylcholinesterase inhibitors. *J. Comput.-Aided Mol. Des.* 2001, *15*, 395–410.
76. Doweyko, A. M. 3D-QSAR illusions. *J. Comput.-Aided Mol. Des.* 2004, *18*, 587–596.
77. Ligneau, X. *et al*. BF2.649 [1-{3-[3-(4-Chlorophenyl) propoxy]propyl} piperidine, hydrochloride], a nonimidazole inverse agonist/antagonist at the human histamine H_3 receptor: preclinical pharmacology. *J. Pharmacol. Exp. Therapeut.* 2007, *320*, 365–375.
78. Stark, H., Kathmann, M., Schlicker, E., Schunack, W., Schlegel, B., Sippl, W. Medicinal chemical and pharmacological aspects of imidazole-containing histamine H_3 receptor antagonists. *Mini Rev. Med. Chem.* 2004, *4*, 965–977.
79. Schlegel, B., Sippl, W., Höltje, H.-D. Molecular dynamics simulations of bovine rhodopsin: influence of protonation states and different membrane-mimicking environments. *J. Mol. Model.* 2005, *12*, 49–64.
80. Steuber, H., Zentgraf, M., La Motta, C., Sartini, S., Heine, A., Klebe, G. Evidence for a novel binding site conformer of aldose reductase in ligand-bound state. *J. Mol. Biol.* 2007, *369*, 186–197.
81. Steuber, H., Zentgraf, M., Gerlach, C., Sotriffer, C. A., Heine, A., Klebe, G. Expect the unexpected or caveat for drug designers: multiple structure determinations using aldose reductase crystals treated under varying soaking and co-crystallisation conditions. *J. Mol. Biol.* 2006, *363*, 174–187.
82. Specker, E., Bottcher, J., Brass, S., Heine, A., Lilie, H., Schoop, A., Muller, G., Griebenow, N., Klebe, G. Unexpected novel binding mode of pyrrolidine-based aspartyl protease inhibitors: design, synthesis and crystal structure in complex with HIV protease. *Chem. Med. Chem.* 2006, *1*, 106–117.
83. Sippl, W. Receptor-based 3D QSAR analysis of estrogen receptor ligands-merging the accuracy of receptor-based alignments with the computational efficiency of ligand-based methods. *J. Comput.-Aided Mol. Des.* 2000, *14*, 559–572.
84. Evers, A., Gohlke, H., Klebe, G. Ligand-supported homology modelling of protein binding-sites using knowledge-based potentials. *J. Mol. Biol.* 2003, *334*, 327–345.

Chapter 29

3D Quantitative Structure–Property Relationships

Thierry Langer and Sharon D. Bryant

I. INTRODUCTION
II. 3D QSAR WORKFLOW
III. 3D QSAR: CONFORMATION ANALYSIS AND MOLECULAR SUPERIMPOSITION
IV. CALCULATION OF 3D MOLECULAR FIELD DESCRIPTORS
V. STATISTICAL TOOLS
VI. ALIGNMENT INDEPENDENT 3D QSAR TECHNIQUES
VII. VALIDATION OF 3D QSAR MODELS
VIII. APPLICATIONS
A. 3D QSAR study on the structural requirements for inhibiting AChE
B. 3D QSAR as a tool to determine molecular similarity
IX. CONCLUSIONS AND FUTURE DIRECTIONS
REFERENCES

I am enough of an artist to draw freely upon my imagination. Imagination is more important than knowledge. Knowledge is limited. Imagination encircles the world.

Albert Einstein (1879–1955)

I. INTRODUCTION

One of the most important goals in computer-assisted ligand optimization is to find a correlation between structural features of molecules and their biological activity, that is, their ability to bind to specific macromolecular targets. In some cases, simple mathematical models may provide a means for identifying the property related to affinity, in other cases a multitude of parameters is necessary to describe the behavior of a compound in a complex biochemical system. In general, such parameters can be derived by forming a relationship between those variables that describe the structural variation within the group of compounds under investigation, and those that describe their bioactivities. This relationship is denoted as quantitative structure–activity relationship (QSAR).[1–3] In traditional QSAR, often referred as 2D QSAR, a so-called *training set* of compounds with known biological activity is analyzed in order to predict the property (in most cases: the biological activity or affinity to a target) of compounds to be synthesized or selected for purchase. Using statistical analysis, one considers the molecular descriptors and the effects of substituents on biological activity. Molecular descriptors are measured or calculated physicochemical properties, such as $\log P$, pK_a, boiling point, melting point, molar refraction, etc.[4] If prepared correctly, this strategy identifies which structural variations are the dominant ones influencing the change in biological properties. Usually, the mathematical models obtained by regression analysis are validated for their predictive power by analyzing their capability to correctly predict the biological properties of compounds belonging to a so-called *test set*, that is, a set of molecules with known biological activity that was not used for generating the initial model.

Traditional QSAR has been used successfully for optimizing congeneric series of lead structures, which has led

Copyright © 2008, Elsevier Ltd
All rights reserved.

to the development of several commercial drugs.[5] However, this approach can have severe limitations. Besides the thermodynamic relationship requirement, where only a series of compounds having a common core fragment should be considered, the configuration and conformational effects are not taken into account, as in general, information regarding the 3D structure of the compounds is not considered at all. On the other hand, during the last three decades, molecular modeling techniques have become extremely popular, in a large part due to the rapid development of powerful computer hardware and related software packages. Such methods can be used easily for computing accurate 3D structures of molecules, and the energies of large numbers of their possible conformations. Moreover, possible interactions between the molecule and its binding site can be studied at different levels of approximation. Thus, from this approach, it became possible to describe each molecule/conformation with a series of theoretically computed parameters that are intrinsically 3D in nature.

As a logical consequence, these methods were introduced in QSAR analyses and a new discipline, 3D QSAR, was born. The major breakthrough in this field was achieved by Cramer *et al.*,[6] who in 1988 described the application of comparative molecular field analysis (CoMFA) for the evaluation of the binding properties of steroids to carrier proteins. This approach, for which interestingly a patent application was filed by Cramer and Wold,[7] and assigned to the software company Tripos,[8] was based primarily on the previous work of Wise *et al.,*[9,10] on a dynamic lattice-oriented molecular modeling system (DYLOMMS) dating back to the early 1980s. Concepts that aided the development of these early 3D QSAR experiments were, for example, Höltje's theoretical models on binding sites,[11–13] and Marshall's active analog approach.[14] It took almost a decade for the scientific community to accept molecular fields instead of physicochemical parameters for the description of ligand–target interactions and without the ongoing tenacity of both Cramer and Wold (who had developed partial least squares (PLS) as a method for analyzing unsymmetrical variable matrices)[15] and the development of appropriate and user-friendly software, probably these methods would have been accepted even later. Two years after the first CoMFA paper in 1988, a 3D QSAR hype started, and the number of papers grew rapidly (Figure 29.1). Note that there were other approaches elaborated at the same time or even earlier, for example, GRID/GOLPE,[16] based on Goodford's GRID program for analyzing protein binding sites,[17] and the chemometric statistical tool GOLPE (generating optimal linear PLS estimations) developed in Clementi's group.[18] As of September 2007, the number of papers dealing with 3D QSAR has reached a sum higher than 2,500 when searching the CAS (Chemical Abstracts Service) online service using the keyword 3D QSAR. The most important journals containing relevant articles about this topic are listed in Table 29.1.

However, the initial hype is over and the number of new 3D QSAR applications has been more or less constant in the last years. One major drawback of 3D QSAR is that it is not applicable to huge datasets containing more than several hundreds of compounds as is possible with high-throughput

FIGURE 29.1 Number of articles containing CoMFA studies published between 1988 and 2007.

screening setups. Instead, techniques utilizing *in silico* or virtual screening, as outlined in other chapters of this book (Chapter 10), are more applicable.

In this chapter, we want to give a short overview on 3D QSAR methods and report how to use them optimally in the context of medicinal chemistry. We will focus primarily on CoMFA as it has been the most successful and widely used approach in this field. For more detailed information, the reader is referred to several reference books[19–21] and reviews.[22–24]

TABLE 29.1 Scientific Journals Containing Articles about 3D QSAR, as Referred by CAS Online September 2007

Journal	Number of 3D QSAR articles
J. Med. Chem.	381
Bioorg. Med. Chem.	204
J. Comput.-Aided Mol. Des.	130
*J. Chem. Inf. Comput. Sci.**	124
Bioorg. Med. Chem. Lett.	64
Quant. Struct.-Act. Relat.	48
Eur. J. Med. Chem.	47
J. Mol. Graph. Model.	45

*Incl. J. Chem. Inf. Model.

II. 3D QSAR WORKFLOW

The aim of any 3D QSAR study is to determine a correlation between a compound's molecular field and its biological activity, that is, the affinity to its target. One of the first steps is to delineate which parts of the molecules form a common core substructure, and which side chains are responsible for specific interactions. These facts will determine for the most part, how the compounds are to be superimposed in the grid for the molecular field calculation. From the user's point of view, a typical 3D QSAR analysis workflow (Figure 29.2) contains several consecutive steps:

1. Training and test set selection
2. Compound preparation (3D structure generation, conformational analysis)
3. Molecular alignment and superimposition
4. Calculation of 3D molecular field descriptors
5. Model generation using correlation analysis
6. Model validation
7. Prediction

For training and test set selection, the same criteria apply as for any other QSAR method, and the reader is referred to literature covering this topic.[25] In the 3D QSAR context, the most critical ones are steps 3 and 4 as they can have dramatic impact on the results of the study, even if the same initial data set is used for model generation. Therefore, these steps will be discussed in detail in the following sections.

The quality and reliability of any 3D QSAR model is strongly dependent on the careful examination of each step within a 3D QSAR analysis. As with any QSAR method, an important point to consider is if the biological activities

FIGURE 29.2 Schematic representation of a CoMFA workflow.

of all compounds studied are of comparable quality. Preferably, biological data has been obtained in the same laboratory under identical conditions. All compounds being tested in a system must have the same mechanism of binding and all inactive compounds must be shown to be truly inactive. Only *in vitro* data should be considered, since only *in vitro* experiments are able to reach a real equilibrium. All other test systems undergo time-dependent changes by multiple coupling to parallel biochemical processes (e.g. membrane permeation). Another critical point is the existence of transport phenomena and diffusion gradients underlying all biological data. One has to bear in mind, that all 3D QSAR approaches were developed to describe only one interaction step in the lifetime of ligands. In all cases, where non-linear phenomena result from drug transport and distribution, any 3D QSAR technique should be applied with caution. The biological activities of the molecules used in a 3D QSAR study should ideally span a range of at least three orders of magnitude. For all molecules under study, the exact 3D structure has to be reported. If no information on the exact stereochemistry of the tested compounds is given (mixtures of enantiomers or diastereomers), then those compounds should be excluded from the 3D QSAR study.

III. 3D QSAR: CONFORMATION ANALYSIS AND MOLECULAR SUPERIMPOSITION

For investigating 3D structure–activity relationships in a traditional 3D QSAR analysis, the compounds must be aligned according to their assumed mode of binding to their target. In the ideal case, a 3D QSAR analysis would be straightforward: superimposition of all the 3D complex structures would serve as input for the molecular field calculations, assuming that the structures while binding to the binding site were determined experimentally for all of the compounds under investigation. However, this situation is far from being the regular case and therefore, the above mentioned requirements pose several problems. Firstly, even if the structure of the target binding site is known, different ligands might bind in different orientations to the target, and therefore the user must select a representative binding orientation. Secondly, if the target structure is not known, the user has to assume a certain binding mode and perform the superimposition according to this assumption. In this case, additionally, the user has to choose a certain conformation in order to have an appropriate starting point for the molecular superimposition. For the majority of cases, neither the bioactive conformation of a compound nor the binding mode of the compound to its receptor is known. This is especially problematic for membrane bound enzymes and integral membrane proteins such as receptors, ion channels, and transporters, for which there are no crystallized structures available. Therefore, assumptions have to be made, in order to generate comparable 3D structures of the compounds under investigation. Normally it is assumed that the ligands, regardless of chemical composition, bind in conformations that reflect complementary steric and electrostatic potential patterns to the target, which, in fact is the basis of the *pharmacophore concept*. A pharmacophore is defined as the ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target structure and to trigger (or to block) its biological response.[26] Thus, the structural features of such a pharmacophore are complementary to the binding features in the binding site. One must, however, keep in mind that this is dependent on the tolerance applied and that a single binding site may recognize multiple pharmacophores.

The conformational flexibility of most ligands, yielding a multitude of possible patterns that the ligand can potentially present to the target, is a major issue for deducing self-consistent molecular alignments. In order to solve this problem, several steps must be considered. One should determine by means of analysis of chemical modifications, the relative importance of each functional group in a ligand. Then, one should consider corresponding properties between different functional groups across a series of compounds. After making these evaluations, constrained conformational analyses can be applied. Conformational analysis using constraints is a computational procedure that forces molecules to assume a conformation as similar as possible to that of a more rigid template molecule in a presumably bioactive conformation. Modern conformational analysis tools provided in OMEGA[27] or Catalyst[28] can reproduce quality bioactive conformations relative to their generated conformational ensembles.[29,30] In cases where pharmacophoric groups are present, they can be used efficiently for defining alignments within large ensembles of conformers.[31] In addition, these groups can be tethered during molecular mechanics minimization routines forcing similar groups to be maximally superimposed while allowing the rest of the molecule to relax to a low strain energy state. Alternatively, field fit minimization,[32] and the SEAL approach[33] are two examples of automated alignment techniques in CoMFA studies[34,35] that utilize algorithms which force potential energy fields of molecules, not atoms, to be as similar as possible. In a recent review, Lemmen and Lengauer present a detailed description of many alternative computational methods for molecular structure alignments.[36]

Another drawback of 3D molecular descriptors worth consideration is that they are sensitive to the position and orientation of the molecular structures in space. Tropsha and Cho have demonstrated[37] that the results of conventional CoMFA are sensitive to the overall orientation of superimposed molecules. They show that for a given alignment analyzed in different orientations, the single q^2 value obtained from standard CoMFA will most likely fall within the region of the highest frequency of q^2. Therefore, they propose cross-validated r^2 guided region selection for CoMFA

studies. On the other hand, the low q^2 value obtained from conventional CoMFA (which, in many cases, will not be reported in the literature) may not necessarily be a result of a poor alignment, but may be caused merely by poor orientation of superimposed molecules. Thus, simple reorientation of the whole set may significantly improve the results.

In any case, the initial alignment of compounds in the series is widely recognized as one of the most difficult and time-consuming steps in 3D QSAR analyses. Moreover, when the ligands are highly flexible, superimposition becomes impractical if not completely subjective. The problem is that no statistical tools can inform the users about a bad superimposition procedure. Usage of automated pharmacophore-based alignments as input for 3D QSAR has been useful in this respect.[38–40] However, the only strong validation method is the model interpretation linked with careful analysis of experimental tests, since no statistical procedure is able to validate the model independently from the superposition performed by the user. Therefore, significant efforts have been devoted to the development of alignment-free 3D molecular descriptors. Such approaches will be discussed in Section VI of this chapter.

IV. CALCULATION OF 3D MOLECULAR FIELD DESCRIPTORS

Molecular interaction fields (MIFs) can be calculated for any molecule of known 3D structure. A MIF describes the spatial variation of the interaction energy between a molecular target and a chosen probe. The target may be a macromolecule or a low-molecular weight compound or a molecular complex. The probe may be a molecule or a fragment of a molecule. The term "field" usually refers to a potential or other scalar property. In fact, molecular fields are derivatives of a potential and therefore are vector quantities. For instance, the molecular electrostatic potentials of molecules may be calculated easily at any position in the surrounding space, resulting in continuous scalar quantities. The derivatives of this potential give the vector field, which is far more complicated to use for further calculations since at each point there are three values (one for each main axis of the Cartesian space) of the field to be considered. In the molecular modeling context, MIFs are useful for establishing QSAR, for example, using the CoMFA or the GRID/GOLPE approach. Fields used for these types of studies represent a discrete type of fields, since they consist of a 3D matrix of scalar values obtained by calculating interaction energies at all grid points of a defined lattice between a probe and a molecule.

MIFs can be applied in many ways, and the reader is referred to Cruciani's recently published book on this topic.[41] In addition to their frequent usage in deriving 3D QSARs for low-molecular weight compounds, they can guide the process of structure-based ligand design, they can be used to dock ligands to macromolecules and to study the structure–activity relationships of macromolecules, and can be applied to the prediction of pharmacokinetic properties, such as in the VolSurf[42] and MetaSite[43] methodology. In the following, we will focus on the application of MIFs in 3D QSAR.

When a MIF is calculated starting with the biomolecular target structure, in order to analyze, for example, a binding site of a protein, regions are identified where certain chemical groups can interact favorably with other chemical groups within the environment. Thus, this procedure would suggest positions in the protein where ligands should place complementary groups. If the MIFs are calculated from the ligands themselves, regions showing a favorable energy of interaction would represent positions where groups of a potential receptor would interact favorably with the ligand. A virtual receptor site can be generated using a set of MIFs for a set of ligands. For a protein binding site, on the other hand, a set of regions that should be "filled" by appropriate groups of a ligand is identified.

For calculating MIFs, so-called probes (an atom or small molecular fragments) are moved around in a rectangular box of grid points in which the target molecules are aligned in order to produce a matrix of interaction energy values. Depending on the computational procedure used, and on the nature of the probe, the obtained interaction energy field may represent total interaction energies (GRID),[17] steric or electrostatic fields (CoMFA),[6–8] molecular lipophilic potentials (CoMPA, MLP),[44,45] hydrophobic interactions (HINT),[46] or electron densities, etc. These fields may be used as point descriptors of the 3D molecular structure and physicochemical behavior of the target molecule, since most properties related to molecular interactions can be represented in such a MIF. In addition, computer-aided graphical analysis of such fields allows simple interpretation through the visualization of regions where the probe interacts most strongly with the target either by attraction or repulsion.

As mentioned, there are several methods available for MIF calculations. In the medicinal chemistry context, the fields obtained using the GRID force field[17] have been found to be the most appropriate and useful, as shown in many application examples.[16] The interaction of small organic molecules with biological receptors is mainly mediated by surface properties such as electrostatic forces, hydrogen-bond formation capability, shape, and hydrophobicity. The potential functions for deriving such interaction energies in GRID have always been calibrated as much as possible by studying experimental measurements, and the calibration is validated by studying how well GRID predicts observed crystal structures, since crystal packing is determined by free energy considerations rather than by enthalpy alone. An advantage of the GRID approach, apart from the large number of chemical probes available, is the use of a 6-4 potential function, which is smoother than the 6-12 form of the Lennard–Jones function, for calculating

the interaction energies at the grid lattice points. The force field also includes entropic terms; GRID therefore can detect the hydrophobic binding regions, which are important for the design of high-affinity ligands, and it can also detect sites for the polar groups, which determine ligand selectivity. More than a hundred different pre-computed probes are available in the program, together with the possibility to build any new chemical probe of interest. Among them, the hydrophilic *water* probe is used for simulating solvation/desolvation processes and hydrogen-bond interactions. The hydrophobic contribution to binding, which is important to consider in order to enhance affinity, is computed at each grid point as $E_{entropy} + E_{LJ} - E_{HB}$ using the hydrophobic *DRY* probe. The term $E_{entropy}$ refers to the ideal entropic component of the hydrophobic effect in an aqueous environment. E_{LJ} measures the induction and dispersion interactions occurring between any pair of molecules, and E_{HB} represents the H-bonding interactions between water molecules and polar groups on the molecular target surface. Additional details on the calculation of MIFs can be found in a recent review by Wade.[47]

Once calculated, the MIFs obtained from a series of compounds, can be visually compared and studied. Naturally, when a large number of compounds is studied, there are limitations in the graphical representation of such fields. Therefore, a statistical analysis using appropriate methods needs to be applied (see Section V of this chapter) in order to correlate the field descriptors with observed variables, for example, the biological activity represented as IC_{50} or as K_i values). It is important to mention that one should use logarithmic transformation of bioactivity values in the data table. This is due to the fact that calculated binding energies correlate logarithmically with experimentally determined binding (or inhibition) constants instead of linearly. Moreover, it has been shown that only enthalpies (ΔH) and not free energies (ΔG) or binding constants can be properly predicted by CoMFA.[48]

V. STATISTICAL TOOLS

It is clear that for an unsymmetrical data matrix that contains more variables (the field descriptors at each point of the grid for each probe used for calculation) than observables (the biological activity values), classical correlation analysis as multilinear regression analysis would fail. All 3D QSAR methods benefit from the development of PLS analysis,[15] a statistical technique that aims to find the multidimensional direction in the X space that explains the maximum multidimensional variance direction in the Y space. PLS is related to principal component analysis (PCA).[49,50] However, instead of finding the hyperplanes of maximum variance, it finds a linear model describing some predicted variables in terms of other observable variables and therefore can be used directly for prediction. Complexity reduction and data simplification are two of the most important features of such tools. PCA and PLS both condense the overall information into two smaller matrices, namely the score plot (which shows the pattern of objects, in this case of compounds) and the loading plot (which shows the pattern of descriptors). Because the interpretation of score and loading plots is simple and straightforward in the chemical sense, PCA and PLS are usually preferred to other non-linear methods, especially when the noise is relatively high.[42] The score and loading plots are interconnected, so that any descriptor change in the loading plot is reflected by changes in the position of compounds in the score plot. Objects located close to each other in the score plot exhibit a higher degree of similarity in their descriptor space than objects located far from each other. On the other hand, in the loading plot, inter-correlation among parameters can easily be detected, since similar parameters group in the vicinity of each other. Observation and interpretation of compound scoring plots is the basis of factorial designs.[51–53]

PCA is a least square method and therefore its results depend on data scaling. The initial variance of a column variable partly determines its importance in the model. In order to avoid the problem of over- or under-representation of variables, column variables are scaled to unit variance before analysis. The column average is then subtracted from each variable, which, from a statistical point of view, corresponds to moving the multivariate system to the center of the data, which becomes the starting point of the mathematical analysis. The same auto-scaling and centering procedures are applied in PLS discriminant analysis.

Once the statistical model is established, predictions for a test set or for new compounds are made by projecting the compound descriptor into the PCA or PLS model. For PCA, this is achieved by calculating the score vector T of descriptors X and average $\bar{x}$ for the new compounds, using the loading P of the model, according to equation (29.1):

$$T = (X - \bar{x})P'(P \cdot P')^{-1} \tag{29.1}$$

Usually, for PLS discrimination, external predictions are made using the following equation (29.2):

$$Y = \bar{y} - \bar{x} \cdot P'(P \cdot P')B \cdot Q + X \cdot P'(P \cdot P')^{-1}B \cdot Q \tag{29.2}$$

where $\bar{y}$ is the Y column average and Q is the loading vector for the y space and B the coefficient between the X and Y spaces.

VI. ALIGNMENT INDEPENDENT 3D QSAR TECHNIQUES

As already outlined before, the most critical step in 3D QSAR analysis is the correct alignment of molecular structures before calculating the MIFs. In order to avoid

the problem that any mistake in the superimposition procedure will affect the rest of the study, novel MIF-derived descriptors were developed that are alignment independent. A recent review about these efforts was published by Pastor.[54] The most prominent examples are the descriptors developed by Cruciani *et al.*, namely the VolSurf descriptors[55,56] and the GRIND (GRid INdependent descriptors).[57]

VolSurf is a computational procedure to produce 2D molecular descriptors from 3D molecular interaction energy grid maps. The basic idea of VolSurf is to compress the information present in 3D maps into a few 2D numerical descriptors that are simple to understand and interpret. VolSurf descriptors are designed specifically for the optimization of *in silico* pharmacokinetic properties. In the standard procedure, interaction fields with a *water* probe and the hydrophobic *DRY* probe are calculated all around the target molecules as in the program GRID. Furthermore, additional polar or charged probes can be used for extracting more specific information, according to the target under investigation. Similar to 2D images, each 3D molecular field map is made of a regular lattice of boxes called voxels, which represents attractive and repulsive forces between an interacting partner and a molecule. Each voxel is defined by a volume, a surface and by an interaction energy value. By contouring the voxels at different energy levels, different images can be obtained. Consequently, the images are used by VolSurf to compute the volumes and the surfaces related to the contouring method selected. In the building phase of volumes, the voxels are grouped by a shape function that assigns the value of 1 to voxels inside an energy range, and 0 to all other voxels. Subsequently, a simple summation over the selected voxels yields back the overall volume for the considered property. The descriptors have a clear chemical meaning and are lattice-independent, and some of them can be projected back into the original 3D-grid map from which they were obtained. VolSurf descriptors have been used successfully in numerous 3D-QSAR studies[58] including research on pharmacokinetic properties of potential drug molecules,[59,60] for stability prediction of thermostable enzymes,[61] and even in biochemical research to name a few.

GRIND,[57] implemented in the software ALMOND[62] represents a new generation of alignment independent descriptors also derived directly from MIFs. The basic idea behind these descriptors is to condense the information present in the field by recognizing a number of highly relevant regions and to describe their spatial distribution on the basis of their mutual distances. For a specific compound, a vector of values is obtained, each one representing the presence or absence of a couple of nodes separated by a certain distance and each one belonging to a different "highly relevant region" in the MIF. The value is set to 0 when the MIF contains no such couple of nodes at the given distance, and when there are one or more couple of nodes the product of the MIF values at these positions are set to positive. This method allows one to condense two kinds of information into the variables: the presence or not in the MIF of a couple of nodes separated by a certain distance and the overall intensity of the MIF at both ends of the distance. The variables so obtained are therefore richer in information and more suitable for analysis using chemometric methods like PCA and PLS. Areas that are recognized as highly relevant regions are those in the MIF where the ligand can establish strongly favorable interactions and therefore represent the position of binding site groups, which can putatively bind the compound. Such areas are grouped into so-called nodes, a procedure that so far needs some human interaction. Nevertheless, the definition of such nodes clearly focuses the application of GRIND toward the description of ligand–target binding. Once the MIF is reduced to a set of nodes representing the most relevant regions, the next step is to describe the spatial position of these nodes neither using their absolute coordinates nor any external reference system. The solution adopted in GRIND is based on describing the node–node distances and the MIF energies represented by such nodes. In order to do this, the node–node distances are first converted into a discrete set of distance ranges or distance "bins," then every couple of selected nodes is analyzed in turn, measuring their mutual distance and assigning them to a certain distance bin. At the end of the analysis, every bin is represented by the couple of nodes for which the product of their MIF is higher, thus representing the more favorable energy at both ends of the line linking the nodes. The result of this analysis is an array of values, one for each distance bin, containing 0 if no couple of nodes is found, and an energy score representing the largest product of MIF, if one or more couples is found. These results can be seen, from a mathematical point of view, as a set of n quantitative-continuous variables, each representing a range of distances. It is possible to represent these values in a correlogram, in which the distances are represented on the horizontal axis and the scaled product of energies on the vertical axis. A characteristic of this geometry representation is that it is completely alignment independent, since the values assigned to every variable depend only on the mutual distance of the nodes and not on the position of the nodes in space. Therefore, if this analysis is carried out for a series of molecules not aligned in space, the variables obtained for every compound do nevertheless have the same meaning and the values obtained can be combined to build a consistent descriptor matrix, without the need to align or otherwise superimpose their structures. The reader can find a detailed description of GRIND in Pastor's recent paper.[54] These descriptors have been successfully used in several 3D QSAR/QSPR[63–66] and molecular similarity determination[67,68] applications. The Anchor-GRIND method introduced by Fontaine *et al.*,[69] uses a specific position of the molecular structure (the "anchor point") to compare the spatial distribution of the MIFs of the substituents. The descriptors produced are more detailed and specific than the original GRIND while still avoiding the bias introduced by the alignment.

Finally, in this context, Hopfinger's 4D QSAR approach should be mentioned.[70–72] In this concept, the flexibility and alignment problem of traditional 3D QSAR is solved by incorporating multiple conformers and multiple alignments per molecule. The descriptors in this type of 4D QSAR analyses are the grid cell (spatial) occupancy measures of the atoms composing each molecule in the training set realized from the sampling of conformation and alignment spaces. Grid cell occupancy descriptors are being generated for any atom type, group, and/or model pharmacophore. A single "active" conformation is then postulated for each compound in the training set and combined with the optimal alignment for use in other molecular design applications, including other 3D QSAR methods. The influence of the conformational entropy of each compound on its activity is estimated and serial use of PLS, regression, and a genetic algorithm (GA) is used to perform data reduction and identify the manifold of top 3D QSAR models for a training set. The unique manifold of 3D QSAR models is determined by computing the extent of orthogonality in the residuals of error among the most significant 3D QSAR models in the general GA population. Receptor independent (RI) 4D QSAR analysis has been used successfully in numerous cases which have been reviewed in a recent paper by Albuquerque *et al.*[73]

VII. VALIDATION OF 3D QSAR MODELS

The final stage of every 3D QSAR analysis should consist in a thorough statistical validation in order to assess the significance of the model and hence its ability to predict biological activities of novel compounds. In most of the 3D QSAR case studies cited in the literature, the leave-one-out (LOO) cross-validation procedure has been used for this purpose. As an output of this procedure, the user obtains the so-called cross-validated q^2 and the standard deviation of error of prediction (SDEP), which are commonly regarded as ultimate criterion of both robustness and the predictive ability of a model. The simplest cross-validation method is LOO, where one object at a time is removed from the dataset and predicted by the model generated. A more robust and therefore more reliable method is the leave-several-out cross-validation. For example, in the leave-20%-out cross-validation, five groups of approximately the same size are generated. Thus, 80% of the compounds are randomly selected for the generation of a model, which is then used to predict the remaining compounds. This operation must be repeated numerous times in order to obtain reliable statistical results. The leave-20%-out or also the more demanding leave-50%-out cross-validation results are much better indicators for the robustness and the predictive ability of a 3D QSAR model than the usually used LOO procedure.[74,75] LOO often yields models that are too optimistic, which later fail when predicting real test set molecules. Despite the known limitations of the LOO procedure reported in the literature, it is not common to test 3D QSAR models for their ability to correctly predict the biological activities of compounds not included in the training set. Still, many authors claim that their models, showing high LOO q^2 values, have high predictive ability in the absence of external validation (for a detailed discussion on this problem see Refs. [76–79]). In contrary to such expectations, however, it has been shown by several studies that a correlation between the LOO cross-validated q^2 value for the training set and the correlation coefficient r^2 between the predicted and observed activities for the test set, does not exist.[76,78] Therefore, it is highly recommended to use demanding cross-validation procedures and external test sets to further validate an established 3D QSAR model.

VIII. APPLICATIONS

A. 3D QSAR study on the structural requirements for inhibiting AChE

To demonstrate the potential of 3D QSAR, a case study is presented here, where a structure-based CoMFA was used for the design of novel acetylcholine esterase (AChE) inhibitors.[80,81] AChE is an enzyme that hydrolyzes the neurotransmitter acetylcholine (ACh) at cholinergic synapses with a turnover rate superior to most other known enzymes.[82] Recent research interest regarding this enzyme is not only due to this high catalytic efficiency, but also due to the broad implications of AChE inhibition on human health, agrochemistry, and chemical agents. For example, Alzheimer's disease (AD) is associated with low *in vivo* levels of acetylcholine; thus, AChE has been the focus of many drug discovery projects aimed at maintaining ACh available via mild or reversible inhibitors such as tacrine and donepezil, etc.[83,84] The cholinergic hypothesis of AD has provided the rationale for the current major therapeutic approach to AD. However, to date, all long-term studies have shown that clinical efficacy declines gradually as a result of either a loss of drug efficacy or the relentless progression of the disease. Thus, interest in the discovery of novel AChE inhibitors is continued since the current AChE inhibitors lack desired therapeutic effects. The availability of the AChE crystal structures of various species in its un/complexed form provides a solid basis for structure-based design of novel AChE inhibitors.[85] Within AchE, two principle binding sites can be found. The catalytic active site is buried at the bottom of a deep gorge in the enzyme. The Ach catalysis reaction is accomplished by a collective interaction of the catalytic triad consisting of Ser203, Glu334, and His447 and nearby residues (e.g. the choline binding site: Trp86). AChE also has a peripheral anionic site (PAS) located near the enzyme surface at the entrance of the active site gorge. In the PAS, residue Trp286 plays

FIGURE 29.3 Example of minaprine-derived AChE inhibitors used in the study by Sippl *et al.*,[80,81,90]

a very important role in ligand binding that affects enzymatic activity through a combination of steric blockade of ligands moving through the gorge and allosteric alteration of the catalytic triad conformation and efficiency.[86] The gorge itself is a narrow hydrophobic channel with a length of 20 Å connecting the PAS site to the active site.[87] An acyl binding pocket consisting of residues Gly122, Trp236, Phe295, Phe297, and Phe338 is responsible for the interaction with the acetyl group of ACh.[88] Early inhibition research focused on ligands binding in the active site (e.g. tacrine, amiridine, etc.). Recent efforts, however, focused on finding novel ligands that bind to both sites in order to obtain more potent reversible inhibitors, selectively favoring the inhibition of AChE rather than the related butyrylcholinesterase (BChE).

The starting point of the present AChE research project was the finding that the morpholine derivative minaprine showed weak inhibition of AchE.[89] Starting with the lead structure minaprine and the available X-ray structures of AChE a variety of minaprine derivatives were developed (Figure 29.3).[90] Using information available from a detailed inspection of the available AChE–inhibitor X-ray structures, the alignment rule for the 3D QSAR was established. Structures of AChE inhibitors were docked into the binding site using the AutoDock[91] procedure in combination with a force-field refinement. This approach yielded good results when docking the AChE inhibitors. AutoDock uses a Lamarckian GA to explore the binding possibilities of a ligand in a binding pocket. The interaction energy of ligand and protein is evaluated using atom affinity potentials calculated on a grid similar to that described by Goodford.[17] The minimized uncomplexed AChE was used as input structure for the docking simulations.

During the docking procedure, all ligand atoms were considered flexible, while protein atoms were kept fixed. The 100 resulting complexes were clustered with a RMSD tolerance of 0.7 Å. In a second step, low-energy complexes were ranked again according to the interaction energy calculated with a more detailed energetic model based on the YETI force field.[92,93] The 20 top-ranked complexes of the AutoDock output were selected. The protein structure was kept fixed during the minimization, whereas the ligand was allowed to change its conformation and position in the binding pocket. During the minimization, the ligand conformation relaxed into a neighboring local energy minimum. For the molecular interaction energy field calculation, the program GRID[17] was used with the *trimethylammonium*, *methyl*, *carbonyl*, *amide*, and *DRY* probes. The authors demonstrated their ability to accurately predict the binding conformation of tacrine, decamethonium, edrophonium, and huperzine. This gave them confidence to use their model to evaluate the binding conformation of new aminopyridazine compounds.

Figure 29.4 shows the predicted position of an aminopyridazine in comparison to the position of decamethonium observed in the corresponding crystal structure. The hydrophobic parts of the aminopyridazine inhibitors interact with various aromatic residues in the binding pocket. The benzyl

FIGURE 29.4 Comparison of the predicted position of the aminopyridazine **3y** (dark-gray) and the X-ray structure of the AChE–decamethonium (gray) complex.

FIGURE 29.5 Receptor-based alignment of all investigated inhibitors obtained by docking analyses. The solvent accessible surface of the binding pocket is displayed.

ring of the inhibitor displays classic π–π stacking with the aromatic ring of Trp84, thereby occupying the binding site for quaternary ligands. The charged nitrogen of the piperidine moiety makes a cation-π interaction with Phe330 and electrostatic interactions with Tyr121. No direct hydrogen bonds were observed between polar groups of the inhibitor and the binding site. A similar binding orientation within the binding pocket was observed for all other inhibitors (Figure 29.5).

FIGURE 29.6 PLS coefficient maps obtained using the water probe (left side) and the methyl probe (right side). Green and cyan fields are contoured at −0.003, yellow and orange fields are contoured at +0.003 (compound **4j** is shown for comparison).

For the final 3D QSAR analysis, the YETI force-field refined docking poses were extracted from the protein environment and were taken as input for a GRID/GOLPE procedure. A predictive 3D QSAR model was obtained after applying the variable-selection strategy incorporated within GOLPE. The significance was tested by applying a variety of validation procedures. The LOO analysis yielded a correlation coefficient with a cross-validated q^2_{LOO} of 0.94 for the water probe and 0.92 for the methyl probe. In addition, the reliability of the model was analyzed by applying leave-20%-out and leave-50%-out cross-validation (100 runs). Both models were shown to be robust, as indicated by high correlation coefficients of $q^2 = 0.91$ (water probe, SDEP = 0.41) and 0.90 (methyl probe, SDEP = 0.44) obtained by using the leave-50%-out cross-validation procedure. The statistical results gave confidence that the derived model could be used for the prediction of novel compounds.

In order to determine which parts of the AChE inhibitors were correlated with variation in activity, the PLS coefficient plots (obtained using the water and methyl probes) were analyzed and compared with amino acid residues of the binding pocket. The plots indicated those lattice points where a particular property significantly contributed and thus explained the variation in biological activity data (Figure 29.6).

The plot obtained with the methyl probe indicated that a region with positive coefficients existed close to the arylpyridazine moiety (region A in Figure 29.6). The coefficients were superimposed with the original GRID field obtained for compound **4j** with the methyl probe. The interaction energies in region A were positive, therefore the decrease in activity was attributed to a steric overlap within this region.

Thus, it should be possible to obtain active inhibitors by reducing the ring size compared to compound **4j** (which is shown in Figure 29.6 together with the PLS coefficient maps). The water probe identified a second interesting field where polar interactions increased activity. It was located above the arylpyridazine moiety in the model (region B in Figure 29.6). After analysis of the entrance of the gorge (the interaction site for the arylpyridazine system), the authors rationalized the design of compounds bearing polar groups. In the calculated AChE–aminopyridazine complexes they observed two polar amino acid residues (Asn280 and Asp285) located at the entrance of the gorge, which could serve as an additional binding site for the substituted arylpyridazine system. To test this hypothesis, several inhibitors possessing polar groups with hydrogen-bond donor and acceptor properties were synthesized and tested. The designed inhibitors were docked into the binding pocket applying the same procedures and their biological activities were predicted using the PLS models. In Table 29.2, the predicted and experimentally determined inhibitor activities are listed for the novel compounds. In general, an excellent agreement between predicted and experimentally determined values was observed, indicated by the low $SDEP_{ext}$ values of 0.44 (water model) and 0.40 (methyl model). The reduction of the size of the aminopyridazine ring system resulted in highly potent inhibitors **4g**–**4i**. The molecules of the second series of designed inhibitors containing polar groups were predicted accurately as well. The gain in activity compared to the non-substituted compound **3y** (Figure 29.3) was moderate, indicating that the potential interaction with the two polar residues at the entrance did not play an important role. Since the two residues are located at the entrance of the binding pocket, it may be possible that these residues make stronger interactions with water molecules than with the protein side chains.

In conclusion, using the receptor-based 3D QSAR strategy Sippl *et al.,*[80,81] were able to design potent novel AChE inhibitors which seemed to interact simultaneously with the cation-π subpocket of the catalytic site and the peripheral site of the enzyme. Further support for their docking study came from the crystal structure of AChE in complex with donepezil.[94] Like their most potent inhibitors, donepezil contains a benzylpiperidine moiety which shows a similar position and orientation in the published crystal structure (Figure 29.7) as the corresponding group in their docking results. The comparison of both AChE–inhibitor complexes revealed that both kinds of inhibitors adopt a comparable conformation in the narrow binding pocket.

B. 3D QSAR as a tool to determine molecular similarity

Within the drug development context, the concept of molecular similarity has proven to be one of the most important tools that can be used to provide new design ideas.[97] Molecular similarity, however, is also a highly complex notion that can only be described with reference to the immediate use for which it is intended and therefore different measures of similarity have to be formulated for each eventual use.[98] In drug design, different notions of molecular similarity are used based on molecular formulae, molecular graphs, molecular skeletons, atom types and positions, conformations, van der Waals surfaces, or molecular fields. Determination of molecular similarity based on MIF will be described in this section.

Basically, molecular similarity can be expressed in terms of shape, electrostatic potential, surface hydrophobicity, hydrogen-bonding capacity, etc. As molecules interact with their binding sites through their molecular fields, it appears justified to define molecular similarity by field comparison, if certain conditions are fulfilled. The major problem in 3D QSAR is the alignment definition, that is, the correct and self-consistent superimposition of all molecular structures under investigation. This remains the main issue if such fields are used for similarity assessment. Therefore, the application of molecular field analysis for similarity determination is limited to those cases where an unambiguous alignment definition is provided. The crucial step is how to analyze the interaction energy matrices. A suitable method has been proposed by Martin *et al.,* within the framework of 3D QSAR.[99] They applied multivariate statistical methods, namely PCA and cluster analysis based on steric potential interaction energy matrices for a CoMFA of shape properties. Since principal properties (PPs) are orthogonal to each other, they are particularly suitable as design variables.[100] Applying criteria of experimental design using PPs as descriptors, one is able to select the most informative combinations of substituents or mole-cules from a series. Moreover, PPs can be used in pairs or triplets to describe substituents linked to each substitution site in a given series of molecules sharing a common skeleton instead of traditional QSAR descriptors that are mimicked in the best possible way. However, it was noted[101,102] that the direct derivation of 3D PPs from interaction energy matrices obtained by CoMFA is not obvious, since in addition to the alignment and conformational flexibility problems, doubts exist concerning the congruency of the descriptor matrix. Clementi *et al.*, address the latter by applying auto- and cross-correlation and covariance (ACC) transforms[103] that have been developed, together with Fourier transforms, to account for the dependencies between consecutive observations. It was determined that PCA on the ACC matrix of a CoMFA field gave results which limited to a certain extent the dependency on the method of orientation of substituents. However, when utilizing this technique, the field descriptor derived PPs of each molecule still depended heavily upon many subjective choices in their derivation, such as selection of the appropriate geometry, alignment of orientation, type of force field, and type of charge calculation. Thus, such scales

TABLE 29.2 Compounds Designed, Synthesized, and Assayed for Their AChE Inhibitory Activity[80,81]

Cpd	Structure	Observed[a]	Predicted[b]
4g		8.00	7.00
4h		7.41	7.62
4i		7.66	7.48
6g		7.24	6.90
6h		7.24	7.05
6i		7.27	7.25
6j		7.14	6.88

[a]Observed pK_i, Torpedo californica.
[b]Predicted pK_i using the GRID water probe.

FIGURE 29.7 Crystal structure of donepezil in AChE (protein data bank (PDB) entry *1eve*, visualization with LigandScout 2.0).[95,96]

should be considered with caution when used in retrieval of information.

Much effort has been spent on the derivation of appropriate descriptors of amino acid properties, and the PPs approach has been used successfully when analyzing peptides via QSAR.[104] Cocchi *et al.*,[105] characterized the 20-coded amino acids by their interaction energies using the program GRID[17] and multivariate data analysis with the aim of extending the amino acid's characterization in the context of the principal properties approach. They used six different probes mimicking various functional groups that could be involved in peptide–peptide interactions (H_2O, —COO—, —CO, $—NH_2$, $—NH^{3+}$, $—CH_3$). PCA of the interaction energies data matrix derived the amino acids PPs and compared the obtained classification with the previously published *z*-scales[106] calculated from a multiproperties matrix containing both experimental data and empirical constants of the amino acids.

Langer[107,108] conducted studies aimed at the multivariate characterization of heteroaromatic moieties using the CoMFA approach together with the Tripos[109] or the GRID force field,[17] respectively. The driving force for these studies was the fact that in medicinal chemistry one of the major problems when dealing with isosteric or bioisosteric replacement[110] in heterocyclic systems is the selection *a priori* of the most promising candidates among several dozens of possible rings. A large number of descriptors were available for such fragments, and PPs for heteroaromatic systems based on both empirical and theoretical data were derived in view of their relevance as building blocks for a large number of compounds of pharmaceutical interest.[111]

FIGURE 29.8 Heteroaromatic residues studied by Langer.[107]

Until that time, descriptors of heteroaromatics, or derived principal properties, respectively, had been measured or calculated only for entire systems without considering differences in the anchoring positions of fragments in a given molecule. It is well known, however, that properties of heteroaromatic moieties may drastically vary upon variation of the substitution position. Therefore, there was a need for appropriate descriptors to describe such effects.

As an initial step,[107] 16 different aromatic ring systems appearing in a total of 37 isomers (Figure 29.8) were examined in order to check the principal utility of molecular similarity characterization using molecular interaction energy fields. All molecules were aligned as shown in Figure 29.9 using a connection bond to a dummy atom located in the origin of a Cartesian coordinate system; the aromatic rings

were placed in the *XY* plane. All statistical calculations were performed within the QSAR module of the SYBYL molecular modeling software.[112] Interaction energies between the heteroaromatic moieties and the probe atoms were calculated at a total of 4,158 grid points with 1 Å spacing in a lattice of $18 \times 21 \times 11$ ($X = -5$ to 12, $Y = -10$ to 10, $Z = -5$ to 5) using the default Lennard–Jones and Coulomb potential functions and the standard Tripos CoMFA probes (the Csp^3 probe was used for calculation of steric interactions and the H^+ probe for calculation of electrostatic interactions, respectively). A PCA (factor analysis without axes rotation) was done on the descriptor matrix and a classification of the heteroaromatic substituents into families was performed using the Sybyl hierarchical clustering procedure of the obtained principal component (PCs). The obtained clustering dendrogram is reproduced in Figure 29.10. In this type of diagram, the most similar compounds clustered together at the lowest levels.

FIGURE 29.9 Alignment of heteroaromatic substituents used in Ref. [107].

It was argued[101,102] that 3D PPs may suffer from major drawbacks when not properly derived. In this particular case, however, the conformational flexibility problem does not exist and the alignment definition assuming a hypothetical binding pocket in which the heteroaromatic moieties would all align in a plane according to the dipole moment vector was straightforward; a possible 180° rotation would just lead to PPs with inverted signs. The potential influence of the substituent parts of the heteroaromatic rings was minimized by the connecting dummy atom. However, a problem was recognized with the parameters of the force-field used: parametrization of sulfur atoms might render heteroaromatic ring systems containing sulfur atoms different from other systems, thus giving rise to different clusters and therefore different possible representative systems. Therefore, the study was extended to include other bicyclic systems[108] and utilize the GRID force-field atom parameters. A total of 72 aromatic moieties (five- and six-membered monocyclic and benzo-fused bicyclic heteroaromatics containing one or two heteroatoms) was analyzed using a total of six GRID multi atom probes (*H_2O*, *alkyl-OH*, *carbonyl-O*, *Csp^3*, *aromatic C*, *NH^{4+}*) that were considered as a representative selection among the variety of the main interaction modes with amino acids in order to mimic possible interactions of the molecule with a putative receptor. The alignment was chosen in a consistent way with the aromatic rings being placed in the *XY* plane in such a way that the dipole moment vectors

FIGURE 29.10 Dendrogram obtained after hierarchical clustering of principal components 1–3 calculated from the CoMFA descriptor matrix. Compounds are numbered according to Figure 29.8. Graphics taken from Ref. [107]

of all compounds were pointing into the same subspace. Interaction energies between the heteroaromatic moieties and the probes were calculated at a total of 3,553 grid points with 1 Å spacing in a lattice of $19 \times 17 \times 11$ Å ($X = -6$ to 12, $Y = -8$ to 8, $Z = -5$ to 5). The first three principal components explaining 78% of the total variance (PC1: 38%, PC2: 31%, PC3: 9%) were extracted and used for further calculations. A classification of the heteroaromatic substituents into families was again performed using a complete linkage hierarchical clustering procedure of the obtained PCs. In fact, the results gained in this case were in better agreement with common chemical knowledge; for example, phenyl was located in the same cluster as 2- and 3-thienyl; the π-electron deficient heteroaromatic moieties 3- and 4-pyridyl were found in the same cluster as 4-pyridazinyl; five-membered π-electron rich heteroaromatics were located in one cluster, like 1-pyrrolyl, 3-pyrrolyl, and 5-thiazolyl.[108] The derived PPs can be used as guidance variables in fractional factorial design. This approach easily can be extended to other heteroaromatics, and to ring systems bearing additional substituents.

IX. CONCLUSIONS AND FUTURE DIRECTIONS

3D QSAR methods have proven useful when optimizing series of compounds in drug development scenarios or agrochemical research applications. Despite all of the limitations and pitfalls, useful information can be obtained if the study is performed carefully and the models are validated. Interestingly, the CoMFA method[6–8] has become the most widely used technique, although a considerable number of other similar approaches are available to the community. Among them, Klebe's *comparative molecular similarity* (CoMSIA) approach[113] has addressed the CoMFA problem of small alignment shifts causing large changes in the molecular field at certain grid points. Instead of grid-based fields, CoMSIA is based on similarity indices that are obtained by using a functional form adapted from the SEAL algorithm.[33] The clear advantage of the CoMSIA method when compared to CoMFA lies in the functions used to describe the molecules studied, as well as the resulting contour maps. The contour maps obtained from CoMSIA are easier to interpret, compared to the ones obtained by the CoMFA approach. The CoMSIA procedure also avoids cut-off values used in CoMFA to restrict potential functions by assuming unacceptably large values. Other 3D QSAR methods that are interesting are Doweyko's *hypothetic active site lattice* (HASL) approach,[114] Silverman's comparative molecular moment analysis (CoMMA),[115] Richard's *self-organizing molecular field analysis* (SoMFA),[116] and Hahn's *receptor surface models* (RSM) approach for 3D QSAR.[117,118] There has been significant effort devoted to the development of alignment independent methods, like Hopfinger's 4D QSAR approach,[70–72] and the methods VolSurf[55,56] and GRIND[57] developed by Cruciani. Vedani has proposed 5D QSAR, which considers hypotheses for changes that might occur in a conformation of a receptor due to ligand binding (induced fit) as a fifth dimension.[119]

One of the trends in drug discovery research over the last decade has been to screen large numbers of compounds in high-throughput assays. Huge compound libraries have been made available using combinatorial chemistry, often, however, containing molecules with unfavorable pharmacokinetic properties. Prediction methods that are able to deal with a large number of molecules will be highly successful if they can be applied to property estimation on a large scale.[120] So far, the 3D QSAR approaches and their related methods have neglected the necessity of automated processing of large quantities of molecules. A great deal of human interaction is necessary to build models and use them for prediction, limiting the number of compounds for which properties could be calculated using this approach. The alignment independent molecular field descriptors based methods might influence the situation and push 3D QSAR forward into the world of high-throughput *in silico* prediction tools.

REFERENCES

1. Hansch, C., Leo, A. *Exploring QSAR. Fundamentals and Applications in Chemistry and Biology*. American Chemical Society: Washington, DC, 1995.
2. Kubinyi, H. QSAR and 3D QSAR in Drug Design. *Drug Discov. Today* 1997, *2*, 457–467.
3. Fujita, T. Recent success stories leading to commercializable bioactive compounds with the aid of traditional QSAR procedures. *Quant. Struct.–Act. Relat.* 1997, *16*, 107–112.
4. Todeschini, R., Consonni, V. *Handbook of Molecular Descriptors*. Wiley-VCH: Weinheim, 2000.
5. Fujita, T. The role of QSAR in drug design. In *Drug Design: Fact or Fantasy?* (Jolles, G., Wolldridge, K. R., Eds). Academic Press: London, 1984, pp. 19–33.
6. Cramer, R. D., III, Patterson, D. E., Bunce, J. D. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J. Am. Chem. Soc.* 1988, *110*, 5959–5967.
7. Cramer, R. D., III, Wold, S. *CoMFA US patent application number. 237491*, August 1988.
8. Cramer, R. D., III, Wold. S. *CoMFA; US patent number 5307287*. Filing date: June 17, 1991, Issue date: April 26, 1994.
9. Wise, M., Cramer, R. D., III, Smith, D., Exman, I. Progress in three-dimensional drug design: the use of real-time colour graphics and computer postulation of bioactive molecules in DYLOMMS. In *Quantitative Approaches to Drug Design* (Deardon, J. C., Ed.). Elsevier: Amsterdam, 1983, pp. 145–146.
10. Cramer, R. D., III, Bunce, J. D. The DYLOMMS method: initial results from a comparative study of approaches to 3D QSAR. In *Pharmacochemistry Library*, Vol. 10, *QSAR Drug Des. Toxicol.*, Elsevier: Amsterdam, 1987, pp. 3–12.
11. Höltje, H.-D., Kier, L. B. Sweet taste receptor studies using model interaction energy calculations. *J. Pharm. Sci.* 1974, *63*, 1722–1725.
12. Höltje, H.-D., Kier, L. B. Structure–activity studies of enzyme substrates using model interaction calculations. *J. Theor. Biol.* 1974, *48*, 197–205.

13. Höltje, H.-D. Quantum chemical experiments on drug receptor complexes. *Pharm. Unserer Zeit.* 1975, *4*, 108–117.
14. Marshall, G. R., Barry, C. D., Bosshard, H. E., Dammkoehler, R. A., Dunn, D. A. The conformational parameter in drug design: the active analog approach. In *Computer-Assisted Drug Design* (Olson, E. C., Christofferesen, R. E., Eds). American Chemical Society: Washington, DC, 1979, pp. 205–226.
15. Wold, S., Ruhe, A., Wold, H., Dunn, W. J. The covariance problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM J. Sci. Stat. Comp.* 1984, *5*, 735–743.
16. Sippl, W. 3D-QSAR using the GRID/GOLPE approach. In *Methods and Principles in Medicinal Chemistry, Molecular Interaction Fields* (Cruciani, G., Ed.), Vol. 27. Wiley-VCH: Weinheim, 2006, pp. 145–170.
17. Goodford, P. J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* 1985, *28*, 849–857. The program GRID is available from Molecular Discovery Ltd., 215 Marsh Road, 1HA5 5NE, Pinner, UK (http://www.moldiscovery.com).
18. Baroni, M., Clementi, S., Cruciani, G., Costantino, G., Riganelli, D., Oberrauch, E. Predictive ability of regression models. Part II: Selection of the best predictive PLS model. *J. Chemometr.* 1992, *6*, 347–356.
19. Kubinyi, H. (Ed.) *3D QSAR in Drug Design. Theory, Methods and Applications*. ESCOM Science Publisher B.V.: Leiden, 1993.
20. Kubinyi, H., Folkers, G., Martin, Y.C., (Eds) *3D QSAR in Drug Design. Ligand–Protein Interactions and Molecular Similarity*. Kluwer/ESCOM: Dordrecht, 1998.
21. Kubinyi, H., Folkers, G., Martin, Y.C., (Eds) *3D QSAR in Drug Design. Recent Advances*. Kluwer/ESCOM: Dordrecht, 1998.
22. Martin, Y. C. 3D QSAR: current state, scope and limitations. *Perspect. Drug Discov. Des.* 1998, *12*, 3–23.
23. Norinder, U. Recent progress in CoMFA methodology and related techniques. *Perspect. Drug Discov. Des.* 1998, *12*, 3–23.
24. Podlogar, B. L., Ferguson, D. M. QSAR and CoMFA: a perspective on the practical application to drug discovery. *Drug Des. Discov.* 2000, *1*, 4–12.
25. Akamatsu, M. Current state and perspectives of 3D QSAR. *Curr. Top. Med. Chem.* 2002, *2*, 1381–1394.
26. Wermuth, C.-G., Ganellin, C. R., Lindberg, P., Mitscher, L. A. Glossary of terms used in medicinal chemistry (IUPAC recommendations 1997). *Annu. Rep. Med. Chem.* 1998, *33*, 385–395.
27. OMEGA is available from OpenEye Scientific Software, 9 Bisbee Court, Suite D Santa Fe, NM 87508, USA (http://www.eyesopen.com).
28. CATALYST is available from Accelrys, Inc. 10188 Telesis Court, Suite 100 San Diego, CA 92121, USA (http://www.accelrys.com).
29. Kirchmair, J., Laggner, C., Wolber, G., Langer, T. Comparative analysis of protein-bound ligand conformations with respect to catalyst's conformational space subsampling algorithms. *J. Chem. Inf. Model.* 2005, *45*, 422–430.
30. Kirchmair, J., Wolber, G., Laggner, C., Langer, T. Comparative performance assessment of the conformational model generators omega and catalyst: a large-scale survey on the retrieval of protein-bound ligand conformations. *J. Chem. Inf. Model.* 2006, *46*, 1848–1861.
31. Wolber, G., Dornhofer, A. A., Langer, T. Efficient overlay of small organic molecules using 3D pharmacophores. *J. Comput.-Aided Mol. Des.* 2006, *20*, 773–788.
32. Clark, M., Cramer, R. D., III, Jones, D., Patterson, D., Simeroth, P. Comparative molecular field analysis (CoMFA): 2. Toward its use with 3D structural databases. *Tetrahedron Comp. Meth.* 1990, *3*, 47.
33. Kearsley, S., Smith, G. An alternative method for the alignment of molecular structures: maximizing electrostatic and steric overlap. *Tetrahedron Comp. Meth.* 1990, *3*, 615.
34. Feuston, B.P., Miller, M.D., Culberson, J.C., Nachbar, R.B., Kearsley, S.K. Comparison of knowledge-based and distance geometry approaches for generation of molecular conformations. *J. Chem. Inf. Comput. Sci.* 2001, 41(3), 754–763.
35. Klebe, G., Mietzner, T., Weber, F. Different approaches toward an automatic structural alignment of drug molecules: applications to sterol mimics, thrombin and thermolysin inhibitors. *J. Comput.-Aided Mol. Des.* 1994, *8*, 751–778.
36. Lemmen, C., Lengauer, T. Computational methods for the structural alignment of molecules. *J. Comput.-Aided Mol. Des.* 2000, *14*, 215–232.
37. Tropsha, A., Cho, S. J. Cross-validated r^2 guided region selection for CoMFA studies. *Perspect. Drug Disc. Des.* 1998, *12/13/14*, 57–69.
38. Langer, T., Hoffmann, R. D. Use of the catalyst program as a new alignment tool for 3D QSAR. In *QSAR and Molecular Modelling: Concepts, Computational Tools and Biological Applications* (Sanz, F., Giraldo, J., Manaut, F., Eds). Prous Science: Barcelona, 1995, pp. 466–469.
39. Palomer, A., Giolitti, A., García, M. L., Cabré, F., Mauléon, D., Carganico, G. Molecular modeling and CoMFA investigations on LTD4 receptor antagonists. In *QSAR and Molecular Modelling: Concepts, Computational Tools and Biological Applications* (Sanz, F., Giraldo, J., Manaut, F., Eds). Prous Science: Barcelona, 1995, pp. 444–450.
40. Langer, T., Hoffmann, R. D. On the use of chemical function-based alignments as input for 3D-QSAR. *J. Chem. Inf. Model.* 1998, *38*, 325–330.
41. Cruciani, G. Molecular Interaction Fields. In *Methods and Principles in Medicinal Chemistry* (Mannhold, R., Kubinyi, H., Folkers, G., Eds), Vol. 27. Wiley-VCH: Weinheim, 2006.
42. Crivori, P., Cruciani, G., Carrupt, P. A., Testa, B. Predicting blood-brain barrier permeation from three-dimensional molecular structure. *J. Med. Chem.* 2000, *43*, 2204–2216.
43. Cruciani, G., Carosati, E., DeBoeck, B., Ethirajulu, K., Mackie, C., Howe, T., Vianello, R. MetaSite: understanding metabolism in human cytochromes from the perspective of the Chemist. *J. Med. Chem.* 2005, *48*, 6970–6979. The program METASITE is available from Molecular Discovery Ltd., 215 Marsh Road, 1HA5 5NE, Pinner, UK (http://www.moldiscovery.com)
44. Floersheim, P., Nozulak, J., Weber, H. P. Exeperience with comparative molecular field analysis. In *Trends in QSAR and Molecular Modeling '92* (Wermuth, C.-G., Ed.). ESCOM: Leiden, 1993, pp. 227–232.
45. Carrupt, P. A., Gaillard, P., Billois, F., Weber, P., Testa, B., Meyer, C., Perez, S. The molecular lipophilicity potential (MLP): a new tool for $\log P$ calculation and docking, and in comparative molecular field analysis (CoMFA). In *Lipophilicity in Drug Action and Toxicology* (Pliska, V., Testa, B., van de Waterbeemd, H., Eds). Wiley-VCH: Weinheim, 1995, pp. 195–215.
46. Kellog, G. E., Abraham, D. J. Key, lock, and locksmith: complementary hydrophobic map predictions of drug structure from a known receptor–receptor structure from known drugs. *J. Mol. Graph.* 1992, *10*, 212–217.
47. Wade, R. C. Calculation and application of molecular interaction fields. In *Molecular Interaction Fields* (Mannhold, R., Kubinyi, H., Folkers, G. Eds), In *Methods and Principles in Medicinal Chemistry* (Cruciani, G., Ed.), Vol. 27, Wiley-VCH: Weinheim, 2006, pp. 27–42.
48. Klebe, G., Abraham, U. On the prediction of binding properties of drug molecules by comparative molecular field analysis. *J. Med. Chem.* 1993, *36*, 70–80.
49. Wold, S., Esbensen, K., Geladi, P. Principal component analysis. *Chemometr. Intellig. Lab. Syst.* 1987, *2*, 37–52.
50. Wold, S., Geladi, P., Esbensen, K., Öhman, J. Multi-way principal components- and PLS-analysis. *J. Chemometr.* 2005, *1*, 41–56.
51. Caliendo, G., Greco, G., Novellino, E., Perissutti, E., Santagada, V. Combined use of factorial design and comparative molecular field analysis (CoMFA): a case study. *Quant.-Struct. Act. Relat. Comb. Des.* 2006, *13*, 249–261.
52. Skagerberg, B., Bonelli, D., Clementi, S., Cruciani, G., Ebert, C. Principal properties for aromatic substituents. A multivariate approach for design in QSAR. *Quant. Struct.–Act. Relat.* 1989, *8*, 32–38.

53. Langer, T. Molecular similarity determination of heteroaromatic ring fragments using GRID and multivariate data analysis. *Quant. Struct.–Act. Relat.* 1996, *15*, 469–474.
54. Pastor, M. Alignment-independent descriptors from molecular interaction fields. In *Molecular Interaction Fields* (Cruciani, G., Ed.), In *Methods and Principles in Medicinal Chemistry* (Mannhold, R., Kubinyi, H., Folkers, G., Eds), Vol. 27, Wiley-VCH: Weinheim, 2006, pp. 117–143.
55. Cruciani, G., Pastor, M., Guba, W. VolSurf: a new tool for the pharmacokinetic optimization of lead compounds. *Eur. J. Pharm. Sci.* 2000, *11*, S29–S39.
56. VOLSURF is available from Molecular Discovery Ltd., 215 Marsh Road, 1HA5 5NE, Pinner, UK (http://www.moldiscovery.com).
57. Pastor, M., Cruciani, G., McLay, I., Pickett, S., Clementi, S. GRid-INdependent descriptors (GRIND): a novel class of alignment-independent three-dimensional molecular descriptors. *J. Med. Chem.* 2000, *43*, 3233–3243.
58. Mannhold, R., Berellini, G., Carosati, E., Benedetti, P. Use of MIF-based VolSurf descriptors in physicochemical and pharmacokinetic studies. In *Molecular Interaction Fields* (Cruciani, G., Ed.), In *Methods and Principles in Medicinal Chemistry* (Mannhold, R., Kubinyi, H., Folkers, G., Eds), Vol. 27, Wiley-VCH: Weinheim, 2006, pp. 173–196.
59. Ottaviani, G., Martel, S., Carrupt, P.-A. *In silico* and *in vitro* filters for the fast estimation of skin permeation and distribution of new chemical entities. *J. Med. Chem.* 2007, *50*, 742–748.
60. Doddareddy, M. R., Cho, Y. S., Koh, H. Y., Kim, D. H., Pae, A. E. *In silico* renal clearance model using classical Volsurf approach. *J. Chem. Inf. Model.* 2006, *46*, 1312–1320.
61. Braiuca, P., Buthe, A., Ebert, C., Linda, P., Gardossi, L. Volsurf computational method applied to the prediction of stability of thermostable enzymes. *Biotechnol. J.* 2007, *2*, 214–220.
62. ALMOND is available from Molecular Discovery Ltd., 215 Marsh Road, 1HA5 5NE, Pinner, UK (http://www.moldiscovery.com).
63. Afzelius, L., Zamora, I., Masimirembwa, C. M., Karlen, A., Andersson, T. B., Mecucci, S., Baroni, M., Cruciani, G. Conformer- and alignment-independent model for predicting structurally diverse competitive CYP2C9 inhibitors. *J. Med. Chem.* 2004, *47*, 907–914.
64. Benedetti, P., Mannhold, R., Cruciani, G., Pastor, M. GBR compounds and mepyramines as cocaine abuse therapeutics: chemometric studies on selectivity using grid independent descriptors (GRIND). *J. Med. Chem.* 2005, *45*, 1577–1584.
65. Cianchetta, G., Singleton, R. W., Zhang, M., Wildgoose, M., Giesing, D., Fravolini, A., Cruciani, G., Vaz, R. J. A pharmacophore hypothesis for P-glycoprotein substrate recognition using GRIND-based 3D-QSAR. *J. Med. Chem.* 2005, *48*, 2927–2935.
66. Caron, G., Ermondi, G. Influence of conformation on GRIND-based three-dimensional quantitative structure–activity relationship (3D-QSAR). *J. Med. Chem.* 2007. ASAP Article 10.1021/jm0704651 S0022–2623(07)00465-7
67. Cruciani, G., Pastor, M., Mannhold, R. Suitability of molecular descriptors for database mining. A comparative analysis. *J. Med. Chem.* 2002, *45*, 2685–2694.
68. Fontaine, F., Pastor, M., Gutierrez de Teran, H., Lozano, J. J., Sanz, F. Use of alignment-free molecular descriptors in diversity analysis and optimal sampling of molecular libraries. *Mol. Divers.* 2003, *6*, 135–147.
69. Fontaine, F., Pastor, M., Zamora, I., Sanz, F. Anchor-GRIND: filling the gap between standard 3D QSAR and the GRid-INdependent descriptors. *J. Med. Chem.* 2005, *48*, 2687–2694.
70. Hopfinger, A. J., Wang, S., Tokarski, J. S., Jin, B., Alburquerque, M. G., Madhav, P. J., Duraiswami, C. Construction of 3D-QSAR models using the 4D-QSAR analysis formalism. *J. Am. Chem. Soc.* 1997, *119*, 10509–10524.
71. Albuquerque, M. G., Hopfinger, A. J., Barreiro, E. J., de Alencastro, R. B. Four-dimensional quantitative structure–activity relationship analysis of a series of interphenylene 7-oxabicycloheptane oxazole thromboxane A2 receptor antagonists. *J. Chem. Inf. Comput. Sci.* 1998, *38*, 925–938.
72. Klein, C. D., Hopfinger, A. J. Pharmacological activity and membrane interactions of antiarrhythmics: 4D-QSAR/QSPR analysis. *Pharm. Res.* 1998, *15*, 303–311.
73. Albuquerque, M. G., Araujo de Brito, M., Ferreira da Cunha, E. F., Bicca de Alencastro, R., Antunes, O. A. C., Castro, H. C., Rodrigues, C. R. Multidimensional-QSAR: beyond the third-dimension in drug design. *Curr. Meth. Med. Chem. Biol. Phys.* 2007, *1*, 91–100.
74. Oprea, T. I., Garcia, A. E. Three-dimensional quantitative structure–activity relationships of steroid aromatase inhibitors. *J. Comput.-Aided Mol. Des.* 1996, *10*, 186–200.
75. Golbraikh, A., Tropsha, A. Beware of q^2!. *J. Mol. Graph. Model.* 2002, *20*, 269–276.
76. Kubinyi, H., Hamprecht, F. A., Mietzner, T. Three-dimensional quantitative similarity-activity relationships (3D QSiAR) from SEAL similarity matrices. *J. Med. Chem.* 1998, *41*, 2553–2564.
77. Golbraikh, A., Shen, M., Xiao, Z., Xiao, Y., Lee, K.-H., Tropsha, A. Rational selection of training and test sets for the development of validated QSAR models. *J. Comput.-Aided Mol. Des.* 2003, *17*, 241–253.
78. Norinder, U. Single and domain made variable selection in 3D QSAR applications. *J. Chemometr.* 1996, *10*, 95–105.
79. Doweyko, A. M. 3D QSAR illusions. *J. Comput.-Aided Mol. Des.* 2004, *18*, 587–596.
80. Sippl, W., Contreras, J. M., Parrot, I., Rival, Y., Wermuth, C. G. Structure-based 3D QSAR and design of novel acetylcholinesterase inhibitors. *J. Comput.- Aided Mol. Des.* 2001, *15*, 395–410.
81. Sippl, W. Development of biologically active compounds by combining 3D QSAR and structure-based design methods. *J. Comput.-Aided Mol. Des.* 2002, *16*, 825–830.
82. Massoulie, J., Pezzementi, L., Bon, S., Krejci, E., Velette, F. M. Molecular and cellular biology of the cholinestereases. *Prog. Neurobiol.* 1993, *41*, 31–39.
83. Crismon, M. L. Tacrine: first drug approved for Alzheimer's disease. *Ann. Pharmacother.* 1994, *28*, 744–751.
84. Barner, E. L., Gray, S. L. Donepezil in Alzheimer's Disease. *Ann. Pharmacother.* 1998, *32*, 70–77.
85. Barril, X., Orozco, M., Luque, F. J. Towards improved acetylcholinesterase inhibitors: a structural and computational approach. *Mini-Rev. Med. Chem.* 2001, *1*, 255–266.
86. Bourne, Y., Taylor, P., Radic, Z., Marchot, P. Structural insights into ligand interactions at the acetylcholinesterase peripheral anionic site. *EMBO J.* 2003, *22*, 1–12.
87. Sussman, J. L., Harel, M., Frolow, F., Oefner, C., Goldman, A., Toker, L., Silman, I. Atomic structure of acetylcholinesterase from torpedo californica: a prototypic acetylcholine-binding protein. *Science* 1991, *253*, 872–879.
88. Harel, M., Quinn, D. M., Nair, H. K., Silman, I., Sussman, J. L. The X-ray structure of a transition state analog complex reveals the molecular origins of the catalytic power and substrate specificity of acetylcholinesterase. *J. Am. Chem. Soc.* 1996, *118*, 2340–2346.
89. Wermuth, C.-G., Schlewer, G., Bourguignon, J.-J., Maghioros, G., Bouchet, M. J., Moire, C., Kan, J. P., Worms, P., Biziere, K. 3-Aminopyridazine derivatives with atypical antidepressant serotonergic, and dopaminergic activities. *J. Med. Chem.* 1989, *32*, 528–537.
90. Contreras, J.-M., Rival, Y., Chayer, S., Bourguignon, J.-J., Wermuth, C.-G. Aminopyridazines as acetylcholinesterase inhibitors. *J. Med. Chem.* 1999, *42*, 730–741.
91. AUTODOCK is available from the Molecular Graphics Laboratory, Department of Molecular Biology, The Scripps Research Institute, MB-5 10550 N. Torrey Pines Rd., La Jolla, CA 92037-1000, USA (http://autodock.scripps.edu/).
92. Vedani, A., Huhta, D. W. A new force field for modeling metalloproteins. *J. Am. Chem. Soc.* 1990, *112*, 269–280.

93. Vedani, A., Dunitz, J. D. Lone-pair directionality of H-bond potential functions for molecular mechanics calculations: the inhibition of human carbonic anhydrase II by sulfonamides. *J. Am. Chem. Soc.* 1985, *107*, 7653–7658.
94. Kryger, G., Silman, I., Sussman, J. L. Structure of acetylcholinesterase complexed with E2020 (Aricept): implications for the design of new anti-Alzheimer drugs. *Struct. Fold. Des.* 1999, *15*, 297–307.
95. Wolber, G., Langer, T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *J. Chem. Inf. Model.* 2005, *45*, 160–169.
96. LIGANDSCOUT is available from Inte:Ligand GmbH, Mariahilferstrasse 74b/11, 1070 Vienna, Austria (http://www.inteligand.com).
97. Dean, P. M. Defining molecular similarity and complementary for drug design. In *Molecular Similarity in Drug Design* (Dean, P. M., Ed.). Blackie Academic & Professional: London, UK, 1995, pp. 1–23.
98. Dean, P. M. Molecular similarity. In *3D QSAR in Drug Design: Theory, Methods, and Applications* (Kubinyi, H., Ed.). ESCOM: Leiden, 1993, pp. 150–172.
99. Lin, T. C., Pavlik, P. A., Martin, Y. C. Use of molecular fields to compare series of potentially bioactive molecules designed by scientists or by computer. *Tetrahedron Comput. Methodol.* 1990, *3*, 723–738.
100. Wold, S., Sjöström, M., Carlson, R., Lundstedt, T., Hellberg, S., Skagerberg, B., Wirkström, C., Öhman, J. Multivariate design. *Anal. Chim. Acta* 1986, *191*, 17–32.
101. Clementi, S., Cruciani, G., Baroni, M., Costantino, G. Series design. In *3D QSAR in drug design: theory, methods, and applications* (Kubinyi, H., Ed.). ESCOM: Leiden, 1993, pp. 567–582.
102. Van de Waterbeemd, H., Clementi, S., Costantino, G., Carrupt, P.-A., Testa, B. CoMFA derived substituent descriptors for structure-property correlations. In *3D QSAR in drug design: theory, methods, and applications* (Kubinyi, H., Ed.). ESCOM: Leiden, 1993, pp. 697–707.
103. Clementi, S., Cruciani, G., Riganelli, D., Valigi, R., Costantino, G., Baroni, M., Wold, S. Autocorrelation as a tool for a congruent description of molecules in 3D QSAR studies. *Pharm. Pharmacol. Lett.* 1993, *3*, 433–438.
104. Hellberg, S., Sjöström, M., Skagerberg, B., Wold, S. Peptide quantitative structure–activity relationships, a multivariate approach. *J. Med. Chem.* 1987, *30*, 1127–1135.
105. Cocchi, M., Johansson, E. Amino acids characterization by GRID and multivariate data analysis. *Quant. Struct.–Act. Relat.* 1993, *12*, 1–8.
106. Hellberg, S., Sjöström, M., Skagerberg, B., Wold, S. On the use of multipositionally varied test series for quantitative structure–activity relationships. *Acta Pharm. Jugosl.* 1987, *37*, 53–65.
107. Langer, T. Molecular similarity determination of heteroaromatics using CoMFA and multivariate data analysis. *Quant. Struct.–Act. Relat.* 1994, *13*, 402–405.
108. Langer, T. Molecular similarity determination of heteroaromatic ring fragments using GRID and multivariate data analysis. *Quant. Struct.–Act. Relat.* 1996, *15*, 469–474.
109. Clark, M., Cramer, R. D., III, Van Opdenbosch, N. Validation of the general purpose TRIPOS 5.2 force field. *J. Comput. Chem.* 1989, *10*, 982–1012.
110. Wermuth, C. G. Molecular variations based on isosteric replacements. In *The practice of medicinal chemistry* (Wermuth, C.-G., Ed.). Academic Press: London, 1996, pp. 203–237.
111. Caruso, L., Katritzky, A. R., Musumarra, G. Classical and magnetic aromaticities as new descriptors for heteroaromatics in QSAR. Part 3. Principal properties for heteroaromatics. *Quant. Struct.–Act. Relat.* 1993, *12*, 146–151.
112. SYBYL is available from Tripos, 1699 South Hanley Road, St. Louis, MO 63144-2319, USA (http://www.tripos.com).
113. Klebe, G., Abraham, U., Mietzner, T. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. *J. Med. Chem.* 1994, *37*, 4130–4146.
114. Doweyko, A. M. The hypothetical active site lattice. An approach to modeling active sites from data on inhibitor molecules. *J. Med. Chem.* 1988, *31*, 1396–1406.
115. Silverman, B. D., Platt, D. E. Comparative molecular moment analysis (CoMMA): 3D-QSAR without molecular superposition. *J. Med. Chem.* 1996, *39*, 2129–2140.
116. Robinson, D. D., Winn, P. J., Lyne, P. D., Richards, W. G. Self-organizing molecular field analysis: a tool for structure–activity studies. *J. Med. Chem.* 1999, *42*, 573–583.
117. Hahn, M. Receptor surface models. 1. Definition and construction. *J. Med. Chem.* 1995, *38*, 2080–2090.
118. Hahn, D., Rogers, D. Receptor surface models. 2. Application to quantitative structure–activity relationships studies. *J. Med. Chem.* 1995, *38*, 2091–2102.
119. Vedani, A., Dobler, M. 5D-QSAR: the key for simulating induced fit? *J. Med. Chem.* 2002, *45*, 2139–2149.
120. Oprea, T. I., Tropsha, A. Cheminformatics and drug discovery. *Drug Disc. Today Technol.* 2006, *3*, 355–356.

Chapter 30

Protein Crystallography and Drug Discovery

Jean-Michel Rondeau and Herman Schreuder

I. PRESENTATION
II. HISTORICAL BACKGROUND
A. The early days of crystallography
B. The current state of the art
C. Past and present contributions to drug discovery
III. EXAMPLES
A. Aliskiren (Tektura™, Rasilez™)
B. Nilotinib
IV. BASIC PRINCIPLES AND METHODS OF PROTEIN CRYSTALLOGRAPHY
A. Crystallization
B. Data collection
C. From diffraction intensities to a molecular structure
D. Information content and limitations of crystal structures
V. PRACTICAL APPLICATIONS
A. Target identification, selection and validation
B. Hit/lead generation
C. Lead optimization
REFERENCES

"If you can look into the seeds of time
And say which grain will grow and which will not Speak then to me … "

Shakespeare, Macbeth[1]

I. PRESENTATION

In spite of the rapid emergence and growing economic importance of biopharmaceuticals (antibodies and other protein drugs such as hormones, growth factors, cytokines, enzymes, etc.), the search for novel, more effective, and safer low molecular weight drugs is still the mainstay of current pharmaceutical research. The target-based approach of modern drug discovery relies heavily on a thorough understanding, at the molecular level, of the target under investigation. Therefore, it is no surprise that structural biology has had a profound influence on pharmaceutical research since its inception during the 1970s.[2] Today, the use of structural information pervades all phases of pre-clinical research: from the initial stage of target identification and selection to the design and validation of suitable *in vitro* assays, the selection of appropriate hit/lead finding strategies, as well as throughout the whole lead optimization phase[3,4] (Figure 30.1a). The impact of structural biology on the daily work of medicinal chemists has been particularly strong.[5,6] Making use of highly complex, 3D structural information in drug design is, however, not a trivial task, not the least because this information must be combined with other constraints, such as synthetic accessibility, absorption, distribution, metabolism, excretion (ADME) properties, toxicology, and intellectual property. Nonetheless, the potential rewards are immense. Structural information not only clarifies structure–activity relationships, reveals binding modes and bioactive conformations, unveils new binding pockets or allosteric binding sites, it also opens new and diverse drug discovery avenues, such as *in silico* screening, the rational design of focused chemical libraries, the *de novo* design of new ligand scaffolds, and so on. Having access to such information provides a strong competitive advantage and makes the professional life of medicinal chemists a highly stimulating, and often very gratifying one too.

X-ray crystallography has played a major role in the structural biology revolution, as the most powerful technique

Copyright © 2008, Elsevier Ltd
All rights reserved.

to decipher the 3D architecture of biological macromolecules at atomic or near atomic resolution. Today, this revolution is still making great strides, in particular with the advent of large scale structural genomics efforts focusing on human proteins of medical importance or on potential drug targets from clinically relevant pathogens. Moreover, advances in protein crystallography technologies and methods have led to a drastic reduction in the time needed to generate crystal structures, in particular when dealing with series of protein–ligand complexes. These developments have not only allowed the gathering of an amazing wealth of structural data, they have also paved the way towards new and important applications of "high-throughput" protein crystallography[7] to drug discovery. In addition to its classical use in structure-based or structure-guided drug design, protein crystallography is now widely applied to hit triaging activities as well as a hit finding technology, notably in the context of fragment-based screening (FBS) (Figure 30.1b). Over the past few years, this latter approach has demonstrated its power for the identification of novel, chemically attractive leads and their successful optimization to highly potent drug candidates.[8]

Biological crystallography can be applied to any macromolecular targets or assemblies, irrespective of their size and complexity, provided that crystals of sufficient quality can be produced. This important *proviso*, in turn, implies that the target under study can be produced at the milligram scale in a stable and homogeneous form. Unfortunately, these pre-requisites are usually not met for many intrinsic membrane proteins and receptors such as G-protein coupled receptors (GPCRs) and ion channels, which constitute a very important and highly successful class of drug targets. However, painstaking work by several laboratories in this field is slowly paying off, with already some notable successes[9] and hopefully many more to come in the near future.[10]

In this chapter, we will describe how crystallographic data can contribute today to the different phases of pharmaceutical research. We will emphasize the strengths but also the technical limitations of protein crystallography, so that any medicinal chemist engaging in a new research program and having access to a structural biology group, can gauge if, and how, his project could potentially benefit from this technology. Also, a brief outline of the basic principles and methods of protein crystallography is provided. Medicinal chemists, in particular those working in the industry, have access to large, public as well as proprietary, depositories of refined crystal structures. To make proper use of these data, it is essential for them to be well aware of the limitations and potential uncertainties associated with X-ray

(a)

(b)

FIGURE 30.1 Panel (a) shows how the different stages of drug discovery research can benefit from the already available structural knowledge. Panel (b) shows how experimental structure determination by X-ray analysis can be integrated into the hit/lead finding, triaging, validation and optimization phases. Structure-based drug design (SBDD).

structures. We also hope that this chapter will contribute to a more effective communication between chemists and their fellow crystallographers.

II. HISTORICAL BACKGROUND

A. The early days of crystallography

Crystallography made its first and notable contributions to the progress of biology and medicine well before the elucidation, by Kendrew and Perutz in 1958–1960, of the first protein structures, myoglobin[11] and hemoglobin.[12] The preparation of "blood crystals" – in fact hemoglobin crystals – was first reported by Hünefeld in 1840.[13] During the second half of the 19th century, this initial observation sparked considerable interest in the crystallization of hemoglobins and other proteins, mainly from plant seeds.[13] This groundwork set the stage for the first major achievement, which took place during the years 1926–1935: the demonstration of the molecular nature of enzymes and viruses through their isolation in crystalline form, by Sumner, Northrop, and Stanley.[13] The second major contribution was made in the early 1950s, when X-ray diffraction photographs of DNA produced by Franklin[14] and Wilkins[15] could be used as a guide by Crick and Watson, ultimately leading to their discovery of the double helical structure of DNA.[16] The next achievement, the determination of the myoglobin and hemoglobin structures by Kendrew and Perutz, revealed, for the very first time, the intricacies of the architecture of proteins, while also shedding light on the molecular basis of sickle cell anemia.[17] Since then, these milestone studies have been followed by a rich crop of other stunning crystallographic feats. To name only a few, the 3D structures of the human common cold virus,[18] the photosynthetic reaction center,[19] the F1-ATP synthase,[20] the proteasome,[21] the nucleosome core particle,[22] the 30S and 50S ribosomal particles,[23–26] the RNA polymerase II,[27] and the potassium channels[28,29] have all been solved using X-ray diffraction methods, in spite of their daunting size and biochemical complexity.

B. The current state of the art

Today, about 50,000 crystal structures are publicly available from the Protein Data Bank (PDB) (see Box 30.3), and hundreds to thousands of proprietary structures are available to chemists working in pharmaceutical companies. Moreover, this universe of crystallographic data is expanding at an ever increasing pace. The determination of the hemoglobin structure by Perutz was the result of a lifetime effort,[17] and, when industrial laboratories first invested in protein X-ray facilities in the early 1980s, the technology was still cumbersome and time consuming. The past three decades have witnessed a series of major technological breakthroughs:[30] powerful X-ray sources (1st, 2nd, and now 3rd generation synchrotron radiation sources, as well as super-bright rotating anode generators), 2D detectors with fast read-out, the advent of cryocrystallography (data collection at 80–120 K), of robotic systems for automated crystal mounting and data collection, new phasing approaches, much improved and largely automated software for data processing, structure determination and refinement, and ever more powerful and affordable computers. Today, the measurement of a complete diffraction data set takes a few minutes only at a synchrotron beamline and not more than a few hours with a good home source. More importantly, in the case of a protein of known 3D structure, the processing of diffraction data sets corresponding to different ligands and the calculation of the corresponding difference electron-density maps can be fully automated and performed within a few hours of CPU time.[31] Software for automated ligand fitting, also in the more complex case of fragment cocktails, have been developed.[32,33] In the context of a fragment screen by X-ray crystallography, full refinement is performed only for interesting hits and requires minimal hands-on time, since only minor adjustments need to be made to the already well-refined starting model. In parallel with the gain in speed, a very substantial improvement in the average resolution and quality of X-ray data sets, as well as in the overall quality of the final refined models, has been achieved. In the recent years, several structures could even be solved at subatomic resolution (0.85 Å resolution or better) including several enzyme inhibitor complexes.[34,35] These studies, which are expected to become more frequent in the forthcoming years, have already provided new and very deep insights into subtle structural features implicated in enzymatic mechanisms, co-factor recognition and inhibitor binding.[35–37]

When it can be applied in the high-throughput mode, the potential value of protein crystallography for other industrial applications such as hit triaging and validation is obvious. Nevertheless, it should be made clear that access to the high-throughput mode is restricted to already established systems, that is those for which good crystals can routinely be prepared. For new projects, the determination of the first X-ray structure and the identification of a crystal form suitable for high-throughput X-ray analyses usually require 3–6 months, but the timelines can shift substantially for particularly difficult targets. In the vast majority of cases, the main bottleneck does not reside in the X-ray analysis itself, but rather in the identification and production of a stable, well-behaved recombinant version of the protein of interest, which is amenable to crystallization. Advances in multi-parallel protein cloning and expression technologies as well as in crystallization robotics have radically transformed the prospects of successfully tackling hitherto intractable targets. However, while the previous generations of X-ray pioneers could select readily available, easily crystallizable proteins as tools for establishing

the techniques, today's crystallographers, and particularly those working in industry, are often faced with difficult to produce, poorly behaved, poorly characterized targets. Such challenging proteins require a lot of biochemical ingenuity and cannot be conquered without a dedicated, appropriately resourced effort in protein production and characterization.

C. Past and present contributions to drug discovery

Drugs with anti-sickling properties were the first drugs ever to be studied in complex with their protein target by means of X-ray analysis.[17] These seminal studies spurred the first attempts at designing improved compounds *via* a structure-based approach.[38] Soon after, the discovery of captopril[39] (Figure 30.2a), an anti-hypertensive agent and the first marketed, orally active inhibitor of the human angiotensin-converting enzyme (ACE), hailed the beginning of a new era for pharmaceutical research: for the very first time, a drug had been rationally designed on the basis of structural information, hence providing the first compelling demonstration of the power of the structure-based approach.

Interestingly, the successful design of captopril used simple chemical concepts guided by a hypothetical, "paper-and-pencil" model of substrate and inhibitor binding to the enzyme active site, that had been inferred from the crystal structure of bovine carboxypeptidase A. The X-ray structure of human ACE became available only in 2003, 25 years after the discovery of the captopril class of drugs. While the crystallographic analysis of the ACE complex with captopril[40] confirmed the designed mode of interaction, it revealed little structural similarity, overall, with carboxypeptidase A. This is not the end of the story, however. Human ACE is now known to be composed of two homologous and functionally active domains, but the inhibition of only one domain appears to be required for achieving the anti-hypertensive effect. Therefore, it is hoped that the detailed structural insights now available on human ACE will soon lead to domain-selective inhibitors with improved efficacy and pharmacological profiles.[41]

Historically, dorzolamide (Figure 30.2b) is the first example of an approved drug which benefited from the complete armament of structure-based design, that is multiple X-ray analyses with the human enzyme target[42] combined with sophisticated molecular modeling studies, including in-depth conformational analyses using *ab initio* quantum chemistry calculations.[43] Dorzolamide is a subnanomolar carbonic anhydrase II inhibitor which was developed in the early 1990s as a topical agent for the treatment of glaucoma. Today, 14 different isoforms of mammalian carbonic anhydrase have been identified, several of which appear to play an important pathophysiological role.[44] Crystallographic information is now available for three carbonic anhydrase isoforms and a large variety of inhibitors. Furthermore, X-ray analyses have shed light on the structural basis for carbonic anhydrase activation by several compound classes.[45] These substantial advances are fueling today's vivid carbonic anhydrase research from which new therapeutic agents are likely to emerge.

The discovery of oseltamivir (Tamiflu™, Figure 30.2d), an inhibitor of *influenza* neuraminidase, is another early

FIGURE 30.2 The first marketed drugs derived from structure-based design. (a) captopril (Capoten™), (b) dorzolamide (Trusopt™), (c) zanamivir (Relenza™), and (d) oseltamivir (Tamiflu™).

example of successful structure-based drug design which still resonates today owing to the current concerns about a potential H5N1 pandemic. Interestingly, long-standing efforts to identify neuraminidase inhibitors via random screening or the rational design of transition-state analogs had failed to produce any potent compound,[46] until the crystal structure of neuraminidase became available in 1983.[47,48] A GRID[49] analysis of the sialic acid complex immediately suggested a simple modification of a known sialic acid analog with low micromolar affinity. Remarkably, two compounds only were synthesized, and both turned out to be extremely potent inhibitors, with K_i values of 50 nM and 0.2 nM, respectively.[50] The most potent compound, zanamivir (Relenza™, Figure 30.2c), became the first marketed neuraminidase inhibitor. Further structure-based design concentrated on the development of an analog with improved stability and lipophilicity. These efforts very quickly resulted in oseltamivir, a second generation, orally bioavailable drug[51] (Figure 30.2d).

These remarkable successes gave momentum to structure-based drug design. Over the past 20 years, this approach has been widely applied to numerous targets[6,52–58] and it has delivered many clinical candidate drugs.[59] With nine approved drugs to date, HIV-1 protease inhibitors[60] form the largest cohort of drugs originating from structure-based design, but kinase inhibitors are catching up at a fast pace. In the following section, two recent examples of drugs derived from structure-based approaches will now be described in more detail.

III. EXAMPLES

A. Aliskiren (Tekturna™, Rasilez™)

Because of its exquisite substrate selectivity and its critical role as the first and rate-limiting enzyme in the renin-angiotensin cascade,[61] renin has been the focus of extensive and particularly tenacious drug discovery efforts in many academic and industrial laboratories. At the outset, the quest for renin inhibitors was inspired by crystallographic data gleaned during previous work on fungal aspartic proteinases.[62–65] In 1989, the structure of human renin became available,[65] followed soon after by the first structure of an inhibitor complex,[66] finally allowing direct structure-based design of renin inhibitors. A wide variety of potent peptide-like inhibitors emerged from these studies, but their further development was hampered by manifold drawbacks such as insufficient oral bioavailability, metabolic degradation, insufficient clinical efficacy or excessive manufacturing costs.[67] To overcome these issues, a new design strategy was gradually developed which aimed at exploiting the contiguous and large hydrophobic cavity formed by the S1 and S3 enzyme subsites (Figure 30.3).[68] The main objective of this bold, unprecedented design concept was to replace the P2-P4 peptide backbone of existing inhibitors by an extended hydrophobic group, tethered to a classical hydroxyethylene-based, dipeptide transition state mimetic, and filling the S1-S3 cavity of the enzyme. In addition, the design strategy sought to maintain a key hydrogen-bonded interaction (to Ser219 in the S3 subsite), previously observed in many complexes with peptidomimetic inhibitors. This strategy led to the discovery of novel, highly potent and selective, non-peptidic renin inhibitors. Subsequent X-ray analyses of several representative compounds in complex with human renin confirmed the validity of the design strategy, and, most intriguingly, revealed the existence of a distinct sub-pocket, directly accessible from the S3 subsite but perpendicular to the substrate binding cleft (Figure 30.3).[68,69] Medicinal chemistry efforts were then concentrated on optimizing *in vitro* potency towards plasma renin, *in vivo* efficacy and duration of action, as well as on further

FIGURE 30.3 Complex of human renin with aliskiren (Tekturna™, Rasilez™). Left: overall view. Right: close-up view of the binding site showing the S3 subpocket where the methoxypropoxy side-chain of aliskiren fits. (the figure was generated with the program Pymol (Delano Scientific) using PDB entry 2v0z).

fine-tuning of the pharmacokinetic properties by modifications of the P1' and P2' groups. This led to the discovery of aliskiren (Figure 30.4a), the first in a novel class of potent, orally active non-peptide renin inhibitor, which was recently approved for the treatment of hypertension.[70,71]

B. Nilotinib (Tasigna™)

The development of protein kinase inhibitors targeting the ATP binding site was initially received with great skepticism, on the ground that it will not be possible to achieve a sufficient level of selectivity to turn them into useful therapeutic agents. In view of the large size of the human kinome[72] (518 genes) and the high conservation of the ATP binding site, this criticism was well founded. However, the discovery of imatinib[73,74] (Glivec®, Gleevec™; Figure 30.4b), an inhibitor of the tyrosine kinase activity of the Bcr-Abl oncogene and an effective frontline therapy for chronic myelogenous leukemia, provided compelling evidence for the viability of this approach. Not surprisingly, soon after this breakthrough protein kinases became the subject of intensive study. Arguably, protein crystallography has made a major contribution to our current understanding of the structure, mechanism, and regulation of protein kinases, as well as to the development of kinase inhibitors.[75–77] Numerous structures of kinase complexes with ATP, ATP analogs or small-molecule inhibitors provided early on a basic understanding of the structural requirements for potency and selectivity, that has guided medicinal chemistry efforts ever since, also in the many cases where the structures of the actual targets were not yet available. Among this first generation of protein kinase inhibitors, imatinib was first to reach the market, but other protein kinase inhibitors have since been approved.

FIGURE 30.4 (a) aliskiren (Tekturna™, Rasilez™), (b) imatinib (Gleevec™), and (c) nilotinib (Tasigna™).

The X-ray structure of the Abl kinase domain in complex with des-methylpiperazinyl imatinib became available in 2000,[78] soon followed by the imatinib complex.[79,80]

The *N*-methylpiperazine moiety of imatinib had been introduced during the lead optimization phase to improve solubility, at a point in time where the exact binding mode of the drug was not known. Unexpectedly, the X-ray analyses revealed that the drug was binding to an inactive conformation of the kinase, with the benzamide and piperazinyl groups accessing a channel at the back of the ATP site (Figure 30.5). A conformational switch of the DFG motif of the kinase was responsible for the formation of this channel, which is therefore also referred to as the "DFG-out" pocket. In this mode of binding, the *N*-methylpiperazine moiety was only partially exposed to solvent and made strong interactions to the kinase.[81] More importantly, several structural features of the inactive state of the Abl kinase were found to play a key role in imatinib binding, and detailed structural comparisons indicated that these features were poorly conserved in other protein kinases, thus shedding light into the selectivity determinants.[81] In addition, these and follow-up structures provided a platform for the analysis of resistance mutants.[81,82]

These studies paved the way for the rational design of a second generation drug effective against a broad range of imatinib-resistant mutants. The design strategy sought to maintain binding to the inactive conformation of the kinase, to ensure that the excellent selectivity profile of the drug will be retained. To this end, four key hydrogen-bonded interactions were kept, but the central amide pharmacophore was replaced by a reverse amide (Figure 30.5). A suitable replacement for the *N*-methylpiperazine moiety was then sought using a focused chemical library approach designed to explore the "DFG-out" pocket.[83] X-ray structures with some representative compounds were determined to further guide the optimization process.[81] This work ultimately led to the identification of nilotinib (Tasigna™), a highly potent, very selective drug which maintains activity against 32 out of 33 imatinib-resistant Bcr-Abl variants.[84–86]

IV. BASIC PRINCIPLES AND METHODS OF PROTEIN CRYSTALLOGRAPHY

A. Crystallization

1. *What are protein crystals?*

Protein crystals (Figure 30.6), like any crystal of organic or inorganic compounds, are regular 3D arrays of identical molecules or molecular complexes (Figure 30.7). Depending on the symmetry of this arrangement (described by the space group), all molecules in a crystal have a limited number of unique orientations with respect to the crystal lattice. The diffraction of all individual molecules adds up to yield intensities which are sufficiently strong to be measured, the crystal lattice thus acting as an amplifier.

One notable difference between crystals of small molecules and macromolecular crystals is the very large solvent content of the latter. Protein crystals typically contain at least 30% and up to 80% (v/v) solvent (in fact, aqueous crystallization buffer).[87] Only a fraction of the protein surface is involved in crystal contacts, the rest being fully solvated, pretty much like in solution. As a consequence, protein crystals are very soft and fragile. But, on the positive side, when they are soaked with low molecular weight ligands, co-factors or substrates, these molecules can diffuse from the surrounding mother liquor into the solvent channels within the crystal. If their binding site is not occluded by crystal contacts, the complex can then be formed *in situ*. Usually, small conformational changes can take place within the crystal lattice without damaging the crystal, and sometimes very large structural changes can be accommodated as well. Therefore, enzyme crystals are very often active as catalysts.

2. *How do we get crystals?*

a. Crystallization methods and technologies

Crystals are produced by slowly driving a concentrated protein solution into a state of supersaturation.[88,89] Under the right conditions, the protein will not form an amorphous

FIGURE 30.5 Complex of human Abl kinase with imatinib (Gleevec™). Left: overall view. Right: close-up view of the binding site, with the amide pharmacophore and the DFG motif in the center, the adenine binding site is on the left, and the "DFG-out" pocket is on the right, where the *N*-methylpiperazinylbenzyl moiety binds (the figure was generated with the program Pymol (Delano Scientific) using PDB entry 2hyy).

FIGURE 30.6 Examples of protein crystals. From left to right: β-secretase inhibitor complex; human farnesyl pyrophosphatase in complex with the nitrogen-containing bisphosphonate drug zoledronic acid; Crystals of the Abl kinase domain in complex with imatinib – *Source*: Courtesy of S.W. Cowan-Jacob, Novartis Institutes for BioMedical Research; crystal of a Cdk2 inhibitor complex.

precipitate but will instead settle in a well-ordered crystalline array. Methods for reducing the protein solubility, and thus achieving a high degree of supersaturation, involve dialyzing away the salt, if the salt has a strong solubilizing (or "salting-in") effect, concentrating a nearly saturated protein solution by evaporation (usually in a hanging or sitting drop setup, see below) and adding "precipitants" such as poly(ethylene glycol) or high salt concentrations, such as ammonium sulfate, which has a strong "salting-out" effect on proteins. Other possibilities, which are less often used, are temperature and pH gradients. The method most often used for crystallization screening is the so-called vapor diffusion method, either in sitting or in hanging drops (Figure 30.8). Nowadays, vapor diffusion experiments are usually performed in 96-well microtiter plates, and are set up with automated dispensing systems.[90,91] Other crystallization techniques have been developed, such as crystallization under oil, crystallization in agarose gels, in capillaries using the free interface diffusion technique, and so on.[88,91] Microfluidics technologies are increasingly being used, either for crystallization screening or, alternatively, for determining phase diagrams.

FIGURE 30.7 Crystal packing of a human thrombin complex. Twelve unit cells with one layer of molecules are shown. By looking carefully, one can see that the two molecules in each unit cell are rotated 180° with respect to each other. Protein crystals used for X-ray diffraction extend into 3D and consist of many layers of molecules. The next layer of thrombin molecules fits into the holes present in the layer shown.

b. Dealing with known targets

When crystallization conditions are already known, 1 mg of protein may be enough to produce a series of crystals with different inhibitors. One should bear in mind, however, that published crystallization protocols are often difficult to reproduce! It is wise, in a first step, to follow as closely as possible the published expression, purification, and crystallization protocols. Particular attention should be paid to the protein construct, since minor changes to the amino acid sequence can have a dramatic influence on the solubility, stability, and crystallization behavior.

c. Tackling novel targets

Obtaining X-ray quality crystals is usually the most difficult and time-consuming step of a new structure determination project, notably in the case of a novel, poorly characterized gene product.[91] Recent advances in crystallization robotics and the provision of numerous commercial crystallization screens have considerably simplified the process and shortened the time needed to set up extensive crystallization screening experiments, while, at the same time, dramatically reducing the amount of material needed. Nowadays, 5–10 mg of a homogeneous protein sample are

BOX 30.1 Some Common Crystallographic Terms

Space group: The group of symmetry operators which describe the symmetry of the crystal. Since biological molecules are optically active, their crystals belong to one of the 65 non-centrosymmetric space groups.

Unit cell: The basic building block of a crystal. The whole crystal can be generated by repeated unit translations of the cell in 3D. The unit cell is characterized by its axes a, b, c and the angles (α, β, γ) between them.

Asymmetric unit: The smallest motif from which the whole unit cell can be generated by applying the symmetry operators of the space group. The asymmetric unit may contain one or more copies of the protein or complex under study. In the case of oligomeric particles, the asymmetric unit may contain one or more complete particles or only one or more subunits if some symmetry axes of the particle coincide with some symmetry axes of the crystal.

Reflection: A diffracted beam of X-rays, characterized by its indices h, k, l, and caused by reflection from the lattice planes making intercepts a/h, b/k, and c/l with the unit cell axes. Each reflection contains information on the entire structure. Reflections occurring at high scattering angles have high indices and carry high resolution information (they correspond to a fine sampling of the structure), while those observed close to the direction of the incident beam have small indices and carry low resolution information (they correspond to a coarse sampling of the structure).

Resolution: The resolution limit corresponds to the highest scattering angle at which reflections can still be measured (*cf.* Box 30.2). Individual atoms can be fully resolved when the resolution is better than 1.0 Å.

sufficient to screen several hundreds to a few thousands of individual crystallization conditions. Nevertheless, this is often not enough to obtain suitable crystals for new and difficult targets.[91] Several protein constructs or protein variants may have to be generated to increase the chance of finding one that is amenable to crystallization. For this reason, a strong and dedicated support in molecular biology, protein expression, and biochemistry is an absolute must for a successful protein crystallography laboratory. Furthermore, in an industrial setting, these activities should be initiated as early as possible to ensure that crystals are available before chemistry activities are started.

Recombinant proteins designed for assay purposes are usually not suitable for crystallization experiments. Because well-diffracting crystals require a high degree of order in the packing of the protein molecules, proteins with fusion partners, long tags, floppy ends, disordered or intrinsically unstructured regions, loosely linked domains,[92] etc., as well as highly glycosylated proteins,[93] are usually not amenable to crystallization. Intrinsically unstructured regions can usually be identified from the amino acid sequence as polypeptide segments with low sequence diversity.[94] Domain boundaries can be pinned down by limited proteolysis or with the help of homology modeling. If tags or fusion partners are needed for enhanced expression and/or ease of purification, then a highly specific protease recognition sequence should be engineered to allow their removal before crystallization.

Glycosylated proteins often give poorly diffracting crystals. To circumvent this problem, several strategies can be tried:[93] the glycosylation can be chopped off enzymatically with PNGase F or other endoglycosidases, glycosylation sites can be mutated, or a non-glycosylated form can be produced using a prokaryotic expression system.

The crystallization of membrane proteins is particularly challenging and is out of the scope of this review.[95] However, it should be emphasized that many transmembrane proteins possess intracellular or extracellular domains which carry the biological function or binding site of interest. Recombinant versions of these soluble domains can frequently be expressed and are usually amenable to crystallization and structure determination.[96]

Other pre-requisites for crystallization are good stability and sufficient solubility (typically 5–15 mg/ml, but some highly soluble proteins may require a much higher concentration, while others can be crystallized at a concentration between 1 and 5 mg/ml). Solubility and stability can sometimes be improved by side-directed mutagenesis. Both random mutagenesis approaches and rational ones have been proposed and successfully used.[92] Alternatively, one may consider using another species instead of the human protein.

Stabilization of the protein target by buffer additives or complex formation is a powerful way to enhance crystallizability.[97] Biophysical techniques such as thermal-shift assays and NMR can be used for the identification of suitable ligands or additives.

Post-translational modifications such as phosphorylation can be a serious issue, since they often give rise to severe conformational heterogeneity. For instance, protein kinases are often produced from eukaryotic expression hosts as a mixture of inactive (unphosphorylated) and active (with one or more phosphorylations) species. One possible workaround consists in mutating the phosphorylation site(s) to glutamate. The negatively charged glutamate side-chain mimics the phosphorylation, thus producing a constitutionally active kinase. Expression in the presence of a potent, cell permeable inhibitor is another strategy which has been found to reduce the heterogeneity of the phosphorylation,

FIGURE 30.8 Principle of the method of vapor diffusion in hanging drops. A droplet (0.1–1 μL) of a solution of protein or protein complex is mixed with crystallization buffer (typically in a 1:1 ratio), suspended above a reservoir containing the crystallization buffer (~0.1–0.5 mL) and allowed to equilibrate. Because of the concentration gradient, water is transferred from the drop to the reservoir via the vapor phase. As a result, the concentrations of protein and precipitant in the drop increase. If a supersaturated state is reached, the protein will then precipitate. The precipitate is usually amorphous, but crystals will form in successful experiments.

while, at the same time, stabilizing one particular conformation.[98] Moreover, enhanced expression levels and purification yields, together with improved sample homogeneity and crystal quality, have been achieved in the case of several protein kinases and nuclear receptors, by adding a ligand during all expression and purification steps.[97,98]

When novel protein targets cannot be produced or crystallized, one should consider the possibility of using a known homolog or anti-target instead. If the binding sites are sufficiently similar, the binding modes of some key compounds or scaffolds can be deciphered and this information fed into the drug design process.

d. Crystallization of protein–ligand complexes

An important aspect of protein crystallography in the context of drug design concerns the determination of protein–ligand complexes.[99] If the ligand is a relatively small molecule, it is often possible to obtain crystals of the complex by soaking crystals of the unliganded protein in an artificial mother liquor containing the ligand. Protein crystals contain solvent channels which are usually large enough to allow the diffusion of the ligand to its binding site. A soaking experiment requires little material (1 μmol of compound is usually plenty), but the solubility of the compound under the crystallization conditions is often an issue. The high protein content of the crystallization drop usually requires ligand concentrations in the range of 0.5–5.0 mM. A typical soaking protocol involves the preparation of a concentrated (50–100 mM) stock solution of the ligand, usually in a suitable organic solvent such as dimethylsulfoxide. This solution is then mixed with crystallization buffer to a final concentration of solvent of up to 5%, and a few microliters of this mixture are added to the crystallization drop (Figure 30.8). Compound purity, or the use of a diastereomeric mixture, may not be an issue if only one component in the sample binds to the protein. However, the chemical structure of the ligand must be known, since ultra-high resolution would otherwise be required for the unambiguous identification of an unknown binder. Likewise, when asymmetric centers are present, it is preferable to know beforehand their stereochemistry. At high resolution (2.0 Å or better), it is frequently possible to assign the absolute configuration on the basis of the electron density, but at a lower resolution ambiguities may be present.

Soaking has some practical advantages, particularly when the crystallization is not robust and good crystals are difficult to prepare. Using this approach, crystallographic analysis can be performed in the high-throughput mode, provided a reasonably short soaking time is sufficient to achieve full occupancy of the binding site. The fast turnaround time afforded by the soaking method ensures fast feedback to modeling and chemistry, and opens up the possibility of using X-ray analyses for hit triaging and validation, as well as for FBS. However, the soaking method also has some drawbacks. Conformational changes induced by ligand binding in solution may be hindered in the crystal, or access to the binding site may be restricted by protein–protein contacts. As a result, the ligand may not bind at all, or it may adopt an artificial mode of binding. Moreover, the diffraction quality can sometimes suffer from the soaking procedure, or the crystals may even crack or dissolve upon soaking. In such cases, gentle cross-linking of the crystals using the method of Lusty[100] may prove useful, but validation of the soaking approach with a co-crystallization experiment would then be worthwhile.

When the crystallization is reasonably fast and robust, co-crystallization is the method of choice and is recommended even in cases where soaking would be possible. In the co-crystallization approach, the complex is formed in solution by mixing an excess of the ligand with the protein, and crystallization experiments are then set up (Figure 30.8). In this way, the risk of artifacts is minimized, but at the expense of speed, particularly when the crystallization is very slow. A further disadvantage is that, for each and every new complex, crystallization conditions may have to be optimized again, or full crystallization screening may be required, since crystallization conditions are sometimes very sensitive to changes in the ligand. Seeding is frequently used for accelerating co-crystallization experiments and improving their reproducibility.

e. The "ligand-fishing" protocol

Biological assays are usually performed under experimental conditions where a very large excess of ligand over the protein is present. This is particularly true for weak ligands, which are assayed at concentrations in the micromolar range while the protein concentration is typically in the nanomolar or subnanomolar range. Medicinal chemists should always bear in mind that weak biological activity may sometimes be due to trace amounts of a highly potent compound "contaminating" an otherwise inactive sample. For instance, an IC_{50} of 10 M could be due to 0.5% of an impurity with a potency of 50 nM. This situation is not so uncommon in medicinal chemistry programs where inactive derivatives are sometimes obtained from very potent precursors. Trace impurities of 1.0% or less are usually not detected by routine analytical techniques, but may give rise to apparent micromolar activity. These impurities will not be detected by crystallography either, if only a small excess of compound (2 to 5-fold) is added to a concentrated protein aliquot, as is usually the case (Figure 30.8). But, if a very large excess of compound (say 500 to 1,500-fold) is added to a diluted sample of protein, then there may be enough active impurity to saturate or nearly saturate the protein. The complex can then be concentrated using standard ultrafiltration techniques and crystallization experiments performed. This procedure, which we like to refer

FIGURE 30.9 The setup for an X-ray experiment. The X-ray generator produces a powerful beam. The monochromator selects X-rays of a single wavelength (1.54 Å for copper targets) and the collimator limits the divergence and diameter of the beam to 0.1–0.3 mm. The incident beam hits the crystal and some X-rays are then diffracted by the crystal. Most X-rays pass straight through and are stopped by a small piece of lead, the beam stop. The diffracted X-rays are detected by a CCD detector or an imaging plate. The goniometer shown here (big black circle) has four rotation axes and allows the crystal to be positioned in any orientation with respect to the X-ray beam.

to as "the ligand-fishing protocol" (Figure 30.8), is a bit more time consuming, and requires larger amounts of compound (5–10 mg), but it has the ability to detect very low amounts (down to approximately 0.1%) of a potent ligand in a mixture. It may prove useful in cases where a weakly active compound, whose structure is at odds with the established structure-activity data, could not be observed using the routine crystallization procedure.

In addition, this protocol may work better than the standard procedure when ligands are very poorly soluble in the crystallization buffer.

B. Data collection

Protein crystals contain on average 50% solvent and, if exposed to air, they dry out, and disintegrate. Moreover, when exposed to high-intensity X-rays at room temperature, they loose their diffraction power very quickly, owing to radiation damage. In order to prolong crystal lifetime and improve data quality, X-ray measurements are now routinely performed at 80–110 K.[101] Crystals are first mounted on 10–20 μ thin nylon loops and then flash frozen by immersion into liquid nitrogen. To prevent the formation of ice crystals, it is often necessary to add to the surrounding mother liquor a cryo-protectant such as glycerol, low molecular weight poly(ethylene glycol), or high salt.

For data collection, one crystal is then placed on a goniometer, a device which controls the rotation of the crystal in the X-ray beam, while temperature is kept at 80–110 K by blowing dry nitrogen over the crystal (Figure 30.9). Large, strongly diffracting crystals can be measured in the laboratory with a rotating anode X-ray generator, but tiny or weakly diffracting crystals must be measured at a synchrotron source, such as the Swiss Light Source (SLS) in Villigen, Switzerland, or the European Synchrotron Radiation Facility (ESRF) in Grenoble, France. During data collection, the crystal is slowly rotated to bring all reflections into diffracting condition (see Box 30.2 about Bragg's law).

The diffraction spots are usually recorded on CCD detectors or on detectors based on imaging plates. The time to collect high quality X-ray data sets from single crystals ranges from a few minutes for good crystals and high-intensity synchrotron radiation to a few days for weakly diffracting crystals and a conventional X-ray generator. The diffraction images from these detectors (Figure 30.10) are fed directly into a computer which produces a list of reflection intensities. Thousands to several hundred thousands of reflections are recorded per crystal, depending on the quality of the crystal and the size of the unit cell.

C. From diffraction intensities to a molecular structure

1. *The diffraction of X-rays by crystals*

a. Light microscopy and X-ray crystallography share the same basic principle

A light microscope allows us to study in great detail small objects like insects or cell slices, but it is physically impossible to resolve any details which are smaller than half the wavelength of the light used. For blue light this limit is about 200 nm. To resolve atomic details, which are on the order of 1–5 Å (0.1–0.5 nm), electromagnetic radiation with

FIGURE 30.10 Example of an X-ray diffraction image. The characteristic pattern of diffraction spots is caused by the fact that X-rays diffracted from different unit cells in the crystal should scatter in phase.

BOX 30.2 Bragg's Law

Observable diffraction is only obtained when waves deflected from adjacent planes reinforce each other, that is when the path difference ABC in the figure is an integer multiple n of the wavelength λ of the X-ray radiation used. This leads to Bragg's equation:

$$n\lambda = 2d \sin\theta$$

This means that to resolve closely spaced planes (small d), we need to measure high angle (large θ) reflections.

a much shorter wavelength than light is required: X-rays. A light microscope and an X-ray set-up share the same basic principle, although the practical implementation is quite different, owing to the different properties of X-rays and visible light.

In a microscope, light from a light source shines on the sample and is scattered in all directions. A set of lenses is used to reconstruct from this scattered light an enlarged image of the original sample. In an X-ray experiment, X-rays from an X-ray source hit the crystal and are scattered in all directions, just as with the light microscope. Unfortunately, no lenses can be made which are able to bring the scattered X-rays into focus to reconstruct an enlarged image of the sample. All the crystallographer can do is to record directly the scattered X-rays (the diffraction pattern, see Figure 30.10) and to use computer programs to reconstruct an enlarged image of the sample.

b. X-rays are scattered by electrons

Although X-rays interact only weakly with matter, they are occasionally absorbed by electrons, which start to oscillate. These oscillating electrons serve as X-ray sources which can send an X-ray photon in any direction. X-ray photons, scattered from different parts of the crystal have to add up constructively in order to produce a measurable intensity. The condition under which the scattered X-rays add up constructively is laid down in Bragg's law, which treats crystals in terms of sets of parallel planes (Box 30.2).

c. The diffraction pattern corresponds to the Fourier transform of the crystal structure

Each diffraction spot is caused by reflection of X-rays by a particular set of planes in the crystal. If the crystal contains layers of atoms with the same spacing and orientation as a particular set of planes which would satisfy Bragg's law (if the set of planes is physically present), the corresponding diffraction spot will be strong. On the other hand, if only few atoms in a crystal correspond to a particular set of planes, the corresponding reflection will be weak. The complicated structure present in the crystal is transformed by the diffraction process into a set of diffraction spots which correspond to sets of planes (more precisely, sinusoidal density waves), just as our ear converts a complicated sound signal into a series of (sinusoidal) tones when we listen to music. This conversion of a complicated function into a series of simple sine and cosine functions is called a Fourier transformation.

2. *The phase problem*

The original function, in our case the electron-density distribution in the crystal, can be reconstructed by performing the inverse Fourier transformation, that is by summing together the corresponding density waves for all reflections (see Figure 30.11). However, in order to make this summation, we need to know not only the amplitude of the density wave, but also its relative position with respect to all other

FIGURE 30.11 The phase problem. The experimental data obtained in an X-ray experiment are the intensities of the reflections. By using an inverse Fourier transform, it is possible to calculate electron-density maps from these intensities. However, it is essential for this calculation to know the phase associated with each reflection. Approximate initial phases can be obtained from heavy-atom derivatives, anomalous dispersion or molecular replacement (see text). More accurate phases can be derived from the refined model, once it has been obtained.

FIGURE 30.12 Close-up view of a protein complex at 1.5 Å resolution showing the initial $2F_o$–F_c electron-density map (magenta mesh, 1.0 σ contour), as well as the initial F_o–F_c map (green mesh, 3.0 σ contour). The ligand has not yet been included in the model. Therefore, it appears as a strong positive difference density in the initial F_o–F_c map.

density waves (the phase). The amplitude can be measured, because it is calculated from the intensity of the corresponding diffraction spot, but there is currently no practical way to measure all the phases directly. This so-called "phase problem" can be solved by one of the following techniques.

a. Multiple isomorphous replacement (MIR)

Crystals are soaked in solutions with "heavy" atom salts (Hg, Pt, Au, etc.), in the hope that a few heavy atoms will bind to some well-defined sites on the protein molecule. The heavy atom positions are then found by analyzing the differences between the diffraction pattern of the native and of the soaked crystals. When two or more suitable heavy atom derivatives are found, phase estimates and an electron-density map can then be calculated.

The MIR approach can be used to solve any protein structure *de novo*, but finding good heavy atom derivatives requires many crystals and is often a cumbersome and lengthy process.

b. Anomalous scattering (AS)

This method makes use of the fact that some inner electrons of the heavier elements have absorption edges in the range of X-ray wavelengths. The method is used to supplement the phase information of a single heavy atom derivative,[102] but also to obtain full phase information from proteins which are labeled with selenomethionine, a selenium-containing amino acid.[103] This method, called "Multiple wavelength Anomalous Dispersion" (MAD) has become very popular for the fast structure determination of novel proteins. Other anomalous methods have recently been proposed. For instance, the "halide-soak" approach uses short soaks in solutions containing 0.5–1.0 M bromine or iodine, and the anomalous signal of the bound halide ions is then exploited to solve the structure.[104] For well-diffracting crystals, it is also possible to use the sulfur anomalous signal from the cysteines and methionines present in the native protein.[105]

The MAD method is performed on a single crystal, but it requires access to tunable radiation (synchrotron source). Moreover, selenomethionine-labeled protein must be produced, purified and crystallized. This is more easily done for proteins which can be expressed in *E. coli*.

c. Molecular replacement

When a suitable model of the unknown crystal structure is available, it can be used to solve the phase problem. Examples are the use of the structure of human thrombin to solve the structure of bovine thrombin, the use of a known antibody fragment to solve the structure of an unknown antibody, or the use of the structure of an enzyme to solve the structure of an inhibitor complex of the same enzyme in a different crystal form. The model is oriented and positioned in the unit cell of the unknown crystal with the use of rotation and translation functions, and the oriented model is subsequently used to calculate phases and an electron-density map.

Molecular replacement is usually straightforward and performed within minutes. However, when only low resolution data and a poor search model are available, model bias can become an issue and experimental phasing may be needed. Nevertheless, the molecular replacement method is extremely useful in the context of structure-based design where it is used to expedite the determination of multiple protein complexes or to analyze, in an automated way, multiple data sets from a fragment-based screen.

3. *Model building and refinement*

Once a first electron-density map is obtained, it is interpreted by the crystallographer. In the case of a MIR(AS) map, a complete model of the protein has to be fitted to the electron density. The Cα atoms are placed first (chain tracing), and subsequently the complete main chain and the side chains are built, a process which has become more and more automated in recent years, notably when high resolution data are available. In the case of molecular replacement, the search model needs to be updated to reflect the molecule present in the crystal. The model is usually of a similar protein and the changes involve the substitution of some amino acids, the introduction of insertions and deletions, the modification of some loops, and so on.

After the (re)building step, the model is refined. Refinement is an iterative procedure which aims at minimizing the differences between the observed diffraction amplitudes (F_o), and the diffraction amplitudes calculated from the model (F_c) while simultaneously optimizing the geometry of the structure. Because of the unfavorable ratio between observations and parameters, a free atom refinement is not possible in protein crystallography, and it is necessary to restrain the bond lengths, valence angles, and dihedral angles towards ideal values (see Section IV. D). Phases calculated from the refined model at the end of each refinement cycle are then used for the calculation of improved electron-density maps, which are again analyzed by the crystallographer to improve the model further. Cycles of refinement and rebuilding are repeated until convergence is reached. The final set of co-ordinates is then ready for deposition with the PDB.[106]

4. *Most used types of electron-density maps*

The direct experimental result of a crystallographic analysis is an electron-density map, while the model is derived from a (subjective) interpretation of this map. It is therefore useful to refer to the original data, the electron density, as often as possible. In the following paragraph, we will discuss the different types of electron-density maps most commonly used.

a. $F_o - F_c$ or difference maps

These maps are obtained after subtracting the calculated structure factors (F_c) from the observed structure factors (F_o), an operation which is, in a first approximation, equivalent to subtracting the calculated electron density from the observed electron density. Features which are present in the "observed" density, but not in the calculated density will give peaks, while atoms present in the model (in the F_c), but not in the "observed" electron density will result in holes (Figure 30.12). These maps are frequently used to detect errors in the model and can also be used to obtain an unbiased electron density of a bound inhibitor, for example, by removing the inhibitor completely from the model. In this case, the resulting electron density for the inhibitor is entirely caused by the experimental data, and not by any model bias present in the phases. These maps are often referred to as "omit maps".

b. $2F_o - F_c$ maps

These are the standard electron-density maps (Figure 30.12). Because of model bias, maps calculated with F_o and model phases tend to show only electron density associated with the model. As discussed above, $F_o - F_c$ maps show everything which is in F_o, but not in the model. By combining an F_o map with an $F_o - F_c$ map, a $2F_o - F_c$ electron-density map is obtained, which shows both electron density for the model and electron density for features which are not yet accounted for in the model, such as bound water molecules, carbohydrates and other molecules associated with the protein.

Several weighting schemes exist to minimize model bias. Examples are the figure-of-merit, σ_A, and maximum-likelihood weighting.

D. Information content and limitations of crystal structures

Most chemists are familiar with X-ray analyses of small molecules, which are typically performed at a resolution better than 0.80 Å. These subatomic resolution studies deliver highly accurate geometric parameters (bond lengths, valence, and dihedral angles), as well as anisotropic displacement parameters ("temperature ellipsoids"). This is made possible by the very favorable observation to parameter ration (typically 50:1) resulting from the ultra-high resolution. Usually, protein crystals do not diffract to atomic or subatomic resolution. The vast majority of X-ray studies are performed at medium to high resolution (say between 3.0 Å and 1.50 Å), where the level of structural details described above is not attainable. In particular, stereochemical parameters such as bond lengths and angles are restrained to standard dictionary values, both for the protein part and any low molecular weight ligand(s), prosthetic group or post-translational modification. The protonation state and the exact orientation of some amino - acid side-chains (His, Asn, Gln) are, whenever possible, inferred from potential H-bonded interactions. Substantially fewer solvent molecules and alternate conformations are observed than in the case of ultra-high resolution studies.[34,35]

1. *Quality indicators*

a. Quality of the experimental data

The quality of a crystal structure cannot be better than the quality of the experimental data it is based upon. The

following criteria are commonly used statistical indicators of the quality of the diffraction data:

(a) *Resolution*. This corresponds to the shortest spacing of planes (d) whose reflections have been used in map calculation and refinement (see Box 30.2). The smaller this spacing, the sharper and the more detailed the electron-density maps will be. The resolution is probably the single most important criterion determining the quality of a crystal structure. At high resolution (better than 2.0 Å) the protein and bound water molecules are well defined, and it is unlikely that the structure will contain any serious errors. At low resolution (2.8–3.5 Å), it is usually not possible to assign bound waters with certainty and significant errors can remain unnoticed due to the problem of model bias.

(b) *Completeness of the data*. One can calculate the total number of reflections to a certain resolution and, ideally, one would like to measure them all. However, for various reasons it is in practice often not possible to measure all reflections. If only a small fraction of the reflections is missing (<10%), and the missing reflections are weak, the electron-density maps will hardly be affected. However, if a significant fraction of the reflections is missing, this may lead to artifacts in the electron-density maps and the problem of model bias will become more severe.

(c) *R-sym*. This reflects the inconsistency of multiple measurements of the same reflection, therefore the lower the *R*-sym the better. *R*-syms below 10% are usually achieved.

b. Quality of the model

The global quality indicators listed below are commonly reported for refined crystal structures:

1. *R-factor*. This is a measure of the disagreement between the observed amplitudes (F_o) and the amplitudes calculated from the model (F_c). Depending on the resolution and quality of the diffraction data, well-refined structures have *R*-factors below 20–25%.
2. *Free R-factor*. Since refinement programs aim at minimizing the difference between observed and calculated amplitudes, that is the *R*-factor, another unbiased indicator is needed to monitor the progress of refinement. Brünger proposed to exclude a subset of reflections from refinement and to use these reflections only for the calculation of "free"*R*-factors.[107] If refinement is progressing correctly the free *R*-factor will drop, but if the model contains serious errors it will remain stalled above ~35%. For correct structures, the free *R*-factor is generally below 30%.
3. *Deviations from ideality of bond lengths and bond angles*. A correctly fitted model is generally not strained. Significant deviations from ideal values for bond lengths and bond angles usually point to problems with the structure. Root-mean-square (r.m.s.) deviations from ideality should not be much larger than 0.02 Å for bond lengths and 3° for bond angles. The bond lengths and angles are biased towards the target values which are used during refinement. Accurate, unbiased values for these parameters can only be derived when ultra-high resolution (0.85 Å or better) is available.
4. *Ramachandran plot*. Because of steric hindrance, only certain combinations of the main-chain dihedral angles ϕ and ψ are "allowed" (Figure 30.13). The protein fold may force some residues to assume unallowed ϕ, ψ values, and this may have functional significance for some active site residues.[108,109] However, if more than a few percent of all the residues have ϕ, ψ values completely outside allowed regions, one should suspect errors in the model.

2. Information content

a. Errors in crystal structures

Serious errors in crystal structures are very rare, and are usually associated with the first structure determination of a novel target, in particular when only low resolution data are available (3.0–5.0 Å). Small errors and inaccuracies, however, are very common and virtually unavoidable. These errors are often underestimated, and small details of crystal structures are frequently overinterpreted by non-crystallographers. Medicinal chemists making use of crystal structures should be well aware of their limitations.[111]

Errors affecting the protein model

One main source of errors in macromolecular crystallography results from our inability to detect and to model "disorder" appropriately,[111] owing to the limited resolution and unfavorable parameter to observation ratio. Crystallographic refinement often attempts to fit a single model to some blurred electron density originating from several distinct, but overlapping conformational states. This may lead to distorted geometry, or to different, but equally valid, interpretations.

A second important source of errors results from the fact that hydrogen atoms cannot be detected and atom types cannot be assigned at the resolution which is typically attainable with most protein crystals (1.5–3.0 Å). This leads to ambiguities in the exact orientation of some groups, such as the side-chain amide of Asn and Gln residues, or the imidazole ring of histidine side-chains.

Errors affecting the ligand

Likewise, the exact orientation of one or more ligand groups can sometimes remain uncertain. The choice of sensible geometric restraints for the refinement of non-standard groups, in particular the ligands, is not always trivial and

FIGURE 30.13 Ramachandran plot. Triangles indicate glycine residues, squares indicate non-glycine residues. The most favored regions are indicated in red, additionally allowed regions in bright yellow and generously allowed regions in pale yellow. Residues with phi-psi angles outside the allowed regions (indicated in red and labeled) are strained. Much larger regions are allowed for glycine residues because they do not have a side-chain. When more than a few percent of the residues have phi-psi outside the allowed regions, one should be suspicious for errors in the structure. The figure was generated with the program Procheck.[110]

constitutes also a potential source of errors.[114] For instance, the nitrogen atom of a tertiary amine bearing one aromatic substituent is usually planar, but it can also be pyramidal. At high resolution (better than 2.0 Å), it may be possible to select, on the basis of the electron density, the appropriate geometric restraints. At lower resolution, it is likely that the refined model will merely reflect the arbitrary choice of geometric restraints.

Errors affecting the solvent model

Water molecules are usually identified on the basis of residual electron-density peaks which meet certain criteria, such as the peak height, the distance and the angle with respect to H-bond donor or acceptor groups. Since atom types and protonation states cannot be determined, a "water" may as well be an hydroxide, an hydroxonium, an ammonium, a sodium, or a magnesium ion. The assignment of metal ions becomes more reliable when the resolution of the data is good enough to reveal the co-ordination sphere, or when the anomalous signal of the metal can be used.

b. Flexibility and temperature factors

Proteins are flexible molecules[112] and they usually retain in the crystalline state a substantial degree of flexibility. The mobility of the atoms in a crystal is expressed in terms of temperature-factors, or B-factors, which are optimized during refinement. The relationship between mean total displacement and B-factors is given in Figure 30.14. The mean displacement of atoms with B-factors in excess of

FIGURE 30.14 Relationship between mean total displacement and temperature factor B. At temperature factors of 60 Å^2 and higher, the displacement becomes larger than 1.5 Å and the electron density becomes very poor (see Figure 30.15). The formula used in the figure is derived from the relationship $B = 8\pi^2 <u^2>$, where $<u^2>$ represents the displacement perpendicular to the diffracting planes. The total mean square displacement $<u^2_{tot}> = 3 <u^2>$.

60 Å^2 is larger than 1.5 Å, which is the length of a carbon–carbon bond. These atoms are generally poorly defined in the electron-density maps (Figure 30.14), although in the case of static disorder these atoms may still have well-defined, albeit weak, electron density. For functional analysis, one should bear in mind that these flexible surface residues are either put in an arbitrary, low-energy conformation, or are deleted from the co-ordinate file. Not taking this

FIGURE 30.15 Long and flexible side-chains (such as Arg, Lys, Glu, and Gln) which are exposed to the solvent often move freely around. As a result, these side-chains have very high temperature factors and are very poorly or not at all defined in the electron-density map. Lys87, located at the surface of human thrombin, is shown as an example. One should take into account that many of such surface residues do not have a well-defined orientation and are either deleted, or as in this case, put in an arbitrary, low-energy conformation.

into account could lead to serious artifacts, especially with electrostatic calculations (Figure 30.15).

3. *False positives and false negatives*

Literally speaking, false positives cannot occur in a protein crystallography experiment. Nonetheless, misinterpretation of some ambiguous electron-density features in terms of a bound ligand would be akin to a false positive. Examples of this kind of errors are extremely uncommon. In the context of a fragment-based screen, the assignment of a not-so-well-defined electron-density blob to one of the small fragments present in the cocktail can sometimes be uncertain. This problem can be compounded by insufficient resolution or data quality, partial occupancy, the presence of multiple and overlapping binding modes, or the competitive binding of several fragments or buffer components to the same site. Automated procedures for ligand identification and fitting are of great help, but careful evaluation of the results by an experienced scientist remains essential. In ambiguous situations, the measurement of control data sets ({cocktail minus presumed binder} and {presumed binder alone}) may be needed.

In contrast to false positives, false negatives are relatively frequent in protein crystallography, in particular when dealing with low to very low affinity hits from high-throughput screening (HTS) or FBS, for reasons which will be explained below. Hence, the failure to observe binding in an X-ray experiment does not necessarily disqualify a compound from being a genuine ligand. Medicinal chemists confronted with the difficult decision to pursue or to abandon dubious, but potentially interesting hits should use a variety of biophysical techniques for hit validation, such as protein NMR,[115] mass spectroscopy, surface plasmon resonance, and microcalorimetry, to avoid too many inconclusive X-ray experiments.[116]

The very high protein and ligand concentrations used in co-crystallization experiments are, in principle, highly favorable for the study of weak complexes. However, protein–protein contacts in the crystal can block access to a ligand binding site (in a soaking experiment) or compete out a ligand (in a co-crystallization experiment). Furthermore, complex formation can be adversely affected by the crystallization conditions (pH, high salt or high concentration of glycerol, low molecular weight poly(ethylene glycol), or organic solvent). In addition, the solubility of some compounds can be exceedingly low under the crystallization conditions. Differential scanning calorimetry experiments or other thermal-shift assays can be used to assess binding under the crystallization conditions.

V. PRACTICAL APPLICATIONS

A. Target identification, selection and validation

Current low molecular weight medicines exploit a fraction of all potential drug targets – a mere 250, according to a recent review.[118] Over the past decade, the large-scale sequencing of whole genomes, including the human genome, has uncovered thousands of previously unknown genes or "open

BOX 30.3 Using PDB Files – Tips and Tricks

Crystal structures are stored in formatted text files called "PDB" files. These files can be freely downloaded from the RCSB Protein Data Bank at http://www.pdb.org. Information on how to search and navigate the PDB is available on the PDB home page.

Tip 1: always download a complete biological unit! PDB files usually contain only the portion of the structure forming the asymmetric unit of the crystal. The asymmetric unit may include only a fraction of the functional biological molecule, for instance a single subunit of a homodimer. In such a case, information derived from viewing a single subunit may be very misleading since binding sites or active sites are sometimes located at the interface between two or more subunits. Although a complete biological unit can be generated with help of crystallographic software, it is also possible to download the corresponding file directly from the PDB (Download files → Biological unit).

Tip 2: check all molecules of the asymmetric unit! Sometimes the asymmetric unit of the crystal contains several copies of the molecule or complex of interest. In such cases, individual copies of the biological unit can be downloaded individually as separate files or all together in the original PDB file. It is very important to inspect them all since significant differences can exist between these molecules due to different crystal contacts, disorder, partial occupancy of a ligand or co-factor, or as a consequence of different conformational states.

Tip 3: do not only look at the 3D *model – check the actual experimental information as well: the electron-density map*. Electron-density maps contain more information than can possibly and accurately be included in an atomic model, even after careful refinement by an experienced crystallographer. For instance, some alternate conformations may not have been modeled (also note that many graphic programs ignore alternate conformations and do not display them!) or some portions of a ligand molecule may be disordered, but co-ordinates for the complete molecule have been included. These and other important details can be revealed by displaying the electron-density map together with the atomic model. For structures which have been deposited with the PDB together with the corresponding diffraction data, electron-density map files can be downloaded from the Uppsala electron-density server[117] at http://eds.bmc.uu.se/eds.

Tip 4: browse through your PDB *file to find out more about its content*. While the 3D structures encoded in the PDB files are best visualized using a graphic program (some interactive viewers are directly accessible from the PDB web pages), bear in mind that PDB files are simple text files which can also be displayed using a text editor. Browsing through PDB files can reveal some important information, notably on the method used to derive the structure (NMR, X-ray, or modeling), some data statistics, the amino acid sequence with comments about engineered residues, the numbering of the protein residues and associated co-factors, ligands and solvent molecules, and more.

reading frames". This avalanche of genomic data has raised hope that many new opportunities for treating diseases will emerge from these and subsequent genetic studies aimed at deciphering the function of the novel genes and their relevance to disease. While it is unclear how many drug targets remain to be discovered,[119,120] it is of utmost importance for drug discovery pipelines that new targets are swiftly identified, judiciously selected and appropriately validated. Genetic and proteomic studies do not provide the full answer to this issue. Structural information, together with chemistry, can also contribute to target identification and validation, while, at the same time, providing critical information for judicious target selection.

1. *Target identification*

a. Structural genomics approaches

Functional annotation of novel gene products used to be mainly based on sequence homologies to previously known proteins. For distant relatives, these homologies are often limited to a few short, but usually characteristic, sequence motifs. Tentative assignments of this kind can now be put on a solid basis by more sophisticated and powerful approaches, such as threading techniques, which aim at verifying the compatibility of a given amino acid sequence with a 3D-fold. However, a substantial fraction of the novel genes unveiled during the course of genome sequencing efforts code for proteins with no apparent relationship to any of the currently known ones. Structural genomic centers are solving the crystal structures of many of these novel proteins. For some of them, the function is immediately revealed by the analysis of the 3D-fold. Unfortunately, the assignment of the function from the X-ray structure is not always straightforward. This situation arises when a new fold is uncovered, but sometimes also when a common fold is found again, since some protein folds are used by nature for many different purposes, either through divergent or convergent evolution. To address this problem, a methodology has been developed that uses a database of 3D active site templates to pin down the biochemical function of a novel protein. For instance, the spatial configuration of the Ser-His-Asp catalytic triad of serine proteases is highly conserved and can be used as a search template to find out if a novel protein may exhibit this kind of catalytic activity.[120] In addition, new software for detailed analyses of protein surfaces are emerging as very promising tools to infer the biochemical function from the 3D structure.[121–124] These developments are of particular interest to pharmaceutical research since the identification, delineation, and comparison of binding sites is an important aspect of the druggability assessment of a potential drug target (see below).

b. Structure-based chemical biology approaches

Finding the target of a known drug is an appealing alternative to the more conventional strategy which consists in finding drugs acting on known targets. When a biologically active compound has been identified, for instance from a phenotypic screen, an emerging technology called "reverse docking", can be used to identify the protein target. In this approach, a 3D database of protein binding sites is searched with a docking algorithm.[125] Candidate protein targets are subsequently validated using a biochemical assay and, ideally, with protein crystallography.[126] This promising approach is currently limited by the shortcomings of today's docking algorithms and associated scoring functions, and by the low coverage of biological space through existing protein structure databases. Nevertheless, its application to target families, such as protein kinases, shows promise for discriminating between the targets and non-targets of a particular inhibitor.[127]

A related approach recently proposed for assigning the function of enzymes of unknown activity consists in structure-based docking of thousands of candidate substrates in their transition-state-like geometries.[128]

2. *Target validation*

Because of the intricacies and highly complex regulation of cellular networks and pathways, it is virtually impossible to predict which particular target will ultimately provide the desired pharmacological effect when blocked specifically by a low molecular weight drug. Moreover, it is becoming apparent that different approaches to the study of protein function in cells, such as gene knockouts, temperature sensitive alleles, RNA interference and low molecular weight compounds often yield conflicting outcomes.[129] These discrepancies may result from the insufficient selectivity of the available inhibitors, or from serious and unwanted biological side-effects caused by a temperature shift or the complete loss of a particular gene product. These shortcomings are circumvented by a new approach called chemical genetics, which exploits structural knowledge to engineer protein mutants exquisitely sensitive to customized inhibitors.[129] In this way, truly selective inhibition of a single protein activity can be achieved without affecting its other biological functions, for instance as a component of a multi-protein complex. Recent applications in the field of protein kinases have demonstrated the exceptional value of this technique for elucidating the cellular role of proteins and validating drug targets.[129,130]

3. *Target selection*

While biology plays a key and legitimate role in the selection of new targets, chemistry must have a strong say too, for pursuing a non-druggable target is a waste of time and resources. If the ultimate goal is the development of a potent, selective and orally active drug, then a binding pocket with suitable properties must be present on the protein target. Potency and selectivity are usually achieved by optimizing the fit of the ligand to its receptor site, while oral bioavailability requires certain criteria to be met, such as Lipinski's rule-of-five.[131] Hence, a druggable target may be defined as a protein with a binding site of suitable size (that can accommodate compounds of MW < 500 Da), of appropriate lipophilicity (that provides hydrophobic surfaces suitable for binding compounds with log $P < 5$) and H-bonding potential (that offers less than 5 H-bonds acceptor sites and less than 10 H-bond donors).[119] Structural information on a potential target and related ones constitutes an important part of the electronic biology resources that can be mined to support the assessment of the druggability.[132] The prospects for achieving selectivity within a target family may be gauged through an analysis of the conservation of the binding site, possibly aided by homology modeling.[133] The presence of allosteric binding sites[134] and the existence of distinct structural conformations[113] can greatly increase the odds on finding a drug.

Most receptors and enzymes possess beautiful binding sites and, for this reason, are likely to remain a rich source of druggable targets.[118,119] In contrast, many protein–protein interactions do not fulfill the above requirements and therefore represent hopeless drug targets. Because of their relevance to many diseases, however, protein–protein interactions are still attracting considerable interest and a few may ultimately turn out to be druggable.[135]

B. Hit/lead generation

1. *Structure-based de novo drug design*

The *de novo* design of novel scaffolds displaying the appropriate steric and chemical complementarity to a receptor site, while, at the same time, fulfilling other criteria such as ease of synthesis and derivatization, intellectual property and drug-likeness, requires a great deal of chemical ingenuity, creativity and expertise, as well as deep insights into the critical interaction patterns underlying molecular recognition by the targeted binding site. Nonetheless, some remarkable examples of *de novo* design have been reported in the literature, and a few selected examples are presented in the Table 30.1.

Usually, the structural insights and inspiration required for *de novo* design emerge from careful scrutiny of multiple X-ray structures of the target in complex with a variety of ligands or tool compounds. In addition, many molecular modeling packages provide sophisticated tools for mapping the spatial and physicochemical properties of a given binding site, and for assisting interactive design; medicinal chemists lured into *de novo* design may thus want to team up with a molecular modeler, to benefit from these tools

TABLE 30.1 Some Selected Examples of D*e novo* Structure-based Design

Target	Description	References
Human thymidylate synthase (TS)	A 5-substituted 1H-benzo[cd]indol-2-one scaffold was designed *de novo* based on a key interaction pattern identified from previous X-ray analyses, and on a GRID analysis of the TS crystal structure. X-ray analysis of the complex with a newly designed, low micromolar derivative revealed unmet interactions, and suggested the replacement of a lactam carbonyl by an amidine functionality, ultimately leading to a low nanomolar TS inhibitor selected for clinical trials.	136
HIV-1 protease (HIV-1 PR)	Non-peptide, C2-symmetric cyclic urea derivatives were designed *de novo* to fit the HIV-1 PR active site. These novel, conformationally-constrained inhibitors feature a diol functionality that interacts with the catalytic aspartates while an urea oxygen mimics a structural water mediating substrate interactions with the enzyme flap. The required absolute configuration at four chiral centers was correctly predicted. Potent, bioavailable clinical candidates emerged from this work.	137
Cdk2	A 5-aryl-1H-pyrazole scaffold was designed *de novo* to fit the ATP pocket of Cdk2. Using this scaffold as a query, a sub-structure search of a corporate compound collection retrieved several candidate molecules. The best compound had an IC_{50} of 1.6 μM. X-ray analysis of the Cdk2 complex revealed a dual mode of binding, of which one exhibited the designed interaction pattern.	138
HIF-1α prolyl hydroxylase (EGLN-1)	The X-ray analysis of EGLN-1 in complex with an isoquinoline derivative inspired the design of a series of pyridine carboxyamide derivatives which could then be further improved by replacing the flexible carboxyamide substituent with a constrained isostere, pyrazole carboxylic acid.	139
HIV-1 reverse transcriptase (HIV-1 RT)	Using crystal structures of wild-type (WT) HIV-1 RT and of three drug resistant variants, several fragments were identified with the program LUDI that were predicted to interact with all four enzyme species. A new scaffold was then constructed by linking two fragments and derivatives were synthesized. One compound was discovered which exhibited an IC_{50} of 3.5 μM against WT HIV-1RT and was more potent against the mutant enzymes than the drug nevirapine.	140

and expertise. However, modeling has its limitations and it is usually not wise to embark in a lengthy synthesis to test a new concept. Rather, it is preferable to retrieve simple derivatives or analogs of a designed template from commercial or corporate compound collections and put them to the test as early as possible. But even the most promising templates are likely to display exceedingly weak potencies during the early stages of the *de novo* design process, so that standard biochemical assays may not be appropriate to evaluate these rudimentary compounds. Protein crystallography may be of great help in such cases, since it can detect high micromolar or even low millimolar binders, provided the solubility of these compounds under the crystallization conditions falls in the same range. Once the design hypothesis is put on a firm structural basis, the subsequent optimization steps are usually more straightforward. To increase the odds on getting a first co-crystal structure, it is best to try as many analogs as possible in the crystallization. A bolder approach may consist in the generation of a small combinatorial library designed around the selected template, with substitution sites and functional groups selected on the basis of structural considerations. This is where *de novo* design meets the FBS approach using target-based compound libraries, described below.

2. *In silico screening*

Structure-based virtual screening approaches (or "high-throughput docking", HTD) are coming of age.[141,142] These computational methods use algorithms to predict possible ligand binding modes and estimate the corresponding binding affinities. Although a general and definitive solution to the current stumbling stones of these technologies – receptor plasticity and reliability of scoring functions – is not in sight, their high-throughput and relatively low cost, combined with their versatility, outweigh these deficiencies. Most importantly, several recent success stories demonstrate that these methods do indeed deliver useful hits[143,144] (Table 30.2). As in the case with other knowledge-based approaches, however, the effectiveness of HTD is usually commensurate to the amount and quality of the structural information that is available for the drug target.[144] Particularly important is a profound understanding of the relevant conformational states and possible induced fit mechanisms of the receptor binding site. Multiple co-crystal structures of the target of interest with different chemotypes, as well as any X-ray structures of related targets, contribute to this understanding and are therefore of great value. Furthermore, it is essential that the most critical interaction sites or binding site "hot spots" are identified.[142] Examples of such key interaction sites include the hinge region of protein kinases, the flap and catalytic aspartates of aspartic proteinases, etc. The information on key interaction patterns, which is often used to eliminate false positives by visual inspection of HTD poses, can also be included in the calculations as pre- or post-processing filters, or as part of tailor-made scoring functions.[145] Sometimes, one or more conserved water molecules have been found to play an important role in ligand recognition and binding.

TABLE 30.2 Some Selected Examples of Successful Structure-based Virtual Screening

Target	Description	References
Mycobacterium tuberculosis dihydrodipicolinate reductase	Non-overlapping chemotypes with low micromolar affinity were found by virtual screening and HTS, showing the complementarity of these two hit finding strategies.	147
AmpC β-lactamase	A competitive, non-covalent inhibitor of AmpC β-lactamase was identified by virtual screening of 200,000 compounds from the ACD, and the predicted binding mode was confirmed by X-ray analysis.	148
Anthrax adenylyl cyclase	Docking of 200,000 compounds from the ACD identified 24 high scoring molecules whose specificity was carefully assessed; one compound was identified which inhibited the adenylyl cyclase from the anthrax pathogen (K_i = 20 μM) without inhibiting the host cyclases.	149
tRNA-guanine transglycosylase	X-ray analysis of a ligand complex led to the discovery of a new, unexpected binding mode involving a main-chain peptide flip, and a water molecule. This observation was then exploited to define a new pharmacophore for virtual screening, leading to the identification of novel sub-micromolar inhibitors.	150
Human casein kinase II	Virtual screening of 400,000 compounds was performed using a homology model of casein kinase II. Out of 12 compounds selected for biochemical testing, a novel, highly potent (IC_{50} = 80 nM), and selective inhibitor was discovered.	151
TGFβ receptor kinase	Alternative scaffolds to a known inhibitor were searched by virtual screening using a shape and pharmacophore-based query. A novel, potent (27 nM) inhibitor was found and its binding mode confirmed by X-ray analysis.	152
Aldose reductase	Hierarchical filters were used to impose the requirement for a carboxylate group isostere and the compatibility with a 3-D pharmacophore model of interaction hot spots. Novel, low micromolar inhibitors of aldose reductase were found.	153
HSP90	Two representative conformational states of HSP90 observed by X-ray crystallography were used for virtual screening. Post-filtering imposed two key hydrogen-bonded interactions to be fulfilled, included one to a conserved water. Two hits with low micromolar IC_{50} values were identified, and a similarity search was used to swiftly generate initial SAR.	154

Since the incorporation of such waters can strongly influence both the docking and the scoring steps, it is wise to search the available crystallographic data for the presence of conserved waters at critical locations within the receptor-binding site.[142,146] Last but not least, it is also important to be aware of the limitations and potential uncertainties of crystal structures which can affect the accuracy of the model used in the virtual screening experiment. These uncertainties include the protonation state of protein residues, which cannot be determined by crystallographic analysis (except at very high resolution, typically 1.0 Å or better), and the exact orientation of some donor/acceptor groups, such as imidazole side-chains and the side-chain amide groups of asparagine and glutamine residues. Moreover, some important protein loops lining a binding site may not have well-defined electron density due to partial disorder (multiple conformations are present in the crystal), or the observed conformation may be influenced by the crystallization conditions or protein–protein contacts.

While the value of experimental 3D structures cannot be overemphasized, homology models may sometimes be of sufficient quality for successful virtual screening in particular for structurally well-characterized target families such as protein kinases (*cf.* examples in Table 30.2).

3. *Fragment-based screening*

To some extent, FBS can be viewed as the experimental solution to the *de novo* design problem: suitable scaffolds that can serve as anchoring molecular modules for subsequent structure-based elaboration are sought experimentally rather than conceived on the basis of structural and chemical inferences, or designed with assistance from modeling software. FBS emerged out of a need to overcome the current shortcomings of existing experimental or computational hit finding approaches. The rationale behind the FBS strategy is well known: because the likelihood of a compound fitting a binding site decreases exponentially as molecular complexity increases,[155] HTS approaches often fail to deliver hits, or provide hits which are difficult to optimize owing to their low ligand efficiency[156,157] and "drug-like" rather than "lead-like" properties,[158] that is high-molecular weight, high log *P*, etc. In contrast, the aim of all FBS approaches is to identify hits with high ligand efficiency among a library of carefully selected, chemically appealing, rule-of-three[159] compliant heterocyclic scaffolds. Because of their small size and of the fact that the entropic penalty associated with the loss of rigid-body translational and rotational freedom upon complex formation is independent of molecular weight,[160,161] small

fragments cannot be expected to display high potency, even when their ligand efficiency is high. Consequently, highly sensitive and, preferably, very robust experimental techniques are needed to detect these weak binders. Historically, NMR has played a pioneering role in the development of FBS,[162] but other technologies are now commonly applied, notably mass spectroscopy, surface plasmon resonance, and, last but not least, protein crystallography.[163] We hope that the reader will not take it amiss if we concentrate below on protein crystallographic applications to FBS. More general information on FBS can be found elsewhere.[164]

a. Validation and optimization of FBS hits

A fundamental aspect of FBS is the requirement for detailed structural information on the binding of the hits before chemistry resources can be allotted to hit optimization activities. This is essential for the following reasons: (i) the biophysical techniques used in FBS detect binding events, as opposed to biochemical or biological activity, and have thus the potential to identify secondary or allosteric binding sites (ii) even when competition experiments are used to ensure binding to a known site, the modeling or docking of small fragments does not usually provide unambiguous results hence the need for additional structural information (iii) the potency of initial FBS hits is typically in the 1 mM–100 μM range. When one has access to a large compound store, it is often possible to dig out analogues with improved potency (say, in the 100–10 μM range) through straightforward substructure or similarity searches. This approach has the further advantage of providing, swiftly and with minimal effort, preliminary structure–activity data (for this reason, it is often referred to as "SAR by inventory"). Nevertheless, in the absence of more detailed structural information, the chemical elaboration of weak to very weak FBS hits into potent leads can be a lengthy and cumbersome process.

A variety of elegant NMR techniques have been devised for inferring structural details on protein–ligand interactions.[115] However, whenever possible, protein crystallography remains the preferred approach for elucidating binding modes with certainty and guiding the hit to lead phase. Unfortunately, experience shows that only a fraction of FBS hits discovered by NMR or other biophysical techniques can be observed by protein crystallography. Possible causes for the high attrition rate of X-ray experiments with low affinity compounds have been discussed in Section IV. D. 3.

b. FBS by X-ray crystallography

When the technical pre-requisites for its application in the "high-throughput" mode are fulfilled, one may consider using protein crystallography as the main FBS technique. Since crystallographic information is usually deemed essential for the subsequent hit optimization phase, using X-ray analysis right away for hit detection can save time and certainly avoids the frustration of finding hits that cannot be observed later by crystallography. Moreover, owing to the particular experimental conditions of the X-ray experiment or due to crystal packing effects, some hits are observed by protein crystallography that are sometimes missed by other technologies.

Crystal requirements

Before an FBS by X-ray campaign can be launched, an initial investment in the preparation of suitable crystals may be needed.[165] Crystals originally used for the first structure determination of a new drug target may not be suited for FBS. They may be difficult to grow, may not diffract well enough, or the binding site of interest may already be occupied by a strong ligand. For FBS by X-ray, it is essential that the crystals diffract to high resolution (better than 2.5 Å, more preferably 2.0 Å or better) and are amenable to soaking, which implies that the targeted binding site is free and accessible. In cases where the crystallization is particularly robust, co-crystallization with the fragment cocktails can be attempted, but this strategy is usually less effective than the soaking approach. High crystal symmetry is not a must, but makes data collection faster. When suitable crystals are not available, it may be necessary to engineer and produce new protein variants.

Library design

Although we leave the responsibility of the library design to the medicinal chemist, we would like to underscore the requirement for selecting fragments with high solubility under the crystallization conditions. The protein concentration in a crystallization experiment is typically in the 0.1–1.0 mM range. Accordingly, the fragments should be well soluble up to a concentration of 1.0–10 mM. This is particularly critical when crystals are grown under high salt conditions. Apolar and aromatic scaffolds should feature one or more solubilizing groups, such as a carboxylic, an ammonium functionality, etc. The risk that electrostatic interactions become the main driving force for binding is largely alleviated under high salt conditions, which strengthen hydrophobic interactions at the expense of the electrostatic ones.

With the current technology, a library of 500–1,000 fragments split up into cocktails of 5–10 compounds can be screened by X-ray crystallography within reasonable timelines. The cocktails should be designed in such a way that each component of a mixture has a distinct shape to allow unambiguous identification of any bound fragment on the basis of the electron density.

Examples

Over the past decade, FBS by X-ray has made notable contributions to the overall success of the FBS approach.

It has provided novel, chemically attractive leads for some notoriously difficult targets, such as β-secretase,[166,167] and these hits could be successfully optimized to highly potent drug candidates, hence fulfilling the initial promise of this approach. Some early as well as more recent success stories of FBS by X-ray can be found in Table 30.3.

4. *Triaging and validation of high-throughput screening* (HTS) *hits*

The application of protein crystallography to the characterization of a few selected HTS screening hits is not new. But the advent of high-throughput crystallography has opened up the possibility of more systematic investigations of HTS results, and X-ray crystallography is now often used by the pharmaceutical industry for the triaging and structural characterization of substantial numbers of HTS hits. To make best use of the available X-ray crystallography resources, however, these should be deployed only after HTS hits have been carefully validated, preferably in an orthogonal assay. HTS hit lists usually comprise many compounds which are immediately discarded by medicinal chemists on the ground of their weak potency, lack of synthetic attractiveness, or potential pharmacological liabilities. It is now possible to analyze many hits by protein crystallography, including some of those which in the past would have been disregarded, and this analysis has the potential to rescue highly valuable information regarding novel binding sites or subsites, alternative binding modes, privileged interaction patterns, protein conformational substates, and so on. This information can then be fed into the structure-based design process, even when some of these hits are not pursued any further. Moreover, among the weak hits many compounds are often "fragment-like" with molecular weights in the 150–250 Da range. It may be of particular interest to investigate this region of chemical space where HTS meets FBS.

C. Lead optimization

Lead optimization can have many different aims, depending on the progress of a particular project. It may be necessary to improve *in vitro* potency further, or *in vitro* potency may be fine, but cellular activity may still be insufficient. Alternatively, fine-tuning of the physicochemical properties of a lead compound may be required to achieve the desired pharmacological profile while maintaining the current level of potency. Toxicity issues or metabolic weak points may need to be addressed. Yet another objective may be to improve the selectivity profile, or, *a contrario*, to gain broad-spectrum activity against a variety of resistance mutants. Good synthetic accessibility, acceptable solubility and strong intellectual property position are additional criteria that must be met by an optimized drug candidate.

Protein crystallography reveals critical details about the topological and physicochemical characteristics of the target binding site, the biologically active conformation of the lead compound and the intermolecular interactions within the complex. The value of this information for guiding the lead optimization process goes without saying. Nevertheless, structure-based design can be a humbling exercise, particularly when some important contributions to the overall binding energy are not understood or are simply overlooked.

1. *Ligand-binding: getting the full picture*

One has to bear in mind that protein crystallography unveils only one side of the equilibrium which gives rise to

TABLE 30.3 Some Selected Examples of Fragment-based Screening Using X-ray Crystallography

Target	Description	References
Human urokinase	From a biased fragment library five novel binders were identified. By combining the best fragment (K_i = 56 μM) with a known inhibitor, a potent urokinase inhibitor (K_i = 370 nM) with improved oral bioavailability was produced.	165
p38α MAP kinase	Two structurally distinct, novel, and potent lead series were derived from structure-guided optimization of two small fragments with IC_{50} values of 1.3 mM and 33 μM, respectively. Fragment growing as well as conjoining of overlapping fragments was used to improve potency, and the mobility of the DFG loop was exploited to achieve selectivity.	168
Cyclic nucleotide phosphodiesterase 4 (PDE4)	A combination of low affinity screening and high-throughput X-ray analysis led to the identification of a pyrazole carboxylic ester scaffold. The design and *in silico* evaluation of small libraries exploring three substitution sites led to a 4,000-fold improvement in potency after only two rounds of synthesis.	169
β-secretase	Using FBS by X-ray, 2-aminoquinoline was identified as a very weak (2 mM), but novel recognition motif for an aspartic protease. A combination of database mining, virtual screening, and X-ray analysis then led to the discovery of a 6-substituted 2-aminopyridine compound with improved affinity which was further optimized to sub-micromolar potency using structure-based design.	166, 167

the formation of the protein–ligand complex from the free partners in solution. The enthalpic and entropic contributions resulting from conformational transitions affecting the protein and the ligand, solvation and desolvation effects, changes in protonation states, and the loss of conformational freedom upon complex formation can all have a dramatic influence on the overall free energy of binding,[170] hence on the observed IC_{50} or K_i values. Furthermore, at the resolution at which most protein X-ray studies are performed, protonation states remain uncertain, and polarization effects and conformational strain cannot be detected or are masked by the application of geometric restraints. Therefore, crystallographic data usually need to be complemented by other experimental or computational investigations to gain a deeper understanding of the key binding determinants. Structure–activity relationships data constitute one main source of information, but enzyme kinetics, isothermal titration calorimetry, conformational analysis, and FBS can all reveal important aspects of what drives ligand binding in a particular system.

2. *Optimizing potency*

Weak compounds which are already big and thus have a very poor ligand efficiency are usually not the best starting points for further chemical elaboration as their optimization will most likely be difficult and time consuming. In many cases, fragments with high ligand efficiency are much better starting points.

The identification of the key interaction sites (or "hot spots") within the binding pocket is a first and essential step when an enhancement in potency is sought.[171] To this end, an experimental, fragment-based approach can be used[172] or computational methods can be utilized.[49,173,174] Interactions with the binding site hot spots should be maximized through the introduction of new substituents or the replacement of functionalities making suboptimum contacts. This process can be expedited by employing a structure-based, focused library approach (e.g. see nilotinib, Section III. B. or PDE4, Table 30.3). Alternatively, software for fragment growing can be tried. With this latter method, low picomolar inhibitors of carbonic anhydrase II have been found.[175]

Protein ligands rarely bind in their lowest energy conformation.[176] When present, unfavorable strain energy should be detected and minimized. An analysis of the conformation of related compounds in the Cambridge data bank can guide this process,[177] and *ab initio* calculations are often useful.[177,178] Compounds requiring minimal conformational reorganization on enzyme binding should be favored. Small molecule ligands frequently adopt an extended conformation in the bound state.[176] Hence, hydrophobic ligands exhibiting a folded conformation in solution may incur a high reorganization energy cost on binding. Introduction of conformational restraints through (macro) cyclization[179] or the introduction of rigid linkers[180] is another strategy that has been successfully used in many cases to minimize entropic penalties. However, retaining a certain degree of flexibility may be important to circumvent the issue of resistance mutations (see below).

3. *Optimizing (pan)-selectivity*

Exquisite selectivity can often be achieved by exploiting binding subsites or pockets adjacent to the main binding site, but which are not involved in the normal biological function of the drug target and are thus poorly conserved in other family members.[68,81] Likewise, taking advantage of the flexibility of the protein by targeting an unusual conformational state is an excellent means to achieve high selectivity. Protein kinase inhibitors provide numerous examples of this kind.[76,77]

One disadvantage of the above approaches is that resistance mutations are more likely to emerge when non-functional states or cavities are used by the drug.[81] Several strategies have been proposed to achieve broad-spectrum activity against a wide range of resistance variants. Optimizing potency as much as possible against the wild type target often limits the emergence of resistance while ensuring that sufficient residual affinity remains for many variants.[181,182] Designing more flexible compounds has also been proposed as a possible strategy for achieving broad-spectrum activity,[183] since rigid molecules selected on the basis of their shape complementarity to the binding site are less likely to adapt to structural changes. The entropic cost of the built-in flexibility needs to be compensated by a larger enthalpic contribution to binding, through an optimization of all available polar interactions. Hence, enthalpic optimization of the binding affinity has been proposed as a better alternative to potency enhancement through hydrophobic binding and rigid fit.[184] In doing so, however, it is important to ensure that the strongest interactions involve residues with a low probability to mutate.[181]

4. *Optimizing* ADME *properties*

A general recipe for turning a potent lead into a real drug candidate does not exist, but some guidelines are available, such as the well-known "rule-of-five".[131] Structure-guided design can aid achieving the right balance between lipophilicity and polar surface area by guiding the introduction or replacement of heteroatoms, polar groups and other solubilizing groups. Essential hydrogen-bonded interactions can be identified as well as dispensable acceptor/donor groups. Minimizing molecular weight is frequently an effective strategy to achieving good oral bioavailability.[177,185] Groups and substituents with low ligand efficiency should be identified, replaced or removed. As already pointed out above, it is often easier to start with a small fragment

TABLE 30.4 Some Selected Examples of Structure-based Lead Optimization

Target	Description	References
Human carbonic anhydrase II	X-ray analyses were combined with in-depth conformational analyses using *ab initio* calculations to clarify the origin of the large variations in potency observed with different optical isomers of thienothiopyran-2-sulfonamide derivatives. Modifications leading to enhanced affinity via the reduction of conformational strain were identified.	43, 178
Human elastase	Good oral bioavailability was achieved by minimizing compound size through iterative structure-based optimization of a peptide lead.	185
Human renin	X-ray analysis of a weak, non-peptide, piperidine-based HTS hit (26 μM) revealed a radically novel mode of binding associated with a large conformational change of the enzyme that had never been observed before. Structure-guided optimization led to a low nanomolar, drug-like inhibitor.	189
Human thrombin	A lipophilic chlorophenyl fragment was identified by fragment-based screening by X-ray as a suitable substitute for the S1 benzamidine group of conventional thrombin inhibitors.	190
HIV-1 protease	Structure-based design aimed at reducing inhibitor size while maintaining sub-nanomolar potency led to the discovery of Amprenavir, a marketed anti-HIV-1 protease drug.	177
HIV reverse transcriptase (HIV RT)	A new generation of non-nucleoside inhibitors against wild-type RT and drug-resistant RT mutants was rationally designed; several compounds were found to outperform existing RT drugs.	182
Src-SH2	A suitable replacement for the phosphotyrosine group of Src-SH2 antagonists was identified using fragment-based screening by crystallography.	191

exhibiting a good ligand efficiency, which can then be grown while keeping the ligand efficiency constant.

Suitable replacements for functionalities which are detrimental to ADME properties can be sought using focused chemical libraries or a FBS approach (e.g. see thrombin and Src-SH2, Table 30.4).

Ultimately, switching to a different chemotype may be the only solution for achieving the desired pharmacological profile. To this end, the determination of multiple X-ray structures with a diverse range of molecular scaffolds or fragments may provide some useful clues. For instance, structural overlays may suggest combining two overlapping fragments into one novel scaffold.[186] Alternatively, the possibility to graft one particular motif from one inhibitor onto another, more attractive scaffold may become apparent.[187,188]

REFERENCES

1. Shakespeare, W. Macbeth. Paris, 1977, Editions Aubier Montaigne.
2. Congreve, M., Murray, C. W., Blundell, T. L. Structural biology and drug discovery. *Drug Discov. Today,* 2005, *10*, 895–907.
3. Scapin, G. Structural biology and drug discovery. *Curr. Pharm. Design,* 2006, *12*, 2087–2097.
4. Tintelnot-Blomley, M., Lewis, R. A. A critical appraisal of structure-based drug design. *IDrugs,* 2006, *9*, 114–118.
5. Lombardino, J. G., Lowe, J. A. A guide to drug discovery: the role of the medicinal chemist in drug discovery – then and now. *Nat. Rev. Drug Discov.* 2004, *3*, 853–862.
6. Williams, S. P., Kuyper, L. F., Pearce, K. H. Recent applications of protein crystallography and structure-guided drug design. *Curr. Opin. Chem. Biol.* 2005, *9*, 371–380.
7. Blundell, T. L., Jhoti, H., Abell, C. High-throughput crystallography for lead discovery in drug design. *Nat. Rev. Drug Discov.* 2002, *1*, 45–54.
8. Hajduk, P. J., Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat. Rev. Drug Discov.* 2006, *6*, 211–219.
9. Werten, P. J. L., Remigy, H. W., de Groot, B. L., Fotiadis, D., Philippsen, A., Stahlberg, H., Grubmuller, H., Engel, A. Progress in the analysis of membrane protein structure and function. *FEBS Lett.* 2002, *529*, 65–72.
10. Lundström, K. Structural genomics for membrane proteins. *Cell. Mol. Life Sci.* 2006, *63*, 2597–2607.
11. Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H., Phillips, D. C. 3-Dimensional model of the myoglobin molecule obtained by X-ray analysis. *Nature,* 1958, *181*, 662–666.
12. Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G., North, A. C. T. Structure of haemoglobin-3-dimensional fourier synthesis at 5.5-A resolution, obtained by X-ray analysis. *Nature,* 1960, *185*, 416–422.
13. McPherson, A. A brief history of protein crystal growth. *J. Cryst. Growth,* 1991, *110*, 1–10.
14. Franklin, R. E., Gosling, R. G. Evidence for 2-chain helix in crystalline structure of sodium deoxyribonucleate. *Nature,* 1953, *172*, 156–157.
15. Wilkins, M. H. F., Seeds, W. E., Stokes, A. R., Wilson, H. R. Helical structure of crystalline deoxypentose nucleic acid. *Nature,* 1953, *172*, 759–762.
16. Watson, J. D., Crick, F. H. C. Molecular structure of nucleic acids – a structure for deoxyribose nucleic acid. *Nature,* 1953, *171*, 737–738.
17. Perutz, M. *Protein Structure. New Approaches to Disease and Therapy*. W.H. Freeman and Company: New York, 1992.
18. Rossmann, M. G., Arnold, E., Erickson, J. W., Frankenberger, E. A., Griffith, J. P., Hecht, H. J., Johnson, J. E., Kamer, G., Luo, M., Mosser, A. G., Rueckert, R. R., Sherry, B., Vriend, G. Structure of a human common cold virus and functional relationship to other picornaviruses. *Nature*, 1985, *317*, 145–153.

19. Deisenhofer, J., Epp, O., Miki, K., Huber, R., Michel, H. Structure of the protein subunits in the photosynthetic reaction center of *Rhodopseudomonas viridis* at 3 Å resolution. *Nature,* 1985, *318*, 618–624.
20. Abrahams, J. P., Leslie, A. G., Lutter, R., Walker, J. E. Structure at 2.8 Å resolution of F1-ATPase from bovine heart mitochondria. *Nature,* 1994, *370*, 621–628.
21. Löwe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W., Huber, R. Crystal structure of the 20S proteasome from the archaeon *T. acidophilum* at 3.4 Å resolution. *Science,* 1995, *268*, 533–539.
22. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å. *Nature,* 1997, *389*, 251–260.
23. Brodersen, D. E., Clemons, W. M., Jr., Carter, A. P., Morgan-Warren, R. J., Wimberly, B. T., Ramakrishnan, V. The structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B on the 30S ribosomal subunit. *Cell,* 2000, *103*, 1143–1154.
24. Ban, N., Nissen, P., Hansen, J., Moore, P. B., Steitz, T. A. The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science,* 2000, *289*, 905–920.
25. Nissen, P., Hansen, J., Ban, N., Moore, P. B., Steitz, T. A. The structural basis of ribosome activity in peptide bond synthesis. *Science,* 2000, *289*, 920–930.
26. Pioletti, M., Schlunzen, F., Harms, J., Zarivach, R., Gluhmann, M., Avila, H., Bashan, A., Bartels, H., Auerbach, T., Jacobi, C., Hartsch, T., Yonath, A., Franceschi, F. Crystal structures of complexes of the small ribosomal subunit with tetracycline, edeine and IF3. *EMBO J.* 2001, *20*, 1829–1839.
27. Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A., Kornberg, R. D. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science,* 2001, *292*, 1876–1882.
28. Doyle, D. A., Morais Cabral, J., Pfuetzner, R. A., Quo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T., Mackinnon, R. The structure of the potassium channel: molecular basis of K^+ conduction and selectivity. *Science,* 1998, *280*, 69–77.
29. Jiang, Y., Lee, A., Chen, J., Cadene, M., Chait, B. T., MacKinnon, R. Crystal structure and mechanism of a calcium-gated potassium channel. *Nature,* 2002, *417*, 515–522.
30. Sharff, A., Jhoti, H. High-throughput crystallography to enhance drug discovery. *Curr. Opin. Chem. Biol.* 2003, *7*, 340–345.
31. Kroemer, M., Dreyer, M. K., Wendt, K. U. APRV – a program for automated data processing, refinement and visualization. *Acta Cryst. D* 2004, *60*, 1679–1682.
32. Mooij, W. T. M., Hartshorn, M. J., Tickle, I. J., Sharff, A. J., Verdonk, M. L., Jhoti, H. Automated protein-ligand crystallography for structure-based drug design. *Chem. Med. Chem.* 2006, *1*, 827–838.
33. Evrard, G. X., Langer, G. G., Perrakis, A., Lamzin, V. S. Assessment of automatic ligand building in ARP/Warp. *Acta Cryst. D,* 2007, *63*, 108–117.
34. Schmidt, A., Lamzin, V. S. Veni, vidi, vici – atomic resolution unraveling the mysteries of protein function. *Curr. Opin. Struct. Biol.* 2002, *12*, 698–703.
35. Petrova, T., Podjarny, A. Protein crystallography at subatomic resolution. *Rep. Prog. Phys.* 2004, *67*, 1565–1605.
36. Howard, E. I., Sanishvili, R., Cachau, R. E., Mitschler, A., Chevrier, B., Barth, P., Lamour, V., Van Zandt, M., Sibley, E., Bon, C., Moras, D., Schneider, T. R., Joachimiak, A., Podjarny, A. Ultrahigh resolution drug design I: details of interactions in human aldose reductase-inhibitor complex at 0.66 Å. *Proteins Struct. Funct. Bioinf.* 2004, *55*, 792–804.
37. El-Kabbani, O., Darmanin, C., Schneider, T. R., Hazemann, I., Ruiz, F., Oka, M., Joachimiak, A., Schulze-Briese, C., Tomizaki, T., Mitschler, A., Podjarny, A. Ultrahigh resolution drug design II: atomic resolution structures of human aldose reductase holoenzyme complexed with Fidarestat and Minalrestat: implications for the binding of cyclic imide inhibitors. *Proteins Struct. Funct. Bioinf.* 2004, *55*, 805–813.
38. Beddell, C. R., Goodford, P. J., Norrington, F. E., Wilkinson, S., Wootton, R. Compounds designed to fit a site of known structure in human haemoglobin. *Br. J. Pharmacol.* 1976, *57*, 201–209.
39. Ondetti, M. A., Cushman, D. W. Design of specific inhibitors of angiotensin-converting enzyme – new class of orally active antihypertensive agents. *Science*, 1977, *196*, 441–444.
40. Natesh, R., Schwager, S. L. U., Evans, H. R., Sturrock, E. D., Acharya, K. R. Structural details on the binding of antihypertensive drugs captopril and enalaprilat to human testicular angiotensin I-converting enzyme. *Biochemistry,* 2004, *43*, 8718–8724.
41. Corradi, H. R., Chitapi, I., Sewell, B. T., Georgiadis, D., Dive, V., Sturrock, E. D., Acharya, K. R. The structure of testis angiotensin-converting enzyme in complex with the C domain-specific inhibitor RXPA380. *Biochemistry,* 2007, *46*, 5473–5478.
42. Eriksson, A. E., Jones, T. A., Liljas, A. Refined structure of human carbonic anhydrase II at 2.0 Å resolution. *Proteins,* 1988, *4*, 274–282.
43. Baldwin, J., Ponticello, G. S., Anderson, P. S., Christy, M. E., Murcko, M. A., Randall, W. C., Schwam, H., Sugrue, M. F., Springer, J. P., Gautheron, P., Grove, J., Mallorga, P., Viader, M.-P., McKeever, B. M., Navia, M. A. Thienothiopyran-2-sulfonamides: novel topically active carbonic anhydrase inhibitors for the treatment of glaucoma. *J. Med. Chem.* 1989, *32*, 2510–2513.
44. Pastorekova, S., Parkkila, S., Pastorek, J., Supuran, C. T. Carbonic anhydrases: current state of the art, therapeutic applications and future prospects. *J. Enzyme Inhib. Med. Chem.* 2004, *19*, 199–229.
45. Briganti, F., Mangani, S., Orioli, P., Scozzafava, A., Vernaglione, G., Supuran, C. T. Carbonic anhydrase activators: X-ray crystallographic and spectroscopic investigations for the interaction of isozymes I and II with histamine. *Biochemistry,* 1997, *36*, 10384–10392.
46. Wade, R. C. 'Flu' and structure-based drug design. *Structure,* 1997, *5*, 1139–1145.
47. Varghese, J. N., Laver, W. G., Colman, P. M. Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution. *Nature,* 1983, *303*, 35–40.
48. Colman, P. M., Varghese, J. N., Laver, W. G. Structure of the catalytic and antigenic sites in influenza virus neuraminidase. *Nature,* 1983, *303*, 41–44.
49. Goodford, P. J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* 1985, *28*, 849–857.
50. von Itzstein, M., Wu, W.-Y., Kok, G. B., Pegg, M. S., Dyason, J. C., Jin, B., Van Phan, T., Smythe, M. L., White, H. F., Oliver, S. W., Colman, P. M., Varghese, J. N., Ryan, D. M., Woods, J. M., Bethell, R. C., Hotham, V. J., Cameron, J. M., Penn, C. R. Rational design of potent sialidase-based inhibitors of influenza virus replication. *Nature,* 1993, *363*, 418–423.
51. Kim, C. U., Lew, W., Williams, M. A., Liu, H., Zhang, L., Swaminathan, S., Bishofberger, N., Chen, M. S., Mendel, D. B., Tai, C. Y., Laver, W. G., Stevens, R. C. Influenza neuraminidase inhibitors possessing a novel hydrophobic interaction in the enzyme active site: design, synthesis, and structural analysis of carbocyclic sialic acid analogues with potent anti-influenza activity. *J. Am. Chem. Soc.* 1997, *119*, 681–690.
52. Matter, H., Schudok, M. Recent advances in the design of matrix metalloprotease inhibitors. *Curr. Opin. Drug Discov. Dev.* 2004, *7*, 513–535.
53. Stout, T. J., Foster, P. G., Matthews, D. J. High-throughput structural biology in drug discovery: protein kinases. *Curr. Pharm. Design,* 2004, *10*, 1069–1082.
54. Johnson, S. L., Pellecchia, M. Structure- and fragment-based approaches to protease inhibition. *Curr. Top. Med. Chem.* 2006, *6*, 317–329.
55. Knapp, M., Bellamacina, C., Murray, J. M., Bussiere, D. E. Targeting cancer: the challenges and successes of structure-based drug design against the human purinome. *Curr. Top. Med. Chem.* 2006, *6*, 1129–1159.
56. Thomas, M. P., McInnes, C. Structure-based discovery and optimization of potential cancer therapeutics targeting the cell cycle. *IDrugs,* 2006, *9*, 273–278.

57. Combs, A. P. Structure-based drug design of new leads for phosphatase research. *IDrugs,* 2006, *10*, 112–115.
58. Wlodawer, A. Rational approach to AIDS drug design through structural biology. *Annu. Rev. Med.* 2002, *53*, 595–614.
59. Hardy, L. W., Malikayil, A. The impact of structure-guided drug design on clinical agents. *Curr. Drug Discov.* 2003(Dec), 15–20.
60. Eder, J., Hommel, U., Cumin, F., Martoglio, B., Gerhartz, B. Aspartic proteases in drug discovery. *Curr. Pharm. Design,* 2007, *13*, 271–285.
61. Ondetti, M. A., Cushman, D. W. Enzymes of the renin-angiotensin system and their inhibitors. *Annu. Rev. Biochem.* 1982, *51*, 283–308.
62. Blundell, T. L., Cooper, J., Foundling, S. I., Jones, D. M., Atrash, B., Szelke, M. On the rational design of renin inhibitors: X-ray studies of aspartic proteinases complexed with transition-state analogues. *Biochemistry,* 1987, *26*, 5585–5590.
63. Greenlee, W. J. Renin inhibitors. *Med. Res. Rev.* 1990, *10*, 173–236.
64. Lunney, E. A., Hamilton, H. W., Hodges, J. C., Kaltenbronn, J. S., Repine, J. T., Badasso, M., Cooper, J. B., Dealwis, C., Wallace, B. A., Lowther, W. T., Dunn, B. M., Humblet, C. Analyses of ligand binding in five endothiapepsin crystal complexes and their use in the design and evaluation of novel renin inhibitors. *J. Med. Chem.* 1993, *36*, 3809–3820.
65. Sielecki, A. R., Hayakawa, K., Fujinaga, M., Murphy, M. E. P., Frazer, M., Muir, A. K., Carilli, C. T., Lewicki, J. A., Baxter, J. D., James, M. N. G. Structure of recombinant human renin, a target for cardio-vascular-active drugs, at 2.5 Å resolution. *Science*, 1989, *243*, 1341–1351.
66. Rahuel, J., Priestle, J.P., Grütter, M.G. The crystal structures of recombinant glycosylated human renin alone and in complex with a transition-state analog inhibitor. *J. Struct. Biol.* 1991, *107*, 227–236.
67. Navia, M. A., Chaturvedi, P. R. Design principles for orally bioavailable drugs. *Drug Discov. Today,* 1996, *1*, 179–189.
68. Rahuel, J., Rasetti, V., Maibaum, J., Rüeger, H., Göschke, R., Cohen, N.-C., Stutz, S., Cumin, F., Fuhrer, W., Wood, J. M., Grütter, M. G. Structure-based drug design: the discovery of novel nonpeptide orally active inhibitors of human rennin. *Chem. Biol.* 2000, *7*, 493–504.
69. Wood, J. M., Maibaum, J., Rahuel, J., Grütter, M. G., Cohen, N.-C., Rasetti, V., Rüeger, H., Göschke, R., Stutz, S., Fuhrer, W., Schilling, W., Rigollier, P., Yamaguchi, Y., Cumin, F., Baum, H.-P., Schnell, C. R., Herold, P., Mah, R., Jensen, C., O'Brien, E., Stanton, A., Bedigian, M. P. Structure-based design of aliskiren, a novel orally effective renin inhibitor. *Biochem. Biophys. Res. Comm.* 2003, *308*, 698–705.
70. Göschke, R., Stutz, S., Rasetti, V., Cohen, N.-C., Rahuel, J., Rigollier, P., Baum, H.-P., Forgiarini, P., Schnell, C. R., Wagner, T., Gruetter, M. G., Fuhrer, W., Schilling, W., Cumin, F., Wood, J. M., Maibaum, J. Novel 2,7-dialky1 substituted 5(S)-amino-4(S)-hydroxy-8-phenyl-octanecarboxamide transition state peptidomimetics are potent and orally active inhibitors of human renin. *J. Med. Chem.* 2007, *50*. 4818–4831.
71. Maibaum, J., Stutz, S., Göschke, R., Rigollier, P., Yamaguchi, Y., Cumin, F., Rahuel, J., Baum, H.-P., Cohen, N.-C., Schnell, C. R., Fuhrer, W., Gruetter, M. G., Schilling, W., Wood, J. M. Structural modification of the P2' position of 2,7-dialkyl substituted 5(S)-amino-4(S)-hydroxy-8-phenyl-octanecarboxamides: the discovery of aliskiren, a potent non-peptide human renin inhibitor active after once daily dosing in marmosets. *J. Med. Chem.* 2007, *50*. 4832–4844.
72. Manning, G., White, D. B., Martinez, R., Hunter, T., Sudarsanam, S. The protein kinase complement of the human genome. *Science,* 2002, *298*, 1912–1934.
73. Zimmermann, J., Buchdunger, E., Mett, H., Meyer, T., Lydon, N. B. Potent and selective inhibitors of the Abl-kinase: phenylaminopyrimidine (PAP) derivatives. *Bioorg. Med. Chem. Lett.* 1997, *7*, 187–192.
74. Capdeville, R., Buchdunger, E., Zimmermann, J., Matter, A. Glivec (STI571, imatinib), a rationally developed, targeted anticancer drug. *Nat. Rev. Drug Discov.* 2002, *1*, 493–502.
75. Scapin, G. Structural biology in drug design: selective protein kinase inhibitors. *Drug Discov. Today,* 2002, *7*, 601–611.
76. Noble, M. E. M., Endicott, J. A., Johnson, L. N. Protein kinase inhibitors: insights into drug design from structure. *Science,* 2004, *303*, 1800–1805.
77. Cowan-Jacob, S. W. Structural biology of protein tyrosine kinases. *Cell. Mol. Life Sci.* 2006, *63*, 2608–2625.
78. Schindler, T., Bommann, W., Pellicena, P., Miller, W. T., Clarkson, B., Kuriyan, J. Structural mechanism for STI-571 inhibition of abelson tyrosine kinase. *Science,* 2000, *289*, 1938–1942.
79. Nagar, B., Bommann, W. G., Pellicena, P., Schindler, T., Veach, D. R., Miller, W. T., Clarkson, B., Kuriyan, J. Crystal structures of the kinase domain of c-Abl in complex with the small molecule inhibitors PD173955 and imatinib (STI-571). *Cancer Res.* 2002, *62*, 4236–4243.
80. Manley, P. W., Cowan-Jacob, S. W., Buchdunger, E., Fabbro, D., Fendrich, G., Furet, P., Meyer, T., Zimmermann, J. Imatinib: a selective tyrosine kinase inhibitor. *Eur. J. Cancer,* 2002, *38*(Suppl. 5), S19–S27.
81. Cowan-Jacob, S. W., Fendrich, G., Floersheimer, A., Furet, P., Liebetanz, J., Rummel, G., Rheinberger, P., Centeleghe, M., Fabbro, D., Manley, P. W. Structural biology contributions to the discovery of drugs to treat chronic myelogenous leukaemia. *Acta Cryst. D,* 2007, *63*, 80–93.
82. Cowan-Jacob, S. W., Guez, V., Fendrich, G., Griffin, J. D., Fabbro, D., Furet, P., Liebetanz, J., Mestan, J., Manley, P. W. Imatinib (STI571) resistance in chronic myelogenous leukemia: molecular basis of the underlying mechanisms and potential strategies for treatment. *Mini Rev. Med. Chem.* 2004, *4*, 285–299.
83. Manley, P. W., Breitenstein, W., Bruggen, J., Cowan-Jacob, S. W., Furet, P., Mestan, J., Meyer, T. Urea-derivatives of STI571 as inhibitors of Bcr-Abl and PDGFR kinases. *Bioorg. Med. Chem. Lett.* 2004, *14*, 5793–5797.
84. Manley, P. W., Cowan-Jacob, S. W., Fendrich, G., Metan, J. Molecular interactions between the highly selective pan-Bcr-Abl inhibitor, AMN107, and the tyrosine kinase domain of Abl. *Blood,* 2005, *106*, 940a.
85. Weisberg, E., Manley, P. W., Breitenstein, W., Bruggen, J., Cowan-Jacob, S. W., Ray, A., Huntly, B., Fabbro, D., Fendrich, G., Hall-Meyers, E., Kung, A. L., Mestan, J., Daley, G. Q., Callahan, L., Catley, L., Cavazza, C., Azam, M., Neuberg, D., Wright, R. D., Gilliland, D. G., Griffin, J. D. Characterization of AMN107, a selective inhibitor of native and mutant Bcr-Abl. *Cancer Cell,* 2005, *7*, 129–141.
86. Weisberg, E., Manley, P., Mestan, J., Cowan-Jacob, S., Ray, A., Griffin, J. D. AMN107 (nilotinib): a novel and selective inhibitor of BCR-ABL. *Br. J. Cancer,* 2006, *94*, 1765–1769.
87. Matthews, B. Solvent content of protein crystals. *J. Mol. Biol.* 1968, *33*, 491–497.
88. Ducruix, A., Giege, R. *Crystallization of Nucleic Acids and Proteins: A Practical Approach*, 2nd Edition. Oxford University Press: Oxford, 1999.
89. McPherson, A. *Crystallization of biological macromolecules*. Cold Spring Harbor Laboratory Press: New York, 1999.
90. Kundrot, C. E. Which strategy for a protein crystallization project?. *Cell. Mol. Life Sci.* 2004, *61*, 525–536.
91. Pusey, M. L., Liu, Z.-J., Tempel, W., Praissman, J., Lin, D., Wang, B.-C., Gavira, J. A., Ng, J. D. Life in the fast lane for protein crystallization and X-ray crystallography. *Prog. Biophys. Mol. Biol.* 2005, *88*, 359–386.
92. Derewenda, Z. S. The use of recombinant methods and molecular engineering in protein crystallization. *Methods,* 2004, *34*, 354–363.
93. Chang, V. T., Crispin, M., Aricescu, A. R., Harvey, D. J., Nettleship, J. E., Fennelly, J. A., Yu, C., Boles, K. S., Evans, E. J., Stuart, D. I., Dwek, R. A., Jones, E. Y., Owens, R. J., Davis, S. J. Glycoprotein structural genomics: solving the glycosylation problem. *Structure,* 2007, *15*, 267–273.
94. Wright, P. E., Dyson, H. J. Intrinsically unstructured proteins: reassessing the protein structure-function paradigm. *J. Mol. Biol.* 1999, *293*, 321–331.
95. Wiener, M. C. A pedestrian guide to membrane protein crystallization. *Methods,* 2004, *34*, 364–372.

96. Minor, D. L., Jr. The neurobiologist's guide to structural biology: a primer on why macromolecular structure matters and how to evaluate structural data. *Neuron,* 2007, *54*, 511–533.
97. McPherson, A., Cudney, B. Searching for silver bullets: an alternative strategy for crystallizing macromolecules. *J. Struct. Biol.* 2006, *156*, 387–406.
98. Strauss, A., Fendrich, G., Horisberger, M. A., Liebetanz, J., Meyhack, B., Schlaeppi, J.-M., Schmitz, R. Improved expression of kinases in Baculovirus-infected insect cells upon addition of specific kinase inhibitors to the culture helpful for structural studies. *Protein Expr. Purif.* 2007, *56*, 167–176.
99. Hassell, A. M., An, G., Bledsoe, R. K., Bynum, J. M., CarterIII, H. L., Deng, S.-J. J., Gampe, R. T., Grisard, T. E., Madauss, K. P., Nolte, R. T., Rocque, W. J., Wang, L., Weaver, K. L., Williams, S. P., Wisely, G. B., Xu, R., Shewchuk, L. M. Crystallization of protein-ligand complexes. *Acta Cryst. D*, 2007, *63*, 72–79.
100. Lusty, C. A gentle vapor-diffusion technique for cross-linking of protein crystals for cryocrystallography. *J. Appl. Cryst.* 1999, *32*, 106–112.
101. Garman, E. F., Owen, R. L. Cryocooling and radiation damage in macromolecular crystallography. *Acta Cryst. D,* 2006, *62*, 32–47.
102. Dauter, Z., Dauter, M., Dodson, E. Jolly SAD. *Acta Cryst. D,* 2002, *58*, 494–506.
103. Ealick, S. E. Advances in multiple wavelength anomalous diffraction crystallography. *Curr. Opin. Chem. Biol.* 2000, *4*, 495–499.
104. Dauter, Z., Dauter, M. Entering a new phase: using solvent halide ions in protein structure determination. *Structure,* 2001, *9*, 21–26.
105. Debreczeni, J.É., Bunkóczi, G., Ma, Q., Blaser, H., Sheldrick, G. M. In-house measurement of the sulfur anomalous signal and its use for phasing. *Acta Cryst. D,* 2003, *59*, 688–696.
106. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., Bourne, P. E. The protein data bank. *Nucleic Acids Res.* 2000, *28*, 235–242.
107. Brünger, A. T., Free, R. Value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature,* 1992, *355*, 472–475.
108. Jia, Z., Vandonselaar, M., Quail, J. W., Delbaere, L. T. J. Active-centre torsion-angle strain revealed in 1.6 Å-resolution structure of histidine-containing phosphocarrier protein. *Nature,* 1993, *361*, 94–97.
109. Chevrier, B., Schalk, C., D'Orchymont, H., Rondeau, J.-M., Moras, D., Tarnus, C. Crystal structure of *Aeromonas proteolytica* aminopeptidase: a prototypical member of the co-catalytic zinc enzyme family. *Structure,* 1994, *2*, 283–291.
110. Laskowski, R. A., MacArthur, M. W., Moss, D. S., Thornton, J. M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* 1993, *26*, 283–291.
111. Davis, A. M., Teague, S. J., Kleywegt, G. J. Application and limitations of X-ray crystallographic data in structure-based ligand and drug design. *Angew. Chem. Int. Ed.* 2003, *42*, 2718–2736.
112. DePristo, M. A., de Bakker, P. I. W., Blundell, T. L. Heterogeneity and inaccuracy in protein structures solved by X-ray crystallography. *Structure,* 2004, *12*, 831–838.
113. Teague, S. Implications of protein flexibility for drug discovery. *Nat. Rev. Drug Discov.* 2003, *2*, 527–541.
114. Kleywegt, G. J. Crystallographic refinement of ligand complexes. *Acta Cryst. D* 2007, *63*, 94–100.
115. Jahnke, W., Widmer, H. Protein NMR in biomedical research. *Cell. Mol. Life Sci.* 2004, *61*, 580–599.
116. Chung. C.-W. The use of biophysical methods increases success in obtaining liganded crystal structures. *Acta Cryst.* 2007, *D63*, 62–71.
117. Kleywegt, G. J., Harris, M. R., Zou, J. Y., Taylor, T. C., Wählby, A., Jones, T. A. The Uppsala electron-density server. *Acta Cryst. D* 2004, *60*, 2240–2249.
118. Overington, J. P., Al-Lazikani, B., Hopkins, A. L. How many drug targets are there?. *Nat. Rev. Drug Discov.* 2006, *5*, 992–996.
119. Hopkins, A. L., Groom, C. R. The druggable genome. *Nat. Rev. Drug Discov.* 2002, *1*, 727–730.
120. Laskowski, R. A., Watson, J. D., Thornton, J. M. Protein function prediction using local 3D templates. *J. Mol. Biol.* 2005, *351*, 614–626.
121. Binkowski, T. A., Joachimiak, A., Liang, J. Protein surface analysis for function annotation in high-throughput structural genomics pipeline. *Protein Sci.* 2005, *14*, 2972–2981.
122. Morris, R. J., Najmanovich, R. J., Kahraman, A., Thornton, J. M. Real spherical harmonic expansion coefficients as 3D shape descriptors for protein binding pocket and ligand comparisons. *Bioinformatics,* 2005, *21*, 2347–2355.
123. Watson, J. D., Sanderson, S., Ezersky, A., Savchenko, A., Edwards, A., Orengo, C., Joachimiak, A., Laskowski, R., Thornton, J. M. Towards fully automated structure-based function prediction in structural genomics: a case study. *J. Mol. Biol.* 2007, *367*, 1511–1522.
124. Gruber, J., Zawaira, A., Saunders, R., Barrett, C. P., Noble, M. E. M. Computational analyses of the surface properties of protein-protein interfaces. *Acta Cryst. D,* 2007, *63*, 50–57.
125. Li, H. L., Gao, Z. T., Kang, L., Zhang, H. L., Yang, K., Yu, K. Q., Luo, X. M., Zhu, W. L., Chen, K. X., Shen, J. H., Wang, X. C., Jiang, H. L. TarFisDock: a web server for identifying drug targets with docking approach. *Nucleic Acids Res*, 2006, *34*, W219–W224.
126. Cai, J. H., Han, C., Hu, T. C., Zhang, J., Wu, D. L., Wang, F. D., Liu, Y. Q., Ding, J. P., Chen, K. X., Yue, J. M., Shen, X., Jiang, H. L. Peptide deformylase is a potential target for anti-*Helicobacter pylori* drugs: Reverse docking, enzymatic assay, and X-ray crystallography validation. *Protein Sci.* 2006, *15*, 2071–2081.
127. Rockey, W. M., Elcock, A. H. Rapid computational identification of the targets of protein kinase inhibitors. *J. Med. Chem.* 2005, *48*, 4138–4152.
128. Hermann, J. C., Marti-Arbona, R., Fedorov, A. A., Fedorov, E., Almo, S. C., Shoichet, B. K., Raushel, F. M. Structure-based activity prediction for an enzyme of unknown function. *Nature,* 2007, *448*, 775–779.
129. Knight, Z. A., Shokat, K. M. Chemical genetics: where genetics and pharmacology meet. *Cell,* 2007, *128*, 425–430.
130. Blair, J. A., Rauh, D., Kung, C., Yun, C.-H., Fan, Q.-W., Rode, H., Zhang, C., Eck, M. J., Weiss, W. A., Shokat, K. M. Structure-guided development of affinity probes for tyrosine kinases using chemical genetics. *Nat. Chem. Biol.* 2007, *3*, 229–238.
131. Lipinski, C. A., Lombardo, F., Dominy, B. W., Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliver. Rev.* 1997, *23*, 3–25.
132. Loging, W., Harland, L., Williams-Jones, B. High-throughput electronic biology: mining information for drug discovery. *Nat. Rev. Drug Discov.* 2007, *6*, 220–230.
133. Blundell, T. L., Sibanda, B. L., Montalvao, R. W., Brewerton, S., Chelliah, V., Worth, C. L., Harmer, N. J., Davies, O., Burke, D. Structural biology and bioinformatics in drug design: opportunities and challenges for target identification and lead discovery. *Phil. Trans. R. Soc. B.* 2006, *361*, 413–423.
134. Hardy, J. A., Wells, J. A. Searching for new allosteric sites in enzymes. *Curr. Opin. Struct. Biol.* 2004, *14*, 706–715.
135. Chène, P. Drug targeting protein-protein interactions. *ChemMedChem* 2006, *1*, 400–411.
136. Varney, M. D., Marzoni, G. P., Palmer, C. L., Deal, J. G., Webber, S., Welsh, K. M., Bacquet, R. J., Bartlett, C. A., Morse, C. A., Booth, C. L. J., Herrmann, S. M., Howland, E. F., Ward, R. W., White, J. Crystal-structure-based design and synthesis of benz[cd]indole-containing inhibitors of thymidylate synthase. *J. Med. Chem.* 1992, *35*, 663–676.
137. Lam, P. Y. S., Jadhav, P. K., Eyermann, C. J., Hodge, C. N., Ru, Y., Bacheler, L. T., Meek, J. L., Otto, M. J., Rayner, M. M., Wong, Y. N., Chang, C.-H., Weber, P. C., Jackson, D. A., Sharpe, T. R., Erickson-Viitanen, S. Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. *Science,* 1994, *263*, 380–384.

138. Furet, P., Meyer, T., Strauss, A., Raccuglia, S., Rondeau, J.-M. Structure-based design and protein X-ray analysis of a protein kinase inhibitor. *Bioorg. Med. Chem. Lett.* 2002, *12*, 221–224.
139. Warshakoon, N. C., Wu, S., Boyer, A., Kawamoto, R., Sheville, J., Bhatt, R. T., Renock, S., Xu, K., Pokross, M., Zhou, S., Walter, R., Mekel, M., Evdokimov, A. G., East, S. Design and synthesis of substituted pyridine derivatives as HIF-1α prolyl hydroxylase inhibitors. *Bioorg. Med. Chem. Lett.* 2006, *16*, 5616–5620.
140. Herschhorn, A., Lerman, L., Weitman, M., Gleenberg, I. O., Nudelman, A., Hizi, A. De novo parallel design, synthesis and evaluation of inhibitors against the reverse transcriptase of human immunodeficiency virus type-1 and drug resistant variants. *J. Med. Chem.* 2007, *50*, 2370–2384.
141. Shoichet, B. K. Virtual screening of chemical libraries. *Nature,* 2004, *432*, 862–865.
142. Klebe, G. Virtual ligand screening: strategies, perspectives and limitations. *Drug Discov. Today.* 2006, *11*, 580–594.
143. Cavasotto, C. N., Orry, A. J. W. Ligand docking and structure-based virtual screening in drug discovery. *Curr. Topics Med. Chem.* 2007, *7*, 1006–1014.
144. Alvarez, J. C. High-throughput docking as a source of novel drug leads. *Curr. Opin. Chem. Biol.* 2004, *8*, 365–370.
145. Jansen, J. M., Martin, E. J. Target-biased scoring approaches and expert systems in structure-based virtual screening. *Curr. Opin. Chem. Biol.* 2004, *8*, 359–364.
146. Zentgraf, M., Steuber, H., Koch, C., La Motta, C., Sartini, S., Sotriffer, C. A., Klebe, G. How reliable are current docking approaches for structure-based drug design? Lessons from aldose reductase. *Angew. Chem. Int. Ed.* 2007, *46*, 3575–3578.
147. Paiva, A. M., Vanderwall, D. E., Blanchard, J. S., Kozarich, J. W., Williamson, J. M., Kelly, T. M. Inhibitors of dihydrodipicolinate reductase, a key enzyme in the diaminopimelate pathway of *Mycobacterium tuberculosis*. *Biochim. Biophys. Acta.* 2001, *1545*, 67–77.
148. Powers, R. A., Morandi, F., Shoichet, B. K. Structure-based discovery of a novel, noncovalent inhibitor of AmpC β-lactamase. *Structure,* 2002, *10*, 1013–1023.
149. Soelaiman, S., Wei, B. Q., Bergson, P., Lee, Y.-S., Shen, Y., Mrksich, M., Shoichet, B. K., Tang, W.-J. Structure-based inhibitor discovery against adenylyl cyclase toxins from pathogenic bacteria that cause anthrax and whooping cough. *J. Biol. Chem.* 2003, *278*, 25990–25997.
150. Brenk, R., Naerum, L., Grädler, U., Gerber, H.-D., Garcia, G. A., Reuter, K., Stubbs, M. T., Klebe, G. Virtual screening for submicromolar leads of tRNA-guanine transglycosylase based on a new unexpected binding mode detected by crystal structure analysis. *J. Med. Chem.* 2003, *46*, 1133–1143.
151. Vangrevelinghe, E., Zimmermann, K., Schoepfer, J., Portmann, R., Fabbro, D., Furet, P. Discovery of a potent and selective protein kinase CK2 inhibitor by high-throughput docking. *J. Med. Chem.* 2003, *46*, 2656–2662.
152. Singh, J., Chuaqui, C. E., Boriack-Sjodin, P. A., Lee, W.-C., Pontz, T., Corbley, M. J., Cheung, H.-K., Arduini, R. M., Mead, J. N., Newman, M. N., Papadatos, J. L., Bowes, S., Josiah, S., Ling, L. E. Successful shape-based virtual screening: the discovery of a potent inhibitor of the type I TGFβ receptor kinase (TβRI). *Bioorg. Med. Chem. Lett.* 2003, *13*, 4355–4359.
153. Kraemer, O., Hazemann, I., Podjarny, A. D., Klebe, G. Virtual screening for inhibitors of human aldose reductase. *Proteins,* 2004, *55*, 814–823.
154. Barril, X., Brough, P., Drysdale, M., Hubbard, R. E., Massey, A., Surgenor, A., Wright, L. Structure-based discovery of a new class of Hsp90 inhibitors. *Bioorg. Med. Chem. Lett.* 2005, *15*, 5187–5191.
155. Hann, M. M., Leach, A. R., Harper, G. Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inf. Comput. Sci.* 2001, *41*, 856–864.
156. Hopkins, A. L., Groom, C. R., Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discov. Today,* 2004, *9*, 430–431.
157. Abad-Zapatero, C. Ligand efficiency indices for effective drug discovery. *Expert Opin. Drug Discov.* 2007, *2*, 469–488.
158. Teague, S. J., Davis, A. M., Leeson, P. D., Oprea, T. The design of leadlike combinatorial libraries. *Angew. Chem. Int. Ed.* 1999, *38*, 3743–3748.
159. Congreve, M., Carr, R., Murray, C., Jhoti, H. A "rule of three" for fragment-based lead discovery?. *Drug Discov. Today,* 2003, *8*, 876–877.
160. Finkelstein, A. V., Janin, J. The price of lost freedom: entropy of bimolecular complex formation. *Protein Eng.* 1989, *3*, 1–3.
161. Murray, C. W., Verdonk, M. L. The consequences of translational and rotational entropy lost by small molecules on binding to proteins. *J. Comput. Aided Mol. Des.* 2002, *16*, 741–753.
162. Hajduk, P. J., Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat. Rev. Drug Discov.* 2007, *6*, 211–219.
163. Rees, D. C., Congreve, M., Murray, C. W., Carr, R. Fragment-based lead discovery. *Nat. Rev. Drug Discov.* 2004, *3*, 660–672.
164. *Fragment-based Approaches in Drug Discovery. Methods and Principles in Medicinal Chemistry* (Jahnke, W.; Erlanson, D. A., Eds.), Vol. 34. Wiley-VCH, Weinheim, Germany 2006.
165. Nienaber, V., Wang, J., Davidson, D., Henkin, J. Re-engineering of human urokinase provides a system for structure-based drug design at high resolution and reveals a novel structural subsite. *J. Biol. Chem.* 2000, *275*, 7239–7248.
166. Murray, C. W., Callaghan, O., Chessari, G., Cleasby, A., Congreve, M., Frederickson, M., Hartshorn, M. J., McMenamin, R., Patel, S., Wallis, N. Applications of fragment-based screening by X-ray crystallography to β-secretase. *J. Med. Chem.* 2007, *50*, 1116–1123.
167. Congreve, M., Aharony, D., Albert, J., Callaghan, O., Campbell, J., Carr, R. A. E., Chessari, G., Cowan, S., Edwards, P. D., Frederickson, M., McMenamin, R., Murray, C. W., Patel, S., Wallis, N. Applications of fragment-based screening by X-ray crystallography to the discovery of aminopyridines as inhibitors of β-secretase. *J. Med. Chem.* 2007, *50*, 1124–1132.
168. Gill, A. L., Frederickson, M., Cleasby, A., Woodhead, S. J., Carr, M. G., Woodhead, A. J., Walker, M. T., Congreve, M. S., Devine, L. A., Tisi, D., O'Reilly, M., Seavers, L. C. A., Davis, D. J., Curry, J., Anthony, R., Padova, A., Murray, C. W., Carr, R. A. E., Jhoti, H. Identification of novel p38α MAP kinase inhibitors using fragment-based lead generation. *J. Med. Chem.* 2005, *48*, 414–426.
169. Card, G. L., Blasdel, L., England, B. P., Zhang, C., Suzuki, Y., Gillette, S., Fong, D., Ibrahim, P. N., Artis, D. R., Bollag, G., Milburn, M. V., Kim, S.-H., Schlessinger, J., Zhang, K. Y. J. A family of phosphodiesterase inhibitors discovered by cocrystallography and scaffold-based drug design. *Nat. Biotechnol.* 2005, *23*, 201–207.
170. Holdgate, G. A., Ward, W. H. J. Measurements of binding thermodynamics in drug discovery. *Drug Discov. Today,* 2005, *10*, 1543–1550.
171. Brenk, R., Klebe, G. "Hot spot" analysis of protein-binding sites as a prerequisite for structure-based virtual screening and lead optimization. *Meth. Prin. Med. Chem.* 2006, *32*, 171–192.
172. Ciulli, A., Williams, G., Smith, A. G., Blundell, T. L., Abell, C. Probing hot spots at protein-ligand binding sites: a fragment-based approach using biophysical methods. *J. Med. Chem.* 2006, *49*, 4992–5000.
173. Gohlke, H., Hendlich, M., Klebe, G. Predicting binding modes, binding affinities and "hot spots" for protein-ligand complexes using a knowledge-based scoring function. *Perspect. Drug Discov. Des.* 2000, *20*, 115–144.
174. Verdonk, M. L., Cole, J. C., Watson, P., Gillet, V., Willett, P. Superstar: improved knowledge-based interaction fields for protein binding sites. *J. Mol. Biol.* 2001, *307*, 841–859.

175. Grzybowski, B. A., Ishchenko, A. V., Kim, C.-Y., Topalov, G., Chapman, R., Christianson, D. W., Whitesides, G. M., Shakhnovich, E. I. Combinatorial computational method gives new picomolar ligands for a known enzyme. *Proc. Natl. Acad. Sci. USA*. 2002, *99*, 1270–1273.
176. Perola, E., Charifson, P. S. Conformational analysis of drug-like molecules bound to proteins: an extensive study of ligand reorganization upon binding. *J. Med. Chem.* 2004, *47*, 2499–2510.
177. Kim, E. E., Baker, C. T., Dwyer, M. D., Murcko, M. A., Rao, B. G., Tung, R. D., Navia, M. A. Crystal structure of HIV-1 protease in complex with VX-478, a potent and orally bioavailable inhibitor of the enzyme. *J. Am. Chem. Soc.* 1995, *117*, 1181–1182.
178. Greer, J., Erickson, J. W., Baldwin, J. J., Varney, M. D. Application of the three-dimensional structures of protein target molecules in structure-based drug design. *J. Med. Chem.* 1994, *37*, 1035–1054.
179. MacPherson, L. J., Bayburt, E. K., Capparelli, M. P., Bohacek, R. S., Clarke, F. H., Ghai, R. D., Sakane, Y., Berry, C. J., Peppard, J. V., Trapani, A. J. Design and synthesis of an orally active macrocyclic neutral endopeptidase 24.11 inhibitor. *J. Med. Chem.* 1993, *36*, 3821–3828.
180. Kim, S.-H., Pudzianowski, A. T., Leavitt, K. J., Barbosa, J., McDonnell, P. A., Metzler, W. J., Rankin, B. M., Liu, R., Vaccaro, W., Pitts, W. Structure-based design of potent and selective inhibitors of collagenase-3 (MMP-13). *Bioorg. Med. Chem. Lett.* 2005, *15*, 1101–1106.
181. Ohtaka, H., Velázquez-Campoy, A., Xie, D., Freire, E. Overcoming drug resistance in HIV-1 chemotherapy: the binding thermodynamics of Amprenavir and TMC-126 to wild-type and drug resistant mutants of the HIV-1 protease. *Protein Science*. 2002, *11*, 1908–1916.
182. Mao, C., Sudbeck, E. A., Ventakatachalam, T. K., Uckun, F. M. Structure-based drug design of non-nucleoside inhibitors for wild-type and drug resistant HIV reverse transcriptase. *Biochem. Pharmacol.* 2000, *60*, 1251–1265.
183. Das, K., Clark, A. D., Jr., Lewi, P. J., Heeres, J., de Jonge, M. R., Koymans, L. M. H., Vinkers, H. M., Daeyaert, F., Ludovici, D. W., Kukla, M. J., De Corte, B., Kavash, R. W., Ho, C. Y., Ye, H., Lichtenstein, M. A., Andries, K., Pauwels, R., de Béthune, M.-P., Boyer, P. L., Clark, P., Hughes, S. H., Janssen, P. A. J., Arnold, E. Roles of conformational and positional adaptability in structure-based design of TMC125-R165335 (etravirine) and related non-nucleoside reverse transcriptase inhibitors that are highly potent and effective against wild-type and drug-resistant HIV-1 variants. *J. Med. Chem.* 2004, *47*, 2550–2560.
184. Velazquez-Campoy, A., Todd, M. J., Freire, E. HIV-1 protease inhibitors: enthalpic versus entropic optimization of the binding affinity. *Biochemistry,* 2000, *39*, 2201–2207.
185. Brown, F. J., Andisik, D. W., Bernstein, P. R., Bryant, C. B., Ceccarelli, C., Damewood, J. R., Jr., Edwards, P. D., Earley, R. A., Feeney, S., Green, R. C., Gomes, B., Kosmider, B. J., Krell, R. D., Shaw, A., Steelman, G. B., Thomas, R. M., Vacek, E. P., Veale, C. A., Tuthill, P. A., Warner, P., Williams, J. C., Wolanin, D. J., Woolson, S. A. Design of orally active, non-peptidic inhibitors of human leukocyte elastase. *J. Med. Chem.* 1994, *37*, 1259–1261.
186. Pierce, A. C., Rao, G., Bemis, G. W. BREED: generating novel inhibitors through hybridization of known ligands. Application to CDK2, P38, and HIV protease. *J. Med. Chem.* 2004, *47*, 2768–2775.
187. Anderson, M., Beattie, J. F., Breault, G. A., Breed, J., Byth, K. F., Culshaw, J. D., Ellston, R. P. A., Green, S., Minshull, C. A., Norman, R. A.,. Pauptit, R. A., Stanway, J., Thomas, A. P., Jewsbury, P. J. Imidazo[1,2a]pyridines: a potent and selective class of cyclin-dependent kinase inhibitors identified through structure-based hydridisation. *Bioorg. Med. Chem. Lett*. 2003, *13*, 3021–3026.
188. Terasaka, T., Kinoshita, T., Kuno, M., Nakanishi, I. A highly potent non-nucleoside adenosine deaminase inhibitor: efficient drug discovery by intentional lead hydridization. *J. Am. Chem. Soc.* 2004, *126*, 34–35.
189. Oefner, C., Binggeli, A., Breu, V., Bur, D., Clozel, J. P., D'Arcy, A., Dorn, A., Fischli, W., Gruninger, F., Guller, R., Hirth, G., Marki, H. P., Mathews, S., Muller, M., Ridley, R. G., Stadler, H., Vieira, E., Wilhelm, M., Winkler, F. K., Wostl, W. Renin inhibition by substituted piperidines: a novel paradigm for the inhibition of monomeric aspartic proteinases?. *Chem. Biol.* 1999, *6*, 127–131.
190. Hartshorn, M. J., Murray, C. W., Cleasby, A., Frederickson, M., Tickle, I. J., Jhoti, H. Fragment-based lead discovery using X-ray crystallography. *J. Med. Chem.* 2005, *48*, 403–413.
191. Lesuisse, D., Lange, G., Deprez, P., Bénard, D., Schoot, B., Delettre, G., Marquette, J.-P., Broto, P., Jean-Baptiste, V., Bichet, P., Sarubbi, E., Mandine, E. SAR and X-ray. A new approach combining fragment-based screening and rational drug design: Application to the Discovery of Nanomolar Inhibitors of Src SH2. *J. Med. Chem.* 2002, *45*, 2379–2387.

Part VI

Chemical Modifications Influencing the Pharmacokinetic Properties

Richard B. Silverman

Section Editor

Chapter 31

Physiological Aspects Determining the Pharmacokinetic Properties of Drugs

Koen Boussery, Frans M. Belpaire and Johan Van de Voorde

I. INTRODUCTION
II. PASSAGE OF DRUGS THROUGH BIOLOGICAL BARRIERS
A. Transcellular drug transport
B. Paracellular drug transport
III. DRUG ABSORPTION
A. Dosage form of the drug
B. GI motility and gastric emptying
C. GI permeability to the drug
D. Perfusion of the GI tract and the first-pass effect
IV. DRUG DISTRIBUTION
A. Plasma protein binding
B. Drug accumulation
C. The blood–brain barrier
V. DRUG ELIMINATION
A. Excretion
B. Biotransformation
VI. SOME PHARMACOKINETIC PARAMETERS AND TERMINOLOGY
A. Plasma concentration–time curve
B. Volume of distribution
C. Clearance
D. Elimination half-life ($T_{1/2}$)
E. Bioavailability
VII. VARIABILITY IN PHARMACOKINETICS
A. Genetic factors
B. Age
C. Drug interactions
D. Disease state
E. Pregnancy
BIBLIOGRAPHY

"To explain all nature is too difficult a task for any one man or even for any one age. 'T is much better to do a little with certainty, and leave the rest for others that come after you, than to explain all things."

Isaac Newton

I. INTRODUCTION

In order to produce its intended effect, a drug must be present at an appropriate concentration in the fluid surrounding the effect site, that is, the biophase. Only rarely can drugs be applied directly to the biophase; in most cases drugs need to be transferred from the site of administration to the biophase. Usually, this translocation involves two steps: *absorption* and *distribution*. During absorption, the drug passes from its site of administration (e.g. the gastrointestinal (GI) tract when the drug is taken orally) into the systemic circulation. Subsequently, the drug is distributed via the circulating blood plasma (the fluid portion of the blood) to the different parts of the organism, including the organ(s) in which the biophase for the drug is localized. Each drug molecule that reaches the target site can add to the intended pharmacological effect of the drug. However, at all times a portion of the drug molecules in the body is also distributed to organs and tissues that account for an irreversible loss of drug molecules from the body (*drug elimination*) by either biotransformation (the conversion of one chemical entity to another) or excretion. This causes a decrease in the concentration of the drug in the body and, consequently, also in the biophase. Figure 31.1 shows a schematic representation of the processes involved in the journey of a drug molecule through the human body.

Copyright © 2008, Elsevier Ltd
All rights reserved.

Figure 31.2 shows a more detailed scheme of the main routes of drug absorption, distribution and elimination. *Pharmacokinetics* is the study of the drug concentrations in the different parts of the organism as a function of time. These concentrations depend on the dose administered and upon the rate and extent of absorption, distribution and elimination.

In the first part of this overview (Sections II, III, IV and V) some physiological aspects of drug absorption, distribution and elimination will be discussed. The second part (Sections VI and VII) will briefly focus on some pharmacokinetic parameters and terminology and on variability in pharmacokinetics. Due to its limited size, this chapter only includes some basic and general information. For more elaborate information the reader is referred to some excellent textbooks in the bibliography.

FIGURE 31.1 Schematic representation of drug absorption, distribution and elimination.

II. PASSAGE OF DRUGS THROUGH BIOLOGICAL BARRIERS

On its journey through the body, a drug needs to cross different biological barriers. These barriers can be a single layer of cells (e.g. the intestinal epithelium), several layers of cells (e.g. in the skin), or the cell membrane itself (e.g. to reach an intracellular receptor). A drug can cross a cell layer either by traveling through the cells (*transcellular* drug transport) or through gaps between the cells (*paracellular* drug transport). The mechanisms by which a drug can cross the cell membrane will be discussed, together with transcellular drug transport.

A. Transcellular drug transport

In order to travel through a cell or to reach a target inside a cell, a drug molecule must be able to traverse the cell membrane. The cell membrane (also called plasma membrane) is a lipid bilayer interspersed with carbohydrates and proteins. Although cell membranes largely vary in their permeability characteristics depending on the tissue, the main mechanisms of drugs passing through the cell membrane are passive diffusion, carrier-mediated processes and vesicular transport.

1. *Passive diffusion*

Passive diffusion is the process by which molecules spontaneously diffuse from a region of higher concentration

FIGURE 31.2 Schematic representation of the main routes of drug absorption, distribution and elimination (GIS = GI system). The red arrows represent the enterohepatic cycle (see Section V.A.2.).

(e.g. outside of the cell) to a region of lower concentration (e.g. inside the cell), and it is the main mechanism for passage of drugs through membranes. Lipid-soluble drugs penetrate the lipid cell membrane with ease, and can pass the cell membrane by passive diffusion. Polar molecules and ionized compounds, on the other hand, partition poorly into lipids and are not able to diffuse through the cell membrane or do so at a much lower rate. Also, large molecules, such as proteins and protein-bound drugs, cannot diffuse through the cell membrane.

Transmembrane diffusion is driven by the concentration gradient of the drug over the cell membrane. The rate of diffusion depends, apart from the lipid/water partition coefficient of the drug (P) and the concentration gradient ($C_{out} - C_{in}$), on membrane properties such as the membrane area (A) and thickness (h), and the diffusion coefficient (D) of the drug in the membrane, according to Fick's law (equation (31.1)).

$$\text{Rate of diffusion} = \frac{\text{DAP}(C_{out} - C_{in})}{h} \qquad (31.1)$$

Many drugs are acidic or basic compounds, which are ionized to a certain degree in aqueous medium. Their degree of ionization depends on their dissociation constant (pK_a) and the pH of the solution, according to the Henderson–Hasselbach equation (Equations (31.2) and (31.3)).

For acidic drugs:

$$\log \frac{\text{Ionized concentration}}{\text{Unionized concentration}} = \text{pH} - pK_a \qquad (31.2)$$

For basic drugs:

$$\log \frac{\text{Unionized concentration}}{\text{Ionized concentration}} = \text{pH} - pK_a \qquad (31.3)$$

Very weak acids with pK_a values higher than 7.5, are essentially unionized at physiological pH values. For these drugs diffusion over the cell membrane is rapid and independent of pH changes within the body, provided the unionized form of the drug is lipid soluble. For acidic drugs with a pK_a value between 3.0 and 7.5, the fraction of unionized drug varies with the changes in pH encountered in the organism. For these drugs the pH of the extracellular environment is critical in determining the diffusion across the cell membrane. For acidic drugs with a pK_a lower than 2.5, the fraction of unionized drug is low at any physiological pH, resulting in very slow diffusion across membranes. A similar analysis can be made for bases.

At the diffusion equilibrium, the concentrations of unionized molecules on both sides of a biological barrier are equal. If the pH on both sides of the barrier is equal, then the concentration of ionized molecules and, consequently, the total concentration of the molecules, will be the same on both sides of the barrier. However, if there is a difference in pH, as e.g. between blood plasma (pH 7.4) and stomach contents (pH 1–3), the concentration of the ionized molecules at equilibrium, and, therefore, the total concentration, will be much higher on one side of the barrier than on the other. This phenomenon is called ion trapping.

2. *Carrier-mediated processes*

Many cell membranes possess specialized transport mechanisms that regulate entry and exit of physiologically important molecules and drugs. Such transport systems involve a carrier molecule, that is, a transmembrane protein that binds one or more molecules and releases them on the other side of the membrane. Such systems may operate passively (without any energy source) and along a concentration gradient; this is called "facilitated diffusion." However, facilitated diffusion seems to play only a minor role in drug transport. An example is the transport of vitamin B_{12} across the GI membrane. Alternatively, the system may spend energy (obtained from the energy rich molecule adenosine triphosphate (ATP) required to pump molecules against a concentration gradient; this mechanism is called "active transport."

At high drug concentrations the carrier sites become saturated, and the rate of transport does not further increase with concentration. Furthermore, competitive inhibition of transport can occur if another substrate for this carrier is present. In recent years, several transporters have been described to be present in various organs and tissues throughout the body and to determine absorption, distribution and elimination of compounds that are substrates for these transporters. Although some transporters mediate the uptake of compounds in the cell (*influx transporters*), others may mediate secretion back out of the cell (*efflux transporters*). Transporters in the intestinal membrane affect the absorption of drugs (see Section III.C.), while transporters in the liver and kidney influence elimination by mediating transport into and out of cells responsible for biotransformation (hepatocytes, see Section V.B.) or excretion (e.g. renal tubule cells in the kidneys, see Section V.A.1.). Furthermore, efflux transporters may limit the penetration of compounds into certain areas of the body, such as the cerebrospinal fluid and blood cells. Chapter 36 of this textbook presents a more elaborate description of drug transport mechanisms.

3. *Vesicular transport*

During vesicular transport the cell membrane forms a small cavity that gradually surrounds particles or macromolecules, thereby internalizing them into the cell in the form of a vesicle or vacuole. Vesicular transport is the proposed process for the absorption of orally administered Sabin polio vaccine and of various large proteins. It is called *endocytosis* when moving a macromolecule into a cell, *exocytosis* when

moving a macromolecule out of a cell, and *transcytosis* when moving a macromolecule across a cell.

B. Paracellular drug transport

Drugs can also cross a cell layer through the small aqueous contact points (*cell junctions*) between cells. This paracellular drug transport can be initiated by a concentration gradient over the cell layer (passive diffusion), or by a hydrostatic pressure gradient across the cell layer (filtration). The size and characteristics of cell junctions widely vary between different barriers to drug transport. For example, the endothelium of glomerular capillaries in the kidney (see Section V.A.1.a.) forms a leaky barrier, which is very rich in intercellular pores. Therefore, this membrane is very permeable and permits filtration of water and solutes. On the other hand, endothelial cells of brain capillaries are sealed together by *tight junctions*, practically eliminating the possibility of paracellular drug transport.

III. DRUG ABSORPTION

Absorption can be defined as the passage of a drug from its site of administration into the systemic circulation. If a drug is administered directly into the systemic circulation by intravenous (i.v.) administration, absorption is not needed. Drugs can be administered by *enteral* and *parenteral* routes. Enteral administration occurs through the GI tract, by contact of the drug with the mucosa in the mouth (buccal or sublingual), by swallowing (oral) or by rectal administration. In parenteral administration the GI tract is bypassed; examples are the i.v. (direct injection in the systemic circulation) and intramuscular (injection into a muscle) routes. Drugs can also be absorbed through the skin or through the mucosa of various organs (e.g. bronchi, nose, and vagina). In some cases, a drug is applied for a local effect, and no absorption is intended (e.g. antacids that neutralize stomach acid). In this chapter, we will describe drug administration by the oral route, which is the most common and popular route of drug dosing. Some characteristics of other common routes of drug administration are listed in Table 31.1. For more details on these various routes, the reader is referred to the bibliography at the end of this chapter (e.g. Ref. [4]).

The *enteral system* consists of the GI tract from the mouth to the anus. With respect to drug absorption after oral dosing, the stomach, the small intestine, and the large intestine (or colon) are the major components (Figure 31.3). The small intestine includes the duodenum, jejunum and ileum. These segments differ from one another anatomically and morphologically, as well as with respect to secretions and pH. As orally administered drugs move through the GI tract, they encounter environments that vary in pH (Figure 31.3), enzyme composition, fluidity of contents, as well as in the surface area available for absorption. The most important area for drug absorption is the duodenum, the upper portion of the small intestine. The small intestine is also the region that is responsible for almost all digestion and absorption of nutrients, and its structural adaptation to this task makes it very suitable for the absorption of drugs. Its length alone provides a large surface area for absorption, and that area is further increased by circular folds, villi (finger-like projections of the intestinal wall), and microvilli (small projections of villi). If the small intestine is thought of as a hollow cylinder, its net increase in total surface area due to these folds, villi and microvilli, is 600-fold (versus that of a smooth cylinder with the same length). The total surface area of the human small intestine is approximately 200 m^2, or the surface area of a doubles tennis court.

The main determinants of the rate and extent of absorption after oral administration are:

- dosage form of the drug
- GI motility and gastric emptying
- GI permeability to the drug
- perfusion of the GI tract
- first-pass effect

A. Dosage form of the drug

For a drug to be absorbed from the GI tract, it has to be dissolved in the aqueous medium of the stomach and the intestine. However, many drugs are administered as tablets or capsules that first have to disintegrate and to release the drug. Therefore, liquid dosage forms are, in general, more rapidly absorbed than solid forms. Disintegration can be controlled, and various drug products have been modified to alter the timing of the release of the active drug from the drug product. The term "controlled release" is used for various types of oral extended release rate dosage forms (such as sustained release, prolonged release) and delayed release rate dosage forms (e.g. enteric coated). One example of the systems for controlled release is the osmotic pump, in which drug delivery is driven by an osmotically controlled device that pumps a constant amount of water through the system, dissolving and releasing a constant amount of drug per time unit.

B. GI motility and gastric emptying

Once a drug is given orally, GI motility tends to move the drug through the GI tract from mouth to anus. The drug rapidly reaches the stomach, which subsequently empties its contents into the small intestine. The residence time of the drug in the stomach varies from a few minutes to several hours and is dependent on a range of factors, such as the volume, viscosity and composition of the stomach content. The surface area of the stomach is limited in comparison with the small intestine, and under normal conditions gastric emptying is rapid.

TABLE 31.1 Common Routes of Drug Administration

1. Parenteral routes
   - Intravenous bolus (i.v.)
     - Direct injection of complete dose into the systemic circulation
     - Complete and instantaneous bioavailability, no need for absorption
     - Often used for immediate effect
     - Main disadvantages: technique requires extensive training; some complications may have serious consequences; formulation must be sterile
   - Intravenous infusion (i.v.)
     - Similar to i.v. bolus, but dose is injected slowly into the systemic circulation at a constant rate (controlled by an infusion pump)
     - Plasma drug levels are more precisely controlled
     - Larger fluid volumes can be injected
   - Intramuscular injection (IM)
     - Injection of a dose into a muscle, from where it is absorbed due to the perfusion of the muscle by blood
     - Easier than i.v. injection
     - Rapid absorption from aqueous solution, slower from non-aqueous (oil) solutions
     - Main disadvantages: irritating drugs may be very painful; different rates of absorption depending on muscle group injected and blood flow
   - Subcutaneous injection (s.c.)
     - Injection of a dose into the s.c. tissue layer immediately beneath the skin
     - Main disadvantage: drug absorption is relatively slow and depends on local blood flow; s.c. tissues are often adipose and poorly perfused
     - Used for insulin replacement therapy in diabetic patients

2. Enteral routes
   - Buccal or sublingual (SL) drug delivery
     - A drug formulation is brought in close contact with the mucous membranes inside the mouth (lining the cheeks (buccal) or under the tongue (SL))
     - No first-pass effects
     - Mainly passive diffusion; only small lipophilic drugs are absorbed
     - Main disadvantage: some drug may be swallowed
   - Oral drug delivery
     - A drug formulation is swallowed, absorption from the GI tract
     - Absorption may vary in rate and extent
     - Safest and easiest route of drug administration
     - Main disadvantage: some drugs may have erratic absorption, be unstable in the GI tract, or be metabolized by liver prior to systemic absorption
   - Rectal drug delivery
     - Absorption from suppository may vary; more reliable absorption from enema (solution)
     - Useful when patient cannot swallow medication (e.g. elderly and very young patients; vomiting patients)
     - Used for both local and systemic effects
     - First-pass metabolism in the liver is partly avoided
     - Main disadvantages: absorption may be erratic; suppository may migrate to different position; some patient discomfort

3. Other routes
   - Transdermal/percutaneous drug delivery
     - Generally, drug absorption through the skin is slow; absorption can be increased by occlusive dressing
     - Permeability of skin varies with condition, anatomic site, age, and gender
     - Easy to use: e.g. patches
     - First-pass metabolism in the liver is avoided
     - Main disadvantage: possible irritation of the skin by patch or drug
   - Intranasal drug delivery
     - Primarily used for local effects, but also developed as a route for systemic effects: absorption through the nasal mucosa
     - First-pass metabolism in the liver is avoided
     - Especially attractive for the delivery of peptides (e.g. desmopressin)
   - Pulmonary drug delivery/inhalation
     - Primarily used for local effects: particle size of drug determines deposition site in respiratory tract
     - Inhaled drugs can be absorbed from their deposition site in various parts of the respiratory tract (large surface area + blood supply)
     - Main disadvantages: may stimulate cough reflex; some drug may be swallowed

FIGURE 31.3 Schematic representation of the GI tract.

Therefore, the stomach's role in drug absorption is – in general – rather modest. Factors that influence gastric emptying can, however, influence the absorption rate of most drugs, but not necessarily the total amount of drug eventually absorbed. For example, consumption of a meal (especially a meal high in fat) reduces gastric emptying. Therefore, a drug taken with food will stay longer in the stomach, which could reduce the absorption rate of that drug. The motility of the intestine can also influence the absorption. It mixes the contents of the duodenum, bringing the drug in close contact with the intestinal wall (the biological barrier the drug needs to cross to be absorbed). When intestinal motility increases, disintegration of the formulation and dissolution of the drug are often accelerated. On the other hand, high motility of the intestinal tract (as in diarrhea) can result in a very short residence time in the small intestine, and less opportunity for adequate absorption as a consequence.

C. GI permeability to the drug

Once a drug is dissolved in the aqueous medium in the GI tract, it can pass into the capillaries of the GI wall. The drug needs to have a reasonable amount of lipid solubility to enable absorption across the lipid membrane, but also a reasonable amount of water solubility is necessary to dissolve in the GI system. A drug that is too lipid soluble will exhibit impaired dissolution in the GI system, which will lead to reduced absorption. On the other hand, a drug that is too water soluble will have adequate dissolution but reduced absorption due to its inability to cross the lipid membrane. General rules for the intestinal absorption of a drug include:

- Small amphiphatic drugs move efficiently through the transcellular route by partitioning into and out of lipid bilayers.
- Small hydrophilic drugs are restricted to the paracellular route, or to aqueous routes that normally absorb nutrients, vitamins or cofactors.
- Peptides and proteins are poorly absorbed intact and require the application of enhancing agents or special uptake mechanisms.
- In general, the permeability for a drug decreases along the intestine, but this is obviously very dependent on the drug and the route of transport.

For acids and bases only the non-ionized molecules can be absorbed. At all physiological pH values weak acids and bases exist mostly in the unionized form and can be absorbed as well from the stomach as from the intestine. In theory, weakly acidic drugs are better substrates for passive diffusion at the pH of the stomach than at that of the intestine. However, the limited residence time of the drug in the stomach and the relatively small surface area of the stomach more than balance the influence of pH in determining the optimal site of absorption. Strong bases such as the quaternary ammonium compounds are, to a large extent, ionized at all physiological pH's, and are hardly absorbed at all.

For many years, the rate and extent of absorption in the small intestine were thought to be determined solely by the lipid/water solubility and membrane permeability characteristics of the drug. While this relatively simplistic model worked for many drugs, there are a number of exceptions to this rule, suggesting other forces are at work within the GI system to control the absorption of drugs. It is now known that a complex system of transporter proteins and metabolic enzymes is present within the GI system. Expression of influx transporters in the intestinal epithelial cells can increase absorption of drugs that are substrates for these transporters, whereas efflux transporters can reduce oral absorption of these drugs. In particular, the impact of the

P-glycoprotein multidrug transporter (P-gp) on drug absorption has been studied extensively. Since P-gp is located on the epithelium of intestinal cells, it can act as a counter-transport pump that transports its substrates back into the intestinal lumen as they begin to be absorbed across the intestinal wall.

D. Perfusion of the GI tract and the first-pass effect

The villi in the duodenal region are highly perfused with a network of capillaries and lymphatic vessels. The capillaries in the villi are fenestrated (i.e. they have large pores) and have a large surface area, so that drugs absorbed from the small intestine can easily enter these capillaries. The drug is then transported to the portal vein and to the liver prior to reaching the systemic circulation. Figure 31.4 shows a schematic representation of the splanchnic circulation, which includes the blood flow through the stomach, small and large intestines, pancreas, spleen and liver. This splanchnic circulation receives about 28% of the cardiac output via the abdominal aorta, a fraction that is significantly elevated for 2–4 h after a meal. Any change in blood flow to the GI tract will affect the rate of drug absorption from the intestinal tract.

FIGURE 31.4 Schematic representation of the splanchnic circulation.

As is clear from Figure 31.4, the liver receives most (approximately 75%) of its blood supply through the portal vein, which carries the venous blood draining from all of the organs in the splanchnic circulation except the liver itself. As a consequence, drugs that are given orally first pass through the liver before being distributed to the rest of the body, that is, before entering the systemic circulation. This is an important issue for some drugs that are highly metabolized by the liver. When administered orally, a substantial fraction of these drugs will be metabolized before reaching systemic circulation. Such a loss when a drug passes through sites of elimination during absorption is known as a *first-pass effect*. Besides the hepatic first-pass effect, biotransformation during absorption can also occur in the lumen of the intestine and by enzymes that are present in the gut wall (Figure 31.5). For example, CYP3A- a major subfamily of Phase I drug metabolizing enzymes in humans (see Section V.B.) – has been shown to be present at high levels in the intestinal wall, which limits oral drug delivery of its substrates.

A first-pass phenomenon may also occur after intraperitoneal and, partially, after rectal administration. It does not occur for parental routes of administration or after buccal or sublingual administration. For some orally administered

FIGURE 31.5 The first-pass effect: an orally administered drug must pass through different potential sites of elimination before it reaches the systemic circulation.

drugs with high lipid solubility, absorption via the lymphatic system is also possible. Drugs are absorbed through the lymphatic vessels under the microvilli. Absorption of drugs through the lymphatic system bypasses the hepatic first-pass effect because drug absorption through the portal vein is avoided (the lymph delivers these substances to the systemic circulation via the thoracic duct).

IV. DRUG DISTRIBUTION

After absorption into the systemic circulation, drugs are distributed to the various organs and tissues in the body. The blood plasma carries the drug molecules to the effect site for drug action, as well as to other tissues where side effects or adverse reactions may occur. The rate and extent of distribution depend on blood flow to different organs, tissue size, binding of drugs to plasma proteins and tissue components, and permeability of tissue membranes. The latter factor is related to the physicochemical properties of the drug, as described above (see Section II). For lipid-soluble drugs tissue membranes represent no barrier, and distribution depends essentially on the perfusion rate of the tissue. For these drugs rapid equilibration occurs between blood and tissues such as lungs, kidney, liver, heart and brain, that is, organs with a high blood flow. Less rapid equilibration is found for skeletal muscle, bone, and adipose tissue, which receive a considerable smaller volume of blood per unit mass. This is called "perfusion limited distribution," since blood flow is the rate-limiting step in the distribution of the drug. In contrast, if drug distribution is limited by the slow passage of drug across the membrane in the tissue, this is called "permeability limited distribution." Tissue uptake of a drug continues until equilibrium is reached between the diffusible form of the drug in the tissue and the blood, that is, until the free concentrations in blood plasma and tissue water are equal. Drugs can be present in tissues in higher concentrations than in blood plasma as a consequence of pH-gradients, but mainly because of a high affinity for that particular type of tissue. This is called *drug accumulation*. On the other hand, drugs can be present in high concentrations in blood plasma due to a high *plasma protein binding*.

A. Plasma protein binding

Many drugs are bound to some extent to plasma proteins. It may be important to know to what extent a certain drug is bound to plasma proteins, since a protein-bound drug is a large complex that cannot easily cross the biological barrier and therefore has a restricted distribution. Furthermore, the protein-bound drug is usually pharmacologically inactive. The plasma protein binding is expressed as "fraction bound," that is, the ratio of bound concentration over total (bound plus free) concentration, or as "percentage bound" if this value is multiplied by 100. The free fraction equals one minus the bound fraction. Many acidic drugs bind to *albumin*, the major component of plasma proteins responsible for reversible drug binding (normal plasma concentrations for albumin range from 35 to 40 g/L). α_1*-Acid glycoprotein* is an acute phase reactant, a group of plasma proteins that changes in concentration following tissue injury or inflammation. It primarily binds to basic drugs such as propranolol and imipramine. The plasma concentration of α_1-acid glycoprotein is low (0.4–1 g/L), but its concentration in plasma rises in inflammation. Binding to other macromolecular components in the blood (including lipoproteins, immunoglobulins and erythrocytes) generally occurs to a much smaller extent.

For most drugs the binding of drugs to plasma proteins is a reversible process with extremely rapid rates of association and dissociation, which can be described by the law of mass action. The degree of binding is determined by affinity (expressed as the association constant), capacity (the number of binding sites per molecule protein), protein concentration and drug concentration. At therapeutic drug concentrations, usually only a small fraction of the available binding sites is occupied; for a given protein concentration the free fraction of the drug is then rather constant and independent of drug concentration. In some instances, the drug concentrations are so high that most binding sites are occupied, and the free fraction becomes concentration-dependent. Concentration-dependent changes in drug binding are most likely to occur with drugs that have a high affinity for the proteins and that are given in large doses, for example, acetylsalicylic acid, phenylbutazone, some penicillins and cephalosporins.

The plasma protein binding of drugs is altered in some physiological and pathological conditions; often as a result of changes in plasma protein concentration or as a result of competition for common binding sites by another (endogenous or exogenous) compound. In various disease states (such as renal failure, liver disease, inflammation), in pregnancy and in the neonatal period, hypoalbuminemia is observed; α_1-acid glycoprotein concentrations rise in inflammatory diseases, stress and malignancy, and fall in liver disease. Free fatty acids bind strongly to albumin. When their concentration in plasma increases due to fasting, exercise or infection, albumin-bound drugs can be displaced from their binding sites. In renal failure, waste products that accumulate in the blood may compete for plasma protein binding. Besides endogenous compounds, other administered drugs may compete for plasma protein binding. Such an interaction is to be expected when the "displacer" is present in the same concentration range as the binding sites at the proteins. This situation may result in a decrease of the binding sites available for the "displaced" drug. The changes in actual free plasma concentration will always be smaller than the changes in free fraction because of redistribution of the displaced drug to the tissues and its more rapid elimination.

B. Drug accumulation

Drugs may accumulate in body tissues because of a high affinity for that particular type of tissue. For example, drugs with a high fat/water partition coefficient are very fat soluble and tend to accumulate in body fat. Accumulation in body fat is important for only a few drugs, mainly because the fat/water partition coefficient is relatively low for most drugs. Morphine, for example, though lipid-soluble enough to cross the blood–brain barrier, has a fat/water partition coefficient of only 0.4 and sequestration of the drug by body fat is of little importance. With thiopentone, on the other hand (fat/water partition coefficient approximately 10), accumulation in body fat is considerable. Accumulation of drugs in body fat is also limited by a low blood supply to body fat – less than 2% of the cardiac output. As blood flow to body fat is limited, drugs are delivered to body fat slowly, and also the equilibrium distribution between fat and body water is approached slowly. As a result, accumulation in body fat is only important when lipid-soluble drugs are given chronically (e.g. benzodiazepines). Only for highly lipid-soluble drugs (e.g. general anesthetics) partition into body fat is also important at first dose. Drugs may also accumulate in tissues by binding reversibly or irreversibly to tissue constituents. For example, tetracyclines, a class of antibiotics, bind with calcium to form an insoluble chelate and therefore accumulate irreversibly in growing bones and teeth.

The body tissues in which a drug accumulates are potential reservoirs for the drug. If a stored drug is in equilibrium with that in plasma and is released as the plasma concentration declines, drug concentration in plasma and at the biophase is sustained, and pharmacological effects of the drug are prolonged. However, if the reservoir for the drug has a large capacity and fills rapidly, a larger initial dose is required to reach a therapeutically effective concentration at the biophase after the first administration.

C. The blood–brain barrier

Delivery of drugs to the brain from the systemic circulation is difficult due to the presence of the so-called blood–brain barrier. This barrier is acting as a self-defense mechanism by preventing the passage of many potentially harmful substances from blood into brain tissue. It is formed by the brain capillary endothelial cells that are sealed together by tight junctions and closely surrounded by processes of large numbers of astrocytes (a type of supporting cells in the brain), thereby eliminating the possibility of paracellular transport. Furthermore, different efflux transporters are present that remove drugs from the brain and transfer them to the systemic circulation. This is why brain penetration of most drugs is markedly restricted. The pathways for drug uptake in the brain are mainly limited to active transport and simple diffusion. As a consequence, only molecules that are either a substrate for an influx transporter or are highly lipid soluble and of relatively low molecular weight can cross the blood–brain barrier. A challenging area of research in drug delivery attempts to make use of influx transporters to develop blood–brain barrier vectors that can improve drug uptake in the brain.

V. DRUG ELIMINATION

Drug elimination refers to the irreversible removal of drug from the body and covers both excretion (i.e. disappearance of unchanged drug from the body) and biotransformation (the process by which the drug is biochemically converted to a metabolite). Excretion of drugs into sweat and tears is quantitatively unimportant. The concentration of some drugs in saliva parallels that in plasma. Therefore, saliva is sometimes a useful biological fluid to determine drug concentrations. However, this is not a real route of excretion since drugs in saliva are swallowed. The excretion of drugs and toxic compounds in milk is of importance in relation to their potential (toxic) effect in the nursing infant. The fact that molecules can also be excreted through the loss of hair, nails and skin is of toxicological and forensic significance, as sensitive methods can detect traces of, for example, toxic metals in hair (arsenic and mercury). General anesthetics, on the other hand, often leave the body in the expired air. However, the major routes of drug elimination are renal excretion and hepatic biotransformation. This chapter will discuss these two processes, as well as biliary excretion.

A. Excretion

1. *Renal excretion*

The kidney functions as a filter, aiming to clear metabolic products and toxins from the blood and to excrete them through the urine. Efficient clearance is promoted by a high blood flow to the kidneys (20% of the total body blood flow for only 0.5% of the total body weight).

The basic functional unit of the kidney is the nephron (Figure 31.6). Blood arriving in the kidney is first filtered in the glomerulus of the nephron. The primary urine formed by filtration flows from the glomerular capsule through the renal tubule and the collecting ducts. The urine drains from the collecting ducts in the renal pelvis and through the ureters in the bladder. The composition of the urine is unaltered after leaving the nephron. Blood that is not filtered in the glomerulus flows through peritubular capillaries situated along the renal tubule. This allows exchange of molecules between the blood and the fluid in the renal tubule (Figure 31.6).

The renal urinary excretion is the ultimate result of three processes: glomerular filtration, tubular reabsorption and active tubular secretion (Figure 31.6).

FIGURE 31.6 Schematic representation of the renal excretion of drugs.

a. Glomerular filtration

Blood flow to the kidneys is about 1.2–1.5 L/min. About 10% of this volume is filtered through the pores in the glomeruli, which amounts to a filtrate (primary urine) of about 125 mL/min or 180 L per 24 h. The pores of the glomerulo-capillary membrane are sufficiently large to permit passage of small molecules and most drug molecules but do not allow passage of blood cells and of large molecules (> 60 kDa), such as plasma proteins. Therefore, drug molecules bound to plasma proteins are not eliminated by glomerular filtration.

b. Tubular reabsorption

More than 99% of the original 180 L of protein-free filtrate is reabsorbed via the tubular cells; only about 1.5 L (per 24 h) is excreted as final urine. Solutes and drugs dissolved in the filtrate can also be reabsorbed. Glucose, for example, is filtered from the blood but is completely reabsorbed in the renal tubule by carriers in the tubular cells. For different drugs, tubular reabsorption varies from almost absent to almost complete. For most drugs, reabsorption is a passive process (passive diffusion). If the tubular wall is freely permeable for the molecule, more than 99% of the filtered molecule will be reabsorbed passively. Drugs with high lipid solubility, and hence high tubular permeability, are therefore excreted slowly. The drugs diffuse from tubular fluid to plasma in accordance to their concentration gradient, lipid/water partition coefficient, degree of ionization and molecular weight. The pH of the urine varies between 4.5 and 7.0, and changes in pH can influence passive reabsorption and thus the excretion of the drug (see the Henderson–Hasselbalch equation in Section II.A.). Acidifying the urine favors the reabsorption of weak acids, such as salicylates, and retards their excretion, whereas the reverse is true for weak bases. Alkalinization of the urine increases the excretion of weak acids. For example, it is possible to accelerate the excretion of phenobarbital (a weak acid) in an intoxicated patient by administration of sodium bicarbonate. On the other hand, the urinary excretion of weak bases is low in alkaline urine.

Increased urine flow by forced intake of fluids or co-administration of a diuretic drug can increase the excretion of some drugs by decreasing the time for drug reabsorption.

c. Active tubular secretion

Considering most blood (90%) leaves the glomerulus unfiltered, most of the drug delivered to the kidneys reaches the peritubular capillaries. Here, drugs can be transferred to the tubular lumen by relatively non-selective carrier systems. These carriers can transport molecules against a concentration gradient from the blood capillaries across the tubular membranes to the tubular fluid. There are at least two active renal secretion systems: one that normally secretes naturally occurring organic acids (such as uric acid) and one that secretes naturally occurring organic bases (such as choline or histamine). Acid drugs such as penicillins, indomethacin and glucuronides are transported by the first system, whereas the second system transports bases such as morphine, procaine and quaternary ammonium compounds. These transport systems can be saturated, and competition for the active transport systems can occur, leading to desirable or undesirable drug interactions. This characteristic has been used to decrease the urinary excretion of penicillin (and thereby prolonging its effect) by co-administering probenicid, another weak organic acid that competes for the acid transport system in the tubulus. Also, the P-glycoprotein multidrug transporter is present in the brush border of the renal tubules and can play a role in the active tubular secretion of exogenous substances. It is involved in tubular secretion of, for example, digoxin, and can be inhibited by quinidine or verapamil. Co-administration of quinidine therefore decreases the renal clearance (Cl_R) of digoxin, leading to an increase in digoxin serum concentrations.

Plasma protein binding does not limit the rate of active tubular secretion, as the affinity of the drugs is much higher for the carrier than for the plasma proteins. Tubular secretion is potentially the most effective mechanism for the elimination of drugs by the kidney. Penicillin, for example, though about 80% protein bound and therefore slowly cleared by glomerular filtration, is almost completely removed from blood by passage through the kidney due to efficient tubular secretion.

2. *Biliary excretion*

The liver can also be considered an excretory organ. The liver is responsible for the formation of the bile fluid that

drains in the gut and is (at least in part) removed along with the feces. The brown color of the feces is due to pigments of the bile. Bile secreted by the hepatocytes of the liver enters bile canaliculi (narrow intercellular canals) that empty into bile ducts that drain the bile fluid into the gallbladder. The gallbladder stores and concentrates the bile, and when the smooth muscle cells in the gallbladder contract, the bile is delivered in the small intestine.

Each day hepatocytes secrete about 1 L of bile, consisting mostly of water, ions, bile salts (important for the absorption of lipids), cholesterol, and bile pigments. Bile formation by hepatocytes requires the active secretion of inorganic and organic solutes into the canalicular lumen, followed by the movement of water. Other solutes can be carried with this movement of water through the tight junctions between hepatocytes.

Some drugs are actively secreted into the bile and pass as such into the intestine. In humans, the molecular weight threshold for appreciable biliary excretion is in the order of 400–500 Da. In order to be excreted into the bile, drugs usually require a strong polar group. Many drugs excreted into bile are metabolites, often glucuronide conjugates.

A drug (and/or its metabolites) entering the intestine through bile may be excreted in the feces. However, it can also be reabsorbed from the intestine and thus undergo "enterohepatic cycling" (*cf.* Figure 31.2). Drug conjugates, for example, glucuronides, can be hydrolyzed in the gut by bacteria, resulting in the liberation and reabsorption of the parent drug. In particular, this has been found for chloramphenicol and for steroids. These compounds may undergo extensive biliary cycling, with final excretion by the kidney.

B. Biotransformation

As was described above, most drugs need to pass through biological membranes to reach their site of action. Therefore, most drugs have lipophilic characteristics and are only partially ionized at the pH values encountered in the organism. These characteristics also favor reabsorption from the renal tubules after glomerular filtration (see Section V.A.1.c.). As a consequence, renal excretion often plays only a modest role in the total elimination of therapeutic agents from the body. For these compounds, biotransformation into metabolites that are more hydrophilic in nature is essential, because it allows their excretion by the kidneys. Biotransformation reactions take place mainly in the liver (hepatic biotransformation), but can also occur in intestinal mucosa, lungs, kidneys, skin, placenta and plasma. Within a given cell, most of the biotransformation enzymes are found in the endoplasmic reticulum (a network of folded membranes inside the cell) and the cytosol (the fluid inside the cell). When liver tissue (or any other tissue) is homogenized and differential centrifugation is applied, the endoplasmic reticulum of the cells breaks up. Fragments of the endoplasmic reticulum then form microvesicles, called *microsomes*. Therefore, the biotransformation enzymes in the endoplasmic reticulum are often referred to as microsomal enzymes.

For hepatic biotransformation to occur, a drug must enter the hepatocytes that contain the biotransformation enzymes. Polar molecules do this more slowly than nonpolar molecules, except where specific transport mechanisms exist. As a consequence, hepatic metabolism is, in general, more important for lipid-soluble drugs than for polar drugs. Renal excretion and biotransformation can therefore be regarded as two additional and synergistic elimination pathways, guaranteeing an efficient elimination of a whole range of hydrophilic as well as lipophilic substances from the body. Biotransformation usually inactivates a drug, but in some cases metabolites with biological activity or toxic properties are formed. For some drugs, the activity may reside wholly in one or more metabolites: drugs that only become active after biotransformation are termed "prodrugs." Prodrugs are sometimes developed to improve absorption, often because of better lipid solubility than the active metabolite. After absorption, these prodrugs are rapidly converted to the active metabolite in the gut wall or in the liver. An example is pivampicillin, an ester of ampicillin, which is rapidly and completely hydrolyzed to ampicillin during absorption.

Two phases can be distinguished in the pathways of biotransformation. *Phase I* involves addition of functionally reactive groups by oxidation, reduction or hydrolysis. These products are sometimes more chemically reactive, and therefore more toxic than the parent drug. The *Cytochrome P450 (CYP) enzyme family* is involved in most (but not all) Phase I reactions. It comprises a large group of enzymes localized in the endoplasmic reticulum of numerous tissues. The CYP-enzymes are grouped in families denoted by an arabic numeral (e.g. the CYP3 family), within which the amino acid sequence homology is higher than 40%. Each P450 family is further divided into subfamilies denoted by a capital letter (e.g. the CYP3A subfamily), with greater than 55% amino acid sequence homology. Finally, another arabic numeral represents the individual enzyme (e.g. CYP3A4). The main CYP enzymes involved in drug metabolism are: CYP1A2, CYP2A6, CYP2C9, CYP2C19, CYP2D6, CYP2E1 and CYP3A4. CYP3A4 accounts for 30% of total P450 enzyme in the liver and is clinically the most important isoenzyme present in the liver. Nearly 50% of all clinically used medications are metabolized by CYP3A4. Phase II biotransformation consists of conjugation of reactive groups, present either in the parent molecule or after Phase I transformation. Phase II conjugates are usually pharmacologically inactive; they are more hydrophilic than the parent compounds and are easily excreted by the kidneys or the bile. Phase I and Phase II reactions often, though not invariably, occur sequentially. Phenytoin, for example, is first hydroxylated by a Phase I reaction and subsequently conjugated with

glucuronic acid. Enzymes involved in Phase I reactions are located primarily in the endoplasmic reticulum, while the Phase II conjugation enzyme systems are mainly cytosolic. For more in-depth information on biotransformation and on biological factors influencing biotransformation, the reader is referred to Chapter 34 of this textbook. For further discussion on toxic metabolites, see Chapter 35.

VI. SOME PHARMACOKINETIC PARAMETERS AND TERMINOLOGY

A. Plasma concentration–time curve

As was described earlier, a drug often reaches its site of action after it is absorbed into the systemic circulation and distributed to both the target tissue and other tissues. Ideally, the drug concentrations in the target tissue should remain above the minimum effective concentration (the lowest concentration that results in the desired pharmacological response) as long as the pharmacological effect is desired, whereas drug concentrations in all tissues should remain below the minimum toxic concentration (the lowest concentration that results in a toxic effect) at all times. As it is often very difficult (if not impossible) to measure drug concentrations in the target tissue, this technique is rarely used to ascertain whether the drug reaches the target tissue at a proper concentration. Instead, the pharmacokinetics of a drug is assessed by measuring drug concentrations at an alternative and more accessible site – the plasma.

A *plasma concentration–time curve* can be obtained by measuring the drug concentration in plasma samples taken at various time intervals after a drug is administered. These concentrations are then plotted against the corresponding time at which the plasma sample was taken. The profile of such a concentration–time curve is, in fact, determined by the complex interplay between the processes described earlier in this chapter: absorption, distribution and elimination of the drug and, more precisely, by the rate at which these processes occur. Usually, (but not always) absorption, distribution and elimination are assumed to be first-order processes, meaning their rate at all times is proportional to the amount of drug involved (e.g. the rate of first-order absorption is at all times proportional to the amount of drug present at the absorption site). As a consequence, concentration–time curves often show an exponential profile.

When a drug is administered as an i.v. bolus, the entire dose of the drug is injected straight into the blood. Therefore, the absorption process is considered to be completed immediately, and the concentration–time profile of the drug in plasma will be determined by the rate of distribution and elimination. When the distribution of the drug is very fast, the plasma concentration–time curve is determined only by the elimination rate and shows a mono-exponential (first-order) decline (a theoretical example is shown in Figure 31.7a;

FIGURE 31.7 Plasma concentration–time curve after i.v. administration of an imaginary drug with a very high distribution rate: (a) linear scale and (b) semi-log scale. $T_{1/2}$ is the elimination half-life as derived from the plasma concentration–time curve (see Section VI.D).

Figure 31.7b shows the same data in a semi-logarithmic graph). For many drugs, however, distribution occurs more slowly and contributes to the profile of the plasma concentration–time curve. After i.v.-bolus administration of such a drug, the plasma concentration–time curve declines bi-exponentially as the sum of two first-order processes: distribution and elimination. A theoretical example of such a plasma concentration–time profile is presented in Figure 31.8a (on a linear scale) and Figure 31.8b (on a semi-log scale). Note the bi-exponential decline in the plasma concentration in Figure 31.8b (both distribution and elimination contribute to the profile), in contrast to the mono-exponential decline in Figure 31.7b (very fast distribution; therefore only elimination contributes to the profile). The plasma concentration–time curve in Figure 31.8b may be divided into two parts: a distribution phase and an elimination phase. The distribution phase is the initial, more rapid, decline that is mainly due to the distribution of drug from plasma to the tissues. Once equilibration between the drug concentrations in plasma and in tissues has occurred, both plasma and tissue concentrations decline in parallel due to elimination. This decline is often referred to as the elimination phase.

When a drug is not injected directly into the plasma, drug absorption from the site of administration also adds to

FIGURE 31.8 Plasma concentration–time curve after i.v. administration of an imaginary drug for which also distribution adds to the profile: (a) linear scale and (b) semi-log scale.

FIGURE 31.9 Plasma concentration–time curve after oral administration of an imaginary drug: (a) linear scale and (b) semi-log scale.

the profile of the plasma concentration–time curve (besides distribution and elimination). Figure 31.9a shows an idealized example of a plasma concentration–time curve after a single oral administration of a drug. Initially, drug concentration at the absorption site is high, and the rate at which the drug is absorbed into the systemic circulation exceeds its rate of elimination from the body. Therefore, the drug concentration in plasma rises and the drug can be distributed to the tissues. As the drug is absorbed into the systemic circulation, its rate of absorption decreases (due to a decrease of the drug concentration at the absorption site), whereas its rate of elimination increases (due to the increase in plasma concentration). As a consequence, the difference between these rates diminishes. However, as long as the rate of absorption exceeds that of elimination, the plasma concentration continues to rise. This rising portion of the curve is often called the *absorption phase*. At peak concentration, both rates are equal. Afterwards, the rate of drug elimination exceeds the rate of its absorption, and the concentration of drug in both the plasma and the tissues starts to decline. This declining part of the curve is often called the *elimination phase*.

The time to reach the peak plasma concentration (t_{max}) is a rough marker for the rate of drug absorption, whereas the peak plasma concentration (C_{max} or maximum drug concentration in the plasma) itself is related to the dose, the rate of absorption and the rate of elimination. The area under the curve (AUC) is related to the total amount of drug that reaches systemic circulation. Figure 31.9b shows the same data on a semi-log scale.

B. Volume of distribution

The volume of distribution is not a "real" volume. It is a proportionality constant, relating the total amount of drug present in the organism to its plasma concentration at the same moment. It is the fluid volume in which the total amount of drug in the body should be dissolved to give rise to the same concentration as measured in the plasma. This calculated volume does not necessarily correspond to an identifiable physiological volume, and can be much larger than the volume of total body water. It is therefore called "apparent" volume of distribution.

Total body water consists of the body fluids within cells (intracellular fluid) and the fluids outside the body cells (extracellular fluid). The extracellular fluid that fills the narrow spaces between cells of tissues is known as *interstitial*

or intercellular fluid, whereas the extracellular fluid within the blood vessels is termed *plasma*. In a normal 70-kg man total body water volume is about 42 L (or 60% of the body weight), consisting of about 3 L plasma, 11 L interstitial fluid, and 28 L intracellular fluid. If a drug is not bound in plasma or tissues and distributes over total body water, the apparent volume of distribution will be 42 L per 70 kg; this is the case, for example, for antipyrine. If a drug is, likewise, not bound in plasma and tissues, but does not penetrate cells, the distribution will be limited to the extracellular fluid (plasma plus interstitial fluid), equaling 14 L. The apparent volumes of distribution of such drugs approximate their true volume of distribution. However, most substances bind to plasma and tissue proteins. For a drug that preferentially binds to plasma proteins, the plasma concentration will be higher than the concentration in the interstitial and intracellular fluid. In this case, the apparent volume of distribution will be smaller than 42 L. If, however, a drug binds preferentially to tissue proteins, the total drug concentration in plasma will be lower than in tissues and the apparent volume of distribution will be larger than 42 L. A typical example is digoxin, which is highly bound in muscle and has an apparent volume of distribution of about 600 L.

Equation (31.4) describes the relationship between apparent volume of distribution, drug binding and anatomical volumes:

$$V_d = V_p + V_T \frac{f_p}{f_T} \tag{31.4}$$

where V_d is the apparent volume of distribution, V_p is the plasma volume, V_T is the extravascular volume (the sum of the interstitial fluid volume and the intracellular fluid volume) and f_p and f_T are the free fractions of drug in plasma and extravascular space, respectively. The apparent volume of distribution increases with increases in anatomical volumes or tissue binding and decreases with increases in plasma or blood binding.

Many acidic drugs, for example, salicylates, sulfonamides, penicillins and anticoagulants, are highly bound to plasma proteins or are not lipophilic enough to distribute into cells. These drugs, therefore, have small volumes of distribution (<20 L). Basic drugs, on the other hand, often exhibit a large apparent volume of distribution because they tend to be highly distributed to tissues, and their plasma concentrations remain relatively low. Table 31.2 shows the apparent volume of distribution for some drugs.

TABLE 31.2 Apparent Volumes of Distribution for Some Drugs in L/kg

0.11	Warfarin	2.1	Cimetidine
0.14	Ibuprofen	3.9	Propranolol
0.17	Salicylic acid	8.0	Digoxin
0.25	Gentamicin	30.0	Imipramine
0.51	Digitoxin	235	Chloroquine
0.70	Atenolol		

C. Clearance

The term *clearance* describes the process of drug elimination from the body or from a single organ regardless of the mechanism involved. It can be defined as the volume of biological fluid, such as blood or plasma, that would have to be completely freed of the drug per unit of time to account for the elimination. The units for clearance are, therefore, milliliters per minute (mL/min) or liters per hour (L/h). As was described earlier, total elimination of a drug from the body may be a result of processes that occur in the kidney, liver and other organs. Clearance by means of these various organs of elimination is additive. The *systemic* or *total body clearance* (Cl_T) is the sum of these respective organ clearances and considers the entire body as a drug-eliminating system from which many elimination processes may occur. As liver and kidneys are the major organs for drug elimination, we will elaborate on renal and hepatic clearance.

1. *Renal clearance*

The efficiency of the renal excretion of a drug is expressed as "renal clearance" (Cl_R). The Cl_R of a drug is the volume of plasma that is cleared of that drug by the kidneys per time unit. Substances such as inulin and creatinine are eliminated by glomerular filtration, are not subject to either tubular secretion or reabsorption, and are not bound to plasma proteins. Their Cl_R in adults with normal renal function will be around 125 mL/min, corresponding to the volume of plasma that is subjected per minute to glomerular filtration. The clearance for inulin or creatinine can therefore be used as an index of the glomerular filtration rate. For substances that are filtered but also actively secreted, Cl_R is higher than 125 mL/min and can be as high as 650 mL/min, which is the total plasma flow through the kidneys. Such values are found for *para*-aminohippuric acid and penicillins, for example. For drugs that are filtered but also reabsorbed or for drugs bound to plasma proteins, clearance values can be lower than 120 mL/min. Relating the Cl_R of a drug to the glomerular filtration rate can therefore provide information on the mechanisms of renal excretion.

The Cl_R of a drug can be calculated by dividing the amount of drug excreted in the urine over a given time interval by the concentration of the drug in blood or plasma at the time corresponding to the midpoint of the urine collection interval.

2. Hepatic clearance

The hepatic clearance (Cl_H) of a drug can be defined as the volume of blood that is cleared of the drug by the liver per unit of time. The pharmacokinetic concept of Cl_H takes into consideration the anatomical and physiological facts that drugs reach the liver through the portal vein and the hepatic artery and leave it through the hepatic vein. Unbound drugs in plasma diffuse through the liver cell membrane to reach the metabolic enzymes. Therefore, at least three major parameters ought to be considered when quantifying drug elimination by the liver: blood flow through the organ (Q), which reflects transport to the liver; free fraction of the drug in blood (f_u) which affects access of the drug to the enzymes; and intrinsic ability of the hepatic enzymes to metabolize the drug, expressed as intrinsic clearance (Cl'_{int}). Cl'_{int} is the ability of the liver to remove a drug in the absence of flow limitations and blood binding. Taking these three parameters into account, the Cl_H can be expressed by the following equation:

$$Cl_H = Q\left[\frac{f_u \cdot Cl'_{int}}{Q + f_u \cdot Cl'_{int}}\right] \tag{31.5}$$

It is obvious that the Cl_H cannot exceed the total volume of blood reaching the liver per unit of time, that is, the liver blood flow Q. The ratio of the Cl_H of a drug to the hepatic blood flow is called the extraction ratio of the drug (E). The value of the extraction ratio can vary between 0 and 1. It is zero when $f_u Cl'_{int}$ is zero, that is, when the drug is not metabolized in the liver. It is 1 when the Cl_H equals the hepatic blood flow (about 1.5 L/min in humans).

When $f_u Cl'_{int}$ is very small in comparison to hepatic blood flow ($f_u Cl'_{int} < Q$), equation (31.5) reduces to the following equation:

$$Cl_H = f_u \cdot Cl'_{int} \tag{31.6}$$

In that case, clearance is not blood-flow-dependent but depends on enzymatic activity and on plasma protein binding. Binding to plasma proteins will limit the elimination. This is called "restrictive elimination." Drugs with a restrictive elimination have a low extraction ratio (<0.3). Examples are antipyrine, phenytoin and warfarin.

When $f_u Cl'_{int}$ is very large in comparison to hepatic blood flow ($f_u Cl'_{int} > Q$), equation (31.5) reduces to the following equation:

$$Cl_H = Q \tag{31.7}$$

In this case, clearance is dependent on hepatic blood flow and independent of Cl'_{int} and f_u. This is called "blood-flow-dependent" or "non-restrictive" elimination (e.g. nitroglycerin, propranolol and lidocaine). Drugs with a non-restrictive elimination have a high extraction ratio (>0.7), and bound, as well as free, molecules are eliminated, since the affinity of the drug for the hepatic enzymes exceeds its affinity for the plasma proteins. As clearance of non-restrictively eliminated drugs depends on hepatic blood flow, a decrease in hepatic blood flow (e.g. in hepatic disease or cardiac failure) will reduce clearance of these drugs.

D. Elimination half-life ($T_{1/2}$)

The elimination half-life ($T_{1/2}$) is the time it takes for the elimination processes to reduce the plasma concentration or the amount of drug in the body by 50%. Elimination half-life is a composite pharmacokinetic parameter determined by both clearance and volume of distribution (V_d), as described by the following equation:

$$T_{1/2} = 0.7\frac{V_d}{Cl} \tag{31.8}$$

Elimination half-life is increased by an increase in volume of distribution or a decrease in clearance, and vice versa. This is because a decrease in the efficiency of elimination (and therefore in clearance) would, of course, cause an increase in the time needed to reduce the plasma concentration by 50%. On the other hand, the larger the volume of distribution, the more the drug is concentrated in the tissues rather than in the plasma. It is, however, the drug in plasma that is exposed to the elimination mechanisms. Therefore, an increase in volume of distribution also increases elimination half-life. For the simplest cases, elimination half-life may be used to make decisions about drug dosage, and can be derived from the plasma concentration–time profile as the time it takes for a random plasma concentration in the elimination phase to be halved (see Figure 31.7). It does not matter at what concentration half-life is measured, as long as it is measured in the mono-exponential elimination phase of the curve. Therefore, the time for the plasma concentration to drop from 10 to 5 mg/L is the same as from 8 to 4 mg/L or from 2 to 1 mg/L. It becomes more complicated when the plasma concentration follows a multiexponential pattern of decline and two or more half-lives can be calculated. This situation is left out of the discussion in this chapter, but the interested reader can refer to the textbooks on pharmacokinetics mentioned in the bibliography.

E. Bioavailability

Bioavailability is (1) the fraction of an administered dose of a drug that reaches the systemic circulation as intact drug (expressed as F) and (2) the rate at which this occurs. As an i.v. dose is injected directly into the systemic circulation, the bioavailability of an i.v. dose is by definition 100% ($F = 1$). For all other routes of administration, bioavailability is determined by the extent of drug absorption (being the result of both drug uptake from the administration site and possible first-pass effects: see Section III.D.), and varies between

0 and 100% ($0 < F < 1$). For example, orally administered morphine has a bioavailability of about 25% due to significant first-pass metabolism in the liver. Therefore, the dose of morphine given orally is usually 3–5 times larger than an i.v. dose of morphine.

The usual method for measuring bioavailability (also called *absolute bioavailability*) of an oral formulation is to give a group of volunteers an i.v. administration of the drug and the oral formulation on separate occasions and to determine the respective area under the plasma concentration–time curves (AUC). Since the AUC is a measure of the total amount of unaltered drug that reaches the systemic circulation (see Section VI.A.), the bioavailability of the oral formulation can subsequently be determined by comparing these respective AUCs, as described by the following equation:

$$\text{Absolute bioavailability} = F = \frac{\text{AUC}_{\text{oral}}/\text{dose}_{\text{oral}}}{\text{AUC}_{\text{i.v.}}/\text{dose}_{\text{i.v.}}} \quad (31.9)$$

For example, if the AUC_{oral} is 25% of the $\text{AUC}_{\text{i.v.}}$, the bioavailability of the oral formulation is 25% ($F = 0.25$).

Sometimes the bioavailability of a new formulation is not assessed against an i.v. formulation but against another (reference) formulation. This is referred to as measuring *relative bioavailability*, and it provides a measure of the relative performance of two formulations (new formulation A and reference formulation B) in getting the drug absorbed into the systemic circulation (see equation (31.10)).

$$\text{Relative bioavailability} = \frac{\text{AUC}_{\text{A}}/\text{dose}_{\text{A}}}{\text{AUC}_{\text{B}}/\text{dose}_{\text{B}}} \quad (31.10)$$

Obviously, the relative bioavailability of a formulation is not equal to F (the fraction of the dose that reaches the systemic circulation), as the absolute bioavailability of the reference formulation might be quite low due to poor absorption and/or first-pass metabolism.

VIII. VARIABILITY IN PHARMACOKINETICS

When a plasma-concentration curve is constructed for different patients that have been given an identical dose of an identical drug, interindividual differences will be noted. In some cases, plasma concentrations in one patient may remain below the minimal effective concentration, whereas the plasma concentration in another patient reaches the minimum toxic concentration. Besides some very obvious causes, such as body weight and body composition, some other factors involved in the interindividual variability in pharmacokinetics are concisely described below.

A. Genetic factors

Studies on identical and non-identical twins have shown that much interindividual pharmacokinetic variability is determined genetically. Pharmacokinetic variability may be caused by *genetic polymorphism* (the situation where several functionally distinct genes are common in a population) in genes involved in drug absorption, distribution and elimination. In recent years, several polymorphisms in genes encoding for transporter proteins have been described. These polymorphisms could alter the absorption, distribution and elimination of compounds that are substrates for these transporters. However, much work remains to be done to understand the clinical implications of these polymorphisms.

Genetic polymorphism of genes involved in drug metabolism is regarded as one of the major sources of variability in pharmacokinetics. On the other hand, renal excretion of drugs does not appear to be prone to genetic polymorphism. The Cl_{R} for any drug tends to be similar in age- and weight-matched healthy subjects. As a consequence, drugs that are predominantly excreted unchanged, tend to show less interindividual variability than extensively metabolized ones.

B. Age

The main reason for age affecting drug action is that renal excretion is less efficient in neonates and elderly people, so that renally cleared drugs commonly produce stronger and more prolonged effects at the extremes of life. Glomerular filtration rate in the newborn (normalized to body surface area) is only about 20% of the rate in adults, and tubular function is also reduced. Accordingly, elimination half-lives of renally eliminated drugs are longer in newborns than in adults. In babies born at term, renal function increases to values similar to those in young adults in less than a week, and continues to increase to a maximum of approximately twice the adult value at 6 months of age. From about 20 years of age, glomerular filtration rate starts to decline slowly, falling by about 25% at 50 years and by 50% at 75 years.

In the developing child, biotransformation of drugs is also altered. Several drugs that are eliminated primarily through hepatic metabolism have exhibited a higher clearance in children than in adults. A number of factors contribute to these changes during development, such as relative liver size (relative to total body size) and the maturation profile of different drug metabolizing enzymes from birth onwards. In neonates, distinct patterns of development in drug metabolizing enzymes have been observed. Some drug metabolizing enzymes show an onset of activity within hours after birth or in the first week of life; others appear to approach full competence only after several months.

C. Drug interactions

The pharmacokinetics of a drug can be influenced by the concurrent administration of another drug that affects its absorption, distribution, metabolism and/or excretion.

The GI *absorption* of drugs can be influenced by agents with a large surface area upon which the drug can be absorbed, bound or chelated; or it can be influenced by agents that alter GI motility and thereby alter the rate or extent of absorption. Drugs that lower local blood flow can slow down the absorption; the addition of adrenaline to local anesthetic injections results in local vasoconstriction and slower absorption of the anesthetic, thus prolonging its local effect at the injection site.

Drug *distribution* can be altered by competition of drugs for plasma protein binding or by displacement of a drug from tissue binding sites. A competition for plasma protein binding can increase the free fraction and temporarily the free concentration of the drug in plasma, and can therefore result in an increased distribution of the drug toward the biophase, the elimination sites and other tissues. Displacement of a drug from tissue binding sites can induce a temporary increase in the plasma concentration, since it allows redistribution of the displaced drug from the tissue toward the plasma. This increase often increases drug elimination so that a new steady state is reached. For some drugs, the temporary rise in plasma concentration before a new steady state is reached may cause toxicity.

Metabolism is the major factor leading to clinically significant drug–drug interactions. Hepatic microsomal drug metabolizing enzymes can be *induced* (leading to higher production of enzyme) after chronic administration of, for example, phenobarbital, phenytoin, rifampicin or St. John's wort. This produces more rapid metabolism of drugs. A very potent enzyme inducer, such as rifampicin, can markedly alter enzyme activity within 48 h after its administration, while for most inducers the maximal effect is obtained only after 7–10 days. An equal or even longer time after stopping the enzyme inducer is required to dissipate the induction. Hepatic microsomal metabolism may also be *inhibited* by exogenous and endogenous compounds, resulting in a slower rate of metabolism. This enzyme inhibition generally occurs more quickly than enzyme induction and may begin as soon as a sufficient hepatic concentration of the inhibitor is achieved. The most common mechanism is competitive inhibition: any two drugs that are metabolized by the same enzyme may compete with each other for binding to the enzyme, thereby slowing down each other's metabolism. Moreover, some drugs act as a competitive inhibitor for a particular enzyme, although they are not metabolized by that particular enzyme. This is the case in humans for quinidine, which selectively inhibits CYP2D6, although it is not metabolized by that enzyme. Also, non-competitive inhibition can occur. In this case, there is no direct competition between the substrate and the inhibitor for the enzyme, but the inhibitor deactivates the enzyme by binding to other parts of the enzyme. For example, macrolide antibiotics, such as erythromycin, are metabolized by CYP3A4 to a reactive metabolic intermediate that forms a stable, inactive complex with the enzyme.

The renal *excretion* of certain drugs that are weak acids or weak bases may be influenced by other drugs that affect urinary pH. This is due to changes in ionization of the drug and thus to alteration of its lipid solubility and the ability to be absorbed back into the blood from the kidney tubule (see SectionV.A.1.b.). Also, active secretion into the renal tubules can be inhibited by concurrent drug therapy, thus increasing serum drug levels and pharmacologic response (see Section V.A.1.c.). A drug can also affect the rate of renal excretion by altering the protein binding and, hence, filtration.

D. Disease state

Several diseases can cause variations in pharmacokinetics. Renal or hepatic insufficiency predisposes to toxicity by causing intense or prolonged drug effects as a result of increased plasma levels following a standard dose regimen. Drug absorption is slowed in conditions causing gastric stasis (e.g. migraine) and may be incomplete in patients with diarrhea or with malabsorption due to diseases of the pancreas or gut or due to edema of ileal mucosa. Nephrotic syndrome is characterized by a heavy loss of proteins in urine (proteinuria) and thus a reduced concentration of albumin in plasma and edema. Edema of intestinal mucosa alters drug absorption, while changes in binding to plasma albumin changes drug disposition. An impaired functioning of the blood–brain barrier occurs in meningitis. Hypothermia (lowered body temperature, often in elderly persons) markedly reduces the clearance of many drugs.

E. Pregnancy

Pregnancy is associated with numerous alterations in physiology that can influence pharmacokinetics. Plasma concentration of albumin is diminished in the mother, resulting in alteration in drug–protein binding. The increase in glomerular filtration, aimed to help excrete the increased amount of waste products, also results in an increased renal elimination of drugs. The blood of the mother and the fetus is separated by the placental barrier (*cf.* blood–brain barrier). This barrier could allow some drugs to be administered to the mother without influencing the fetus. The placental barrier is, however, rapidly crossed by lipophilic molecules that can have effects on the fetus. For example, some drugs are known to cause abnormal development of the fetus (teratogenic effect). When transferred to the fetus, drugs are

usually slowly eliminated by the fetus; drug metabolizing enzymes in the fetal liver are less active than in adults, and elimination through fetal kidney is not efficient since fetal urine drains in the amniotic fluid, which is swallowed by the fetus.

BIBLIOGRAPHY

The following textbooks and reviews were used to compile this chapter. These references can offer a good start for the interested reader looking for more elaborate and/or more in depth information.

Textbooks on pharmacokinetics:

1. Birkett, D. J. *Pharmacokinetics Made Easy*, Revised edition. McGraw-Hill: Australia, 2002.
2. Rowland, M., Tozer, T. N. *Clinical Pharmacokinetics: Concepts and Applications*, 3rd Edition. Lippincott Williams & Wilkins: Philadelphia, PA, 1995.
3. Shargel, L., Wu-Pong, S., Yu, A. B. C. *Applied Biopharmaceutics and Pharmacokinetics*, 5th Edition. The McGraw-Hill Companies: New York, 2005.
4. Washington, N., Washington, C., Wilson, C. G. *Physiological Pharmaceutics, Barriers to Drug Absorption*, 2nd Edition. Taylor and Francis: London, 2001.

General textbooks on physiology and pharmacology:

5. Boron, W. F., Boulpaep, E. L. *Medical Physiology*, Updated Edition. Elsevier Saunders, Philadelphia: Philadelphia, PA, 2005.
6. Ganong, W. F. *Review of Medical Physiology*, 22nd Edition. McGraw-Hill: New York, 2005.
7. Hardman, J. G., Limbird, L. E., Molinoff, P. B., Ruddon, R. W. *Goodman & Gilman's The Pharmacological Basis of Therapeutics*, 9th Edition. McGraw-Hill: New York, 1996.
8. Rang, H. P., Dale, M. M., Ritter, J. M., Flower, R. *Rang & Dale's Pharmacology*, 6th Edition. Churchill Livingstone: New York, 2007.
9. Tortora, G. J., Derrickson, B. *Principles of Anatomy and Physiology*, 11th Edition. John Wiley & Sons: Chichester, 2006.

Reviews on selected topics:

10. Beringer, P. M., Slaughter, R. L. Transporters and their impact on drug disposition. *Ann. Pharmacother.* 2005, *39*, 1097–1108.
11. Daugherty, A. L., Mrsny, R. J. Transcellular uptake mechanisms of the intestinal epithelial barrier – Part one. *PSTT* 1999, *2*(4), 144–151.
12. Mizuno, N., Niwa, T., Yotsumoto, Y., Sugiyama, Y. Impact of drug transporter studies on drug discovery and development. *Pharmacol. Rev.* 2003, *55*(3), 425–461.
13. Zhang, Y., Benet, L. Z. The gut as a barrier to drug absorption. *Clin. Pharmacokinet.* 2001, *40*(3), 159–168.

Chapter 32

Biotransformation Reactions and Their Enzymes

Bernard Testa

I. INTRODUCTION
II. FUNCTIONALIZATION REACTIONS
A. Enzymes catalyzing functionalization reactions
B. Reactions of carbon oxidation and reduction
C. Oxidation and reduction of N- and S-containing moieties
D. Reactions of hydration and hydrolysis
III. CONJUGATION REACTIONS
A. Introduction
B. Methylation
C. Sulfonation
D. Glucuronidation
E. Acetylation
F. Conjugation with coenzyme A and subsequent reactions
G. Conjugation reactions of glutathione
IV. BIOLOGICAL FACTORS INFLUENCING DRUG METABOLISM
V. CONCLUDING REMARKS
REFERENCES

"On fait la science avec des faits comme une maison avec des pierres; mais une accumulaiton des faits n'est pas plus une science qu'un tas de pierres n'est une maison."

"Science is made of facts as a house with stones; but an accumulation of facts is no more a science than a heap of stones is a house."
Henri Poincaré (La Science et l'Hypothèse)

I. INTRODUCTION

Drug metabolism, and more generally xenobiotic metabolism, has become a major pharmacological science with particular relevance to biology, therapeutics and toxicology. As such, it is also of great importance in medicinal chemistry since it influences the deactivation, activation, detoxification and toxification of most drugs. This chapter is therefore written for medicinal chemists and will offer structured knowledge rather than encyclopedic information. Readers who wish to go further in the study of xenobiotic metabolism may consult various books[1–16] and reviews.[17–24]

The metabolism of drugs and other xenobiotics is often a biphasic process in which the compound may first undergo a functionalization reaction (phase I reaction) of oxidation, reduction or hydrolysis. This introduces or unveils a functional group such as a hydroxyl or amino group suitable for coupling with an endogenous molecule or moiety in a second metabolic step known as a conjugation reaction (phase II reaction). In a number of cases, phase I metabolites may be excreted prior to conjugation, while many xenobiotics can be directly conjugated. Furthermore, reactions of functionalization may follow some reactions of conjugation, e.g. some conjugates are hydrolyzed and/or oxidized prior to their excretion.

The major function of xenobiotic metabolism is the elimination of physiologically useless compounds, some of which may be harmful; witness the many toxins produced by plants. This evolutionary function justifies the designation of detoxification originally given to reactions of xenobiotic metabolism. However, numerous xenobiotics and even a marked number of drugs are known to yields toxic

Copyright © 2008, Elsevier Ltd
All rights reserved.

FIGURE 32.1 The metabolism of propranolol (**1**) in humans, accounting for more than 90% of the dose; Gluc = glucuronide(s); Sulf = sulfate(s).[25]

metabolites, a situation known as toxification (see Chapter 35). In pharmacological terms, a drug may or may not have active metabolites. The former case is rather frequent, especially with phase I metabolites. Drugs that do not yield any active metabolite are, for example, the soft drugs. The reverse case is that of prodrugs (see Chapter 38), namely, inactive therapeutic agents whose clinical activity is due to an active metabolite.

The present chapter aims at laying the foundations of drug metabolism by offering an analytical view of the field. In other words, the focus will be on metabolic reactions, the target groups they affect, and the enzymes by which they are catalyzed. Further information can be found in the accompanying Chapters 33, 35 and 38. Also, the analytical information presented here needs to be complemented by a synthetic view as provided by metabolic schemes. These show at a glance the competitive and sequential reactions undergone by a given drug, and they bring logic and order in what may appear as a random presentation. As an example of a metabolic scheme, Figure 32.1 presents the biotransformation of propranolol (**1**) in humans.[25] There are relatively few studies as comprehensive and clinically relevant as this one, which remains as current today as it was when published in 1985. Indeed, over 90% of a dose was accounted for and consisted mainly in products of oxidation and conjugation. The missing 10% may represent other minor and presumably quite numerous metabolites, e.g. those resulting from ring hydroxylation at other positions or from the progressive breakdown of glutathione (GSH) conjugates.

II. FUNCTIONALIZATION REACTIONS

Reactions of functionalization comprise oxidations (electron removal, dehydrogenation, and oxygenation), reductions (electron addition, hydrogenation, and removal of oxygen), and hydrations/dehydrations (hydrolysis and addition or removal of water). The reactions of oxidation and reduction are catalyzed by a very large variety of oxidoreductases,

TABLE 32.1 A Survey of Oxidoreductases Playing a Role in Drug Metabolism (listed in the order they appear in the text).[21,23,26,30]

Enzymes	EC numbers	*Gene root* (or *gene*) and major human enzymes
Cytochrome P450	Mainly EC 1.14.13 and 1.14.14.1	*CYP* (see Table 34.2)
Flavin-containing monooxygenases	EC 1.14.13.8	*FMO* (FMO1–FMO5)
Monoamine oxidases	EC 1.4.3.4	*MAO* (MAO-A and MAO-B)
Copper-containing amine oxidases	EC 1.4.3.6	*AOC* (DAO and SSAO)
Aldehyde oxidase	EC 1.2.3.1	*AOX1* (AO)
Xanthine oxidoreductase	EC 1.17.1.4 and 1.17.3.2	*XOR* (XDH and XO)
Various peroxidases	EC 1.11.1.7 and 1.11.1.8	e.g. *EPO* (EPO), *MPO* (MPO) and *TPO* (TPO)
Prostaglandin G/H synthase	EC 1.14.99.1	*PTGS* (COX-1 and COX-2)
Alcohol dehydrogenases	EC 1.1.1.1	*ADH* (ADH1A, 1B and 1C, ADH4, ADH5, ADH6 and ADH7)
Aldehyde dehydrogenases	EC 1.2.1.3 and 1.2.1.5	*ALDH* (e.g. ALDH1A1, 1A2 and 1A3, 1B1, 1B2, 3A1, 3A2, 3B1, 3B2, 8A1 and 9A1)
Aldo-keto reductases	In EC 1.1.1 and 1.3.1	*AKR* (e.g. ALR1, ALR2, DD1, DD2, DD3, DD4, AKR7A2, 7A3 and 7A4)
Carbonyl reductases	EC 1.1.1.184	*CBR* (CR1, CR3)
Quinone reductases	EC 1.6.5.2 and 1.10.99.2	*NQO* (NQO1 and NGO2)

while various hydrolases catalyze hydrations. A large majority of the enzymes involved in xenobiotic functionalization are briefly reviewed below before we examine in some detail the metabolic reactions and pathways of functionalization.

A. Enzymes catalyzing functionalization reactions

1. *Monooxygenases*

Monooxygenation reactions are of major significance in drug metabolism and are mediated mainly by two enzymes that differ markedly in their structure and properties, namely the cytochromes P450 (CYPs) and the flavin-containing monooxygenases (FMOs) (see Table 32.1).[7,8,17–19,21,23,26–31] Among these, the most important as far as xenobiotic metabolism is concerned are the CYPs, a very large group of enzymes belonging to heme-coupled monooxygenases. Cytochrome P450 is the major drug-metabolizing enzyme system, playing a key role in detoxification and toxification and is of additional significance in medicinal chemistry because several CYP enzymes are drug targets, e.g. thromboxane synthase (CYP5A1) and aromatase (CYP19A1). The CYPs are encoded by the *CYP* gene superfamily and are classified in families and subfamilies. A total of 57 human *CYP* genes are known to date. The three CYP families mostly involved in xenobiotic metabolism are CYP1–CYP3 (Table 32.2). The relative importance of the major human CYPs is given in Table 32.3. The endobiotic-metabolizing CYPs are in families 4, 5, 7, 8, 11, 17, 19–21, 24, 26, 27, 39, 46 and 51.

An understanding of the regiospecificity and broad reactivity of cytochrome P450 requires a presentation of its catalytic cycle (Figure 32.2). The enzyme in its ferric (oxidized) form exists in equilibrium between two spin states, a hexacoordinated low-spin form whose reduction requires a high-energy level, and a pentacoordinated high-spin form. Binding of the substrate to enzyme induces a shift to the reducible high-spin form (reaction **a**). The first electron then enters the enzyme–substrate complex (reaction **b**), reducing the enzyme to its ferrous form, which has a high affinity for diatomic gases such as CO (a strong inhibitor of cytochrome P450) and dioxygen (reaction **c**). The cycle continues with a second electron entering *via* either F_{P1} or F_{P2} and reducing the ternary complex (reaction **d**). Electron transfer

TABLE 32.2 The Human *CYP* gene Superfamily: A Table of the Families and Subfamilies of Gene Products.

Families	Subfamilies	Human gene products
CYP 1 Family (aryl hydrocarbon hydroxylases; xenobiotic metabolism; xenobiotic-inducible)		
	CYP 1A	CYP1A1, CYP1A2
	CYP 1B	CYP1B1
CYP 2 Family (xenobiotic metabolism; constitutive and xenobiotic-inducible)		
	CYP 2A	CYP2A6, CYP2A7, CYP2A13
	CYP 2B	CYP2B6
	CYP 2C	CYP2C8, CYP2C9, CYP2C18, CYP2C19
	CYP 2D	CYP2D6
	CYP 2E	CYP2E1
	CYP 2F	CYP2F1
	CYP 2J	CYP2J2
	CYP 2R	CYP2R1
	CYP 2S	CYP2S1
	CYP 2U	CYP2U1
	CYP 2W	CYP2W1
CYP 3 Family (xenobiotic and steroid metabolism; steroid-inducible)		
	CYP 3A	CYP3A4, CYP3A5, CYP3A7 (fetal CYP enzyme), CYP3A43

TABLE 32.3 Levels and Variability of Human CYP Enzymes Involved in Drug Metabolism.[17]

CYP	Level in liver (percent of total)	Variability range	Percent of drugs being substrates
1A1			3
1A2	*ca.* 13	*ca.* 40-fold	10
1B1	<1		1
2A6	*ca.* 4	*ca.* 30- to 100-fold	3
2B6	<1	*ca.* 50-fold	4
2C	*ca.* 18	25- to 100-fold	25
2D6	up to 2.5	>1000-fold	15
2E1	up to 7	ca. 20-fold	3
3A4	up to 28	ca. 20-fold	36

within the ternary complex generates bound peroxide anion (O_2^{2-}). The bound peroxide anion is split by the addition of two protons, liberating H_2O (reaction **e**). The remaining oxygen atom is equivalent to an oxene species, a neutral electrophilic species having only six electrons in its outer shell, stabilized by electron density from the iron. This is the reactive form of oxygen that will attack the substrate. The binary enzyme–product complex dissociates, thereby regenerating the initial state of cytochrome P450 (reaction **f**).

FIGURE 32.2 Catalytic cycle of cytochrome P450 associated with monooxygenase reactions. [Fe^{3+}] = ferricytochrome P450; hs = high spin; ls = low spin; [Fe^{2+}] = ferrocytochrome P450; F_{P1} = flavoprotein 1 = NADPH-cytochrome P450 reductase; F_{P2} = NADH-cytochrome b_5 reductase; cyt b_5 = cytochrome b_5; XH = substrate (modified from[23]).

TABLE 32.4 A Survey of Hydrolases Playing a Role in Drug Metabolism (listed in the order they appear in the text).[12,22,24]

Classes of hydrolases	Examples of enzymes (with some *gene roots* and human enzymes)
EC 3.1.1: Carboxylic ester hydrolases	EC 3.1.1.1: Carboxylesterase (*CES*) CES1A1, CES2, CES3 EC3.1.1.2: Arylesterase (*PON*, see 3.1.8.1) EC 3.1.1.8: Cholinesterase (*BCHE*)
EC 3.1.2: Thiolester hydrolases	EC 3.1.2.20: Acyl-CoA hydrolase
EC 3.1.3: Phosphoric monoester hydrolases	EC 3.1.3.1: Alkaline phosphatase (*ALP*) EC 3.1.3.2: Acid phosphatase (*ACP*)
EC 3.1.6: Sulfuric ester hydrolases	EC 3.1.6.1: Arylsulfatase
EC 3.1.8: Phosphoric triester hydrolases	EC 3.1.8.1: Paraoxonase (*PON*) PON1, PON2, PON3 EC 3.1.8.2: Diisopropyl-fluorophosphatase
EC 3.2: Glycosylases	EC 3.2.1.31: β-Glucuronidase (*GUSB*)
EC 3.3.2: Ether hydrolases	EC 3.3.2.9: Microsomal epoxide hydrolase (*EPHX1*) mEH EC 3.3.2.10: Soluble epoxide hydrolase (*EPHX2*) sEH
EC 3.4.11: Aminopeptidases	EC 3.4.11.1: Leucyl aminopeptidase (*LAP*)
EC 3.4.13 and 3.4.14: Peptidases acting on di- and tri-peptides	EC 3.4.14.5: Dipeptidyl-peptidase IV (*DPP4*)
EC 3.4.16–3.4.18: Carboxypeptidases	EC 3.4.16.2: Lysosomal Pro-Xaa carboxypeptidase EC 3.4.17.1: Carboxypeptidase A (*CPA*)
EC 3.4.21–3.4.25: Endopeptidases	EC 3.4.21.1: Chymotrypsin (*CTRB*) EC 3.4.24.15: Thimet oligopeptidase (*THOP*)
EC 3.5.1: Hydrolases acting on linear amides	EC 3.5.1.4: Amidase EC 3.5.1.39: Alkylamidase
EC 3.5.2: Hydrolases acting on cyclic amides	EC 3.5.2.1: Barbiturase EC 3.5.2.2: Dihydropyrimidinase (*DPYS*) EC 3.5.2.6: β-Lactamase

2. *Other oxidoreductases*

Other oxidoreductases that can play a major or less important role in drug oxidation are (Table 32.1):

- Monoamine oxidases, which are essentially mitochondrial enzymes.
- The broad group of copper-containing amine oxidases, which contain diamine oxidase (DAO) and semicarbazide-sensitive amine oxidases (SSAO).
- The cytosolic molybdenum hydroxylases, namely aldehyde oxidase and xanthine oxidoreductase, which exist in a dehydrogenase form (XDH) and an oxidase form (XO).
- Various peroxidases, such as eosinophil peroxidase (EPO), myeloperoxidase (MPO) and thyroid peroxidase (TPO) (note that several cytochrome P450 enzymes have been shown to have peroxidase activity).
- Prostaglandin G/H synthase, which is able to use a number of xenobiotics as cofactors in a cooxidation reaction.

Dehydrogenases/reductases involved in reactions of oxidation (dehydrogenation) and/or reduction (hydrogenation) are:

- Alcohol dehydrogenases (ADH), which are zinc enzymes found in the cytosol of the mammalian liver and in various extrahepatic tissues.
- Aldehyde dehydrogenases (ALDH), a large superfamily of enzymes produced by 19 human genes in 11 families and 13 subfamilies.
- The aldo-keto reductases (AKR), a complex superfamily of enzymes, which includes aldehyde reductases (ALR) and dihydrodiol dehydrogenase (DD).
- Carbonyl reductases (CR) and quinone reductases (NQO).

3. *Hydrolases*

Hydrolases constitute a very complex ensemble of enzymes, many of which are known or suspected to be involved in xenobiotic metabolism (Table 32.4).[12,22,24]

Relevant enzymes among the serine hydrolases include carboxylesterases, arylesterases, cholinesterase and a number of serine endopeptidases [EC 3.4.21]. The role of arylsulfatases, phosphatases, β-glucuronidases, epoxide hydrolases and some endopeptidases is also significant.

B. Reactions of carbon oxidation and reduction

As is usual, we distinguish here between reactions targeting sp^3-, sp^2- and sp-carbon atoms.[21,23]

1. *sp^3-Carbon atoms*

The most important reactions of oxidation of sp^3-carbon atoms are schematized in Figure 32.3. In the simplest cases, a non-activated carbon atom in an alkyl group undergoes CYP-catalyzed hydroxylation (Figure 32.3a). The penultimate position is a preferred site of attack, but hydroxylation can also occur at the terminal position or at another position in case of steric hindrance or with some highly regiospecific cytochromes P450. Dehydrogenation by dehydrogenases can then yield a carbonyl derivative that is either an aldehyde or a ketone. Note that these reactions may involve not only metabolites, but also xenobiotic alcohols and are reversible since dehydrogenases catalyze the reactions in both directions. And while a xenobiotic ketone is very seldom oxidized further, aldehydes are good substrates for aldehyde dehydrogenases and lead irreversibly to carboxylic acid metabolites. A classical example is that of ethanol, which in the body exists in redox equilibrium with acetaldehyde, this metabolite being rapidly and irreversibly oxidized to acetic acid.

There is a known regioselectivity in CYP-catalyzed hydroxylations for carbon atoms adjacent (alpha) to an unsaturated system or a heteroatom such as N, O or S. In the former cases (Figure 32.3b), hydroxylation can easily be followed by dehydrogenation (not shown). In the latter

FIGURE 32.3 Major reactions of oxidation involving an sp^3-carbon in substrate molecules (a): CYP-catalyzed hydroxylation of an alkyl group, followed by reversible dehydrogenation to a carbonyl, and followed for aldehydes by irreversible oxidation to a carboxylic acid. (b): The priviledged CYP-catalyzed hydroxylation of benzylic, allylic, and propargylic positions. (c): CYP-catalyzed hydroxylation alpha to a heteroatom leads to spontaneous C-heteroatom cleavage (i.e. N-dealkylation or deamination, O-dealkylation and S-dealkylation). (d): CYP-catalyzed reactions of oxidative dehalogenation.

cases (Figure 32.3c), however, the hydroxylated metabolite is usually unstable and undergoes a rapid, post-enzymatic elimination. Depending on the substrate, this pathway produces a secondary or primary amine, an alcohol or a thiol, while the alkyl group is released as an aldehyde or a ketone. Such reactions of deamination and N-dealkylation constitute a very common and frequent pathway as far as drug metabolism is concerned, since it underlies some well-known metabolic reactions of N-C cleavage discussed later. Reactions of O- and S-dealkylation occur readily, but are less frequent than N-dealkylations due to a limited occurrence of target groups in drug molecules.

Aliphatic carbon atoms bearing one or more halogen atoms (mainly chlorine or bromine) can be similarly metabolized by hydroxylation and loss of HX to dehalogenated products (Figure 32.3d). In all of the CYP-catalyzed carbon-hydroxylations in Figure 32.3, the iron-bound oxene (see above) acts by a mechanism known as oxygen rebound, whereby an H atom is exchanged for an OH group. In simplified terms, the oxene atom attacks the substrate by cleaving a C-H bond and removing the hydrogen atom (hydrogen radical). This forms an iron-bound $HO^{\bullet}$ species and leaves the substrate as a C-centered radical. In the last step, the iron-bound $HO^{\bullet}$ species is transferred to the substrate.

Dehalogenation reactions can also proceed reductively or without change in the oxidation state. The former reactions involve replacement of a halogen by a hydrogen *via* an intermediate radical, which may have toxicological significance. Carbon tetrachloride (**2**) offers a telling example of the metabolic fate of halogenated compounds. Indeed, this toxic compound undergoes CYP-catalyzed reductive dehalogenation whose mechanism is schematized in Figure 32.4. Due to its electronegativity, the substrate reoxidizes the ferroheme after its first reduction. The intermediate carbon tetrachloride anion so formed immediately breaks down by loss of a chloride anion and formation of the reactive trichloromethyl radical (**3**). One of the mechanisms of toxicity of the latter is by formation of a peroxyl radical (**4**), which causes oxidative stress and lipid peroxidation.

2. *sp^2- and sp-Carbon atoms*

Reactions at sp^2-carbons are characterized by their own pathways, catalytic mechanisms, and products (Figure 32.5). Thus, the oxidation of aromatic rings generates a variety of (usually stable) metabolites. Their common precursor is often a reactive epoxide (Figure 32.5a), which can either be hydrolyzed by epoxide hydrolase to a dihydrodiol or rearrange under proton catalysis to a phenol. The production of a phenol is a very common metabolic reaction for drugs containing one or more aromatic rings, e.g. the metabolism of propranolol (Figure 32.1). The *para*-position is the preferred position of hydroxylation for unsubstituted phenyl rings, but the regioselectivity of the reaction becomes more complex with substituted phenyl or with other aromatic rings.

Dihydrodiols are seldom observed, as are catechol metabolites produced by their dehydrogenation, catalyzed by dihydrodiol dehydrogenase. The further oxidation of phenols and phenolic metabolites to a catechol or hydroquinone is also possible, the rate of reaction and the nature of products depending on the ring and on the nature and position of its substituents. In a few cases, catechols and hydroquinones have been found to undergo further oxidation to quinones by two single-electron steps. The intermediate in this reaction is a semiquinone. Both quinones and semiquinones are reactive, in particular toward biomolecules, and have been implicated in many toxification reactions. For example, the high toxicity of benzene in bone marrow is believed to be due to the oxidation of catechol and hydroquinone catalyzed by myeloperoxidase.

The oxidation of diphenols to quinones is reversible, a variety of cellular reductants are able to mediate the reduction of quinones either by a two-electron mechanism or by two single-electron steps. The two-electron reduction can be catalyzed by carbonyl reductase and quinone reductase, while cytochrome P450 and some flavoproteins act by single-electron transfers. The non-enzymatic reduction of quinones can occur, for example, in the presence of $O_2^{\bullet -}$ or some thiols such as GSH.

FIGURE 32.4 The CYP-catalyzed reductive dehalogenation of carbon tetrachloride (**2**) leading to the reactive trichloromethyl radical (**3**). The latter reacts, among others, with molecular oxygen to form a peroxyl radical (**4**).

FIGURE 32.5 Major functionalization reactions involving an sp^2- or sp-carbon in substrate molecules. These reactions are oxidations (oxygenations and dehydrogenations), reductions (hydrogenations) and hydrations, plus some postenzymatic rearrangements. They have as target sites aromatic rings (a), carbon–carbon double bonds (b) and carbon–carbon triple bonds (c).

Olefinic bonds in xenobiotic molecules can also be targets of cytochrome P450-catalyzed epoxidation (Figure 32.5b). In contrast to arene oxides, the resulting epoxides are fairly stable and can be isolated and characterized. But like arene oxides, they are substrates of epoxide hydrolase to yield dihydrodiols. This is exemplified by carbamazepine, whose 10,11-epoxide is a major and pharmacologically active metabolite in humans, and is further metabolized to the inactive dihydrodiol.[32]

The few drugs that contain an acetylenic moiety are also targets for cytochrome P450-catalyzed oxidation (Figure 32.5c). Oxygenation of the C-C triple bond yields an intermediate, which, depending on the substrate, can react in a number of ways, for example, binding covalently to the enzyme or forming a highly reactive ketene whose hydration produces a substituted acetic acid.

C. Oxidation and reduction of N- and S-containing moieties

The main metabolic reactions of oxidation and reduction of N- or S-containing functional groups in organic molecules are summarized in Figure 32.6. The reactions of oxidation are catalyzed mainly by cytochromes P450 and/or flavin-containing monooxygenases, whereas the enzymes catalyzing reductions are not always characterized and can be CYPs, NADPH-CYP reductase, and a variety of other reductases.[16,21,23]

Nitrogen oxygenation is a straightforward metabolic reaction of tertiary amines (Figure 32.6b), both aliphatic and aromatic. Numerous drugs undergo this reaction; the resulting *N*-oxide metabolite is more polar and hydrophilic than the parent compound. Identical considerations apply to pyridines and analogous aromatic azaheterocycles. These reactions can be reversible. Secondary and primary amines (Figure 32.6b and c) also undergo N-oxygenation, the first isolable metabolites are hydroxylamines. Again, reversibility is documented. These compounds can be aliphatic or aromatic amines, and the same metabolic pathway occurs in secondary and primary amides while tertiary amides appear resistant to N-oxygenation. The oxidation of secondary amines and amides usually stops at the hydroxylamine-hydroxylamide level. As opposed to secondary amines, primary amines (Figure 32.6c) can sometimes be further oxidized to nitroso metabolites, but oxidation of the latter metabolites to nitro compounds does not seem to occur *in vivo*. In contrast, aromatic nitro compounds are readily reduced to primary amines as shown. This is the case for numerous chemotherapeutic drugs such as metronidazole.[23]

A limited number of drugs contain a sulfur atom. Thus, thiol compounds (Figure 32.6d) can be oxidized to sulfenic acids, sulfinic acids, and finally to sulfonic acids. Depending on the substrate, the pathway is mediated

FIGURE 32.6 Major functionalization reactions involving N- or S-containing moieties in xenobiotics. The reactions shown here are mainly oxidations (oxygenations and dehydrogenations) and reductions (deoxygenations and hydrogenations).

by CYP and/or FMO. Another route of oxidation of thiols is to disulfides, usually by dehydration between a thiol and a sulfenic acid. The metabolism of sulfides (thioethers) is rather straightforward (Figure 32.6e). Besides S-dealkylation reactions discussed earlier, these compounds can also be oxygenated by monooxygenases to sulfoxides and then to sulfones. Sulfoxides can be reduced to sulfides, whereas the formation of sulfones is irreversible.

Thiocarbonyl compounds (Figure 32.6f) are also substrates of monooxygenases, forming S-monoxides (sulfines) and then S-dioxides (sulfenes). As a rule, these metabolites cannot be identified as such due to their reactivity. Thus, S-monoxides rearrange to the corresponding carbonyl by expelling a sulfur atom, a reaction known as oxidative desulfuration and occurring in thioamides and thioureas (e.g. thiopental). As for the *S*-dioxides, they are strong electrophiles that react very rapidly with nucleophiles and particularly with nucleophilic sites in biological macromolecules. This covalent binding results in the formation of adducts of toxicological significance. Such a mechanism is believed to account for the carcinogenicity of a number of thioamides.

Some other elements besides carbon, nitrogen and sulfur can undergo metabolic redox reactions. The direct oxidation of oxygen atoms in phenols and alcohols is well documented for some substrates. Thus, the oxidation of secondary alcohols by some peroxidases can yield a hydroperoxide and ultimately a ketone. Some phenols are known to be oxidized by cytochrome P450 to a semiquinone and ultimately to a quinone. A classical example is that of the anti-inflammatory drug acetaminophen, a minor fraction of which is oxidized by CYP2E1 to a highly reactive and toxic quinone imine,[21] Additional elements of limited significance in medicinal chemistry able to enter redox reactions are silicon, phosphorus, arsenic and selenium, among others.

D. Reactions of hydration and hydrolysis

The two terms of hydrolysis and hydration both imply bond breakage with addition of a molecule of H_2O. In this text, we prefer to apply the term hydrolysis to the

FIGURE 32.7 Examples of hydrolytic reactions. (a): The diester cocaine (**5**). (b): The diimide thalidomide (**6**).

cleavage of esters (carboxylesters, lactones, inorganic esters), amides (e.g. carboxamides, sulfamates, phosphoamides and lactams) and glycosides. In contrast, the term hydration will be restricted to epoxides, although the enzymes catalyzing this reaction are also classified as hydrolases (see Table 32.4). More extensive treatments of hydrolases and their metabolic reactions can be found elsewhere.[12,22,24]

1. *Esters and amides*

The hydrolysis of esters and amides (including peptides) is important in medicinal chemistry (see Chapters 36 and 38). Here, two drugs are presented to illustrate these two chemical classes. (-)-*Cocaine* (**5**) has two ester groups whose hydrolysis (Figure 32.7a) is a route of detoxification, which accounts for as much as 90% of the dose in humans.[12,33,34] Three human enzymes are now known to be involved in the hydrolysis of cocaine. One is the liver carboxylesterase hCE-1 which catalyzes the hydrolysis of the methyl ester group. As for the benzoyl ester goup, it is hydrolyzed by the liver carboxylesterase hCE-2 and serum cholinesterase.

Among amides, we turn our attention to a cyclic analog, namely the infamous thalidomide (**6**), which bears two imide rings. Parallel to rapid inversion of configuration and very low rates of hydroxylation, thalidomide is rapidly hydrolyzed to ring-opened products (Figure 32.7b).[35,36] All four imide bonds of the molecule are susceptible to hydrolytic cleavage at pH > 6, and the reactions are non-enzymatic and base-catalyzed. The two main urinary metabolites in humans are shown here, each accounting for about 30–50% of a dose.

2. *Epoxides*

The overall reaction catalyzed by epoxide hydrolases (Table 32.4) is the addition of a water molecule to an epoxide. Together with GSH conjugation, hydration is a major pathway in the inactivation and detoxification of arene oxides. As a rule, these are good substrates of microsomal epoxide hydrolase yielding *trans*-dihydrodiols (Figure 32.5a). In phenyl and naphthyl rings, the proton-catalyzed isomerization of epoxides to phenols is an extremely fast reaction which markedly reduces the likelihood of the epoxide from being hydrated by epoxide hydrolase. This chemical instability decreases for chemicals with three or more fused rings, but such compounds are no longer of medicinal interest. Yet despite the high reactivity of benzene epoxides, the characterization of a dihydrodiol metabolite has been achieved for a limited number of phenyl-containing drugs, and particularly for neurodepressant drugs, such as hypnotics (e.g. glutethimide) and antiepileptics (e.g. ethotoin and phenytoin).

Alkene oxides are generally quite stable chemically, indicating a much reduced chemical reactivity compared to arene oxides. Under physiologically relevant conditions, they have little capacity to undergo rearrangement reactions and are resistant to uncatalyzed hydration. In contrast, they are often good substrates of epoxide hydrolases yielding diols (Figure 32.5b). A well-known example is that of the anticonvulsant drug carbamazepine. Thus, the 10,11-epoxide and the 10,11-dihydrodiol are urinary metabolites in humans and rats given the drug.

III. CONJUGATION REACTIONS

A. Introduction

Conjugation reactions (also known as phase II reactions) are of critical significance in the metabolism of endogenous compounds, as witnessed by the impressive battery of enzymes that have evolved to catalyze them. Conjugation is also of great importance in the biotransformation of

xenobiotics, involving parent compounds or metabolites thereof.[19,22,37] Conjugation reactions of xenobiotics and their metabolites are characterized by a number of criteria:

A) Substrates are coupled covalently to an endogenous molecule or moiety …
B) … which is polar (hydrophilic) and …
C) … has a molecular weight (MW) of approximately 100–300 Da.
D) Conjugation reactions are catalyzed by enzymes known as transferases …
E) … they involve a cofactor that binds to the enzyme in close proximity to the substrate and carries the endogenous molecule or moiety to be transferred.

It is important from a biochemical and practical viewpoint to note that criterion A is essential to define conjugation reactions of xenobiotics. In contrast, criteria B to E are neither sufficient nor necessary. They are not sufficient, since in hydrogenation reactions (i.e. typical reactions of functionalization) the hydride is also transferred from a cofactor (NADPH or NADH). And they are not necessary, since all the above criteria suffer from some important exceptions mentioned below. For example, some GSH conjugations are non-enzymatic. A survey of transferases involved in drug metabolism is offered in Table 32.5.

B. Methylation

Reactions of methylation imply the transfer of a methyl group from the cofactor *S*-adenosyl-*L*-methionine (SAM). A number of methyltransferases are able to methylate xenobiotic molecules and metabolites.[38] They target the *ortho*-diphenol moiety (COMT), N-atoms in aromatic azaheterocycles such as pyridines (NNMT), primary arylamines (INMT), endocyclic secondary amines (HNMT) and a variety of thiols including heterocyclic ones (TMP, TPMT). O-Methylation (Figure 32.8a) is a common reaction of compounds containing a catechol moiety (**7**), with a usual regioselectivity for the *meta*-position. The substrates can be xenobiotics and particularly drugs, L-dopa being a classical example. More frequently, however, O-methylation occurs as a late event in the metabolism of aryl groups, after they have been oxidized to catechols. A toxicologically relevant reaction of N-methylation (Figure 32.8b) is that of theophylline (**8**) to yield caffeine.[39] This reaction is not seen in adult humans, but is effective in neonates (5–10% of a dose of theophylline), where it causes unwanted side-effects.

C. Sulfonation

Sulfonation reactions (also less correctly known as sulfation reactions) consist in a sulfate being transferred from the cofactor 3′-phosphoadenosine 5′-phosphosulfate (PAPS) to the substrate under catalysis by a sulfotransferase. Sulfotransferases, which catalyze a variety of physiological reactions, are soluble enzymes.[22,37,40–42] The most significant for drug metabolism are listed in Table 32.5. The sulfate moiety in PAPS is linked to a phosphate group by an anhydride bridge whose cleavage is exothermic and supplies enthalpy to the reaction. The nucleophilic —OH or —NH— site in the substrate will react with the leaving SO_3^- moiety, forming an ester sulfate or a sulfamate. Some of these conjugates are unstable under biological conditions and will form electrophilic intermediates of considerable toxicological significance.

The sulfoconjugation of alcohols leads to metabolites of different stabilities. Endogenous hydroxysteroids (i.e. cyclic secondary alcohols) form relatively stable sulfates, while some secondary alcohol metabolites of allylbenzenes (e.g. safrole and estragole) form highly genotoxic carbocations. In contrast to alcohols, phenols form stable sulfate esters. The reaction is usually of high affinity (i.e. rapid), but the limited availability of PAPS restricts the amount of conjugates being produced. Typical drugs undergoing limited sulfonation (Figure 32.8c) include paracetamol and diflunisal (**9**). Aromatic hydroxylamines and hydroxylamides are good substrates for some sulfotransferases and yield reactive sulfate esters. In contrast, significantly more stable products are obtained upon N-sulfoconjugation of amines. An intriguing and rare reaction of conjugation (Figure 32.8d) occurs for minoxidil (**10**). This drug is an *N*-oxide, and the actual active form responsible for the different therapeutic effects is the *N*,*O*-sulfate ester.

D. Glucuronidation

Glucuronidation is a major and very frequent reaction of conjugation. It involves the transfer to the substrate of a molecule of glucuronic acid from the cofactor uridine-5′-diphospho-α-*D*-glucuronic acid (UDPGA). As listed in Table 32.5, the enzyme catalyzing this reaction consists of a number of proteins coded by genes of the *UGT* superfamily.[22,37,43–46] The human UDPGA transferases (UGT) known to metabolize xenobiotics are the products of two gene families, *UGT1* and *UGT2*. Glucuronic acid exists in UDPGA in the 1α-configuration, but the products of conjugation are β-glucuronides (**11**) (Figure 32.9). This is due to the mechanism of the reaction being a nucleophilic substitution with inversion of configuration. Indeed, all functional groups able to undergo glucuronidation are nucleophiles, a common characteristic they share despite their great chemical variety.

O-Glucuronidation is a frequent metabolic reaction of xenobiotic phenols and alcohols, yielding polar metabolites excreted in urine and/or bile. An important example is that of morphine (**12**), which is conjugated on its phenolic and secondary alcohol groups to form the 3-*O*-glucuronide (a weak opiate antagonist) and the 6-*O*-glucuronide (a strong

TABLE 32.5 A Survey of Transferases (EC 2).[22,37]

Cofactor	Examples of enzymes (with some *gene roots* and human enzymes)
S-Adenosyl-L-methionine (SAM)	*Methyltransferases* (EC 2.1.1) EC 2.1.1.6: Catechol O-methyltransferase (*COMT*) EC 2.1.1.1: Nicotinamide N-methyltransferase (*NNMT*) EC 2.1.1.8: Histamine N-methyltransferase (*HNMT*) EC 2.1.1.28: Noradrenaline N-methyltransferase (*PNMT*) EC 2.1.1.49: Arylamine N-methyltransferase, indolethylamine N-methyltransferase (*INMT*) EC 2.1.1.9: Thiol S-methyltransferase (*TMT*) EC 2.1.1.67: Thiopurine S-methyltransferase (*TPMT*)
3′-Phosphoadenosine 5′-phosphosulfate (PAPS)	*Sulfotransferases* (EC 2.8.2) (*SULT*) EC 2.8.2.1: Aryl sulfotransferase (SULT1A1, 1A2 and 1A3) EC 2.8.2.4: Estrogen sulfotransferase (SULT1E1) EC 2.8.2.14: Bile salt sulfotransferase (SULT2A1) EC 2.8.2.2: Alcohol sulfotransferase (SULT2B1) EC 2.8.2.15: Steroid sulfotransferase EC 2.8.2.18: Cortisol sulfotransferase EC 2.8.2.3: Amine sulfotransferase (SULT3)
Uridine-5′-diphospho-α-D-glucuronic acid (UDPGA)	UDP-*Glucuronosyltransferases* (2.4.1.17) (*UGT*) Subfamily UGT1: UGT1A1, 1A3, 1A4, 1A5 to 1A10 Subfamily UGT2A: UGT2A1 to 2A3 Subfamily UGT2B: UGT2B4, 2B7, 2B10, 2B15, 2B17, 2B28 Subfamily UGT3A: UGT3A1, 3A2 Subfamily UGT8A: UGT8A1
Acetylcoenzyme A (AcCoA)	*Acetyltransferases* EC 2.3.1.5: N-Acetyltransferase (*NAT*) NAT1 and NAT2 EC 2.3.1.56: Aromatic-hydroxylamine O-acetyltransferase EC 2.3.1.118: N-Hydroxyarylamine O-acetyltransferase
Coenzyme A (CoA)	*Acyl-CoA synthetases* EC 6.2.1.1: Short-chain fatty acyl-CoA synthetase (*ACSS*) EC 6.2.1.2: Medium-chain acyl-CoA synthetase EC 6.2.1.3: Long-chain acyl-CoA synthetase (*ACSL*) EC 6.2.1.7: Cholate-CoA ligase EC 6.2.1.25: Benzoyl-CoA synthetase
Xenobiotic acyl-Coenzyme A	*Acyltransferases* EC 2.3.1.13: Glycine N-acyltransferase (*GLYAT*) EC 2.3.1.71: Glycine N-benzoyltransferase EC 2.3.1.14: Glutamine N-phenylacetyltransferase EC 2.3.1.68: Glutamine N-acyltransferase EC 2.3.1.65: Cholyl-CoA glycine-taurine N-acyltransferase (*BAAT*)
(Glutathione)	*Glutathione S-transferases* (EC 2.5.1.18) (*GST*) Microsomal GST superfamily (homotrimers): *MGST*: GST1–GST3 Cytoplasmic GST superfamily (homodimers, and a few heterodimers): *GST A*: Alpha class, GST A1-1, A1-2, A2-2, A3-3, A4-4, A5-5 *GST K*: Kappa class, GST K1-1 *GST M*: Mu class, GST M1-1, M2-2, M3-3, M4-4, M5-5 *GST O*: Omega class, GST O1-1, O2 *GST P*: Pi class: GST P1-1 *GST T*: Theta class, GST T1-1, T2 *GST Z*: Zeta class, GST Z1

FIGURE 32.8 A few examples of reactions of methylation (a and b) and sulfonation (c and d). The substrates are catechols (**7**), theophylline (**8**), diflunisal (**9**) and minoxidil (**10**).

opiate agonist), respectively.[47] O-Glucuronidation is often in competition with O-sulfonation with the latter reaction predominating at low doses and the former at high doses. In biochemical terms, glucuronidation is a reaction of low affinity and high capacity while sulfonation displays high affinity and low capacity.

An important pathway of O-glucuronidation is the formation of acyl-glucuronides. Substrates are anti-inflammatory arylacetic acids and aliphatic acids such as valproic acid (**13**). These metabolites are quite reactive, rearranging to positional isomers and binding covalently to plasma and seemingly also tissue proteins.[48] Thus, acyl glucuronide formation cannot be viewed solely as a reaction of inactivation and detoxification.

Second in importance to *O*-glucuronides are the *N*-glucuronides formed from amides and amines. The reaction has special significance for antibacterial sulfanilamides such as sulfadimethoxine (**14**) since it produces highly water-soluble metabolites that show no risk of crystallizing in the kidneys. For amines, there are a number of observations that pyridine-type nitrogens can be N-glucuronidated, as is illustrated by the *N*-glucuronide of nicotine (**15**, X = 2H) and cotinine (**15**, X = O). Another reaction of significance is the N-glucuronidation of lipophilic basic tertiary amines. More and more drugs of this type (e.g. antihistamines and neuroleptics) are found to undergo this reaction to a marked extent in humans, as illustrated by the *N*-glucuronide of imipramine (**16**).[22]

E. Acetylation

The major enzyme system catalyzing acetylation reactions is arylamine *N*-acetyltransferase. Two enzymes have

FIGURE 32.9 The generic structure of *O*-β-*D*-glucuronides is shown as **11**. Morphine (**12**) exemplifies the glucuronidation of phenols and alcohols, whereas valproic acid (**13**) forms an acylglucuronide. Sulfadimethoxine (**14**) exemplifies the N-glucuronidation of amides. The zwitterionic *N*-glucuronides are those of nicotine (**15**, X = 2H), cotinine (**15**, X = O) and imipramine (**16**). The arrows point to the target sites.

been characterized, NAT1 and NAT2, the latter showing considerably reduced levels in slow acetylators (i.e. subjects expressing a mutated NAT2 protein).[49–51] Two other activities, aromatic-hydroxylamine *O*-acetyltransferase and *N*-hydroxyarylamine *O*-acetyltransferase, are also involved in the acetylation of aromatic amines and hydroxylamines (Table 32.5). The substrates of acetylation are mainly amines of medium basicity. Very few basic amines (primary or secondary) of medicinal interest have been reported to form N-acetylated metabolites and when they did, the yields were low. In contrast, a large variety of primary aromatic amines are N-acetylated. Thus several drugs such as sulfonamides and *para*-aminosalicylic acid (**17**) (Figure 32.10) are acetylated to large extents, not to mention various carcinogenic amines such as benzidine. The same is true of hydrazines and hydrazides, such as isoniazid (**18**).

F. Conjugation with coenzyme A and subsequent reactions

The reactions described in this subsection all have in common the fact that they involve xenobiotic carboxylic acids (R—COOH) forming an acyl-CoA thioester (R—CO—S—CoA) as the metabolic intermediate and as a cofactor. The reaction requires ATP and is catalyzed by various acyl-CoA synthetases also known as acyl-CoA ligases (Table 32.5) of overlapping substrate specificity. The acyl-CoA conjugates thus formed are seldom excreted, but they can be isolated and characterized relatively easily in *in vitro* studies. In the present context, the interest of acyl-CoA conjugates is their further transformation by a considerable variety of pathways[22,37,52–54] as summarized in Table 32.6.

Amino acid conjugation is a major route for a number of small aromatic acids and involves the formation of an amide bond between the xenobiotic acyl-CoA and the amino acid. Glycine is the amino acid most frequently used for conjugation as illustrated with the formation of salicyluric acid (**20**) from salicylic acid (**19**) (Figure 32.10), although some glutamine and taurine conjugates have also been characterized in humans. The enzymes catalyzing these transfer reactions are various *N*-acyltransferases listed in Table 32.5.

Incorporation of xenobiotic acids into lipids forms highly lipophilic metabolites that may burden the body as long retained residues. In the majority of cases, *triacylglycerol*

FIGURE 32.10 A few examples of reactions of acetylation (a), and of some reactions consecutive to the formation of xenobiotic acyl-CoA conjugates (b). The substrates are salicylic acid (**19**), (*R*)-ibuprofen (**21**), and valproic acid whose intermediate acyl-CoA thioester is shown here (**22**). The arrows point to the target sites.

TABLE 32.6 Metabolic Consequences of the Conjugation of Xenobiotic Acids with Coenzyme A.[37,52–54]

Initial reaction	R—COOH → R—CO—S—CoA → •
Subsequent metabolic options	• Hydrolysis (futile cycle) • Formation of amino acid conjugates (glycine, glutamic acid, taurine, …) • Formation of hybrid triglycerides • Formation of cholesteryl and bile acid esters • Formation of acyl-carnitines • Unidirectional chiral inversion of arylpropionic acids (profens) • Dehydrogenation and β-oxidation • 2-Carbon chain elongation

analogs or *cholesterol esters* are formed. One telling example is that of ibuprofen (**21** in Figure 32.10), a much used anti-inflammatory drug whose (*R*)-enantiomer forms hybrid triglycerides detectable in rat liver and adipose tissues.

In addition, the lesser active (*R*)-ibuprofen enantiomer (experts called it the distomer, see Chapter 26) and a few other arylpropionic acids (i.e. profens) undergo an intriguing metabolic reaction such that the (*R*)-enantiomer is converted to the (*S*)-enantiomer, while the reverse reaction is negligible. This unidirectional chiral inversion is thus a reaction of bioactivation.[55]

In some cases, acyl-CoA conjugates formed from xenobiotic acids can also enter the physiological pathways of fatty acids catabolism or anabolism. A few examples are known of xenobiotic alkanoic and arylalkanoic acids undergoing two-carbon chain elongation or two-, four- or even six-carbon chain shortening. In addition, intermediate metabolites of β-oxidation may be seen, as illustrated in Figure 32.10 with valproic acid, whose acyl-CoA intermediate (**22**) is a substrate for some first steps of β-oxidation.[56]

G. Conjugation reactions of glutathione

1. *Introduction*

Glutathione (**23** in Figure 32.11GSH) is a thiol-containing tripeptide of major significance in the detoxification and toxification of drugs and other xenobiotics. In the body, it exists in a redox equilibrium between the reduced form (GSH) and an oxidized form (GS-SG). The metabolism of

FIGURE 32.11 (a): The structure of GSH (**23**), mercapturic acids (**24**) and further degradation products [22]. (b): The metabolism of acetaminophen (**25**) to its GSH conjugate *via* the intermediate quinone-imine **26**. (c): The conjugation of arene oxides (**27**) to mercapturic acids. (d): The GSH-mediated toxification of haloalkenes (**28**) to thioketenes (**29**).

GSH (i.e. its synthesis, redox equilibrium and degradation) is quite complex and involves a number of enzymes.[37,57–59] Glutathione reacts in a variety of ways, one of which is its redox capacity. Indeed, GSH can reduce peroxides (a reaction catalyzed by glutathione peroxidase) and organic nitrates; in its GSSG form, GSH can oxidize the superoxide anion-radical. Of major significance in detoxification reactions is the capacity of GSH (and other endogenous thiols including albumin) to scavenge free radicals, in particular reactive oxygen species (ROSs, e.g. $R^{\bullet}$, $HO^{\bullet}$, $HOO^{\bullet}$, $ROO^{\bullet}$). As such, GSH and other thiols have a critical role to play in cellular protection.

In this writing, we focus on the conjugation reactions of GSH as catalyzed by glutathione transferases. The glutathione transferases are multifunctional proteins coded by two multigene superfamilies (Table 32.5).[60–64] Seven classes are now known in humans. The conjugating reactivity of GSH is due to its thiol group (pK_a 9.0), which makes it a highly effective nucleophile. This nucleophilic character is greatly enhanced by deprotonation to a thiolate. In fact, an essential component of the catalytic mechanism of glutathione transferases is the marked increase in acidity (pK_a decreased by 2–3 units) experienced by the thiol group upon binding of GSH to the active site of the enzyme.[61] As a result, GSTs transfer GSH to a very large variety of electrophilic groups; depending on the nature of the substrate, the reactions can be categorized as nucleophilic substitutions or nucleophilic additions. And with compounds of sufficient reactivity, these reactions can also occur non-enzymatically.[61,65]

Once formed, GSH conjugates may be excreted as such (they are best characterized *in vitro* or in the bile of laboratory animals), but they usually undergo further biotransformation prior to urinary or fecal excretion. Cleavage of the glutamyl and cysteinyl residues by peptidases leave a cysteine conjugate, which is further N-acetylated by cysteine-*S*-conjugate *N*-acetyltransferase (EC 2.3.1.80) to yield an *N*-acetylcysteine

conjugate. The latter type of conjugates are known as mercapturic acids (**24** in Figure 32.11). This, however, does not imply that the degradation of unexcreted GSH conjugates must stop at this stage, since cysteine conjugates can be substrates of cysteine-S-conjugate β-lyase (EC 4.4.1.13) to yield thiols (R—SH). These, in turn, can rearrange, be oxidized or be S-methylated and then S-oxygenated to yield thiomethyl conjugates (R—S—Me), sulfoxides (R—SO—Me) and sulfones (R—SO_2—Me).

2. *Substrates of glutathione transferases*

Most known cases of GSH conjugation are nucleophilic attacks at electron-deficient carbon atoms, but attack of a nitrogen atom (e.g. in an aromatic nitroso group) or a sulfur atom (in thiols) is also documented. Nucleophilic additions can involve metabolites arising from oxidation reactions, but they can also occur as primary metabolic reactions. Frequent cases of GSH addition are to α,β-unsaturated carbonyls; a typical xenobiotic substrate is the toxin acrolein (CH_2═CH—CHO). Attack occurs at the activated CH_2 group. Quinones (*ortho*- and *para*-) and quinone imines are structurally very similar to α,β-unsaturated carbonyls. The reaction has physiological significance since endogenous metabolites such as quinone metabolites of estrogens are conjugated to GSH. A medicinal example is provided by the toxic quinone imine metabolite (**26** in Figure 32.11) of acetaminophen (**25**). Its GSH conjugate is not excreted as such in humans dosed with the drug, but as the mercapturic acid. The reaction is one of major detoxification, the quinone imine being extremely hepatotoxic and resulting in liver necrosis, liver failure and even death when produced at levels and rates that over-saturate the GSH conjugation pathway. Nevertheless, the GSH conjugation of quinones and quinone imines is not always a reaction of detoxification, as some of these conjugates are known to undergo further transformations leading to reactive products.[66]

An important role of GSH is in the conjugation of arene oxides, particularly those that rearrange slowly to the phenol and are poor substrates of epoxide hydrolase. The first reaction is again a nucleophilic addition to the epoxide (**27** in Figure 32.11). The resulting non-aromatic conjugate then dehydrates to an aromatic GSH conjugate, followed by a cascade leading to the mercapturic acid as also shown in Figure 32.11c. This is a common reaction of metabolically produced arene oxides, as documented for naphthalene and numerous drugs and xenobiotics containing a phenyl moiety. Note that the same reaction can also occur readily for epoxides of olefins.

Glutathione conjugations occurring by a mechanism of nucleophilic substitution (including addition-elimination) are documented for a number of industrial xenobiotics as well as drugs. This is the case for compounds having an activated alkyl moiety, for example the —CH_2Cl group of nitrogen mustards, which yields conjugates with structure —CH_2SG. Haloalkenes (**28** in Figure 32.11) are a special group of substrates of GS-transferases since they may react with GSH either by substitution to form an alkene conjugate as shown, or by addition to form an alkane conjugate (not shown). Formation of mercapturic acids occurs as for other GSH conjugates, but in both routes S-C cleavage of the *S*-cysteinyl or *N*-acetyl-*S*-cysteinyl conjugates by renal β-lyase yields thiols of significant toxicity. Indeed, these thiols rearrange by hydrohalide expulsion to form highly reactive thioketenes (**29**) and/or thioacyl halides.[67]

TABLE 32.7 Biological Factors Affecting Xenobiotic Metabolism.[69,70]

Inter-individual factors	Animal species	
	Genetic and ethnic factors (genetic polymorphism)	
	Sex	
Intra-individual factors	Physiological changes	Age
		Biological rhythms
		Pregnancy
	Pathological changes	Disease
		Stress
	External influences	Nutrition
		Enzyme induction by drugs and other xenobiotics
		Enzyme inhibition by drugs and other xenobiotics

IV. BIOLOGICAL FACTORS INFLUENCING DRUG METABOLISM

A variety of genetic, physiological and pathological factors influence xenobiotic metabolism and hence the wanted and unwanted activities associated with a drug. In this very brief section, we restrict ourselves to a conceptual overview and invite readers to personal study.[9,10,11,13,14,17–19,68]

The major difference is between inter-individual and intra-individual factors that influence drug metabolism (Table 32.7). The former remain constant throughout the life span of an organism and are the expression of its genome. In other words, they are written in the genome.[69] In contrast, intra-individual factors vary depending on time (age, even time of day), pathological states or external factors (nutrition, pollutants, drug treatment).[70]

Inter-individual factors are species differences, genetic differences between organisms in a given species (including genetic polymorphism and ethnic differences), and sex-related differences. Species differences are well-known since the beginnings of the science of drug metabolism. In contrast, genetic differences (pharmacogenetics) have gained significance only during the last 2–3 decades. Sex-related differences are well-documented in laboratory rodents, but

until recently have received comparatively less attention in humans.

Intra-individual factors are those that change as a function of time for a given organism. Among physiological factors, the first is obviously age, with major differences seen in humans between infants, children, adults and elderly persons. Biological rhythms include the 24-hour cycle (circadian rhythm), the 28-day cycle in women (the lunar cycle), and the yearly cycle; their study is known as chronopharmacology. Pregnancy is another physiological factor well studies in laboratory animals. For obvious reasons, little is known in humans. Pathological factors are the many diseases (e.g. fever, infections, cardiac or renal conditions) and abnormal conditions (e.g. stress) that perturb physiological equilibria and indirectly influence the body's reponse to drugs. And finally, we find the external influences. Differences due to different diets (nutritional factors) remain modest. In contrast, enzyme induction and inhibition play a major role in increasing or decreasing (often strongly) the biotransformation of numerous drugs. This is a huge and critical issue in drug-drug interactions, whose study and significance keep growing with new drugs entering the market.

Note that several among the factors listed above influence not only biotransformation, but can also affect absorption, distribution and excretion by interacting with transporters and therapeutic effects by influencing drug targets.

V. CONCLUDING REMARKS

Our objective in writing this chapter was to present structured data, namely, a reasoned classification of metabolic reactions and their enzymes. In this way, the vast diversity of metabolic reactions and of xenobiotic-metabolizing enzymes ceases to be a vague notion (a heap of stones) and can begin to be grasped as a structured whole made of many interacting parts.

The implicit objective of the chapter is to warn medicinal chemists against the danger of over-simplification. For too many drug discoverers, biotransformation begins and ends with CYP-catalyzed oxidations. As a result of this narrow view, little is done computationally and experimentally before the development phases to obtain a comprehensive view of the biotransformation of lead and preclinical candidates. One can only wonder about the proportion of metabolism-related side-effects that could have been avoided during clinical phases had medicinal chemists been more attentive to potential toxification reactions and drug-drug interactions caused by non-CYP reactions.

REFERENCES

1. Testa, B., Krämer, S.D. The biochemistry of drug metabolism. VHCA, Zurich, and Wiley-VCH, Weinheim; Vol. 1, 2008, pp. 319; Vol. 2, 2009, in press.
2. Timbrell, J. A. *Principles of Biochemical Toxicology*, 2nd Edition. Taylor & Francis: London, 1991.
3. Silverman, R. B. *The Organic Chemistry of Drug Design and Drug Action*, 2nd Edition. Academic Press: San Diego, 2004.
4. Kauffman, F. C. (Ed.) *Conjugation-Deconjugation Reactions in Drug Metabolism and Toxicity*. Springer Verlag: Berlin, 1994.
5. Welling, P. G., Balant, L. P. (Eds) *Pharmacokinetics of Drugs*. Springer: Heidelberg, 1994.
6. Anders, M. W., Dekant, W. (Eds) *Conjugation-Dependent Carcinogenicity and Toxicity of Foreign Compounds*. Academic Press: San Diego, 1994.
7. Testa, B. *The Metabolism of Drugs and Other Xenobiotics – Biochemistry of Redox Reactions*. Academic Press: London, 1995.
8. Ortiz de Montellano, P. R. (Ed.) *Cytochrome P450. Structure, Mechanism, and Biochemistry*, 2nd Edition. Plenum Press: New York, 1996.
9. Woolf, T. F. (Ed.) *Handbook of Drug Metabolism*. Dekker: New York, 1999.
10. Ioannides, C. (Ed.) *Enzyme Systems that Metabolise Drugs and Other Xenobiotics*. Wiley: Chichester, 2002.
11. Rodrigues, A. D. (Ed.) *Drug-Drug Interactions*. Dekker: New York, 2002.
12. Testa, B., Mayer, J. M. *Hydrolysis in Drug and Prodrug Metabolism – Chemistry, Biochemistry and Enzymology*. Wiley-VHCA: Zurich, 2003.
13. Boullata, J., Armenti, V. T. (Eds). *Handbook of Drug-Nutrient Interactions*. Humana Press: Totowa, NJ, 2004.
14. Lash, L. H. (Ed.) *Drug Metabolism and Transport*. Humana Press: Totowa, NJ, 2005.
15. Phillips, I. R., Shephard, E. A. (Eds). *Cytochrome P450 Protocols*. Humana Press: Totowa, NJ, 2006.
16. Utrecht, J. P., Trager, W. F. *Drug Metabolism – Chemical and Enzymatic Aspects*. Informa: New York, 2007.
17. Rendic, S. Summary of information on human CYP enzymes: Human P450 metabolism data. *Drug Metab. Rev.* 2002, *34*, 83–448.
18. Pelkonen, O. Human CYPs: *in vivo* and clinical aspects. *Drug Metab. Rev* 2002, *34*, 37–46.
19. Testa, B., Soine, W. Principles of drug metabolism. In *Burger's Medicinal Chemistry and Drug Discovery, 6th Edition* (Abraham, D. J., Ed.), Volume 2. Wiley-Interscience: Hoboken, NJ, 2003, pp. 431–498.
20. Testa, B., Krämer, S. D. The biochemistry of drug metabolism-An introduction. Part 1: Principles and overview. *Chem. Biodiver.* 2006, *3*, 1053–1101.
21. Trager, W. F. Principles of drug metabolism 1: Redox reactions. In *ADME-Tox Approaches* (Testa, B., van de Waterbeemd, H., Eds.), Vol. 5. In *Comprehensive Medicinal Chemistry* (Taylor, J. B., Triggle, D. J., Eds.), 2nd Edition. Elsevier: Oxford, 2007, pp. 87–132.
22. Testa, B. Principles of drug metabolism 2: Hydrolysis and conjugation reactions. In *ADME-Tox Approaches* (Testa, B., van de Waterbeemd, H., Eds.), Vol. 5. In *Comprehensive Medicinal Chemistry* (Taylor, J. B., Triggle, D.J., Eds.), 2nd Edition. Elsevier: Oxford, 2007, pp. 133–166.
23. Testa, B., Krämer, S. D. The biochemistry of drug metabolism – an introduction. Part 2: Redox reactions and their enzymes. *Chem. Biodiver.* 2007, *4*, 257–405.
24. Testa, B., Krämer, S. D. The biochemistry of drug metabolism – an introduction. Part 3: Reactions of hydrolysis and their enzymes. *Chem. Biodiver.* 2007, *4*, 2031–2122.
25. Walle, T., Walle, U. K., Olanoff, L. S. Quantitative account of propranolol metabolism in urine of normal man. *Drug Metab. Dispos.* 1985, *13*, 204–209.
26. Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (IUBMB). Enzyme Nomenclature (www.chem.qmul.ac.uk/iubmb/enzyme).
27. Nelson, D. R., Zeldin, D. C., Hoffman, S. M. G., Maltais, L. J., Wain, H. M., Nebert, D. W. Comparison of cytochrome P450 (*CYP*) genes from the mouse and human genomes, including nomenclature recommendations for genes, pseudogenes and alternative splice variants. *Pharmacogenetics* 2004, *14*, 1–18.

28. Nelson, D. R. Cytochrome P450 homepage (http://drnelson.utmem.edu/CytochromeP450.html).
29. Directory of P450-containing Systems. (http://www.icgeb.org/~p450srv/).
30. Brenda: The Comprehensive Enzyme Information System. (www.brenda.unikoeln.de).
31. ExPASy Proteomics Server (www.expasy.org).
32. Rambeck, B., May, T., Juergens, U. Serum concentrations of carbamazepine and its epoxide and diol metabolites in epileptic patients: the influence of dose and comedication. *Ther. Drug Monit.* 1987, *9*, 298–303.
33. Jindal, S. P., Lutz, T. Mass spectrometric studies of cocaine disposition in animals and humans using stable isotope-labeled analogues. *J. Pharmacol. Sci.* 1989, *78*, 1009–1014.
34. Brzezinski, M. R., Spink, B. J., Dean, R. A., Berkman, C. E., Cashman, J. R., Bosron, W. F. Human liver carboxylesterase hCE-1: binding specificity for cocaine, heroin, and their metabolites and analogs. *Drug Metab. Dispos.* 1997, *25*, 1089–1096.
35. Reist, M., Carrupt, P. A., Francotte, E., Testa, B. Chiral inversion and hydrolysis of thalidomide: mechanisms and catalysis by bases and serum albumin, and chiral stability of teratogenic metabolites. *Chem. Res. Toxicol.* 1998, *11*, 1521–1528.
36. Lu, J., Helsby, N., Palmer, B. D., Tingle, M., Baguley, B. C., Kestell, P., Ching, L. M. Metabolism of thalidomide in liver microsomes of mice, rabbits, and humans. *J. Pharmacol. Exp. Ther.* 2004, *310*, 571–577.
37. Testa, B., Krämer, S. D. The biochemistry of drug metabolism – an introduction. Part 4: Reactions of conjugation and their enzymes. *Chem. Biodiver.* 2008. in press.
38. Weinshilboum, R. M., Otterness, D. M., Szumlanski, C. L. Methylation pharmacogenetics: catechol *O*-methyltransferase, thiopurine methyltransferase, and histamine *N*-methyltransferase. *Annu. Rev. Pharmacol. Toxicol.* 1999, *39*, 19–52.
39. Ginsberg, G., Hattis, D., Russ, A., Sonawane, B. Physiologically based pharmacokinetic (PBPK) modeling of caffeine and theophylline in neonates and adults: implications for assessing children's risks from environmental agents. *J. Toxicol. Environ. Health A* 2004, *67*, 297–329.
40. Nagata, K., Yamazoe, Y. Pharmacogenetics of sulfotransferase. *Annu. Rev. Pharmacol. Tox.* 2000, *40*, 159–176.
41. Banoglu, E. Current status of the cytosolic sulfotransferases in the metabolic activation of promutagens and procarcinogens. *Curr. Drug Metab.* 2000, *1*, 1–30.
42. Kauffman, F. C. Sulfonation in pharmacology and toxicology. *Drug Metab. Rev.* 2004, *36*, 823–843.
43. Tukey, R. H., Strassburg, C. P. Human UDP-glucuronosyltransferases: metabolism, expression, and disease. *Annu. Rev. Pharmacol. Toxicol.* 2000, *40*, 581–616.
44. Fisher, M. B., Paine, M. F., Strelevitz, T. J., Wrighton, S. A. The role of hepatic and extrahepatic UDP-glucuronosyltransferases in human drug metabolism. *Drug Metab. Rev.* 2001, *33*, 273–297.
45. Bock, K. W. Vertebrate UDP-glucuronosyltransferases: functional and evolutionary aspects. *Biochem. Pharmacol.* 2003, *66*, 691–696.
46. Maruo, Y., Iwai, M., Mori, A., Sato, H., Takeuchi, Y. Polymorphism of UDP-glucuronosyltransferase and drug metabolism. *Curr. Drug Metab.* 2005, *6*, 91–99.
47. Milne, R. W., Nation, R. L., Somogyi, A. A. The disposition of morphine and its 3- and 6-glucuronide metabolites in humans and animals, and the importance of the metabolites to the pharmacological effects of morphine. *Drug Metab. Rev.* 1996, *28*, 345–472.
48. Spahn-Langguth, H., Benet, L. Z. Acyl glucuronides revisited: is the glucuronidation process a toxification as well as a detoxification mechanism? *Drug Metab. Rev.* 1992, *24*, 5–48.
49. Grant, D. M., Blum, M., Meyer, U. A. Polymorphisms of *N*-acetyltransferase genes. *Xenobiotica* 1992, *22*, 1073–1081.
50. Agundez, J. A. G., Olivera, M., Martinez, C., Ladero, J. M., Benitez, J. Identification and prevalence study of 17 allelic variants of the human NAT2 gene in a white population. *Pharmacogenetics* 1996, *6*, 423–428.
51. Kawamura, A., Graham, J., Mushtaq, A., Tsiftsoglou, S. A., Vath, G. M., Hanna, P. E., Wagner, C. R., Sim, E. Eukaryotic arylamine *N*-acetyltransferase: investigation of substrate specificity by high-throughput screening. *Biochem. Pharmacol.* 2005, *69*, 347–359.
52. Knights, K. M., Drogemuller, C. J. Xenobiotic-CoA ligases: kinetics and molecular characterization. *Curr. Drug Metab.* 2000, *1*, 49–66.
53. Kasuya, F., Igarashi, K., Fukui, M. Participation of a medium chain acyl-CoA synthetase in glycine conjugation of the benzoic acid derivatives with the electron-donating groups. *Biochem. Pharmacol.* 1996, *51*, 805–809.
54. Brugger, R., Reichel, C., Garcia Alia, B., Brune, K., Yamamoto, T., Tegeder, I., Geissinger, G. Expression of rat liver long-chain acyl-CoA synthetase and characterization of its role in the metabolism of *R*-ibuprofen and other fatty acid-like xenobiotics. *Biochem. Pharmacol.* 2001, *61*, 651–656.
55. Mayer, J. M., Testa, B. Pharmacodynamics, pharmacokinetics and toxicity of ibuprofen enantiomers. *Drug. Fut.* 1997, *22*, 1347–1366.
56. Gopaul, V. S., Tang, W., Farrell, K., Abbott, F. S. Amino acid conjugates: metabolites of 2-propylpentanoic acid (valproic acid) in epileptic patients. *Drug Metab. Dispos.* 2003, *31*, 114–121.
57. Sies, H. Glutathione and its role in cellular functions. *Free Rad. Biol. Med.* 1999, *27*, 916–921.
58. Dickinson, D. A., Forman, H. J. Cellular glutathione and thiols metabolism. *Biochem. Pharmacol.* 2002, *64*, 1019–1026.
59. Pompella, A., Visvikis, A., Paolicchi, A., De Tala, V., Casini, A. F. The changing faces of glutathione, a cellular protagonist. *Biochem. Pharmacol.* 2003, *66*, 1499–1503.
60. Hayes, J. D., Pullford, D. J. The glutathione *S*-transferase supergene family: regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance. *Crit. Rev. Biochem. Mol. Biol.* 1995, *30*, 445–600.
61. Armstrong, R. N. Structure, catalytic mechanism, and evolution of the glutathione transferases. *Chem. Res. Toxicol.* 1997, *10*, 2–18.
62. Sheenan, D., Meade, G., Foley, V. M., Dowd, C. Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. *Biochem. J.* 2001, *360*, 1–16.
63. Jowsey, I. R., Thomson, A. M., Flanagan, J. U., Murdock, P. R., Moore, G. B. T., Meyer, D. J., Murphy, G. J., Smith, S. A., Hayes, J. D. Mammalian class sigma glutathione *S*-transferases: catalytic properties and tissue-specific expression of human and rat GSH-dependent prostaglandin D2 synthases. *Biochem. J.* 2001, *359*, 507–516.
64. Schnekenburger, M., Morceau, F., Duvoix, A., Delhalle, S., Trentesaux, C., Dicato, M., Diederich, M. Increased glutathione *S*-transferase P1-1 expression by mRNA stabilization in hemin-induced differentiation of K562 cells. *Biochem. Pharmacol.* 2004, *68*, 1269–1277.
65. Ketterer, B. The role of nonenzymatic reactions of glutathione in xenobiotic metabolism. *Drug Metab. Rev.* 1982, *13*, 161–187.
66. Monks, T. J., Lau, S. S. Biological reactivity of polyphenolic-glutathione conjugates. *Chem Res. Toxicol.* 1997, *10*, 1296–1313.
67. Anders, M. W. Glutathione-dependent bioactivation of haloalkanes and haloalkenes. *Drug Metabol. Rev.* 2004, *36*, 583–594.
68. Totah, R. A., Rettie, A. E. Principles of drug metabolism 3: Enzymes and tissues. In *ADME-Tox Approaches* (Testa, B., van de Waterbeemd, H., Eds.), Vol. 5. In *Comprehensive Medicinal Chemistry* (Taylor, J. B., Triggle, D. J., Eds.), 2nd Edition. Elsevier: Oxford, 2007, pp. 167–191.
69. Krämer, S. D., Testa, B. The biochemistry of drug metabolism – an introduction. Part 6: Inter-individual factors affecting drug metabolism. *Chem. Biodiver.* 2008. in press.
70. Krämer, S. D., Testa, B. The biochemistry of drug metabolism – an introduction. Part 7: Intra-individual factors affecting drug metabolism. *Chem. Biodiver.* 2008, *16*. in press.

Chapter 33

Biotransformations Leading to Toxic Metabolites: Chemical Aspect

Anne-Christine Macherey and Patrick M. Dansette

I. HISTORICAL BACKGROUND
II. INTRODUCTION
III. REACTIONS INVOLVED IN THE BIOACTIVATION PROCESS
A. Oxidation
B. Oxidative stress
C. Reduction
D. Substitutions: hydrolysis and conjugation
E. Eliminations
F. Further biotransformations leading to the ultimate toxicant
IV. EXAMPLES OF METABOLIC CONVERSIONS LEADING TO TOXIC METABOLITES
A. Acetaminophen
B. Tienilic acid
C. Halothane
D. Valproic acid
E. Troglitazone
V. CONCLUSION
REFERENCES

"La matière demeure et la forme se perd."
"The matter remains and the form is lost."

Pierre de Ronsard

I. HISTORICAL BACKGROUND

As drugs are usually foreign chemicals, history of concern for the biotransformations of drugs leading to toxic metabolites formation is intrinsically linked to the history of xenobiotic metabolism studies. The International Society for the Study of Xenobiotics (ISSX) website (http://www.issx.org) presents an overview of the field history where some key figures may be pointed out.

One is probably Richard Tecwyn Williams who introduced the Phase I and II classification of xenobiotics metabolism reactions. Although his emblematic book[1] was called "Detoxication mechanisms," he estimated that, in some cases, metabolism may increase toxicity. He also considered that this "bioactivation" may occur during the Phase II reactions (usually considered as detoxication reactions), and not only that of Phase I (functionalization reactions).

Quite at the same time, Bernard Brodie studied the antimalarial atabrine (quinacrine) metabolism in order to avoid the toxic side effects of the drug. He also developed some new analytical methodologies, necessary for metabolic studies. Then he put together a group of researchers (including Julius Axelrod, James Gillette and many others) in this field, and they published many studies of great importance related to drug metabolism, most famous probably concerning acetaminophen. Among these works, these scientists developed the covalent binding theory concept, which provides an explanation for the toxic side effects of drugs. Following the work of James and Elizabeth Miller on covalent binding of polycyclic aromatic hydrocarbon electrophilic metabolites on DNA in the 1940s, Brodie *et al.* suggested that *in vivo* bioactivation may lead to the formation of electrophilic entities, which are capable of linking with biological macromolecules, thus inducing disturbances in cellular functions.

The discovery of mixed function oxidases during the 1950s and the characterization of cytochrome P450 by Omura and Sato[2] were a "revolution" in the field of

Copyright © 2008, Elsevier Ltd
All rights reserved.

xenobiotic metabolism. Remmer discovered that cytochrome P450 may be induced by phenobarbital, and Conney characterized the induction with 3-methyl cholanthrene. These works[3] were of great importance for mechanistic studies of drug metabolism. A new step was done in 1999 with the first crystallization of a mammalian cytochrome P450 by Johnson, which provided new perspectives in safer drug design.

Induction of cytochrome P450 synthesis suggests that xenobiotics may exert an effect on the genome: the use of genomics and proteomics represents a new challenge for predictive toxicology in drug design.

FIGURE 33.1 Indirect toxicity.

II. INTRODUCTION

Toxicity is the result of the more or less harmful action of chemicals on a living organism. Toxicology, the study of toxicity, is situated at the border of chemistry, biology, and in some cases, physics. Molecular toxicology tries to elucidate the mechanisms by which chemicals exert their toxic effects. Because many foreign chemicals enter the body in inert but unexcretable forms, biotransformations are an important aspect of the fate of xenobiotics.[4,5] In the case of drugs, metabolic conversions may be required for therapeutic effect ("prodrugs"; see Chapter 36 for a detailed discussion of prodrugs). In other cases, metabolism results in a loss of the biological activity. Sometimes, biotransformations produce toxic metabolites. The last process is called toxification or bioactivation. It should be emphasized that the general principles of pharmacology embrace the occurrence of toxic events: although biotransformation processes are often referred to as detoxification, the metabolic products are, in a number of cases, more toxic than the parent compounds. For drugs, whether biotransformations lead to the formation of toxic metabolites or to variations in therapeutic effects depends on intrinsic (such as the genetic polymorphism of some metabolism pathways) and extrinsic (such as the dose, the route or the duration) factors. The biochemical conversions are usually of an enzymatic nature and yield reactive intermediates, which may be implicated in the toxicity as far as the final metabolites. The primary events, which constitute the beginning of the toxic effect may result, after metabolism, from an inhibition of a specific (and in most cases enzymatic) cellular function, an alkylating attack or an oxidative stress.

With regard to the toxicity arising from metabolites ("indirect toxicity"), three cases may be distinguished (Figure 33.1):

A. Biotransformation begins with the transient formation of a reactive intermediate, whose lifetime is long enough to allow an attack on cellular components. This occurs when a reactive intermediate (generally radicals or electrophiles such as a carbonium ion) is formed and reacts rapidly with cellular macromolecules (such as unsaturated lipids, proteins, nucleic acids…), thus leading to their degradation and finally to cellular necrosis.

B. The first step of the metabolic process yields a primary metabolite, which can, in some cases, accumulate in the cell and react with cellular components before being transformed.

C. The final metabolites, when in excess, may accumulate and react with cellular macromolecules.

Usually, metabolic conversions are divided into two major types of reactions (see chapter 32 for a detailed discussion of metabolic biotransformations). Phase I reactions, or functionalization reactions, involve the introduction of a polar functionality such as a hydroxyl group into the xenobiotic structure. During Phase II reactions, this group is subsequently coupled (or conjugated) with an endogenous cofactor, which contains a functional group that is usually ionized at physiological pH. This ionic functional group facilitates active excretion into the urinary and/or hepatobiliary system. The elimination by transport mechanism is sometimes also called "Phase III."

Because bioactivation is mainly an activation of xenobiotics to electrophilic forms, which are entities capable of reacting irreversibly with tissue nucleophiles, biotransformations leading to toxic metabolites are in most cases Phase I reactions. But Phase II reactions may also give rise to toxic phenomena, for example, when conjugation produces a toxic metabolite, or when it is responsible for a specific target organ toxicity by acting as a delivery form to particular sites in the body where it is hydrolyzed and exerts a localized effect. Also, the final toxic metabolite may be formed by combinations of several Phase I and Phase II reactions. Because of the increasing understanding of drug metabolizing enzymes, some authors[6] claim that Williams "Phases I and II" classification is now inaccurate and even misleading. Pointing out the fact that Williams only introduced the classification at the end of his book and did not use it in his monograph, they consider it would be now wiser to avoid using any special category.

III. REACTIONS INVOLVED IN THE BIOACTIVATION PROCESSES

During the biotransformations affecting xenobiotics, five major kinds of chemical reactions may occur: oxidations (by far the most important), reductions, hydrolysis, substitutions, and eliminations. As Phase I and II reactions are part of this classification, each class of reactions can give rise to toxic metabolites.

A. Oxidation

Several enzymatic systems are involved during the oxidative transformations of xenobiotics. Whether substances act upon one enzyme rather than another depends not only on its specific function, but also on the electromolecular environment. The most important is the microsomal drug metabolizing system known as cytochrome P450 (CYP) monooxygenase, which is localized mainly in the liver and is involved in most biological oxidations of xenobiotics.[7–9] Those include *C*-, *N*- and *S*-oxidations, *N*-, *O*- and *S*-dealkylation, deaminations, and certain dehalogenations. Under anaerobic conditions, it can also catalyze reductive reactions. The CYP monooxygenase system is a multienzymatic complex constituted by the CYP hemoprotein, the flavoprotein enzyme NADPH CYP reductase, and the unsaturated phospholipid phosphatidylcholine. The isoforms involved in xenobiotic metabolism are membrane bound enzymes situated in the endoplasmic reticulum. After cell lysis for *in vitro* studies, they are found in the microsomal fraction. There are numerous isoforms (more than 6,000 known in all species). Thus a nomenclature based on their sequence similarity has been designed, and they are classified in families and subfamilies: for instance CYP3A4 is the major human CYP, CYP is for cytochrome P450, 3 for the number of the family (more than 40% sequence identity), A for the letter of the subfamily (more than 55% sequence identity) and 4 the number in the subfamily. The human genome shows 57 complete CYP sequences plus a number of pseudo-genes. The CYPs involved in xenobiotic metabolism[9] (about 15) belong to families 1 to 4. The catalytic mechanism of CYP involves a formal $(FeO)^{3+}$ complex formed by the elimination of H_2O from the iron site after two electrons have been added (Figure 33.2).

FIGURE 33.2 Catalytic cycle of cytochrome P450 (CYP) monooxygenase.

Another oxidative enzyme is the FAD-containing monooxygenase, which is capable of oxidizing nucleophilic nitrogen, sulfur, and organophosphorus compounds. The flavoprotein binds NADPH, oxygen and then the substrate. The oxidized metabolite is released, followed by NADP. Alcohol dehydrogenase and aldehyde dehydrogenase catalyze the oxidation of a variety of alcohols and aldehydes into aldehydes and acids, respectively, in the liver. Xanthine oxidase oxidizes several purine derivatives such as theophylline. Monoamine oxidase (MAO) and diamine oxidase convert amines into alkyl or aryl aldehydes by oxidation of the amine to an imine followed by subsequent hydrolysis. Peroxidases are oxidative enzymes, which couple the reduction of hydrogen peroxide and lipid hydroperoxides to the oxidation of other substrates. This co-oxidation is responsible for the production of reactive electrophiles from aromatic amines (e.g. the highly carcinogenic benzidine), phenols, hydroquinones, polycyclic aromatic hydrocarbons, etc.

The oxidation reactions can be described in terms of a rather common chemistry that involves the abstraction of either a hydrogen atom or a nonbonded (or π) electron by the iron-oxo porphyrin complex (Figure 33.3). The high-valent complex electronic configuration is unknown, but is usually written as Fe^{V}=O.

The one-electron oxidation yields transient radicals (Figure 33.4), which are transformed into more stable forms.

These radicals can incorporate an oxygen atom by abstraction of a hydroxyl group from the CYP iron-oxo species. This yields an oxidized derivative that may be sometimes more toxic than the parent compound or susceptible to further metabolic conversions. Free radicals may also

FIGURE 33.3 CYP oxidation process.

bind to the site of their formation, thus leading to inhibition or inactivation of the enzyme. When the radical is not efficiently controlled by the iron, it may leave the active site. The subsequent released radical is able to produce damage to unsaturated fatty acids, thus leading to lipid peroxidation and destruction of the cellular structure. Another mode of the radical stabilization is a second one-electron oxidation, which consists of the loss of another electron. The fate of free radicals is now extensively studied because of their great capacities for forming covalent bonds with cellular macromolecules.[10–12]

1. C-H *bond oxidations*

These oxidations, which are usually catalyzed by CYP monooxygenases, produce hydroxylated derivatives.[13] When the C-H bond is located in the α position to a heteroatom (such as O, S, N, halogen), the α hydroxylated derivative obtained is usually unstable and may be further oxidized or cleaved (Figure 33.5).

The antibiotic chloramphenicol is oxidized by CYP monooxygenase to chloramphenicol oxamyl chloride formed by the oxidation of the dichloromethyl moiety of chloramphenicol followed by elimination of hydrochloric acid[14] (Figure 33.6). The reactive metabolite reacts with the ε-amino group of a lysine residue in CYP[15] and inhibits the enzymatic reaction progressively with time. This type of inhibition is a time-dependent inhibition or a mechanism-based inhibition or inactivation, and the substrate involved historically has been called a suicide substrate because the enzymatic reaction yields a reactive metabolite, which destroys the enzyme.[16]

In the case of chloroform, the unstable trichloromethanol loses hydrochloric acid and forms phosgene, which is very reactive (Figure 33.7).[17]

FIGURE 33.4 One-electron oxidation.

FIGURE 33.5 C-H bond oxidation in the α-position to a heteroatom.

Tertiary amines containing at least one hydrogen on the α carbon may either be *N*-oxidized (leading to an *N*-oxide in the case of tertiary amines), or *C*-oxidized, thus leading to a carbinolamine. The latter, usually being unstable, splits into a secondary amine and an aldehyde moiety (Figure 33.8). Several electron transfer mechanisms have been proposed.[7–9]

During the oxidation of nitrosamines, the hydroxylated derivative formed cleaves spontaneously into highly reactive metabolites capable of alkylating nucleophilic sites in the cellular components.

FIGURE 33.6 Metabolic activation of chloramphenicol.

FIGURE 33.7 Oxidation of chloroform.

FIGURE 33.8 Oxidation of a tertiary amine.

FIGURE 33.9 Oxidation of aflatoxin B_1.

2. *Unsaturated bond oxidations*

Double bonds are oxidized by CYP monooxygenases into epoxides, which are generally very reactive. Epoxides are considered responsible for the toxicity of the unsaturated compounds.

The hepatocarcinogenicity of aflatoxin B_1 (AFB_1) is known to be due to the epoxide (AFB_1-oxides) formed, which binds directly with the N-7 atom of a guanine molecule in DNA (Figure 33.9).[18]

Aromatic chemicals are metabolized into unstable arene-oxides, which, as epoxides, are comparable to potentially equivalent electrophilic carbocations. These metabolites react easily with thiol groups derived from proteins, leading, for example, to hepatotoxicity. Bromobenzene seems to target a large group of functionally diverse hepatic proteins, as demonstrated recently in a proteomic analysis.[19] The chemical is oxidized (Figure 33.10) into a 3,4-epoxide, which does not exhibit mutagenic or carcinogenic activity, but reacts nonenzymatically with liver proteins and produces hepatic necrosis.[20] The isomeric 2,3-epoxide rearranges very quickly to 2-bromophenol and is less toxic. A secondary CYP-catalyzed oxidation to hydroquinone and benzoquinone also can occur. In this alternative pathway, conjugation with glutathione can lead to the formation of products, which may elicit their toxicity elsewhere than the liver and especially in the kidney.[20]

3. N-*oxidations*

Tertiary amines are transformed into *N*-oxides (generally less toxic), but primary and secondary amines are oxidized into hydroxylated derivatives (hydroxylamines). This oxidation is responsible for the hepatotoxicity and mutagenicity of acetamino-2-fluorene (Figure 33.11).[21]

Nitrenium ions may occur during bioactivation of aromatic amines and amides, which are usually *N*-oxidized into *N*-hydroxylated derivatives. By sulfation or esterification followed by elimination of the newly formed leaving group, the latter may be transformed into highly reactive nitrenium ions. In the case of aromatic nitrenium ions they are in equilibrium with their tautomeric aromatic carbocations, which react with cellular nucleophilic macromolecules (nucleic acids, etc.).

4. H*eteroatom oxidations*

Heteroatoms such as nitrogen or sulfur are oxidized at their nonbonded peripheric electrons as described for thiophene (Figure 33.12a).[22] Thiophene is oxidized to thiophene sulfoxide, which is unstable and dimerizes spontaneously to thiophene *S*-oxide dimers through a Diels–Alder reaction.[23,24] They also react with nucleophiles like the thiol group of glutathione or proteins, giving glutathione or protein adducts. In addition, thiophenes are oxidized to unstable thiophene epoxides, which rearrange spontaneously to thiolenones as found recently for 2- and 3-phenylthiophenes (Figure 33.12b). In fact, there is a competition between *S*-oxidation (sulfoxide pathway) and double bond oxidation (epoxide pathway). In the presence of glutathione, adducts formed from both reactive intermediates have been found, in addition to thiophene *S*-oxide dimers and the thiolenones tautomers of hydroxythiophenes.[25,26]

Halogenated aromatic compounds may also be oxidized by CYP monooxygenases, yielding hypervalent halogenated compounds.

B. Oxidative stress

Oxidative stress has been defined as a disturbance in the pro-oxidant–antioxidant balance in favor of the pro-oxidant state resulting from alterations in the redox state of the cell. The stepwise reduction of oxygen into superoxide anion,

FIGURE 33.10 Metabolism of bromobenzene.

FIGURE 33.11 *N*-oxidation of acetamino-2-fluorene.

FIGURE 33.12 (a) oxidation of thiophene, (b) oxidation of 2-phenylthiophene.

hydrogen peroxide, hydroxyl radical, and finally water, which accounts for about 5% of the normal oxygen reduction (versus 95% by means of the mitochondrial electron transport chain), may be increased by the redox cycling of some xenobiotics such as quinones or nitro-aromatic derivatives. These compounds are susceptible to one-electron reduction, which yields radical structures that may be back-oxidized to the parent compound. During this reoxidation, oxygen is reduced into superoxide anion. The oxygen reduction products are highly reactive entities that attack all the cellular components, especially when their normal degradation systems (superoxide dismutase, glutathione peroxidase, catalase) are overburdened. The polyunsaturated lipids are especially sensitive to these attacks because they are susceptible to a membrane-degrading peroxidation.

C. Reduction

Reductive biotransformations of several compounds such as polyhalogenated, keto, nitro and azo derivatives, are catalyzed by a variety of enzymes that differ according to the substrates and the species. The liver CYP-dependent drug metabolizing system is capable of reducing *N*-oxide, nitro and azo bonds, whereas the cytosolic nitrobenzene reductase activity is mainly due to CYP reductase, which transforms nitrobenzene into its hydroxylamino derivative. NADPH cytochrome c reductase is also able to catalyze the reduction of nitro compounds. These metabolic conversions may also be brought about by gastrointestinal anaerobic bacteria.

Reductive processes that occur during the metabolism of xenobiotics involve either one-electron reduction or a two-electron transfer.

Ionic reduction using a hydride occurs *in vivo* during the reduction catalyzed by NADH or NADPH enzymes, whereas one-electron reduction releases a radical structure, which may contribute to the toxic effect. Figure 33.13 illustrates the biotransformations affecting the anthracycline antitumor drug daunomycin.[27] Recent studies suggest that nitric oxide synthases may contribute to the cardiotoxicity, probably because of their structural similarities with CYP reductase.[28]

FIGURE 33.13 Biotransformations of daunorubicin.

1. *Reduction of polyhalogenated compounds*

Some polyhalogenated compounds, such as CCl_4, $BrCCl_3$ and halothane (CF_3-CHBrCl), when in the presence of the reduced form of CYP, may undergo one-electron reduction[29,30] (Figure 33.14), which leads to a radical that may be transformed by different pathways.

The radical formed may add directly on the unsaturated lipid bonds or initiate an unsaturated lipid peroxidation or undergo another one-electron reduction. The last reaction yields a carbene that can form a complex with the iron of the reduced form of CYP. Reduction of polyhalogenated compounds gives rise to several reactive intermediates, such as radicals, carbenes and peroxides, whose participation in the toxic effect varies greatly.[13]

2. *Reduction of nitro compounds*

The different steps of the biotransformations that produce a primary amine from an aromatic nitro compound involve a nitro radical-anion, a nitroso derivative, a nitroxyl radical, a hydroxylamine and then the primary amine (Figure 33.15).

Each of these different intermediates may contribute to the toxicity. Hydroxylamines are often responsible for methemoglobinemia,[31] whereas mutagenic and carcinogenic activity may be due to the combination of nitro radical-anion, nitroso derivatives or esterified hydroxylamine (such as sulfate derivatives) with cellular macromolecules.

Carcinogenicity may also be the result of the oxidative stress subsequent to the formation of oxygen–reduction products (superoxide anion, hydrogen peroxide, hydroxyl radical)

FIGURE 33.14 Reduction of polyhalogenated compounds.

FIGURE 33.15 Reductive biotransformation of nitro arene compounds.

during redox cycling of the nitro radical-anion, which restores the parent nitro compound.

3. *Reduction of azo compounds*

Azo compounds are susceptible to reduction, first to hydrazo intermediates, which are reductively cleaved into the appropriate amines. It has been proposed[32] that the first step, as with nitro compounds, is the formation of an azo-anion radical.

D. Substitutions: hydrolysis and conjugation

Among substitution reactions, ester and amide hydrolysis are common, and often operate during detoxification processes. Both specific enzymatic and chemical hydrolysis may occur. Acid-catalyzed reactions may occur in the stomach and the kidney, whereas base-catalyzed reactions may be assisted by the alkaline pH of the intestine.

Phase II, or conjugation reactions, are also substitution reactions, which proceed by means of an endogenous and generally activated electrophile. In mammals, five major conjugation reactions of xenobiotics exist and are mediated by transferase enzymes. Acid compounds, through their acyl-CoA ester, may also be conjugated with amino acids such as glycine, glutamine, and taurine. The specificity for the endogenous agent is high, but the specificity for the xenobiotic is broader.

To a great extent, conjugation produces excretable and nontoxic metabolites and thus is referred to as detoxification, but exceptions exist in each class of conjugation reaction. A more-in-depth discussion of Phase II metabolism can be found in Chapter 32.

1. *Glucuronic acid conjugation*

This substitution involves the transfer of a glucuronic acid from uridine diphosphate glucuronic acid (UDPGA) to a functional group in the xenobiotic substrate. The group may be a hydroxyl, carboxylic acid, amino or sulfur functional group. Glucuronides are never directly implicated in toxicity but are sometimes responsible for target-organ toxicity. Aromatic amines may be converted in the liver into hydroxylamine *O*-glucuronides, which are excreted in the urine and broken down in the bladder (if its pH is acidic) to liberate the proximate hydroxylamine carcinogen.

2. *Sulfation*

Sulfate conjugation gives a polar and ionized conjugate by means of the esterification of a hydroxyl group with sulfate ion (transferred from 3′-phosphoadenosine-5′-phosphosulfate or PAPS). The reaction is catalyzed by a hydrosoluble sulfotransferase. Sulfation sometimes gives rise to reactive intermediates that may undergo further reactions to yield electrophilic metabolites. In the case of 2-acetaminofluorene, the *O*-sulfate moiety is a facile leaving group, and this cleavage produces nitrenium ions, which act as alkylating agents for DNA (Figure 33.11).

3. *Acetylation*

Acetylation is a very common metabolic reaction, which occurs with amino, hydroxyl or sulfhydryl groups. The acetyl group is transferred from acetyl-Coenzyme A, and the reaction is catalyzed by acetyltransferases. An important aspect of this kind of substitution is the genetic polymorphism of one acetyltransferase in humans, who are divided into fast and slow acetylators. In a few cases, the conjugates are further metabolized to toxic compounds, as is seen with isoniazid. Some evidence exists that acetylation of the antitubercular isoniazid leads to enhanced hepatotoxicity of the drug.[33,34] Acetylation followed by hydrolysis and CYP-dependent oxidation yields free acetyl radicals[35] or acylium cations, which may acetylate the nucleophilic macromolecule functions (Figure 33.16).

4. *Glutathione conjugation*

Substitution reactions of xenobiotics with glutathione are the most important and contribute efficiently to detoxification. Nevertheless, in some cases, such as vicinal dihalogenated compounds, glutathione conjugation produces monosubstituted derivatives, which may cyclize into a highly electrophilic episulfonium ion (Figure 33.17).[36]

5. *Methylation*

Methylation is rarely of quantitative importance in the metabolism of xenobiotics. The methyl group is transferred from the nucleotide *S*-adenosyl-L-methionine (SAM) by means of a methyltransferase. The functional groups that undergo methylation include primary, secondary and tertiary amines, pyridines, phenols, catechols, thiophenols. The azaheterocycle pyridine is metabolized to the *N*-methylpyridinium ion, which is more toxic than pyridine itself[37] (Figure 33.18). The binding properties of the ionized metabolite are disturbed by the loss of its hydrophobic feature, resulting from the polarity inversion.

E. Eliminations

Eliminations of hydrogen and a halogen occur sometimes during the metabolism of halogenated xenobiotics and lead to an alkene. The double bond may be oxidized into an epoxide by means of oxidative enzyme systems as discussed above. Dehydrogenation, dehydrochlorination and dechlorination are (with oxidation) the different metabolic pathways of the γ-isomer of the insecticide hexachlorocyclohexane (lindane).[38]

F. Further biotransformations leading to the ultimate toxicant

Other reactions must be mentioned beside the major reactions described above. These reactions may be responsible for the transformation of a toxic metabolite into the ultimate toxicant.[39] Rearrangements and cyclizations are examples of reactions involved in these processes. In the case of the solvent hexane (Figure 33.19), the toxic metabolite, 2, 5-hexanedione, is formed by four successive oxidations of the molecule. The condensation of the γ-dicetone with the lysyl amino group of a neurofilament protein is followed by a Paal–Knorr cyclization reaction. This is the initial process that explains the hexane-induced neurotoxicity.[40] A further auto-oxidation of the *N*-pyrrolyl derivatives leads to the cross-linking of the axonal intermediate filament proteins and the subsequent occurrence of peripheral neurotoxicity.[41]

FIGURE 33.16 Bioactivation of isoniazid.

FIGURE 33.17 Bioactivation to episulfonium ion.

FIGURE 33.18 Bioactivation of pyridine.

Analogous pyrrolyl derivatives are also found as furan metabolites. Furans are oxidized by CYP to reactive furan-epoxides, which rearrange to ene-dial or ene-keto-aldehyde metabolites (Figure 33.20).[26,42,43] After reaction with thiols and amines like lysine, they form stable pyrrolic derivatives. This first depletes the cell of glutathione then creates cross-links in proteins and toxicity.

FIGURE 33.19 Bioactivation of hexane.

IV. EXAMPLES OF METABOLIC CONVERSIONS LEADING TO TOXIC METABOLITES

The formation of toxic metabolites and/or intermediates during the metabolism of drugs may occur by a considerable variety of pathways that are mediated by several enzyme systems. The following five examples do not represent an exhaustive list of the bioactivation processes, but are samples of original, significant and/or well-known drugs whose biotransformations lead to toxic compounds by the main types of reactions discussed above. Two of them (acetaminophen and tienilic acid) are CYP-mediated oxidations. Halothane acts through both oxidative and reductive biotransformations. Valproic acid is toxic through its elimination product. The toxicity of troglitazone seems to involve two distinct metabolic pathways, leading to both alkylating and oxidative stresses.

FIGURE 33.20 Bioactivation of furans.

A. Acetaminophen

The analgesic acetaminophen (4-hydroxyacetanilide, paracetamol) exhibits hepatotoxicity when administered in very high doses (approximately 250 mg/kg in rat and about 13 g for a 75 kg human).[44] The metabolite responsible is known to be the *N*-acetyl-*p*-benzoquinone imine (NAPQI) (Figure 33.21).[45]

The formation of NAPQI may proceed via CYP2E1,[46] but also via peroxidases such as prostaglandin hydroperoxidase.

The most commonly described mechanism proposes that metabolic activation occurs through *N*-oxidation of acetaminophen to *N*-hydroxyacetaminophen followed by dehydration to NAPQI (Figure 33.22).[47]

However, it seems that *N*-hydroxyacetaminophen is not a major intermediate in the oxidation of acetaminophen. The formation of NAPQI probably proceeds by two successive one-electron oxidations[48] (Figure 33.23).

During the first step, a one-electron oxidation yields a phenoxy radical (Ar-O$^{\bullet}$).[49] The presence of the radical was supported by fast flow ESR spectroscopy in the presence of horseradish peroxidase. In the second one-electron oxidation, the phenoxy radical is oxidized to NAPQI. As described in Figure 33.21, the highly electrophilic NAPQI may easily react with glutathione or protein thiol groups according to a Michael-type addition. The attack of liver protein thiol

FIGURE 33.21 Biotransformation pathway of acetaminophen.

FIGURE 33.22 Oxidation of acetaminophen according to the "*N*-hydroxyacetaminophen pathway."

groups and the subsequent adduct formation is frequently mentioned in the mechanism of acetaminophen hepatotoxicity. In mice, a number of proteins were identified such as glyceraldehyde-3-phosphate dehydrogenase,[50] calreticulin and the thiol: protein disulfide reductases Q1 and Q5[51] and this number is increasing with the advances of proteomics.[52]

Another hypothesis for the mechanism of toxicity is supported by the oxidative potency of NAPQI, but still suffers from lack of evidence.[53] NAPQI is a good oxidant for thiols of cellular components and pyridine nucleotides. Moreover, it may undergo a redox cycling with formation of superoxide anion by means of an oxygen one-electron reduction (Figure 33.24).

The stepwise reduction of oxygen produces hydrogen peroxide, and finally, a hydroxyl radical, which is a strong oxidant implicated in cellular oxidative stress. This oxidative stress causes glutathione depletion, a disruption of the cellular calcium regulation and modifications of cellular proteins, thus

FIGURE 33.23 Oxidation of acetaminophen by means of the phenoxy radical.

FIGURE 33.24 Redox cycling of *N*-acetyl-*p*-benzoquinone imine.

leading to cell death. Some biochemical parameters related to necrotic and apoptotic processes are affected in acetaminophen-exposed PC12 cells transfected with CYP2E1.[54,55]

It therefore appears that both covalent (e.g. alkylation) and noncovalent (e.g. oxidative stress) interactions play a major role in the pathogenesis of acute lethal cell injury caused by NAPQI.[56] At present, it is not possible to identify which of these two interactions is the critical event in initiating acetaminophen hepatotoxicity, even if some authors suggest that the characteristic features of oxidative stress are more likely the consequences of damage mediated by protein adduction.[57]

B. Tienilic acid

Tienilic acid is a uricosuric diuretic drug that may cause immunoallergic hepatitis in 1 in 10,000 patients, a side effect that resulted in its withdrawal from the market. The immunoallergic hepatitis was associated with the appearance of circulating antireticulum antibodies called anti-LKM2 antibodies, which are directed toward a liver endoplasmic reticulum protein.[58,59] From these observations, the mechanism of the immunotoxicity associated with the prolonged use of tienilic acid was elucidated by the Mansuy group.[60–62]

Tienilic acid is oxidized in the liver by CYP monooxygenase to 5-hydroxytienilic acid, which is the major urinary metabolite (about 50% in human). In humans, the bioactivation of tienilic acid depends on CYP2C9. This isoform is one of the major forms of CYP in the human liver. This oxidation occurs through an electrophilic intermediate capable of alkylating very specifically the CYP[61,63,64] leading to its inactivation. This mechanism-based inactivation is also observed with many xenobiotics such as alkenes with terminal unsaturation, alkynes, strained cycloalkylamines, 4-alkyldihydropyridines, benzodioxoles, and some tertiary amines.[16,65] The irreversible binding of the compound with CYP leads to an immune response and to generation of antibodies against both the modified protein and its native form. In fact, the autoantibodies anti-LKM2 present in hepatitis patients recognize CYP2C9 both as native protein and as modified protein. In addition, patient sera contain antibodies to tienilic acid-modified proteins. It has also been demonstrated in a rat model that tienilic acid modified CYP2C9 is exported to the plasma membrane of hepatocytes[66] and has been shown that tienilic acid treated rabbit hepatocytes, when first incubated with anti-LKM2, were lysed by human NK cells.[67] Thus, it is hypothesized that appearance of tienilic acid bound proteins on the hepatocyte surface triggers their cytolysis.

In the case of tienilic acid, the electrophilic reactive species is unknown. This is either a thiophene sulfoxide, as has been demonstrated for its 3-isomer,[68] or a thiophene epoxide (Figure 33.25). In both cases the electrophilic character of the intermediate is enhanced by the presence of an activating 2-keto group. In any event, this electrophilic species reacts with the enzyme CYP2C9 where it is produced and inactivates it efficiently (one inactivation event every 13 turnover).[61] The covalent binding of tienilic acid to CYP2C9 has been directly observed by mass spectrometry.[69] This reaction occurs in all patients with active CYP2C9 using this drug; however, very few produce anti-LKM2 and have hepatitis, which suggests some specificity in their immune response.

C. Halothane

Halothane is a widely used anesthetic drug that occasionally results in severe hepatitis. About 60–80% of the dose is eliminated in unmetabolized form during the 24 h following administration to patients. This compound is metabolized in the presence of CYP monooxygenase CYP2E1 according to the two main pathways[13] depicted in Figure 33.26.

FIGURE 33.25 Tienilic acid biotransformation to reactive intermediates and stable metabolites.

The major biotransformation pathway involves an oxidative step with introduction of an oxygen atom and subsequent formation of halohydrin. The unstable halohydrin loses hydrobromic acid to yield trifluoroacetyl chloride, which in turn is hydrolyzed to trifluoroacetic acid. This final metabolite is found in the urine.[70]

In conditions of low levels of oxygen, a reductive pathway (10%) is enhanced and yields a free radical intermediate characterized as 1-chloro-2,2,2-trifluoroethyl radical. Another one-electron reduction produces the 1-chloro-2,2,2-trifluoroethyl carbanion, which may undergo two possible kinds of eliminations.

One is the abstraction of a fluoride ion according to a E1cB elimination, which yields 1-chloro-2,2-difluoroethylene. This metabolite is eliminated by exhalation. Early studies suggested that a second elimination process might be an α-elimination of a chloride ion, which produces trifluoromethylcarbene,[71] but this was later reconsidered.[72] It was hypothesized that a carbene complex with the FeII in the active site might lead to inactivation of the CYP, but this inactivation is now thought to be due to the formation of an iron-σ-alkyl complex derived from the 1-chloro-2,2,2-trifluoroethyl radical.

The initially formed 1-chloro-2,2,2-trifluoroethyl radical may also cause a radical attack of polyunsaturated lipids, which produces 1-chloro-2,2,2-trifluoroethane. This mechanism is similar to the pathway described with the trichloromethyl radical formed during the one-electron reduction of carbon tetrachloride (Figure 33.14). The trichloromethyl radical may initiate a peroxidation of unsaturated lipids from the membrane with subsequent liberation of chloroform.

Several studies have demonstrated that halothane hepatotoxicity is mainly due to an immune reaction toward modified proteins of the liver. In fact, these proteins are trifluoroacetylated on their ε-NH_2-lysyl residue by the trifluoroacetyl chloride formed during the oxidative metabolism of halothane.[73,74] The product of the reaction can act as a foreign

FIGURE 33.26 The major metabolic pathways of halothane.

epitope, and the drug–protein conjugate, called neoantigen, elicits an immune response toward the liver[75] (Figure 33.27).

A related fluorocarbon used in air conditioning systems, HCFC 1,2,3, is metabolized to the same acyl halide and was recently implicated in an epidemic of liver disease in nine workers of a Belgian factory.[76] All patients had serum antibodies to trifluoroacetylated proteins.

D. Valproic acid

Valproic acid is an anticonvulsant agent used for the therapy of epilepsy, which occasionally results in hepatotoxicity in young children. The toxicity is characterized by mitochondrial damage, impairment of fatty acid β-oxidation and lipid accumulation.

It has been proposed that hepatotoxicity is a consequence of the further biotransformation of the valproic acid metabolite 2-propyl-4-pentenoic acid (also called Δ^4VPA).[77]

As depicted in Figure 33.28, Δ^4VPA is not formed by dehydration of 4- or 5-hydroxy valproic acids, which are, with the glucuronide conjugate, the major metabolites of valproic acid.[78] The mechanism is proposed to involve an initial hydrogen abstraction to generate a transient free radical intermediate. It has been demonstrated that the carbon-centred radical was localized at the C4 position. The radical undergoes both recombination (which yields 4-hydroxy valproic acid) and elimination (which produces

FIGURE 33.27 Biotransformation of halothane to trifluoroacetyl chloride and the subsequent binding to protein.

FIGURE 33.28 Bioactivation of valproic acid to Δ^4VPA.

the unsaturated derivative Δ^4VPA). The formation of these metabolites is catalyzed in rat by CYP4B1[79] and in human by CYP2C9.[9] Δ^4VPA is a hepatotoxic and strong teratogenic compound in animal models. In addition to that metabolic pathway, valproic acid undergoes biotransformation leading to *(E)*Δ^2VPA, which is devoid of embryotoxic effect in rodents.[80]

Further biotransformations of Δ^4VPA involve both the liver microsomal CYP enzymes and the fatty acid β-oxidation pathway (Figure 33.29). The mixed-function-oxidase system metabolizes the unsaturated metabolite to a γ-butyrolactone[81] derivative through a chemically reactive entity that is a mechanism-based inhibitor of CYP. The alkylation of the prosthetic heme by means of the radical occurs prior to formation of the epoxide.[82] Thus, the epoxide is not involved in the CYP inhibition.

The β-oxidation cycle activates Δ^4VPA to its Coenzyme A derivative and, through sequential steps of β-oxidation, yields the Coenzyme A ester of 3-oxo-2-propyl-4-pentenoic acid.[83] This final metabolite is believed to be a reactive electrophilic species that alkylates 3-ketoacyl-CoA thiolase (the terminal enzyme of β-oxidation) by means of a Michael-type addition through nucleophilic attack at the olefinic terminus.[84] Oxidative stress may also be implicated, at least in part, in valproic acid hepatotoxicity, as suggested by experimental data on the effect of the drug on reactive oxygen species.[85]

E. Troglitazone

Troglitazone ((±)-5-[4-(6-hydroxy-2,5,7,8-tetramethyl-chroman-2-ylmethoxy) benzyl]-2,4-thiazolidinedione) is an oral insulin sensitizer belonging to the thiazolidinedione class of compounds used for the treatment of type II diabetes. Its withdrawal from the US market was the consequence of the recent occurrence of hepatic failure leading sometimes to death.

It was first demonstrated that troglitazone is metabolized mainly to sulfate and glucuronide conjugates.[86] Also

FIGURE 33.29 Bioactivation of Δ^4VPA.

troglitazone is an inducer of CYP3A.[87] The mechanism of toxicity is still unclear, but seems to proceed according to two distinct pathways. This is supported by the demonstration that incubation of troglitazone with P450 isoforms in the presence of glutathione give rise to at least five GSH conjugates.[88,89] Identification of these adducts provided evidence for the two pathways described in Figures 33.30 and 33.31.

As described in Figure 33.30, oxidative cleavage of the thiazolidinedione ring probably generates highly electrophilic α-ketoisocyanate and sulfenic acid intermediates. This CYP3A mediated oxidation would afford a reactive sulfoxide intermediate, which undergoes a spontaneous ring opening.

The second pathway (Figure 33.31) consists of a CYP3A-mediated[90] one-electron oxidation of the phenolic hydroxyl group leading to an unstable hemiacetal, which opens spontaneously to form the quinone metabolite. This undergoes thiazolidinedione ring oxidation according to the pathway shown in Figure 33.30. Alternatively, a CYP-mediated hydrogen abstraction may occur on the phenoxy radical, leading to an *o*-quinone methide derivative.

It is now well established that troglitazone undergoes several metabolic transformation mediated by CYP3A4, leading to numerous electrophilic species.[91] Thus toxicity

FIGURE 33.30 Oxidation of the thiazolidinedione ring of troglitazone.

FIGURE 33.31 Oxidation of the chromane ring of troglitazone.

acts probably both by covalent binding to hepatic proteins and oxidative stress through a redox cycling process. The implication of the thiazolidinedione moiety is less likely since the more recent drugs of this series seems devoid of toxicity. Recent studies using mitochondrial manganese superoxide dismutase partially deficient mice also suggested that genetic deficiencies may be, at least partially, responsible for the liver failure in troglitazone-treated patients.[92,93]

V. CONCLUSION

In the foregoing discussion, it has been emphasized that almost all metabolic reactions are capable of producing reactive metabolites. This bioactivation yields toxic compounds that may act directly or indirectly[56] (Figure 33.32). The emergence of toxicity may be the outcome of the interactions of metabolites or reactive intermediates with biological targets such as cellular macromolecules. Some compounds exhibit their toxicity by inducing the generation of reactive oxygen species, thus producing alterations in the redox state of the cell. Often, covalent bonds are formed during a phenomenon that may be referred to as "alkylating stress." Bioactivation of drugs followed by drug protein adduction is then considered as a key sequence in the occurrence of toxic side effects.[94] As the precise damages of adducts on cellular functions are not fully understood, the formation of electrophilic metabolites is to be

FIGURE 33.32 Alkylating and oxidative stresses.

avoided in drug design. Proteome profiling (proteomics) may help to identify and compare proteins implicated in alkylating stress due to drugs, but this field remains to be developed and methods are to be validated. The specific inhibition of an enzyme by its own substrate is a peculiar feature of alkylating stress. Determination and monitoring of drug protein adducts have important implications in drug development, for example, in identifying CYP3A4 inactivation, since this CYP isoform is responsible for the metabolism of about 50% of the therapeutic drugs.[95] Thus medicinal chemists have set a threshold of acceptable covalent binding when developing a new drug. For example, this value for covalent binding levels to liver proteins was less than 50 pmol-equiv/mg protein under standard conditions at Merck[96] and can be subject to discussion on a case-to-case basis. This target represents about 1/20th of the level of binding for model hepatotoxins. Often the molecule can be modified to decrease this type of unwanted reaction without losing too much pharmacological activity.[96–98]

Such a variety of mechanisms makes it difficult to point at molecular functions susceptible to produce toxic effects through bioactivation. However, some major toxophoric groups may be highlighted (Table 33.1). They may be implicated in acute or chronic toxicity. These patterns must be of particular concern in drug design. A number of recent papers on these matters have been published on how to avoid those toxic events in drug design.[99–101]

TABLE 33.1 Some Major Toxophoric Groups and Their Bioactivation Mechanisms

Toxophoric group	Bioactivation mechanism
Azocompounds Acetamides Aromatic/heterocyclic amines Nitro compounds	Nitrenium ions, tautomeric carbonium ions
Nitroaromatic compounds	Radical formation/oxidative stress
Bromoarenes	Arene oxide formation
Ethinyl	Ketene formation/heme destruction
Furanes	Furane epoxide and ene-dial formation
Pyrroles	Pyrrole oxide
Nitrogen mustard	Aziridium ions
Nitroso compounds Hydrazines	Diazonium ions/heme adduct/radical formation
Nitrosamines	Carbenium ions/DNA alkylation
Polyhalogenated compounds	Radical and carbene formation/episulfonium with GSH
Quinone	Semiquinone radical formation/oxidative stress/thiol trapping
Thioamides	Thiourea formation
Thiophene	Thiophene sulfoxide or thiophene epoxide formation
Vinyl	Epoxidation/heme destruction

Generally, the formation of toxic metabolites is not the only pathway of biotransformation, and the overall metabolism is constituted toward detoxication and bioactivation processes. The toxic metabolites are themselves often further detoxified. The duality between a beneficial detoxication phenomenon (metabolism, drug resistance) and the occurrence of a toxic effect represents the cost for adaptability of metabolic enzymes to the diversity of xenobiotics. For those interested, a recent review applies the above chemistry to predict drug safety.[102]

REFERENCES

1. Williams, R. *Detoxication Mechanisms*, 2nd Edition. John Wiley & Sons: New York, 1959, p. 796.
2. Omura, T., Sato, R. The carbon monoxide-binding pigment of liver microsomes. I. evidence for its hemoprotein nature. *J. Biol. Chem.* 1964, *239*, 2370–2378.
3. Snyder, R. Microsomal enzyme induction. *Toxicol. Sci.* 2000, *55*, 233–234.

4. Sipes, G., Gandolfi, A. J. Biotransformation of toxicants. In *Casarett and Doull's Toxicology: The Basic Science of Poisons* (Amdur, M. O., Doull, J., Klaassen, C. D., Eds), 4th Edition. Pergamon Press: New York, 1991, pp. 88–126.
5. Alvares, A. P., Pratt, W. B. Pathways of drug metabolism. In *Principles of Drug Action: The Basis of Pharmacology* (Pratt, W. B., Taylor, P., Eds), 3rd Edition. Churchill Livingstone: New York, 1990, pp. 227–300.
6. Josephy, P. D., Guengerich, F. P., Miners, J. O. "Phase I and phase II" drug metabolism: terminology that we should phase out? *Drug Metab. Rev.* 2005, *37*, 575.
7. Guengerich, F. P. Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity. *Chem. Res. Toxicol.* 2001, *14*, 611–650.
8. Guengerich, F. P. Cytochrome P450 oxidations in the generation of reactive electrophiles: epoxidation and related reactions. *Arch. Biochem. Biophys.* 2003, *409*, 59–71.
9. Isin, E. M., Guengerich, F. P. Complex reactions catalyzed by cytochrome P450 enzymes. *Biochim. Biophys. Acta* 2007, *1770*, 314–329.
10. Mason, R. P., Chignell, C. F. Free radicals in pharmacology and toxicology – selected topics. *Pharmacol. Rev.* 1981, *33*, 189–211.
11. Aust, S. D., Chignell, C. F., Bray, T. M., Kalyanaraman, B., Mason, R. P. Free radicals in toxicology. *Toxicol. Appl. Pharmacol.* 1993, *120*, 168–178.
12. Singal, P. K., Petkau, A., Gerrard, J. M., Hrushovetz, S., Foerster, J. Free radicals in health and disease. *Mol. Cell. Biochem.* 1988, *84*, 121–122.
13. Anders, M. W., Pohl, L. R. Halogenated alkanes. In *Bioactivation of Foreign Compounds* (Anders, M. W., Ed.), Academic Press: Orlando, FL, 1985, pp. 284–315.
14. Pohl, L. R., Nelson, S. D., Krishna, G. Investigation of the mechanism of the metabolic activation of chloramphenicol by rat liver microsomes. Identification of a new metabolite. *Biochem. Pharmacol.* 1978, *27*, 491–496.
15. Halpert, J. R., Miller, N. E., Gorsky, L. D. On the mechanism of the inactivation of the major phenobarbital-inducible isozyme of rat liver cytochrome P-450 by chloramphenicol. *J. Biol. Chem.* 1985, *260*, 8397–8403.
16. Fontana, E., Dansette, P. M., Poli, S. M. Cytochrome P450 enzymes mechanism based inhibitors: common sub-structures and reactivity. *Curr. Drug Metab.* 2005, *6*, 413–454.
17. Cresteil, T., Beaune, P., Leroux, J. P., Lange, M., Mansuy, D. Biotransformation of chloroform by rat and human liver microsomes; *in vitro* effect on some enzyme activities and mechanism of irreversible binding to macromolecules. *Chem. Biol. Interact.* 1979, *24*, 153–165.
18. Benasutti, M., Ejadi, S., Whitlow, M. D., Loechler, E. L. Mapping the binding site of aflatoxin B1 in DNA: systematic analysis of the reactivity of aflatoxin B1 with guanines in different DNA sequences. *Biochemistry* 1988, *27*, 472–481.
19. Koen, Y. M., Gogichaeva, N. V., Alterman, M. A., Hanzlik, R. P. A proteomic analysis of bromobenzene reactive metabolite targets in rat liver cytosol *in vivo*. *Chem. Res. Toxicol.* 2007, *20*, 511–519.
20. Rietjens, I. M., den Besten, C., Hanzlik, R. P., van Bladeren, P. J. Cytochrome P450-catalyzed oxidation of halobenzene derivatives. *Chem. Res. Toxicol.* 1997, *10*, 629–635.
21. Verna, L., Whysner, J., Williams, G. M. 2-Acetylaminofluorene mechanistic data and risk assessment: DNA reactivity, enhanced cell proliferation and tumor initiation. *Pharmacol. Ther.* 1996, *71*, 83–105.
22. Dansette, P. M., Thang, D. C., el Amri, H., Mansuy, D. Evidence for thiophene-*S*-oxide as a primary reactive metabolite of thiophene *in vivo*: formation of a dihydrothiophene sulfoxide mercapturic acid. *Biochem. Biophys. Res. Commun.* 1992, *186*, 1624–1630.
23. Treiber, A., Dansette, P. M., ElAmri, H., Girault, J. P., Ginderow, D., Mornon, J. P., Mansuy, D. Chemical and biological oxidation of thiophene: preparation and complete characterization of thiophene *S*-oxide dimers and evidence for thiophene *S*-oxide as an intermediate in thiophene metabolism *in vivo* and *in vitro*. *J. Am. Chem. Soc.* 1997, *119*, 1565–1571.
24. Treiber, A., Dansette, P. M., Mansuy, D. Mechanism of the aromatic hydroxylation of thiophene by acid-catalyzed peracid oxidation. *J. Org. Chem.* 2002, *67*, 7261–7266.
25. Dansette, P. M., Bertho, G., Mansuy, D. First evidence that cytochrome P450 may catalyze both *S*-oxidation and epoxidation of thiophene derivatives. *Biochem. Biophys. Res. Commun.* 2005, *338*, 450–455.
26. Dalvie, D. K., Kalgutkar, A. S., Khojasteh-Bakht, S. C., Obach, R. S., O'Donnell, J. P. Biotransformation reactions of five-membered aromatic heterocyclic rings. *Chem. Res. Toxicol.* 2002, *15*, 269–299.
27. Gaudiano, G., Koch, T. H. Redox chemistry of anthracycline antitumor drugs and use of captodative radicals as tools for its elucidation and control. *Chem. Res. Toxicol.* 1991, *4*, 2–16.
28. Garner, A. P., Paine, M. J., Rodriguez-Crespo, I., Chinje, E. C., Ortiz De Montellano, P., Stratford, I. J., Tew, D. G., Wolf, C. R. Nitric oxide synthases catalyze the activation of redox cycling and bioreductive anticancer agents. *Cancer Res.* 1999, *59*, 1929–1934.
29. Butler, T. C. Reduction of carbon tetrachloride *in vivo* and reduction of carbon tetrachloride and chloroform *in vitro* by tissues and tissue constituents. *J. Pharmacol. Exp. Ther.* 1961, *134*, 311–319.
30. Mico, B. A., Branchflower, R. V., Pohl, L. R. Formation of electrophilic chlorine from carbon tetrachloride – involvement of cytochrome P-450. *Biochem. Pharmacol.* 1983, *32*, 2357–2359.
31. Gross, P. Biologic activity of hydroxylamine: a review. *Crit. Rev. Toxicol.* 1985, *14*, 87–99.
32. Mason, R. P. Redox cycling of radical anion metabolites of toxic chemicals and drugs and the Marcus theory of electron transfer. *Environ. Health Perspect.* 1990, *87*, 237–243.
33. Grant, D. M., Hughes, N. C., Janezic, S. A., Goodfellow, G. H., Chen, H. J., Gaedigk, A., Yu, V. L., Grewal, R. Human acetyltransferase polymorphisms. *Mutat. Res.* 1997, *376*, 61–70.
34. Timbrell, J. A., Mitchell, J. R., Snodgrass, W. R., Nelson, S. D. Isoniazid hepatoxicity: the relationship between covalent binding and metabolism *in vivo*. *J. Pharmacol. Exp. Ther.* 1980, *213*, 364–369.
35. Sinha, B. K. Activation of hydrazine derivatives to free radicals in the perfused rat liver: a spin-trapping study. *Biochim. Biophys. Acta* 1987, *924*, 261–269.
36. Weber, G. L., Steenwyk, R. C., Nelson, S. D., Pearson, P. G. Identification of *N*-acetylcysteine conjugates of 1,2-dibromo-3-chloropropane: evidence for cytochrome P450 and glutathione mediated bioactivation pathways. *Chem. Res. Toxicol.* 1995, *8*, 560–573.
37. D'Souza, J., Caldwell, J., Smith, R. L. Species variations in the *N*-methylation and quaternization of [^{14}C]pyridine. *Xenobiotica* 1980, *10*, 151–157.
38. Videla, L. A., Barros, S. B., Junqueira, V. B. Lindane-induced liver oxidative stress. *Free Radic. Biol. Med.* 1990, *9*, 169–179.
39. Miller, E. C., Miller, J. A. Mechanisms of chemical carcinogenesis. *Cancer* 1981, *47*, 1055–1064.
40. DeCaprio, A. P., Strominger, N. L., Weber, P. Neurotoxicity and protein binding of 2,5-hexanedione in the hen. *Toxicol. Appl. Pharmacol.* 1983, *68*, 297–307.
41. Genter St Clair, M. B., Amarnath, V., Moody, M. A., Anthony, D. C., Anderson, C. W., Graham, D. G. Pyrrole oxidation and protein cross-linking as necessary steps in the development of gamma-diketone neuropathy. *Chem. Res. Toxicol.* 1988, *1*, 179–185.
42. Chen, L. J., Hecht, S. S., Peterson, L. A. Characterization of amino acid and glutathione adducts of *cis*-2-butene-1,4-dial, a reactive metabolite of furan. *Chem. Res. Toxicol.* 1997, *10*, 866–874.
43. Baer, B. R., Rettie, A. E., Henne, K. R. Bioactivation of 4-ipomeanol by CYP4B1: adduct characterization and evidence for an enedial intermediate. *Chem. Res. Toxicol.* 2005, *18*, 855–864.
44. Thomas, S. H. Paracetamol (acetaminophen) poisoning. *Pharmacol. Ther.* 1993, *60*, 91–120.
45. Dahlin, D. C., Miwa, G. T., Lu, A. Y., Nelson, S. D. *N*-acetyl-*p*-benzoquinone imine: a cytochrome P-450-mediated oxidation product of acetaminophen. *Proc. Natl. Acad. Sci. USA* 1984, *81*, 1327–1331.
46. Chen, W., Koenigs, L. L., Thompson, S. J., Peter, R. M., Rettie, A. E., Trager, W. F., Nelson, S. D. Oxidation of acetaminophen to its toxic

quinone imine and nontoxic catechol metabolites by baculovirus-expressed and purified human cytochromes P450 2E1 and 2A6. *Chem. Res. Toxicol.* 1998, *11*, 295–301.
47. Mitchell, J. R., Jollow, D. J., Gillette, J. R., Brodie, B. B. Drug metabolism as a cause of drug toxicity. *Drug Metab. Dispos.* 1973, *1*, 418–423.
48. Ramakrishna Rao, D. N., Fischer, V., Mason, R. P. Glutathione and ascorbate reduction of the acetaminophen radical formed by peroxidase. Detection of the glutathione disulfide radical anion and the ascorbyl radical. *J. Biol. Chem.* 1990, *265*, 844–847.
49. Fischer, V., West, P. R., Harman, L. S., Mason, R. P. Free-radical metabolites of acetaminophen and a dimethylated derivative. *Environ. Health Perspect.* 1985, *64*, 127–137.
50. Dietze, E. C., Schafer, A., Omichinski, J. G., Nelson, S. D. Inactivation of glyceraldehyde-3-phosphate dehydrogenase by a reactive metabolite of acetaminophen and mass spectral characterization of an arylated active site peptide. *Chem. Res. Toxicol.* 1997, *10*, 1097–1103.
51. Zhou, L., McKenzie, B. A., Eccleston, E. D., Jr, Srivastava, S. P., Chen, N., Erickson, R. R., Holtzman, J. L. The covalent binding of [^{14}C]acetaminophen to mouse hepatic microsomal proteins: the specific binding to calreticulin and the two forms of the thiol:protein disulfide oxidoreductases. *Chem. Res. Toxicol.* 1996, *9*, 1176–1182.
52. Welch, K. D., Reilly, T. P., Bourdi, M., Hays, T., Pise-Masison, C. A., Radonovich, M. F., Brady, J. N., Dix, D. J., Pohl, L. R. Genomic identification of potential risk factors during acetaminophen-induced liver disease in susceptible and resistant strains of mice. *Chem. Res. Toxicol.* 2006, *19*, 223–233.
53. Rosen, G. M., Singletary, W. V., Jr, Rauckman, E. J., Killenberg, P. G. Acetaminophen hepatotoxicity. An alternative mechanism. *Biochem. Pharmacol.* 1983, *32*, 2053–2059.
54. Holownia, A., Mapoles, J., Menez, J. F., Braszko, J. J. Acetaminophen metabolism and cytotoxicity in PC12 cells transfected with cytochrome P4502E1. *J. Mol. Med.* 1997, *75*, 522–527.
55. Dai, Y., Cederbaum, A. I. Cytotoxicity of acetaminophen in human cytochrome P4502E1-transfected HepG2 cells. *J. Pharmacol. Exp. Ther.* 1955, *273*, 1497–1505.
56. Nelson, S. D., Pearson, P. G. Covalent and noncovalent interactions in acute lethal cell injury caused by chemicals. *Annu. Rev. Pharmacol. Toxicol.* 1990, *30*, 169–195.
57. Josephy, P. D. The molecular toxicology of acetaminophen. *Drug Metab. Rev.* 2005, *37*, 581.
58. Homberg, J. C., Andre, C., Abuaf, N. A new anti-liver-kidney microsome antibody (anti-LKM2) in tienilic acid-induced hepatitis. *Clin. Exp. Immunol.* 1984, *55*, 561–570.
59. Dansette, P. M., Bonierbale, E., Minoletti, C., Beaune, P. H., Pessayre, D., Mansuy, D. Drug-induced immunotoxicity. *Eur. J. Drug Metab. Pharmacokinet.* 1998, *23*, 443–451.
60. Beaune, P., Dansette, P. M., Mansuy, D., Kiffel, L., Finck, M., Amar, C., Leroux, J. P., Homberg, J. C. Human anti-endoplasmic reticulum autoantibodies appearing in a drug-induced hepatitis are directed against a human liver cytochrome P-450 that hydroxylates the drug. *Proc. Natl. Acad. Sci. USA* 1987, *84*, 551–555.
61. Lopez-Garcia, M. P., Dansette, P. M., Mansuy, D. Thiophene derivatives as new mechanism-based inhibitors of cytochromes P-450: inactivation of yeast-expressed human liver cytochrome P-450 2C9 by tienilic acid. *Biochemistry* 1994, *33*, 166–175.
62. Lecoeur, S., Bonierbale, E., Challine, D., Gautier, J. C., Valadon, P., Dansette, P. M., Catinot, R., Ballet, F., Mansuy, D., Beaune, P. H. Specificity of *in vitro* covalent binding of tienilic acid metabolites to human liver microsomes in relationship to the type of hepatotoxicity: comparison with two directly hepatotoxic drugs. *Chem. Res. Toxicol.* 1994, *7*, 434–442.
63. Mansuy, D. Molecular structure and hepatotoxicity: compared data about two closely related thiophene compounds. *J. Hepatol.* 1997, *26*(Suppl 2), 22–25.
64. Lopez Garcia, M. P., Dansette, P. M., Valadon, P., Amar, C., Beaune, P. H., Guengerich, F. P., Mansuy, D. Human-liver cytochromes P-450 expressed in yeast as tools for reactive-metabolite formation studies. Oxidative activation of tienilic acid by cytochromes P-450 2C9 and 2C10. *Eur. J. Biochem.* 1993, *213*, 223–232.
65. Masubuchi, Y., Horie, T. Toxicological significance of mechanism-based inactivation of cytochrome p450 enzymes by drugs. *Crit. Rev. Toxicol.* 2007, *37*, 389–412.
66. Robin, M. A., Maratrat, M., Le Roy, M., Le Breton, F. P., Bonierbale, E., Dansette, P., Ballet, F., Mansuy, D., Pessayre, D. Antigenic targets in tienilic acid hepatitis. Both cytochrome P450 2C11 and 2C11-tienilic acid adducts are transported to the plasma membrane of rat hepatocytes and recognized by human sera. *J. Clin. Invest.* 1996, *98*, 1471–1480.
67. Neuberger, J., Williams, R. Immune mechanisms in tienilic acid associated hepatotoxicity. *Gut* 1989, *30*, 515–519.
68. Valadon, P., Dansette, P. M., Girault, J. P., Amar, C., Mansuy, D. Thiophene sulfoxides as reactive metabolites: formation upon microsomal oxidation of a 3-aroylthiophene and fate in the presence of nucleophiles *in vitro* and *in vivo*. *Chem. Res. Toxicol.* 1996, *9*, 1403–1413.
69. Koenigs, L. L., Peter, R. M., Hunter, A. P., Haining, R. L., Rettie, A. E., Friedberg, T., Pritchard, M. P., Shou, M., Rushmore, T. H., Trager, W. F. Electrospray ionization mass spectrometric analysis of intact cytochrome P450: identification of tienilic acid adducts to P450 2C9. *Biochemistry* 1999, *38*, 2312–2319.
70. Harris, J. W., Pohl, L. R., Martin, J. L., Anders, M. W. Tissue acylation by the chlorofluorocarbon substitute 2,2-dichloro-1,1,1-trifluoroethane. *Proc. Natl. Acad. Sci. USA* 1991, *88*, 1407–1410.
71. Mansuy, D., Nastainczyk, W., Ullrich, V. The mechanism of halothane binding to microsomal cytochrome P450. *Naunyn Schmiedebergs Arch. Pharmacol.* 1974, *285*, 315–324.
72. Ahr, H. J., King, L. J., Nastainczyk, W., Ullrich, V. The mechanism of reductive dehalogenation of halothane by liver cytochrome P450. *Biochem. Pharmacol.* 1982, *31*, 383–390.
73. Pohl, L. R. An immunochemical approach of identifying and characterizing protein targets of toxic reactive metabolites. *Chem. Res. Toxicol.* 1993, *6*, 786–793.
74. Kenna, J. G., Neuberger, J., Williams, R. Evidence for expression in human liver of halothane-induced neoantigens recognized by antibodies *in sera* from patients with halothane hepatitis. *Hepatology* 1988, *8*, 1635–1641.
75. Pohl, L. R., Kenna, J. G., Satoh, H., Christ, D., Martin, J. L. Neoantigens associated with halothane hepatitis. *Drug Metab. Rev.* 1989, *20*, 203–217.
76. Hoet, P., Graf, M. L., Bourdi, M., Pohl, L. R., Duray, P. H., Chen, W., Peter, R. M., Nelson, S. D., Verlinden, N., Lison, D. Epidemic of liver disease caused by hydrochlorofluorocarbons used as ozone-sparing substitutes of chlorofluorocarbons. *Lancet* 1997, *350*, 556–559.
77. Baillie, T. A. Metabolic activation of valproic acid and drug-mediated hepatotoxicity. Role of the terminal olefin, 2-*n*-propyl-4-pentenoic acid. *Chem. Res. Toxicol.* 1988, *1*, 195–199.
78. Rettie, A. E., Rettenmeier, A. W., Howald, W. N., Baillie, T. A. Cytochrome P-450-catalyzed formation of delta 4-VPA, a toxic metabolite of valproic acid. *Science* 1987, *235*, 890–893.
79. Rettie, A. E., Sheffels, P. R., Korzekwa, K. R., Gonzalez, F. J., Philpot, R. M., Baillie, T. A. CYP4 isozyme specificity and the relationship between omega-hydroxylation and terminal desaturation of valproic acid. *Biochemistry* 1995, *34*, 7889–7895.
80. Kassahun, K., Baillie, T. A. Cytochrome P-450-mediated dehydrogenation of 2-n-propyl-2(E)-pentenoic acid, a pharmacologically active metabolite of valproic acid in rat liver microsomal preparations. *Drug Metab. Dispos.* 1993, *21*, 242–248.
81. Prickett, K. S., Baillie, T. A. Metabolism of unsaturated derivatives of valproic acid in rat liver microsomes and destruction of cytochrome P-450. *Drug Metab. Dispos.* 1986, *14*, 221–229.
82. Ortiz de Montellano, P. R., Yost, G. S., Mico, B. A., Dinizo, S. E., Correia, M. A., Kumbara, H. Destruction of cytochrome P-450 by 2-isopropyl-4-pentenamide and methyl 2-isopropyl-4-pentenoate: mass spectrometric characterization of prosthetic heme adducts and

nonparticipation of epoxide metabolites. *Arch. Biochem. Biophys.* 1979, *197*, 524–533.

83. Rettenmeier, A. W., Gordon, W. P., Prickett, K. S., Levy, R. H., Baillie, T. A. Biotransformation and pharmacokinetics in the rhesus monkey of 2-n-propyl-4-pentenoic acid, a toxic metabolite of valproic acid. *Drug Metab. Dispos.* 1986, *14*, 454–464.
84. Rettenmeier, A. W., Prickett, K. S., Gordon, W. P., Bjorge, S. M., Chang, S. L., Levy, R. H., Baillie, T. A. Studies on the biotransformation in the perfused rat liver of 2-n-propyl-4-pentenoic acid, a metabolite of the antiepileptic drug valproic acid. Evidence for the formation of chemically reactive intermediates. *Drug Metab. Dispos.* 1985, *13*, 81–96.
85. Chang, T. K., Abbott, F. S. Oxidative stress as a mechanism of valproic acid-associated hepatotoxicity. *Drug Metab. Rev.* 2006, *38*, 627–639.
86. Kawai, K., Kawasaki-Tokui, Y., Odaka, T., Tsuruta, F., Kazui, M., Iwabuchi, H., Nakamura, T., Kinoshita, T., Ikeda, T., Yoshioka, T., Komai, T., Nakamura, K. Disposition and metabolism of the new oral antidiabetic drug troglitazone in rats mice and dogs. *Arzneimittelforschung* 1997, *47*, 356–368.
87. Ramachandran, V., Kostrubsky, V. E., Komoroski, B. J., Zhang, S., Dorko, K., Esplen, J. E., Strom, S. C., Venkataramanan, R. Troglitazone increases cytochrome P-450 3 A protein and activity in primary cultures of human hepatocytes. *Drug Metab. Dispos.* 1999, *27*, 1194–1199.
88. Kassahun, K., Pearson, P. G., Tang, W., McIntosh, I., Leung, K., Elmore, C., Dean, D., Wang, R., Doss, G., Baillie, T. A. Studies on the metabolism of troglitazone to reactive intermediates *in vitro* and *in vivo*. Evidence for novel biotransformation pathways involving quinone methide formation and thiazolidinedione ring scission. *Chem. Res. Toxicol.* 2001, *14*, 62–70.
89. Prabhu, S., Fackett, A., Lloyd, S., McClellan, H. A., Terrell, C. M., Silber, P. M., Li, A. P. Identification of glutathione conjugates of troglitazone in human hepatocytes. *Chem. Biol. Interact.* 2002, *142*, 83–97.
90. Yamazaki, H., Shibata, A., Suzuki, M., Nakajima, M., Shimada, N., Guengerich, F. P., Yokoi, T. Oxidation of troglitazone to a quinone-type metabolite catalyzed by cytochrome P-450 2C8 and P-450 3A4 in human liver microsomes. *Drug Metab. Dispos.* 1999, *27*, 1260–1266.
91. Smith, M. T. Mechanisms of troglitazone hepatotoxicity. *Chem. Res. Toxicol.* 2003, *16*, 679–687.
92. Jaeschke, H. Troglitazone hepatotoxicity: Are we getting closer to understanding idiosyncratic liver injury? *Toxicol. Sci.* 2007, *97*, 1–3.
93. Ong, M. M., Latchoumycandane, C., Boelsterli, U. A. Troglitazone-induced hepatic necrosis in an animal model of silent genetic mitochondrial abnormalities. *Toxicol. Sci.* 2007, *97*, 205–213.
94. Zhou, S., Chan, E., Duan, W., Huang, M., Chen, Y. Z. Drug bioactivation, covalent binding to target proteins and toxicity relevance. *Drug Metab. Rev.* 2005, *37*, 41–213.
95. Yang, X. X., Hu, Z. P., Chan, S. Y., Zhou, S. F. Monitoring drug-protein interaction. *Clin. Chim. Acta.* 2006, *365*, 9–29.
96. Evans, D. C., Watt, A. P., Nicoll-Griffith, D. A., Baillie, T. A. Drug-protein adducts: an industry perspective on minimizing the potential for drug bioactivation in drug discovery and development. *Chem. Res. Toxicol.* 2004, *17*, 3–16.
97. Tang, C., Subramanian, R., Kuo, Y., Krymgold, S., Lu, P., Kuduk, S. D., Ng, C., Feng, D. M., Elmore, C., Soli, E., Ho, J., Bock, M. G., Baillie, T. A., Prueksaritanont, T. Bioactivation of 2,3-diaminopyridine-containing bradykinin B1 receptor antagonists: irreversible binding to liver microsomal proteins and formation of glutathione conjugates. *Chem. Res. Toxicol.* 2005, *18*, 934–945.
98. Doss, G. A., Baillie, T. A. Addressing metabolic activation as an integral component of drug design. *Drug Metab. Rev.* 2006, *38*, 641–649.
99. Baillie, T. A., Cayen, M. N., Fouda, H., Gerson, R. J., Green, J. D., Grossman, S. J., Klunk, L. J., LeBlanc, B., Perkins, D. G., Shipley, L. A. Drug metabolites in safety testing. *Toxicol. Appl. Pharmacol.* 2002, *182*, 188–196.
100. Obach, R. S., Walsky, R. L., Venkatakrishnan, K. Mechanism-based inactivation of human cytochrome p450 enzymes and the prediction of drug–drug interactions. *Drug Metab. Dispos.* 2007, *35*, 246–255.
101. Kalgutkar, A. S., Obach, R. S., Maurer, T. S. Mechanism-based inactivation of cytochrome P450 enzymes: chemical mechanisms, structure–activity relationships and relationship to clinical drug–drug interactions and idiosyncratic adverse drug reactions. *Curr. Drug Metab.* 2007, *8*, 407–447.
102. Guengerich, F. P., MacDonald, J. S. Applying mechanisms of chemical toxicity to predict drug safety. *Chem. Res. Toxicol.* 2007, *20*, 344–369.

Chapter 34

Drug Transport Mechanisms and their Impact on the Disposition and Effects of Drugs

Jean-Michel Scherrmann

I. INTRODUCTION
II. BIOLOGY AND FUNCTION OF TRANSPORTERS
A. Modes of active transport
B. Genes and classification
C. Basic structure
D. Distributions and properties of transporters in tissues
III. TRANSPORTERS IN DRUG DISPOSITION
A. ABC transporters
B. SLC transporters
IV. ROLES OF TRANSPORTERS IN DRUG PHARMACODYNAMICS AND TOXICOLOGY
A. Intestinal absorption
B. Liver and hepatic clearance
C. Blood barriers and tissue distribution
D. Kidney and renal clearance
V. CONCLUSION
ACKNOWLEDGMENTS
REFERENCES

Science may set limits to knowledge but should not set limits to imagination.

Bertrand Russell

I. INTRODUCTION

The exchange of solutes between body compartments depends, to a considerable extent, on the properties of the body that allow easy communication between tissues and compartments via pores and fenestra on the walls of the blood vessel or gap junctions between the cells of many epithelia. These features allow free solutes to move in both directions through biological membranes by the so-called paracellular pathway. But the organs of the body and pharmacological targets at the biophase are not readily accessible to exogenous molecules because of the integrity of the lipid bilayer membranes that protect the interiors of cells. Some physiological barriers like the blood–brain barrier (BBB), the blood-placenta barrier, and the blood–testis barrier are so impermeable that solutes can only cross the lipid bilayer by a transcellular pathway. It has been established for some time that only small non-polar, uncharged molecules like oxygen, carbon dioxide, water and ethanol can easily diffuse through membranes if there is an appropriate gradient, whereas charged small ions like sodium and potassium and molecules like glucose (180 Da) cross membranes considerably less readily than water. As the delivery of many polar molecules, such as anions and cations, vitamins, sugars, nucleosides, amino acids, peptides, bile acids and porphyrins, to cells is essential for life, essential transporter proteins anchored in the lipid bilayer have evolved to permit their exchange between cells and their environment.

Pharmacokinetics is now challenged by the growing importance of transporters, a relatively new and potentially major factor in drug absorption, distribution, metabolism and excretion (the ADME process). Several years ago, passive diffusion was the main advanced process by which xenobiotics were believed to move through body membranes. The

Copyright © 2008, Elsevier Ltd
All rights reserved.

FIGURE 34.1 Schematic representations of the A, D, M, E processes of drugs and xenobiotics in virtual biological systems (body, organ, cell …) illustrating the fate of a drug where membrane permeation is either mediated by passive diffusion "diffusional pharmacokinetics" or by a combination of passive diffusion and active transport processes "vectorial pharmacokinetics." The coordinated activities between transporters and metabolizing enzymes has led to identify these sequential events by Phases 0, I, II and III (X = drug or xenobiotic; CYPs = cytochromes P450; CEs = conjugation enzymes).

recent intrusion of drug transporters means that there is no single mechanism by which drugs permeate through the membranes. The presence of transporters at membranes that facilitate the movement of solutes into cells – influx (import) transporters, and of transporters that remove substances from the cytosol of cells – efflux (export) transporters, modulates the traditional theory of "diffusional pharmacokinetics" toward "vectorial pharmacokinetics" in which ADME processes are more deterministically governed (Figure 34.1).

We can now prepare a fairly complete list of drug transporters, the tissues in which they occur and function, how they are regulated or mutated, and the clinical relevance of their presence in normal and diseased tissues.

II. BIOLOGY AND FUNCTION OF TRANSPORTERS

A. Modes of active transport

Several types of transporters have been identified; they differ in their energy source and the direction of transport. The primary active transport systems are coupled to an energy source like the hydrolysis of adenosine triphosphate (ATP) by ion pumps (ATPases), and the ABC transporters move their substrate in one specific direction; movement is independent of the solute concentration gradient. They are primary transporters because no additional biochemical step is needed for solute transport. The second group of transporters are the co-transporters; these use a voltage and/or ion gradient to transport both ions and solutes together. They are uniporters when only one species is transported and symporters when both species are transported in the same direction, whereas the antiporters transport solutes and ions in opposite directions. The H^+ ion is the most common form of energy in prokaryotes, while Na^+ is more frequently encountered in eukaryotic cells. Other sources of energy involve HCO_3^-, glutathione (GSH) and the dicarboxylate, α-ketoglutarate (α-KG). Some of them only transport solutes in a direction defined by the solute concentration gradient with the so-called facilitative diffusion pathway. The main role of the Na^+, K^+-ATPase system is to activate the cascades of multiple co-transport processes. These co-transporters are also called secondary or even tertiary transporters because the machinery of ion transport must be activated by one or two pumps before solute transport occurs (Figure 34.2).

B. Genes and classification

Following the determination of the prokaryotic and eukaryotic genomes, it was predicted that 15% of the 23,000 genes in the human genome code for transport proteins – nearly 3,500 transporters. They are clustered in several superfamilies and only the members of three superfamilies are presently known to affect drug transport. They are the ABCs, the SLCs and the multidrug and toxin extrusion (MATE) transporters. ABC proteins are widespread in all organisms, from bacteria to mammals, with about 600 referenced transporters, but only 48 genes have been identified in humans and no more than around nine ABCs have been shown to affect drug pharmacokinetics and pharmacodynamics.[1] The SLC family, which may have about 2,000 members, is presently known to have 46 families, including 475 transporter genes with documented transport functions.[2] The MATE emerged very recently, and only two proteins (MATE1 and MATE2) are presently known to efflux drugs in mammals, whereas 861 related sequences are found in the three kingdoms of living organisms (Eukarya, Archaea and Eubacteria).[3] The transporters have not yet been completely assigned to superfamilies, families and subfamilies. The Human Genome Organisation

FIGURE 34.2 ABC transporters (green) that transport the substrate (S) in one defined direction are called primary transporters because no other additional biochemical step than the ATP hydrolysis or GSH co-transport is needed for (S) transport. SLC transporters (3, pink) need the activation by one or two ion transporters before S transport occurs. In this model, 1, is the Na^+, K^+-ATPase and 2, the Na^+, H^+ antiporter providing the H^+ driving force for S transport by 3.

(HUGO) Nomenclature Committee Database has provided a list of genes and defined the families of these transporters (see http://www.gene.ulc.ac.uk/nomenclature/). MATE1 and MATE2 were recently ranked by the HUGO as members of the SLC family (SLC47) and not as an independent superfamily. This review uses the HUGO as the primary reference for identifying genes and proteins. Human proteins (genes) are shown in capitals [e.g. ABCB1 (*MDR1*)], while rat and mouse proteins (genes) are indicated by an initial capital followed by small letters (proteins) and small letters (genes) [e.g. ABCC1 (*mdr1*)].

C. Basic structure

Transporters are integral membrane proteins that typically have 12 transmembrane domains (TMDs), although some have 6, 8, 10, 11, 13 or even 17 TMDs. The TMDs are folded in α-helical structures within the membrane and linked at both sides by amino acid sequences floating in the internal or external cell environment. The amino acids in the external loop domains are frequently *N*-glycosylated, while those of the intracellular loops of SLC, ABC and MATE proteins bear phosphorylation sites; the ABCs also have one or two ATP-binding domains. The 3D structure of TMDs is a crown shape, and they look like a channel allowing communication between the two fluid spaces separated by the lipid bilayer (Figure 34.3). Many SLC and MATE transporters have 300–800 amino acid residues and a molecular mass of 40–90 kDa, while the ABC transporters are larger, with 1,200–1,500 residues and a mass of 140–190 kDa. The amino acid sequence can be used to classify transporters by family and subfamily according to the degree of amino acid homology. For example, a transporter protein is assigned to a specific SLC family if it has an amino acid sequence that is at least 20–25% identical to those of other members of that family. A new nomenclature system was recently proposed that is based on the classification of drug metabolizing enzymes; the transport proteins of a superfamily are arranged in clusters of families (⩾40% identity) and subfamilies (⩾60% identity). Amino acid sequences are also extremely helpful for assessing the effect of a single mutation within the sequence that can change the conformation of the transport protein and alter its transport functions.

D. Distributions and properties of transporters in tissues

1. *Cellular and subcellular distribution*

There are about 200 types of cells in human tissues and all their plasma membranes and the membranes of their organelles contain transporters. The drug transporters at the organelles may well become most important in the future. This was recently documented during a dramatic Phase II trial in which the nucleoside antiviral fialuridine (FIAU) caused the death of subjects as a result of severe toxicity

FIGURE 34.3 Topological models inserted in the lipid bilayer membrane of the best-characterized ABC and SLC transporters that transport drugs and toxicants. (A) Model of MDR1 (P-gp) resembles MRP4, 5, 8 and (B) model of the human *SLCO,* the organic anion polypeptide transporter (OATP). Features common to all members of the OCT, OAT, OATP transporter family include 12 transmembrane spanning domains (TMDs represented by yellow rectangular bars) with intracellular N and C termini. NBD, nucleotide binding domain; *N*-glycosylation sites (indicated by branches) are present on extracellular protein loops. Cytoplasmic (IN) and extracellular (OUT) orientations are indicated.

including hepatotoxicity, pancreatis, neuropathy or myopathy. These toxic events were clearly linked to mitochondrial damage due to the transport of FIAU into mitochondria by an SLC transporter in the human mitochondrial membrane. Very little attention has been paid to the processes regulating transport across the endoplasmic reticulum (ER) membrane, although they are required for the activities of intraluminal UDP-glucuronosyltransferases (UGTs). It is now evidenced that the presence of multiple ER glucuronide transporters of different specificities in the ER membrane allow the import of the aglycone substrate and UDP-glucuronic acid and the exit of the conjugated end products, which are impermeable bulky polar and charged molecules, to lipid bilayers (Figure 34.4). The size and shape of the aglycone are critical determinants of transporter specificity, rather than the glucuronic acid moiety and hydrophobicity. The fact that glucuronide transport in the ER membrane is independent of ATP and GSH suggests that the translocation is mainly mediated by several SLC transporters. The presence of transporters in mitochondria, the ER and other constituents of the cell cytosol opens the possibility that the intracellular kinetic trafficking of xenobiotics and their metabolites may be mediated by active transport processes.

2. *Polarized expression of transporters in epithelia and vectorial transport*

The location of transporters at the cell plasma membrane is a critical issue because most of the cells involved in the A, D, E pharmacokinetic processes are polarized. Hence their apical (luminal) and basolateral (abluminal) membranes do not have the same populations of transporters (Figure 34.4). The same transporter is rarely found at both the apical and basolateral membranes. But most of the ABC and SLC transporters are located at either the apical or the basolateral epithelial membranes, and their location helps to define the direction of substrate transport and the resulting pharmacokinetic event.

For example, some SLC on the sinusoidal (basolateral) membrane of hepatocytes take up organic anions, while the ABC on the apical membranes of bile canicular cells excrete them. The combined activities of these two transporters thus results in the vectorial transport of drugs from the blood to the bile. Similarly, the basolateral transporters of the kidney tubular cells act in a coordinated, vectorial manner with apical transporters to secrete organic cations (OCs) from the blood to the urine.

FIGURE 34.4 A schematic proposal for cellular and subcellular distribution of ABC (green) and SLC (pink) transporters on the apical and basolateral membranes of an epithelial cell. SLC import and ABC export at the apical membrane of the unchanged xenobiotic (X), and ABC export of glucuronated-X (X-OG) at the basolateral membrane; subcellular events include mitochondrial uptake of X by a SLC transporter and X-OG efflux from the lumen of the ER where cytochrome P450 (CYP) and UDP-glucuronosyltransferase (UGT) produce X-OH and X-OG metabolites.

3. *Coordination between transporters and metabolic enzymes*

Drug metabolism was considered to be one of the main processes for removing xenobiotics prior to the emergence of transporters. The cytochrome P450 (CYP) isoenzymes catalyse the first step of biotransformation; this function was called Phase I metabolism, while the subsequent conjugation step was called Phase II metabolism. We now know that these two phases occur in specialized cells like the hepatocytes and enterocytes and that they are preceded and followed by two other phases controlled by transporters. Efflux or influx transporters reduce or increase the uptake of substrates, and these actions help to regulate the amounts of a xenobiotic reaching the enzyme binding sites or the rate at which the metabolites produced are eliminated. The first step has been called "Phase O" and the second "Phase III," indicating a close relationship between transporters and enzymes (Figures 34.1 and 34.4). They provide the cell with a suite of processes that may operate in parallel and in series. This integrated biological function of combined transport and metabolic processes is strongly supported by the presence of common regulation pathways that act via similar nuclear receptors, such as PXR, RXR and others, to induce or repress the genes encoding enzymes and transporters.

4. *Polyspecific transport and inhibition*

The substrate specificities of transporters are often very broad, as indicated by the many overlaps of substrates and inhibitors, much like the specificity of the drug metabolism enzymes. Thus, probenecid was initially known to produce many drug interactions by blocking the secretion of many drugs by the kidney, including the penicillins and the antiviral Tamiflu®. Probenecid is today known to be a polyspecific inhibitor of several ABCs and SLCs. Thus, all ionized chemicals, peptides and nucleosides that cannot diffuse freely across membranes are very likely to interact with one or more transporters.

5. *Transport kinetics and variability*

As each transporter has a limited capacity like the metabolism enzymes, it can be saturated by substrate concentrations greater than its K_m. The K_m of transporters can vary from nM to mM values, and the risk of saturating transport will depend on the amount of substrate in the transporter environment. Transport can also be inhibited in a competitive or non-competitive manner, in the same way as the drug metabolizing enzymes, so that transporters can promote drug–drug interactions that were initially thought to be due to the drug metabolizing enzymes alone. *In vitro* transporter assays are increasingly being used to assess the potential risks of drug–drug interactions mediated by transporters. The *in vitro* inhibition constant (K_i) can be measured and used to predict changes in the clearance or systemic exposure by measuring the area under the curve (AUC) (see Chapter 31 for a discussion of AUC). Transport kinetics may also depend on the amount of transporter,

which will depend on the actions of drugs, nutrients and disease states on the nuclear receptor pathways mentioned above. The most recent area of variation concerns the presence of genetic polymorphisms. However, studies on the pharmacogenetics of most drug transporters have only recently begun.

III. TRANSPORTERS IN DRUG DISPOSITION

This section covers only those transporters that influence the A, D or E of drugs and xenobiotics. About 40 transporters belonging to the ABC and SLC superfamilies are presently known to influence the pharmacokinetics, pharmacodynamics and toxicity of drugs and xenobiotics. They are often classified according to the chemical nature of their substrates. Hence, they translocate organic anions or cations, peptides or nucleosides. Most of them were first named according to their specific chemical substrate, such as the organic cation transporters (OCT) or the organic anion transporters (OAT), before they were named using the HUGO nomenclature rules.

A. ABC transporters

Most of the ABC transporters in eukaryotes move compounds from the cytoplasm to the outside of cells. This means that they are frequently called efflux pumps. Phylogenetic analysis has grouped the eukaryotic ABC genes into seven subfamilies (A–G). Only three of these subfamilies, B, C and G, contain transporters that influence drug disposition.[4]

1. ABCB *subfamily*

ABCB1, also called P-glycoprotein (P-gp) is a 170-kDa protein that was the first human ABC transporter cloned. It is responsible for the MDR phenomenon that occurs with such anticancer agents as the anthracyclines, vinca alkaloids and taxanes. P-gp is the product of two MDR genes in humans, MDR1 and MDR2 (also called MDR3), and only the MDR1 protein is involved in the MDR phenotype. Two genes, mdr1a and 1b, result in a similar MDR phenotype in rodents. P-gp is present mainly on the apical membrane of many secretory cells, including those of the intestine, liver, kidney, and adrenal gland. In the placenta, P-gp is found on the apical surface of syncytiotrophoblasts, where it can protect the fetus from toxic xenobiotics. P-gp is also abundant on hematopoietic stem cells, where it may protect the cells from toxins, and on the luminal surface of endothelial cells forming physiological barriers like the blood-testis and BBB. P-gp transports not only antineoplastic agents, but also a wide variety of structurally dissimilar substrates; they are mostly hydrophobic compounds that are either neutral or positively charged and are presented to the transporter binding sites directly from the lipid bilayer. The range of substrates that P-gp recognizes overlaps those of the main drug metabolizing enzymes CYP3A4/5. These enzymes are known to metabolize about 50% of the drugs currently on the market. Immunosuppressive agents (cyclosporin A and its analog PSC833), cardiac glycosides (digoxin), protease inhibitors (saquinavir, indinavir), antibiotics (rifampicin), calcium channel blockers (verapamil) and quinoline (quinidine) have all been found to interact with P-gp as both substrates and inhibitors.

2. ABCC *subfamily*

Although P-gp is considered to be the major transporter responsible for drug export at the plasma membrane of many cells, MDR phenotypes that are not P-gp-mediated have been linked to several MRPs or ABCCs. At least five of them, MRP1, MRP2, MRP3, MRP4, MRP5 – and recently MRP8 – are likely to be involved in mediating drug resistance and affecting drug pharmacokinetics. Although several molecules are substrates of P-gp when the unconjugated cationic (vincristine, doxorubicin) and neutral (etoposide) compounds are transported, these MRPs preferentially transport anions (like many Phase II metabolites of drugs) conjugated to GSH, glucuronate or sulfate. The MRP4 and MRP5 proteins mainly confer resistance to cyclic nucleosides and purine analogs. They transport substrates by a different mechanism from P-gp; there may even be multiple mechanisms that include co-transport with GSH. All these isoforms are concentrated on specific areas of polarized cells, like the epithelial cells of the gut and kidney, and probably also in the endothelial cells of brain microvessels. MRP2, MRP4 and MRP8 are, like P-gp, found in the apical (luminal) membrane, while MRP1, MRP3, MRP4 and MRP5 are found in the basolateral (abluminal) membrane.

3. ABCG *subfamily*

There are presently four known human members of the G subfamily, ABCG1, ABCG2, ABCG5 and ABCG8. Three of them (ABCG1, ABCG5 and ABCG8) are all implicated in lipid transport. ABCG2 is important for drug resistance and drug disposition. It was cloned independently by three different groups and called BCRP, mitoxantrone-resistance protein (MXR), and placenta-specific ABC protein (ABCP) before it was designated ABCG2. This second member of the G subfamily confers resistance to anticancer agents like mitoxantrone, topotecan and irinotecan and flavopiridol but not to paclitaxel, cisplatin or vinca alkaloids. BCRP also actively transports structurally diverse organic molecules, both conjugated and unconjugated, such as SN38, the metabolite of irinotecan and its glucuronide conjugate SN38-G, estrone-3-sulfate, 17β-E2G, DHEAS and organic anions

like methotrexate. Other BCRP substrates include tyrosine kinase inhibitors like imatinib mesylate (Gleevec®), which may also be a potent inhibitor of BCRP, nucleotide reverse transcriptase inhibitors like zidovudine (AZT), its active metabolite AZT5′-monophosphate, lamivudine (3TC) and the proton pump inhibitor pantoprazole. BCRP can transport chemical toxins such as pheophorbide a, a breakdown product of dietary chlorophyll that is phototoxic, and the small heterocyclic amine carcinogen PhIP that causes mammary and prostate cancers.

BCRP inhibitors include the fungal toxin derivative fumitremorgin C (FTC) and several dietary flavonoids. BCRP lies primarily in the plasma membrane and at the apical membrane of polarized epithelia, as does P-gp. High concentrations of BCRP are found in the placental syncytiotrophoblasts, the apical membrane of the epithelium of the small intestine, the membranes of liver canaliculi and at the luminal surface of the endothelial cells of the brain microvessels that form the BBB. Thus, ABCG2 is found mainly in organs that are important for absorption (proximal part of the small intestine), distribution (placenta and the BBB) and elimination (liver, kidney and small intestine). BCRP has been found in stem cells where it protects them from cytotoxic substrates. BCRP was also recently shown to secrete drugs or toxins into milk. BCRP lies in the apical membrane of the mammary gland alveolar epithelial cells, at the main site of milk production. The milk-to-plasma ratios of several drugs such as acyclovir, cimetidine and nitrofurantoin, were found to be high even before they were known to be BCRP substrates. The secretion of xenobiotics into milk by BCRP is puzzling because this function exposes the suckling infant to a range of drugs and toxins.

B. SLC transporters

The system for naming members of the SLC superfamily differs somewhat from the ABC nomenclature. The genes are usually named using the root symbol *SLC*, followed by a number corresponding to the family (e.g. *SLC22*, solute carrier family 22), the letter A and finally the number of the individual transporter (e.g. *SLC22A2*). But there may be differences between families. The *SLC21* family encoding the organic anion-transporting (OATP) proteins has been reclassified as a superfamily with families and subfamilies much like the classification of drug metabolizing enzymes. The gene symbol then becomes *SLCO* (i.e. the "21" and the "A" have been replaced by the letter "O" for organic transporter) and the "OATP" symbol has been kept for protein nomenclature (e.g. *SLCO1A2* for the gene and OATP1A2 for the protein).

1. OATP (SLC21/SLCO) *transporters*

A total of 11 human OATPs have been identified to date. The OATPs were originally identified as uptake transporters, although some may function primarily in efflux.[5] The driving force for OATP-mediated transport is still not clear, but it is independent of ATP or sodium gradients. There is experimental evidence that bidirectional transmembrane transport can be mediated by anion (HCO_3^-) or GSH exchange. It has now been shown that many OATPs are polyspecific OAT with partially overlapping substrate specificities for a wide range of solutes including bile salts, the organic dye bromosulfopthalein (BSP), steroid conjugates (DHEAS, 17βE2G estrone-3-sulfate (E-3-S)), thyroid hormones, neuroactive peptides ((D-penicillamine 2,5) enkephalin (DPDPE), Leu-enkephalin and deltorphin II) and numerous drugs and toxins, such as the cardiotonic digoxin, the angiotensin-converting enzyme inhibitors enalapril and temocaprilat, the 3-hydroxy-3methylglutaryl coenzyme A (HMG-CoA) reductase inhibitor pravastatin. OATP substrates are mainly high molecular weight (>450 Da) amphiphatic molecules, mostly bound to albumin, that have a steroid nucleus or linear and cyclic peptides. Most OATPs, mainly those of the OATP1 family, are found in many tissues and are thought to be part of the body's detoxification system, helping to remove xenobiotics from the systemic circulation (e.g. drug uptake into hepatocytes). The rat Oatp1a1, Oatp1a4, Oatp1b2 and human, OATP1B3, OATP1B1 and OATP2B1 are all found in the sinusoid membrane of hepatocytes, where they are responsible for the uptake of xenobiotics for hepatic clearance. The hepatic OATPs may play a strategic role in drug–drug interactions and hepatotoxicity. For example, rifampicin is a potent inhibitor of both OATP transporters and CYP3A4. Thus, giving rifampicin with OATP substrates may reduce hepatic first-pass clearance and increase the bioavailability of an intrahepatically active drug like pravastatin and decrease its efficacy. On the other hand, induction of OATP gene expression could increase the hepatic uptake and the total body clearance of the substrate.

2. OCT (SLC22) *transporters*

The OCTs include three potential-sensitive proteins (OCT1, OCT2, OCT3) and three H^+-driven transporters of carnitine and/or cations (OCTN1, OCTN2 and CT2, also known as *SLC22A4, SLC22A5* and either *FLIPT2* or *SLC22A16*, respectively.[6] Both OCT1 and OCT2 are found primarily in the major excretory organs (kidney and liver) and to a smaller extent in the intestine and the brain, while OCT3 is much more widely distributed.

All three OCTs recognize a variety of OCs, including endogenous bioactive amines like acetylcholine, choline, epinephrine, norepinephrine, dopamine, and serotonin, and drugs like cimetidine, quinine, quinidine, prazosin, desipramine, verapamil and morphine. The nitrogen moiety of the weak bases bears a net positive charge at physiological pH, allowing them to interact electrostatically with the binding sites of the OCTs. The "type 1" and "type 2"

classifications of OCs were developed to study their uptake by the liver. Type 1 OCs are small (60–350 Da), monovalent hydrophilic compounds such as tetraethylammonium (TEA) and the parkinsonian neurotoxin 1-methyl-4-phenylpyridinium (MPP^+). In contrast, type 2 OCs are usually bulkier (>500 Da; e.g. anthracyclines) and polyvalent (e.g. d-tubocurarine). This classification helps to differentiate the mechanisms by which they are transported across polarized cells. Type 2 OCs are believed to diffuse across the basolateral membrane and to be exported across the apical membrane by MDR1. In contrast, the basolateral entry of type 1 OCs involves one or more transporters, including OCT1, OCT2 and OCT3, and their efflux at the apex may be mediated by OCTN1 and OCTN2 or MATE1 and MATE2. The OCTs generally mediate the bidirectional transport of substrate molecules, and this depends mainly on the membrane potential and not directly on the transmembrane gradients of Na^+ or H^+. Unlike the OCTs, which have a common energy-supply mechanism, the OCTNs differ markedly in their mode of action. OCTN1 supports electroneutral OC/H^+ exchange, OCTN2 supports both Na^+-dependent co-transport (e.g. carnitine) and electrogenic-facilitated diffusion (e.g. TEA and type 1 OCs), and OCTN3 mediates the electrogenic transport of carnitine. OCTN3 and CT2 are present only in the testes of mice and humans, where transported carnitine improves sperm quality and fertility. OCTN1 is most abundant in the kidney (at the apical membrane of tubule cells), small intestine, bone marrow, fetal liver, but not in the adult liver. OCTN2 is mainly found in the heart, placenta, skeletal muscle, kidney and pancreas. Both OCTN1 and OCTN2 have a low affinity for MPP^+, cimetidine and TEA, and OCTN2 plays a major role in carnitine homeostasis.

3. OAT (SLC22) *transporters*

As their name implies, small organic anions (300–500 Da) possess a net negative charge at physiological pH and their transepithelial transport into the negatively charged environment of the cell requires energy. The OATs (*SLC22* family) are found mainly in cells playing a critical role in the excretion and detoxification of xenobiotics. There are six members of the OAT family (OAT1, OAT2, OAT3, OAT4, OAT5 and URAT1), present mainly in the liver, kidney, placenta, brain capillaries and choroid plexus.[7]

The OAT proteins play a critical role in the excretion and detoxification of a wide variety of drugs, toxins, hormones and neurotransmitter metabolites. A number of common non-steroid anti-inflammatory drugs (NSAID), including acetyl salicylate and salicylate, acetaminophen, diclofenac, ibuprofen, ketoprofen, indomethacin, and naproxen, are substrates of one or more OAT isoforms, so that there can be significant interactions between NSAIDs and other drugs. The β-lactam antibiotics (penicillins, cephalosporins and penems) and the antiviral nucleosides adefovir, cidofovir, acyclovir and AZT are also substrates of one or more OAT isoforms and are actively excreted in the urine. Toxins like chlorinated phenoxyacetic acid herbicides, mercuric conjugates, cadmium and ochratoxin A are also transported either into the renal tubule cells or hepatocytes via the OAT network, and this predisposes these tissues to nephrotoxicity or hepatotoxicity.

4. PEPT1 (SLC15A1) *and* PEPT2 (SLC15A2) *transporters*

PEPT1 and PEPT2 translocate dipeptides and tripeptides and their pharmacological importance is due to their ability to transport a wide variety of peptide-mimetic drugs, such as β-lactam antibiotics of the cephalosporin and penicillin classes and drugs like captopril, enalapril and fosinopril. Other drugs include the dopamine D2 receptor antagonist sulpiride and the peptidase inhibitor bestatin.[8] PHT1 and PHT2 transport histidine and certain di- and tripeptides, but their location on the cell or lysosomal membranes remains as questionable as their implication in pharmacotherapy. PEPT1 is the low-affinity (mM range) high-capacity transporter that is mainly found in the apical membranes of enterocytes in the small intestine, in renal proximal tubule cells of the S1 segment, and in bile duct epithelial cells. In contrast, PEPT2 is a high-affinity (μM range) low capacity transporter that is more widely distributed in the apical membranes of kidney tubule cells of the S2 and S3 segments, brain astrocytes and epithelial cells of the choroid plexus. They are involved in the uptake of their substrates, leaving a basal transporter(s) to account for the exit. This basal transporter could be PHT1 and/or PHT2, or the amino acid transporters of the *SLC1* and *SLC7* families. Both PEPT1 and PEPT2 can mediate the renal reabsorption of the filtered compounds in kidney tubules, whereas PEPT2 may be responsible for the removal of brain-derived peptide substrates from the cerebrospinal fluid via the choroid plexus. The pharmaceutical relevance of these peptide transporters is closely linked to the design of drug delivery strategies mediated by the intestinal PEPT1. One successful approach has been to produce peptide derivatives of parent compounds as substrates for PEPT1. The pharmacophoric pattern for the transporter includes the rules that the peptide bond is not a prerequisite for a substrate and that 5′-amino acid esterification, mostly using L-valine or L-alanine, markedly improves recognition by PEPT1. This prodrug strategy was used to improve the bioavailability of oral enalapril from 3–12% to 60–70% for the ester enalaprilat, which resembles the structure of the tripeptide Phe-Ala-Pro. The oral bioavailability of the nucleoside antiviral acyclovir (22%) was similarly improved by adding a valine residue to give valacyclovir (70%). Current studies on the regulation of PEPT1 and PEPT2 synthesis in inflammatory intestinal diseases may provide helpful information on the variations in bioavaibility of oral PEPT1 drug substrates.

5. CNT (SLC28) *and* ENT (SLC29) *transporters*

The members of the human *SLC28* and *SLC29* families catalyse mainly the transport of purine and pyrimidine nucleosides.[9] Hydrophilic nucleosides, like the purine adenosine, are important signaling molecules that control both neurotransmission and cardiovascular activity. They are also precursors of nucleotides, the constitutive elements of DNA and RNA and are the basic elements of a variety of antineoplastic and antiviral drugs. The *SLC28* proteins in the apical membranes of polarized cells work in tandem with the *SLC29* proteins found in the basolateral membrane. The *SLC28* family consists of three dependent concentrative nucleoside transporters that differ in their substrate specificities. CNT1 transports naturally occurring pyrimidine nucleosides plus the purine adenosine. Several antiviral analogs, like AZT, lamivudine (3TC) and ddc, are substrates of CNT1. The cytotoxic cytidine analogs, cytarabine (AraC) and gemcitabine (dFdc), are also transported by CNT1. CNT1 is primarily found at the apical membrane of epithelial cells, including those of the small intestine, kidney and liver. Human CNT2 is widely distributed in the kidney, liver, heart, brain, intestine, skeletal muscle, pancreas and placenta. CNT2 transports purine nucleosides and uridine. Pharmaceutical substrates include the antiviral didanosine (ddI) and ribavirin.

Human CNT3 has, like CNT2, a wide tissue distribution with high concentrations in the pancreas, bone marrow and mammary gland. CNT3 is broadly selective and transports both purine and pyrimidine nucleosides in a 2:1 Na^+ nucleoside coupling ratio – in contrast to the 1:1 ratio employed by CNT1 and CNT2. CNT3 transports several anticancer nucleoside analogs including cladrabine, dFdc, fludarabine and zebularine.

The human *SLC29* family has four members. ENT1 and ENT2 can both transport adenosine, but differ in their abilities to transport other nucleosides and nucleobases. ENT1 is almost ubiquitous in human and rodent tissues and transports purine and pyrimidine nucleosides with K_m values of from 50 μM (adenoside) to 680 μM (cytidine). The antiviral drugs ddC and ddI are also poorly transported. ENT2 is present in a wide range of tissues including the brain, heart, pancreas, prostate, and kidney and is particularly abundant in skeletal muscle. ENT2 differs from ENT1 in that it can also transport nucleobases like hypoxanthine and AZT. ENTs also mediate the uptake and efflux of several nucleoside drugs because of their bidirectional transport property.

6. MATE *transporters*

The MATE transporters are involved in MDR, preferentially to OCs, via a H^+ or Na^+-coupled antiport mechanism. In humans, two genes encode MATE1 and MATE2. MATE1 is ubiquitous throughout the body, but is most abundant in the luminal membrane of the urinary tubules and bile canaliculi in the liver. By contrast, MATE2 is specific to the kidney brush border membranes. Both MATEs are responsible for the final step in the excretion of metabolic waste and xenobiotic OCs with very diverse chemical structures by the kidney and liver.

MATE1 was shown to transport TEA and 1-methyl-4-phenyl pyridinium (MPP^+). MATE2 also transports multiple OCs including TEA, MPP^+, cimetidine and metformin.[3]

IV. ROLES OF TRANSPORTERS IN DRUG PHARMACOKINETICS, PHARMACODYNAMICS AND TOXICOLOGY

Transporters are now recognized to be as important as the metabolizing enzymes in the modulation of the main steps controlling the fate and action of xenobiotics in the body. They affect all the main pharmacokinetic events like the oral bioavailability, distribution and clearance of substrates. They are presently known to modulate the active drug concentration in all biophases and influence the effects of drugs.

A. Intestinal absorption

Both influx and efflux transporters modulating drug absorption are present in the epithelium of the various segments of the intestine[10] (Figure 34.5). PEPT1, OATP1A2, OATP2B1, OATP3A1 and OATP4A1 are all found on the apical membrane and mostly import substrates from the lumen into the circulation. PEPT1 is the best-characterized drug transporter in the small intestine of mammals and is widely used to improve the absorption of poorly absorbed oral drugs using a prodrug strategy. ABC transporters, including MDR1, MRP2 and BCRP, are also present on the apical membrane, where they either limit the intestinal uptake of their substrates or contribute to the active secretion of drugs from the blood to the intestinal lumen. For example, the antineoplastic agent, paclitaxel, a P-gp substrate, is poorly absorbed when taken orally by humans – only 5% is bioavailable, but when it is administered with the P-gp competitor, cyclosporin A, its bioavailability is increased to 50%. The roles of basolateral transporters are much less well known. Oct1 is present on the basolateral sides of cells, and studies using Oct1 knockout mice indicate that Oct1 is important for the secretion of OCs into the lumen of the small intestine.

The intestinal transporters are not uniformly distributed along the crypt-villus axis. Many of those implicated in the absorption of drugs, like PEPT1, MDR1, BCRP, MRP2 and MRP3, are villus-specific. This restriction of transporters to the villus is also correlated with the presence of CYP3A in intestinal cells, suggesting coordinated Phase 0 and I activities of MDR1 and CYP3A4 in the so-called intestinal first-pass effect. A major concern is the way their densities vary along the gastrointestinal tract. For example,

FIGURE 34.5 Distribution of the main drug ABC (green) and SLC (pink) transporters on the apical and basolateral membranes of the human intestinal enterocytes. All apical transporters (except MRP1) lie at the top of the villi. They define active absorption and/or secretion of drugs and xenobiotics through the intestinal epithelium.

MRP3, is the most abundant ABC protein throughout the human intestine, except for the terminal ileum where MDR1 is most abundant. Similarly, the concentration of MDR1 increases from the duodenum to the colon, whereas BCRP is found throughout the small intestine and colon, and MRP2 is most prevalent in the duodenum and becomes undetectable toward the terminal ileum and colon. These diverse distributions of the intestinal transporters may have dramatic pharmaceutical consequences. The pharmaceutical form of an oral drug can vary from a simple solution to a solid controlled-release complex, and this can influence the gastrointestinal site (stomach, duodenum, jejunum, ileum or colon) at which the active compound is released. Such differences may also influence the efficacy of the carrier-mediated transports, as these may vary from one region of the intestine to another. The great risk of saturating active transport is that it can also affect the kinetics of drug absorption. This can occur when a large amount of drug is rapidly dissolved in the intestinal lumen, ready to be absorbed by a relatively small area of intestine. Active transport can be saturated by a relatively high concentration of substrate, so shifting absorption toward diffusion. Here, too, the properties of the oral preparation, like its rate of dissolution, may influence the contribution of active transport to drug absorption.

B. Liver and hepatic clearance

Hepatic clearance is a combination of metabolic (Phase I and II) and biliary clearance. As previously indicated, hepatocytes can take up drugs by diffusion or active transport (Phase 0)[11] (Figure 34.6). The basolateral membrane OAT include OAT2, OAT4, OATP1A2, -1B1, -1B3, and -2B1, the OCT1 and the Na-taurocholate co-transporting polypeptide NTCP (SLC10A1). They are responsible for the uptake by the liver of a wide variety of drugs because of their broad, overlapping substrate specificities. Phase III, which follows Phases 0, I and II, results in the elimination of the intact drug and/or metabolite(s) via efflux transporters on the apical and basolateral membranes. The hepato-biliary transporters include several ABC proteins (MDR1, MDR3, MRP2, BSEP and BCRP and MATE1) that are the main mediators of the excretion of numerous endogenous conjugated and unconjugated bile salts and drugs via the bile. Phase III also includes the efflux of compounds from hepatocytes back into the systemic circulation via basolateral membrane efflux transporters. Some of the OATPs, OATs and OCT1 are bidirectional and may facilitate efflux, but the main exporters are the ABC proteins, which transport a wide range of glucuronides, and sulphated and GSH conjugates. The main ones are MRP1 and MRP3, whose synthesis is readily induced, and the cyclic nucleoside transporters MRP4 and MRP5.

This huge network of hepato-biliary transporters can give rise to variations in drug disposition between individuals by modulating the uptake or the exit of drugs and their metabolites from hepatocytes. A change in hepatic uptake may have clinical consequences. It may modulate the pharmacological activity of drugs that act via the intrahepatocellular transduction pathways, it may cause hepatotoxicity, or give rise to drug–drug interactions. The concentration of the cholesterol-lowering HMG-CoA inhibitors in hepatocytes must be adequate for their pharmacological activity, and most of the statins – like pravastatin, simvastatin, lovastatin, cerivastatin and pitavastatin – enter hepatocytes via OATP1B1, and to a lesser degree via OAT1B3. Recently identified genetic polymorphisms like the SLCO1B3

FIGURE 34.6 Distribution of the main drug ABC (green) and SLC (pink) transporters on the basolateral (sinusoïd) and apical (bile canaliculus) membranes of the human hepatocytes. SLC transporters at the basolateral membrane mainly define active hepatic uptake whereas ABC transporters at the basolateral and bile canaliculus membranes efflux drugs and their metabolites in blood or bile, respectively.

haplotype *17 are associated with reduced uptake of statins by the liver and lower concentrations in hepatocytes; they thus have less effect on cholesterol synthesis. Large scale clinical studies are needed to confirm the impact of OATP1B1 polymorphisms on the considerable variation between individuals to therapy with hypolipidemic agents. Transporters can also mediate hepatotoxicity. For example, the sulfate conjugate of the antidiabetic troglitazone can cause troglitazone hepatotoxicity by inhibiting OATP1B1 and OATP1B3. These hepatic impacts of the basolateral transporters have their counterpart at the apical pole. The multiple ABC transporters may also be responsible for variable drug disposition. For example, giving patients on digoxin the P-gp inhibitor verapamil decreases the biliary clearance of digoxin by 43% and increases its plasma concentration by 44%.

C. Blood barriers and tissue distribution

The tissue distribution of a drug can be affected by transporters because they lie on the luminal or abluminal membranes of the endothelial cells of the tissue blood vessels, or on the membranes of the specific cells of the underlying organ. The transporters on the membranes of the blood vessels may be several key physiological components of the blood barriers throughout the human body if tight junctions seal adjacent cells and prevent the paracellular exchange of solutes. In contrast, solutes can freely communicate between extracellular spaces when blood vessels are fenestrated, as in the liver sinusoids, and transporters on the plasma membranes of the tissue cells (e.g. the hepatocyte membranes) become the first barrier regulating the import and export of solutes. Several organs, including the brain, nose, eyes, testes, prostate and placenta, are protected by endothelial barriers that contain extensive networks of transporters. Figure 34.7 illustrates the luminal and abluminal distributions of several transporters at the BBB.[12] The two ABC proteins MDR1 and BCRP are most abundant on the luminal side of the endothelial cells and are most important for protecting the brain from numerous xenobiotics. Few SLCs have been characterized in the rat BBB, except for the important network of SLC transporters that allows the blood–brain exchange of amino acids and sugars. Rat Oatp1a4 is found on both the luminal and abluminal membranes of the brain capillaries. The human isoform OATP1A2 is also present, but its membrane location has not been determined. Both OATPs can mediate uptake or efflux transport because of their bidirectional transport characteristic. The members of the SLC22, OAT3, OCTN2 and URAT1, have been found in the BBB. OAT3 is abluminal and effluxes benzylpenicillin, cimetidine, PAH and several acidic metabolites of neurotransmitters from the brain to the blood. The luminal position of URAT1 enables this vectorial transport of the OAT3 substrates. OCTN2, which is believed to be luminal, can simultaneously transport carnitine into the brain and efflux OCs from the brain to the blood. Here, too, drug transporters on the membranes of physiological barriers or on specific membranes of the tissue cells can affect drug distribution and consequently the fraction of the drug available for binding to intracellular receptors or other biological targets.

D. Kidney and renal clearance

The presence of a drug in the urine is the net result of filtration, secretion and reabsorption. Filtration occurs by passive glomerular filtration of unbound plasma solutes, whereas secretion and reabsorption are generally carrier mediated. They can occur in the proximal tubule, which has three segments (S1, S2 and S3), the loop of Henle, the distal tubule and the collecting tubule. These specific anatomical and functional regions of the kidney must be carefully considered, just like the regions of the intestine, because region-specific distributions of transporters define

FIGURE 34.7 Distribution of the main drug ABC (green) and SLC (pink) transporters on the abluminal (facing brain extracellular fluid) and luminal membranes of the brain microvessel endothelial cells constituting the BBB.

FIGURE 34.8 Distribution of the main drug ABC (green) and SLC (pink) transporters on the basolateral (peritubular fluid) and apical (glomerular filtrate) membranes of the kidney proximal tubule cells. Active secretion and reabsorption help to define the overall renal clearance of drugs and xenobiotics.

their action in renal clearance. Most renal transporters lie on the apical and basolateral membranes of the proximal tubule cells, with fewer on the epithelial membrane of the other components of the nephron. The resulting vectorial transport from the peritubular capillaries to the tubule lumen or vice versa can produce either secretion or reabsorption. Figure 34.8 shows the locations of the major drug transporters in the human proximal renal tubule cells.[13] Multiple SLC transporters at their basolateral membrane (close to peritubular capillaries) mediate drug uptake into the tubule cells. Although, by nature, bidirectional, the direction of the transmembrane driving gradients favors tubular uptake rather than the efflux of organic anions and cations. Organic anions enter these cells via OAT1, OAT2 and OAT3 – and probably via OATP1A2 and OATP4C1, which was recently identified and transports digoxin and methotrexate. OCs are similarly transported by OCT1, OCT2 and OCT3; the efflux transporters MRP1, MRP3, MRP5 and MRP6 mediate their efflux back into the systemic circulation. At the apical membrane, OAT4 and URAT1, OCTN1 and OCTN2 can mediate drug transport with bidirectional properties, either secretion or reabsorption. For example, OCTN2 secretes OCs and reabsorbs zwitterions. OATP1A2, PEPT1 and PEPT2 mediate the reabsorption of their substrates from the tubule lumen.

The ABC transporters MDR1, BCRP MRP2 and MRP4 are also present on the apical membrane and efflux compounds by secretion. As indicated above, transporters are not evenly distributed along the nephron; MDR1, MRP2, MRP4 and MRP6 are found mainly within the three segments

of the proximal tubule; MRP3 lies in the distal convoluted tubule; and MRP1 is found in the epithelial cells of the loop of Henle and the distal and collecting duct tubule cells, but not in proximal tubule cells. The regional distributions of the SLC transporters are also specific. OAT1 is found only on the basolateral membrane of the S2 segment cells of the proximal tubule, whereas OAT3 is present on the cells of the S1, S2 and S3 segments. This transporter network can be responsible for drug–drug interactions, nephrotoxicity and drug efficacy mediated by the reabsorptive and secretory capacities of the kidney. If the renal clearance of a drug is equal to or more than the overall body clearance, renal transporters can be important in clinical efficacy or toxicity. For example, the cephalosporin antibiotics are primarily eliminated via the kidney. Creatinine clearance is normally 100–140 mL/min, but the renal clearance of cephalosporins is 16.8–469 mL/min, suggesting that some of them, like cefotaxine and cefadroxil, are excreted into the urine by tubular secretion, whereas others, like ceftriaxone and cefazodin, whose renal clearance is less than that of creatinine, are reabsorbed. OAT1, OAT2 and OAT3 are located on the basolateral side of the proximal tubule and mediate the uptake of most of the cephalosporins into the proximal tubule from the peritubular capillary. The apical OAT4 mediates both the uptake (reabsorption from the tubular lumen) and the efflux (secretion) of these anionic antibiotics. Like the basolateral transporters of hepatocytes, which can modulate the pharmacological activity of drugs acting via intrahepatocyte targets or induce hepatotoxicity, the basolateral OATs can make some cephalosporins cause nephrotoxicity, which may lead to acute proximal tubular necrosis. This toxicity is mainly due to the accumulation of cephalosporin in the renal cortex because of the lack of efficient vectorial transtubular transport. This transport-mediated nephrotoxicity also results in the adverse effect of cisplatin and related drugs via their basolateral uptake in the proximal tubule by OCT2, and the toxic effects depend on the platinium complex used, as does the structure-dependent nephrotoxicity of cephalosporins. Nephrotoxicity also limits the use of the nucleoside phosphonates, adefovir and cidofovir, in the treatment of human immunodeficiency virus. The toxicity of these drugs appears to be a function of both OAT1-mediated proximal tubular accumulation and decreased efflux at the luminal membrane by MRP2. A small dose of the OAT1 inhibitor, probenecid, may reduce the nephrotoxicity of cidofovir. The use of transporter inhibitors to reduce nephrotoxicity suggests that drug–drug interactions affecting anionic and cationic drugs can be mediated via competition at the basolateral and luminal tubular transporters. Multiple drug–drug interactions have been reported with probenecid and cimetidine, and there have been fatal cases with methotrexate and NSAID following the inhibition of the basolateral OAT1 and OAT3. Finally, renal transporters can be critical for the action of diuretics. Tubular secretion is the main route by which diuretics act in the kidney and are excreted. The diuretic drugs like the thiazides, the loop diuretics bumetanide and furosemide, and the carbonic anhydrase inhibitors are all competitive inhibitors of the renal OATs, although their affinities and specificities vary.

V. CONCLUSION

The recent expansion of information on drug transporters in pharmacokinetics had added a new layer of complexity to our understanding of the mechanisms underlying the absorption, distribution and elimination of drugs. New transporters undoubtedly remain to be identified at the plasma membranes of both cells and organelles. Nevertheless, it is clear that drug transporters are significant determinants of variations in drug responsiveness between individuals, drug–drug interactions, drug-induced organ toxicities and diseases. Detailed knowledge of genetic polymorphisms in transporters and how they affect transporter function will help to optimize drug therapies and identify unknown, residual factors that influence subject-to-subject variations.[14]

Transporters are now an integral part of the drug discovery and development processes. They are attractive markers in the creation of drugs that are readily absorbed and accurately targeted. The incorporation of transport properties into structure–activity models should help medicinal chemists design more efficient, safer new medicines.

ACKNOWLEDGMENTS

The author is most grateful to Ms. Elisabeth Niel for providing illustrations and technical support, and to Dr. Owen Parkes for reviewing the English text.

REFERENCES

1. Dassa, E., Bouige, P. The ABC of ABCS: a phylogenetic and functional classification of ABC systems in living organisms. *Res. Microbiol.* 2001, *152*, 211–229.
2. Hediger, M. A., Romero, M. F., Peng, J. B., Rolfs, A., Takanaga, H., Bruford, E. A. The ABCs of solute carriers: physiological, pathological and therapeutic implications of human membrane transport proteins: introduction. *Pflug. Arch.* 2004, *447*, 465–468.
3. Omote, H., Hiasa, M., Matsumoto, T., Otsuka, M., Moriyama, Y. The MATE proteins as fundamental transporters of metabolic and xenobiotic organic cations. *Trends Pharmacol. Sci.* 2006, *27*, 587–593.
4. Leslie, E. M., Deeley, R. G., Cole, S. P. Multidrug resistance proteins: role of P-glycoprotein, MRP1, MRP2, and BCRP (ABCG2) in tissue defense. *Toxicol. Appl. Pharmacol.* 2005, *204*, 216–237.
5. Hagenbuch, B., Meier, P. J. Organic anion transporting polypeptides of the OATP/ SLC21 family: phylogenetic classification as OATP/ SLCO superfamily new nomenclature and molecular/functional properties. *Pflug. Arch.* 2004, *447*, 653–665.
6. Jonker, J. W., Schinkel, A. H. Pharmacological and physiological functions of the polyspecific organic cation transporters: OCT1, 2, and 3 (SLC22A1-3). *J. Pharmacol. Exp. Ther.* 2004, *308*, 2–9.

7. Sweet, D. H. Organic anion transporter (Slc22a) family members as mediators of toxicity. *Toxicol. Appl. Pharmacol.* 2005, *204*, 198–215.
8. Daniel, H., Rubio-Aliaga, L. An update on renal peptide transporters. *Am. J. Physiol. Renal Physiol.* 2003, *284*, F885–F892.
9. Pastor-Anglada, M., Cano-Soldado, P., Molina-Arcas, M., Lostao, M. P., Larrayoz, I., Martinez-Picado, J., Casado, F. J. Cell entry and export of nucleoside analogues. *Virus Res.* 2005, *107*, 151–164.
10. Takano, M., Yumoto, R., Murakami, T. Expression and function of efflux drug transporters in the intestine. *Pharmacol. Ther.* 2006, *109*, 137–161.
11. Chandra, P., Brouwer, K. L. The complexities of hepatic drug transport: current knowledge and emerging concepts. *Pharm. Res.* 2004, *21*, 719–735.
12. Ohtsuki, H. New aspects of the blood–brain barrier transporters; its physiological roles in the central nervous system. *Biol. Pharm. Bull.* 2004, *27*, 1489–1496.
13. Shitara, Y., Sato, H., Sugiyama, Y. Evaluation of drug–drug interaction in the hepatobiliary and renal transport of drugs. *Annu. Rev. Pharmacol. Toxicol.* 2005, *45*, 689–723.
14. Ho, R. H., Kim, R. B. Transporters and drug therapy: implications for drug disposition and disease. *Clin. Pharmacol. Ther.* 2005, *78*, 260–277.

Chapter 35

Strategies for Enhancing Oral Bioavailability and Brain Penetration

Brian C. Shook and Paul F. Jackson

I. INTRODUCTION
II. ENHANCING ORAL BIOAVAILABILITY
A. Metabolic stability
B. Structural rigidity
C. pK_a
D. Hydrogen bond interactions
E. Miscellaneous
III. ENHANCING BRAIN PENETRATION
A. Metabolic stability
B. pK_a
C. Log P
D. Hydrogen bond interactions
IV. CONCLUSION
REFERENCES

I. INTRODUCTION

One of the most time-consuming portions of a medicinal chemistry program is attempting to increase the bioavailability of a lead compound. There are several strategies available, depending on what the target organ is. In terms of increasing brain levels, the lipophilicity and molecular weight of the compound are usually the first focus. In addition, the overall polar surface area (PSA) is taken into account. With regard to increasing oral bioavailability, the guidelines described by Lipinski offer a good general start to the problem.[1] These guidelines, coupled with early metabolic profiling, provide a rationale for designing a path forward. This chapter summarizes some recent examples of structural modifications that have increased the bioavailability of compounds. The first section deals with efforts to increase oral bioavailability, while the second section describes examples in which structural modifications lead to increases in brain penetration.

II. ENHANCING ORAL BIOAVAILABILITY

A. Metabolic stability

A series of changes to a group of platelet-derived growth factor antagonists (**1**) identified isopropoxyl as the R substituent that most increased stability.[2] This compound displayed good *in vitro* activity, but potency decreased when assayed *in vitro* in the presence of plasma. In addition, the *in vivo* efficacy was poor. This suggested that the compound is undergoing plasma protein binding. The 6,7–di-(methoxyethoxy) group present in compound **1** was found to be a metabolic liability and was also not optimal for good solubility, as a result modifications were needed in these positions in order to identify the lead compound. Compound **2** was prepared, and although the *in vitro* potency decreased slightly, the compound showed no plasma protein binding while maintaining similar stability to compound **1**. Finally, compound **3** was prepared, and it was shown that the one-carbon extension exhibited both a 5-fold increase in potency and a 3-fold increase in metabolic stability. The biggest advantage of compound **3**, when compared to compound **2**, is seen when comparing the PK data. The oral bioavailability in dog was 14% for compound **2** and 98% for compound **3**, while both compounds have identical *in vivo* half-lives. The bioavailability in monkeys also showed a similar trend.

Compound **4** is a potent gonadotropin release hormone antagonist,[3] and as such could be useful in treating prostate cancer and endometriosis. This compound, however, showed low oral bioavailability and high clearance in dogs. The clearance was attributed to metabolism at the 6-position of the piperidine, and, as a result, a 6-trifluoromethyl piperidine analog (**5**) was prepared. This change resulted in a 3-fold increase in oral bioavailability and a 2-fold decrease in clearance rate.

Copyright © 2008, Elsevier Ltd
All rights reserved.

1

2

3

4

5

A series of macrocyclic 3-aminopyrrolidinones was designed and found to act as potent farnesyltransferase inhibitors.[4] Although **6** has good oral bioavailability ($F = 47\%$), it has a short *in vivo* half-life and a high clearance rate. The 1-benzyl substituent on the pyrrolidinone in **6** was suggested as a site of metabolism that may account for the short half-life. Optimization showed that a more rigid, less metabolically susceptible naphthalene substituent, as in **7**, resulted in a 21% increase in bioavailability, a 5-fold increase in half-life, and a 9-fold decrease in clearance.

6

7

Compound **8** represents a potent and selective molecular scaffold for growth hormone secretagogue receptor (GHS-R) antagonists.[5] Unfortunately, due to metabolic instability of the amide substituent, this compound has very poor bioavailability in rats ($F = 4\%$). The isoxazole ring was found to act as a suitable linker that holds the aromatic ring and amide substituent in the appropriate orientation required for binding. Therefore, **9** was prepared and showed modest activity allowing for further optimization. Efforts then focused on quaternizing the α-carbon of the amide to stabilize it against hydrolysis. The resulting carbamate (**10**) showed increases in metabolic stability *in vivo* that translated into reasonable bioavailability ($F = 19\%$) and improved clearance over the initial scaffold.

8

9

10

B. Structural rigidity

As can be seen from the following examples, the bioavailability of a compound can often be increased by modifications

that increase the structural rigidity of the molecule. For example, a series of thrombin inhibitors was prepared and optimized in the P2 pocket.[6] Structure–activity relationship (SAR) studies showed that replacing the flexible *N*-methyl alanine moiety present in **11** with the more rigid proline derivative **12** gave a 6-fold increase *in vitro* activity, a 5-fold increase *in vivo* efficacy, and enhanced oral bioavailability.

Compound **13** is representative of a series of potent and selective human cytomegalovirus inhibitors (HCMV).[7] This particular series has limited utility due to little or no oral bioavailability. Dihydroquinoline analog **14**, however, maintains *in vitro* activity comparable to **13,** but has acceptable bioavailability. Interestingly, the corresponding tetrahydroquinoline analog of **14** does not show any HCMV activity *in vitro*, suggesting that **14** has the optimal amount of flexibility, planarity, and stability for *in vitro* and *in vivo* potency. Further optimization of this series led to **15**, which has a >7-fold increase in *in vitro* potency with oral bioavailability equal to 23% and 60% when dosed at 5 mg/kg and 50 mg/kg, po, respectively.

C. pK_a

A strategy often used to increase the bioavailability of a compound is to modify the pK_a of basic functionalities in a molecule. For example, a series of 1-(2-naphthyl)-1 *H*-pyrazole-5-carboxylamides was developed as factor Xa inhibitors based on the structure of SN429 (**16**).[8] Although SN429 (**16**) is a very potent factor Xa inhibitor *in vitro*, the compound has poor bioavailability ($F = 4\%$) in dogs with a short *in vivo* half-life. Replacement of the basic amidino moiety with a naphthyl substituent produced **17**, which was less potent *in vitro* but, showed improved PK properties with 25% bioavailability and a 3-fold increase in half-life. Compound **17** was optimized further to increase potency, revealing that fluoro analog **18** increased potency while maintaining bioavailability comparable to that of **17**.

Compound **19** is a potent benzazepinone growth hormone secretagogue that suffers from poor bioavailability in dog ($F = 2\%$).[9] The molecule exists in a zwitterionic form that is presumably the cause for the poor bioavailability. Two key structural modifications were required to improve the PK profile. The acidic tetrazole was replaced with a number of amide and urea substituents before identifying the key methyl urea moiety, which resulted in an 8-fold increase in activity over **19**. Further, shortening the amino acid side chain decreased the pK_a of the amine from 9.2 to 7.5. The combination of changes provided compound **20**, which resulted in a 60-fold increase in activity over parent compound **19** and increased oral bioavailability in dogs to 24%.

D. Hydrogen bond interactions

Compound **21** is representative of a class of β_3-adrenergic receptor agonists.[10] The poor bioavailability ($F = 4\%$) of

21 in monkeys was significantly increased (F = 23%) by changing the 3-pyridyl analog to 2-pyridyl analog **22**. The increase in bioavailability is attributed to the increased intra-molecular H-bonding interactions, thereby minimizing inter-molecular interactions that limit oral absorption. Interestingly, the bioavailability in monkeys increases to 56% when **21** is tied into the corresponding morpholine analog (**23**).

A series of pyridopyrimidine analogs was prepared as potent adenosine kinase inhibitors.[11] Studies found that the C2' substituent on the pyridine ring did not bind to a specific site of the enzyme, but rather projected outside the binding pocket. This substituent did not alter the *in vitro* activity, and most of the metabolism occurred at that position. A number of analogs were prepared, but morpholine analog **24** was the first to show an acceptable PK profile. Compound **24** exhibited good *in vitro* and *in vivo* potencies, but suffered from moderate oral bioavailability (F_{rat} = 23%; F_{dog} = 62%) and a short half-life. The oxygen atom appeared to be necessary for obtaining *in vivo* activity and was further optimized. Spiral ketal derivative **25** increased rat bioavailability (F = 45%), but dog bioavailability was decreased (F = 39%). Finally, the 4-OTHP oxime derivative (**26**) gave the desired bioavailability in both rats and dogs (F_{rat} = 46%; F_{dog} = 100%), as well as increased half-life by 4-fold in both species.

A series of phenoxyphenyl sulfone *N*-formylhydroxylamines was identified that contained potent and selective matrix metalloproteinase (MMP) inhibitors.[12] Compound **27** displayed remarkable selectivity for both MMP-2 and MMP-9 over the undesired MMP-1. Unfortunately, **27** had no oral bioavailability. SAR modifications revealed that diol **28** maintained ideal selectivity, but suffered from low plasma exposure in monkey and a short *in vivo* half-life. Further modification showed that acetonide **29** had the desired selectivity for MMP-2 and MMP-9 and also displayed an optimal PK profile in monkeys with a plasma concentration 18-fold higher than **28**. Oral bioavailabilities for **29** were >70% in rat, dog, and monkey.

E. Miscellaneous

Noscapine (**30**) is a naturally occurring phthalideisoquinoline alkaloid obtained from opium that has weak anticancer properties.[13] Subtle changes to the corresponding phenol (**31**) and aniline (**32**) resulted in a >50-fold increase in activity in an *in vitro* tubulin polymerization assay. Moreover, the phenol shows a modest 1.5-fold increase in oral bioavailability compared to noscapine (**30**).

A series of cyclobutenediones (**33**) was prepared and found to act as potent CXCR2 and CXCR1 antagonists.[14] Optimization of this series revealed significant increases in plasma (rat) concentration when the size of the alkyl substituent in the 4-position of the furan is decreased from *i*-Pr (AUC = 5.4 μM) to ethyl (AUC = 17.4 μM) to methyl (AUC = 32.6 μM) after dosing at 10 mg/kg, po.

Inducible nitric oxide synthase (iNOS) is a target to potentially treat inflammatory conditions.[15] A series of 2-iminohomopiperidinium salts has been identified as selective inhibitors of iNOS. Although regioisomers **34** and **35** show similar selectivity profiles, isomer **35** is significantly more potent *in vitro*. Not surprisingly, **35** also showed superior *in vivo* activity, while isomer **34** was not orally active. Even though the bioavailabilties were not calculated, it seems clear that simply changing from one regioisomer to another has a dramatic effect on oral bioavailability.

Oral bioavailability is sometimes unpredictable, and does not fit the paradigms described above. For example, some epimers have essentially identical activity *in vitro*.[16] The bioavailabilities, however, are completely different: in dogs the *S*-isomer has 25% bioavailability, and the *R*-isomer has 0%. The literature has a large number of examples where subtle changes will often completely change the PK profile without a clear rational.

III. ENHANCING BRAIN PENETRATION

A. Metabolic stability

A series of dihydropyrrolo[2,3-d] pyrimidines was prepared as potent corticotropin-releasing factor-1 receptor antagonists.[17] Compound **37** exhibits moderate oral bioavailability (F = 21%) and a brain-to-plasma (B/P) ratio of 0.6. Compound **38**, on the other hand, has high oral bioavailability (F = 86%) and a B/P ratio of 2.3. The increases in bioavailability and B/P ratio are attributed to the 4-fold decrease in clearance exhibited by **38**, compared to **37**. In general, dialkylamine or cyclic aliphatic amine substituted pyrimidines had higher plasma clearance and volumes of distribution than compounds bearing aromatic substituents, which, in turn, resulted in compounds having better oral bioavailability and B/P ratios.

Compound **39** represents a series of aminopiperidine indazoles that are antagonists of the melanin-concentrating hormone receptor-1.[18] Indazole **39** demonstrated high exposure in the brain and had an excellent B/P ratio of 6.5. Alkylation of the *N*-1 nitrogen resulted in compound **40** having comparable *in vitro* potencies to **39**, but only moderate brain concentrations and a suboptimal B/P ratio of 0.4. Further, when the *N*-1 nitrogen was acylated to give **41**, a drastic decrease in brain penetration and half-life were observed, despite having comparable *in vitro* potencies when compared to **39** and **40**. Indazole **41** was rapidly cleared from the plasma with concentrations below the levels of quantitation. Presumably, this is a major reason for decreased brain penetration.

A series of quinazolinone inhibitors of poly(ADP-ribose) polymerase-1 (PARP-1) was prepared.[19] It was shown that, in general, compounds linked with the cyclopentene **43** were more metabolically stable and had lower clearance rates than the corresponding compounds linked with linear alkyl linkers, as is **42**. Being able to block the primary site of metabolism also translated into having more efficacious compounds *in vivo* with superior B/P ratios. Compound **42** had a B/P ratio of 0.4 after oral administration, while cyclopentene **43** had a B/P ratio of 4.1.

A series of quinuclidines was synthesized as agonists of α-7 neuronal nicotinic acetylcholine receptors.[20] Benzofuran **44** exhibited good brain penetration with a B/P ratio of 1.2. Interestingly, the benzofuran regioisomer **45** had a B/P ratio of 3.0. Because benzofurans have potential to undergo metabolic activation and generate reactive metabolites, they were examined in the reactive metabolite assay (RMA).[21] Compound **44** was negative, while **45** tested positive in the RMA. Further studies revealed that both **44** and **45** have a modest accumulation in the brain. Therefore, pyridine analog **46** was prepared to further attenuate any metabolic liability (negative in RMA). It showed no accumulation in the brain while maintaining a good B/P ratio of 1.5.

B. pK_a

A series of carbazole analogs was prepared as neuropeptide Y (NPY) antagonists.[22] Compound **47** was found to have a high affinity for NPY, but had poor permeability properties *in vivo* with a B/P ratio of 0.1. The pK_a for **47** was calculated to be 11.0 and thought to be a major cause for the poor PK profile. Efforts to decrease the pK_a led to **48**, which had a $pK_a = 9.7$ and a better B/P ratio of 0.8. Finally, by decreasing the pK_a even further to 7.9, compound **49** gave very good brain penetration and had a B/P ratio of 4.2.

C. Log P

To increase brain penetration, attention is often focused on a compound's lipophilicity. There are several reviews that discuss this, and a variety of equations has been described that can aid in optimizing the lipophilicity of a lead series.[23] As a first approximation, a log P measurement is usually employed to guide the SAR. For example, a series of 3-phenylpyrazolo[1,5-a] pyrimidines was identified as potent corticotrophin-releasing factor-1 antagonists.[24] Compound **50** is a very potent antagonist that has poor water solubility and high lipophilicity (log P = 5.9). Attempts to reduce the lipophilicity were first accomplished by changing the 2, 4-dichlorophenyl analog **50** to the 2,4-dimethoxyphenyl analog **51**, which resulted in a lowering of log P to 4.2. Further optimization revealed that changing the cyclopropylmethyl to the methoxyethyl substituent in **52** resulted in additional lowering of lipophilicity (log P = 3.2), which is more ideal for brain penetration. Analog **52** possessed moderate bioavailability (F = 31%) and good brain penetration after oral administration, with a B/P ratio of 1.2.

A series of dopamine D_3-receptor antagonists was prepared.[25] Changing from the 1*H*-pyrimidin-2-one (X = N) compound (**53**) to the 1*H*-pyridin-2-one (X = CH) analog (**54**) increased the log P value 1.1 units and decreased PSA by 13 angstroms, which resulted in greater brain penetration and a B/P ratio of 3.7.

Compounds **55** and **56** are potent dopamine D_3-receptor antagonists, and both have good oral bioavailability ($F \sim 45\%$).[26] The sulfone analog (**55**), however, shows essentially no brain penetration, while the 4-F analog (**56**) has a good B/P ratio of 1.0. The higher brain permeability of **56** is, presumably, a result of increased lipophilicity and reduction in the number of hydrogen bond acceptors.

Celecoxib (**57**) and rofecoxib (**58**) are marketed COX-2 inhibitors.[27] Celecoxib has a relatively poor B/P ratio of 0.1 compared to refecoxib, which has a B/P ratio of 0.8. Presumably, changing from the sulfonamide to the methyl sulfone removes a hydrogen bond donor and allows for better brain penetration. Keeping the methyl sulfone intact, a new pyrazolo[1,5-b]pyradazine COX-2 inhibitor (**59**) has been developed which shows an enhanced B/P ratio of 1.5. The ethoxyl and new pyrazolo[1,5-b]pyradazine moieties help lower log P and allow for better brain penetration.

59

D. Hydrogen bond interactions

As seen in the sections above, altering the basicity of a molecule can increase bioavailability. This is a common strategy used to increase the brain penetration of a compound. For example, compounds **60** and **61** are bis-aminopyrrolidine urea derived small molecule antagonists of the melanin-concentrating hormone receptor-1.[28] Although both compounds exhibit somewhat high clearance rates, the clearance of **60** is 4-fold lower than its *N*-methyl analog (**61**). Interestingly, **61** has a higher brain concentration and B/P ratio (3.4) compared to its *des*-methyl counterpart **60** (B/P = 0.6). The discrepancy is attributed to the fewer H-bond donors of the tertiary *N*-methyl analog **61** compared to that of the secondary amine analog **60**. In general, this series has shown that secondary amines typically had low brain penetration, and that conversion to a tertiary amine gave much better blood-brain barrier (BBB) penetration.

A series of propanoic acids was identified as containing potent allosteric potentiators of the metabotropic glutamate receptor-2.[29] Phenol **62** displayed good potency, but had low plasma levels and very little brain penetration (B/P < 0.03). The phenol group in **62** was alkylated to give ethyl ether **63**, which had good brain penetration and a B/P ratio of 1.2. The aqueous solubility of ether **63** is much higher than its phenol counterpart (**62**), which, in turn, leads to higher plasma levels. Interestingly, tetrazole **64** had very high plasma levels, but essentially no brain penetration (B/P < 0.03).

Compounds **65** and **66** were prepared as a series of 5-HT_6-receptor antagonists.[30] Both indoles exhibited good potency and PK profiles, but had differences in their brain-to-blood ratios. Capping the indole nitrogen with a methyl group (**66**) reduced the number of hydrogen bond donors and resulted in a brain-to-blood ratio of 2.6, compared to 0.7 for indole **65**.

Sulfamide **67** is a potent γ-secretase inhibitor that suffers from very low brain concentrations after oral dosing.[31] Changing the sulfamide ring to a sulfone ring (**68**) decreased the hydrogen bond donors and PSA, which, in turn, significantly improved its PK profile. The sulfone analog (**68**) showed a 7.5-fold increase in plasma levels, a 16-fold increase in brain concentration, and a 3-fold increase in B/P ratio, compared to sulfamide **67**.

Indoles **69** and **70** were prepared and found to exhibit potent NK_1-receptor antagonist activity.[32] To improve brain penetration, the indole NH was alkylated to give **70** which reduced polarity and removed the hydrogen bond donor.

Compound **71** is a potent 5-HT_6-receptor antagonist that suffers from poor brain penetration and has a B/P ratio of 0.1.[33] Efforts to decrease the molecular weight and hydrogen bond acceptors led to a new benzo[1,4]oxazine (**72**) that maintains *in vitro* potency and has similar structural features (i.e. piperazine and sulfonamide) to **71**. Compound **72** has an improved PK profile with enhanced brain penetration and a B/P ratio of 1.8.

IV. CONCLUSION

As can be seen from the examples above, there are a number of strategies routinely used to try to increase oral bioavailability of a compound. Often several of these are used without success before the optimal approach is found for a given chemical series. In terms of brain bioavailability, it is common to focus first on log*P* calculations and minimization of molecular weight to obtain increased brain bioavailabilities. The examples provided here will give the reader useful starting points for optimization programs.

REFERENCES

1. Lipinski, C. A., Lombardo, F., Dominy, B. W. *et a*l. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development setting. *Adv. Drug Deliver. Rev.* 1997, *23*, 3–25.
2. Pandey, A., Volkots, D. L., Seroogy, J. M. *et al*. Identification of orally active, potent, and selective 4-piperazinylquinazolines as antagonists of the platelet-derived growth factor receptor tyrosine kinase family. *J. Med. Chem.* 2002, *45*, 3772–3793.
3. Jiang, J., DeVita, R. J., Goulet, M. T. *et al*. Syntheses and structure-activity relationship studies of piperidine-substituted quinolones as nonpeptide gonadotropin releasing hormone antagonists. *Bioorg. Med. Chem. Lett.* 2004, *14*, 1795–1798.
4. Bell, I. M., Gallicchio, S. N., Abrams, M. *et al*. 3-Aminopyrrolidinone farnesyltransferase inhibitors: design of macrocyclic compounds with improved pharmacokinetics and excellent cell potency. *J. Med. Chem.* 2002, *45*, 2388–2409.
5. Zhao, H., Xin, Z., Liu, G. *et al*. Discovery of tetralin carboxamide growth hormone secretagogue receptor antagonists via scaffold manipulation. *J. Med. Chem.* 2004, *47*, 6655–6657.
6. Lange, U. E. W., Baucke, D., Hornberger, W. *et al*. Orally active thrombin inhibitors. Part 2: Optimization of the P2-moiety. *Bioorg. Med. Chem. Lett.* 2006, *16*, 2648–2653.
7. Chan, L., Stefanac, T., Turcotte, N. *et al*. Design and evaluation of dihydroisoquinolines as potent and orally bioavailable human cytomegalovirus inhibitors. *Bioorg. Med. Chem. Lett.* 2000, *10*, 1477–1480.
8. Jia, Z. J., Wu, Y., Huang, W. *et al*. Design, synthesis and biological activity of novel non-amidine factor Xa inhibitors. Part 1: P_1 structure-activity relationships of the substituted 1-(2-naphthyl)-1*H*-pyrazole-5-carboxylamides. *Bioorg. Med. Chem. Lett.* 2002, *12*, 1651–1655.
9. DeVita, R. J., Bochis, R., Frontier, A. J. *et al*. A potent, orally bioavailable benzazepinone growth hormone secretagogue. *J. Med. Chem.* 1998, *41*, 1716–1728.
10. Stearns, R. A., Miller, R. R., Tang, W. *et al*. The pharmacokinetics of a thiazole benzenesulfonamide β_3-adrenergic receptor agonist and it analogs in rats, dogs, and monkeys: improving oral bioavailability. *Drug Metab. Dispos.* 2002, *30*, 771–777.
11. Zheng, G. Z., Lee, C. H., Pratt, J. K. *et al*. Pyridopyrimidine analogues as novel adenosine kinase inhibitors. *Bioorg. Med. Chem. Lett.* 2001, *11*, 2071–2074.
12. Wada, C. K., Holms, J. H., Curtin, M. L. *et al*. Phenoxyphenyl sulfone *N*-formylhydroxylamines (retrohydroxamates) as potent, selective, orally bioavailable matrix metalloproteinase inhibitors. *J. Med. Chem.* 2002, *45*, 219–232.
13. Anderson, J. T., Ting, A. E., Boozer, S. *et al*. Identification of novel and improved antimitotic agents derived from noscapine. *J. Med. Chem.* 2005, *48*, 7096–7098.
14. Chao, J., Taveras, A. G., Chao, J. *et al*. C(4)-alkyl substituted furanyl cyclobutenediones as potent, orally bioavailable CXCR2 and CXCR1 receptor antagonists. *Bioorg. Med. Chem. Lett.* 2007, *17*, 3778–3783.
15. Hansen, D. W., Jr., Peterson, K. B., Trivedi, M. *et al*. 2-Iminohomopiperidinium salts as selective inhibitors of inducible nitric oxide synthase (iNOS). *J. Med. Chem.* 1998, *41*, 1361–1366.
16. Hansen, T. K., Ankersen, M., Hansen, B. S. *et al*. Novel orally active growth hormone secretagogoues. *J. Med. Chem.* 1998, *41*, 3705–3714.
17. Arban, R., Benedetti, R., Bonanomi, G. *et al*. Cyclopenta[*d*]pyrimidines and dihydropyrrolo[2,3-*d*]pyrimidines as potent and selective corticotrophin-releasing factor-1 receptor antagonists. *ChemMedChem.* 2007, *2*, 528–540.
18. Vasudevan, A., Souers, A. J., Freeman, J. C. *et al*. Aminopiperidine indazoles as orally efficacious melanin concentrating hormone receptor-1 antagonists. *Bioorg. Med. Chem. Lett.* 2005, *15*, 5293–5297.
19. Hattori, K., Kido, Y., Yamamoto, H. *et al*. Rational design of conformationally restricted quinazolinone inhibitors of poly(ADP-ribose)polymerase. *Bioorg. Med. Chem. Lett.* 2007, *17*, 5577–5581.
20. Wishka, D. G., Walker, D. P., Yates, K. M. *et al*. Discovery of *N*-[(3*R*)-1-azabicyclo [2.2.2]oct-3-yl]furo[2,3-*c*]pyridine-5-carboxamide, an agonist of the α7 nicotinic acetylcholine receptor, for the potential treatment of cognitive deficits in schizophrenia: synthesis and structure-activity relationship. *J. Med. Chem.* 2006, *49*, 4425–4436.
21. Soglia, J. R., Harriman, S. P., Zhao, S. *et al*. The development of a higher throughput reactive intermediate screening assay incorporating microbore liquid chromatography electrospray ionization tandem mass spectrometry and glutathione ethyl ester as an in vitro conjugating agent. *J. Pharm. Biomed. Anal.* 2004, *36*, 105–116.
22. Leslie, C. P., Di Fabio, R., Bonetti, F. *et al*. Novel carbazole derivatives as NPY Y1 antagonists. *Bioorg. Med. Chem. Lett.* 2007, *17*, 1043–1046.
23. (a) Su, Y., Sinko, P. J. Drug delivery across the blood–brain barrier: why is it difficult? How to measure and improve it? *Expert Opin. Drug Delivery*, 2006, *3*, 419–435. (b) Liu, X., Chen, C. Strategies to optimize brain penetration in drug discovery. *Curr. Opinion in Drug Discovery and Delivery*, 2005, *8*, 505–517.
24. Chen, C., Wilcoxen, K. M., Huang, C. Q. *et al*. Optimization of 3-phenylpyrazolo[1,5-*a*]pyrimidines as potent corticotrophin-releasing factor-1 antagonists with adequate lipophilicity and water solubility. *Bioorg. Med. Chem. Lett.* 2004, *14*, 3669–3673.
25. Geneste, H., Amberg, W., Backfisch, G. *et al*. Synthesis and SAR of highly potent and selective dopamine D_3-receptor antagonists: variations on the 1*H*-pyrimidin-2-one theme. *Bioorg. Med. Chem. Lett.* 2006, *16*, 1934–1937.
26. Stemp, G., Ashmeade, T., Branch, C. L. *et al*. Design and synthesis of *trans*-*N*-[4-[2-(6-cyano-1,2,3,4-tetrahydroisoquinolin-2-yl)ethyl]

cyclohexyl]-4-quinolinecarboxamide (SB-277011): a potent and selective dopamine D_3 receptor antagonist with high oral bioavailability and CNS penetration in the rat. *J. Med. Chem.* 2000, *43*, 1878–1885.

27. Beswick, P., Bingham, S., Bountra, C. *et al*. Identification of 2,3-diarylpyrazolo[1,5-*b*]pyridazines as potent and selective cyclooxygenase-2 inhibitors. *Bioorg. Med. Chem. Lett.* 2004, *14*, 5445–5448.
28. Hudson, S., Kiankarimi, M., Rowbottom, M. W. *et al*. Synthesis and structure-activity relationships of retro bis-aminopyrrolidine urea (rAPU) derived small-molecule antagonists of the melanin-concentrating hormone receptor-1 (MCH-R1). Part 2. *Bioorg. Med. Chem. Lett.* 2006, *16*, 4922–4930.
29. Cube, R. V., Vernier, J. M., Hutchinson, J. H. *et al*. 3-(2-Ethoxy-4-{4-[3-hydroxy-2-methyl-4-(3-methylbutanoyl)-phenoxy]butoxy}phenyl)propanoic acid: a brain penetrant allosteric potentiator at the metabotropic glutamate receptor 2 (mGluR2). *Bioorg. Med. Chem. Lett.* 2005, *15*, 2389–2393.
30. Ahmed, M., Briggs, M. A., Bromidge, S. M. *et al*. Bicyclic heteroarylpiperazines as selective brain penetrant 5-HT_6 receptor antagonists. *Bioorg. Med. Chem. Lett.* 2005, *15*, 4867–4871.
31. Shaw, D., Best, J., Dinnell, K. *et al*. 3,4-Fused cyclohexyl sulfones as γ-secretase inhibitors. *Bioorg. Med. Chem. Lett.* 2006, *16*, 3073–3077.
32. Dinnell, K., Chicchi, G. G., Dhar, M. J. *et al*. 2-Aryl indole NK_1 receptor antagonists: optimisation of the 2-aryl ring and the indole nitrogen substituent. *Bioorg. Med. Chem. Lett.* 2001, *11*, 1237–1240.
33. Zhao, S. H., Berger, J., Clark, R. D. *et al*. 3,4-Dihydro-2*H*-benzo[1,4]oxazine derivatives as 5-HT_6 receptor antagonists. *Bioorg. Med. Chem. Lett.* 2007, *17*, 3504–3507.

Chapter 36

Designing Prodrugs and Bioprecursors

Camille G. Wermuth

I. INTRODUCTION
II. THE DIFFERENT KINDS OF PRODRUGS
A. Definitions and classifications
B. The carrier prodrug principle
C. The bioprecursor-prodrug principle
D. Other categories of prodrugs
E. Practical applications of prodrug design
III. CARRIER PRODRUGS: APPLICATION EXAMPLES
A. Improvement of the bioavailability and the biomembrane passage
B. Site-specific delivery
C. Prolonged duration of action
IV. PARTICULAR ASPECTS OF CARRIER PRODRUG DESIGN
A. Use of cascade prodrugs
B. Codrugs
C. Soft drugs
D. Carrier prodrugs: conclusion
V. BIOPRECURSOR PRODRUGS: APPLICATION EXAMPLES
A. Oxidative bioactivations
B. Reductive bioactivations
C. Mixed bioactivation mechanism
D. Reactions without change in the state of oxidation
VI. DISCUSSION
A. Bioprecursors versus carrier prodrugs
B. Existence of mixed-type prodrugs
VII. DIFFICULTIES AND LIMITATIONS
VIII. CONCLUSION
REFERENCES

"La façon de donner vaut mieux que ce que l'on donne."
The manner of giving counts more than what one gives.
Pierre CORNEILLE, Le Menteur, Acte 1, Scene 1.

I. INTRODUCTION

The bioavailability of molecules exclusively screened through *in vitro* assays can be low. Because of the polarity of the functional groups present in the molecule, they may be poorly absorbed or incorrectly distributed. They may also, as a result of their vulnerability, be the subject of early metabolic destructions, such as first-pass effects or any other kind of degradation leading to a short biological half-life. For such molecules, *in vivo* administration is limited to the parenteral route, and their clinical usefulness is thus restricted. Sometimes an adequate pharmaceutical formulation (micro-encapsulation, sustained-release or enterosoluble preparations) can overcome these drawbacks, but often the galenic formulation is inoperant, and a chemical modification of the active molecule is necessary to correct its pharmacokinetic insufficiencies. This *chemical formulation* process, whose objective is to convert an interesting active molecule into a clinically acceptable drug, often involves a so-called prodrug design.[1–5]

II. THE DIFFERENT KINDS OF PRODRUGS

A. Definitions and classifications

Initially, the term prodrug was introduced by Albert to describe "any compound that undergoes biotransformation prior to exhibiting its pharmacological effects."[6] Such a broad definition includes accidental historic prodrugs (salicin

Copyright © 2008, Elsevier Ltd
All rights reserved.

FIGURE 36.1 The carrier prodrug principle.[11]

and salicylic acid), active metabolites (imipramine and desmethylimipramine) and compounds intentionally prepared to improve the pharmacokinetic profile of an active molecule. From this point of view, the term "drug latentiation" proposed by Harper[7] is more appropriate for prodrug design, as it indicates that there is an intention. Drug latentiation is defined as "the chemical modification of a biologically active compound to form a new compound that, upon *in vivo* enzymatic attack, will liberate the parent compound." Even this definition is too broad, and an attentive survey of the specialized literature led us to divide prodrugs into two main classes: carrier prodrugs and bioprecursors.[3,8,9]

C*arrier prodrugs* result from a temporary linkage of the active molecule with a transport moiety that is frequently of lipophilic nature. A simple hydrolytic reaction cleaves this transport moiety at the correct moment (e.g. pivampicillin, bacampicillin,). Such prodrugs are, *per se,* less active than the parent compounds or even inactive. The transport moiety (carrier group) is chosen for its non-toxicity and its ability to ensure the release of the active principle with efficient kinetics.

Bioprecursors do not imply a temporary linkage between the active principle and a carrier group, but result from a molecular modification of the active principle itself. This modification generates a new compound, which is a substrate for metabolizing enzymes, leading to a metabolite that is the expected active principle. This approach exemplifies the active metabolite concept in a provisional way (e.g. sulindac, fenbufen, acyclovir, losartan).

B. The carrier prodrug principle

The carrier prodrug principle (Figure 36.1) consists of "the attachment of a carrier group to the active drug to alter its physicochemical properties and then the subsequent enzyme attack to release the active drug moiety.[7]" "Prodrugs can thus be viewed as drugs containing specialized non-toxic protective groups used in a transient manner to alter or eliminate undesirable properties in the parent molecule.[2]"

A well-designed carrier prodrug satisfies the following criteria:[1,10]

1. The linkage between the drug substance and the transport moiety is usually a covalent bond.
2. As a rule the prodrug is inactive or less active than the parent compound.
3. The linkage between the parent compound and the transport moiety must be broken *in vivo*.
4. The prodrug, as well as the *in vivo* released transport moiety, must be non-toxic.
5. The generation of the active form must take place with rapid kinetics to ensure effective drug levels at the site of action and to minimize either direct prodrug metabolism or gradual drug inactivation.

An example of prodrug design taking into account these criteria is found in orally active ampicillin derivatives.[12–14] Ampicillin is one of the main β-lactam antibiotics. It is widely used as a broad spectrum antibiotic, but it suffers from poor absorption when administered orally: only about 40% of the drug is absorbed. In other words, to achieve the same clinical efficiency and the same blood levels, one must give two to 3 times more ampicillin by mouth than by intramuscular injection. The clinical tolerance of orally given ampicillin may be affected, with the non-absorbed part of the drug destroying the intestinal flora. Therefore, numerous attempts have been made to improve these poor absorption properties.

Figure 36.2 represents two prodrugs of ampicillin: pivampicillin and bacampicillin. They both result from the esterification of the polar carboxylic group with a lipophilic, enzymatically labile ester. The main properties of these prodrugs can be summarized as follows:

1. The absorption of these compounds is nearly quantitative (98–99%).
2. The generation of free ampicillin in the blood stream is rapid (less than 15 minutes).
3. The released carrier molecules are formaldehyde and pivalic acid (trimethyl acetic acid) for pivampicillin and acetaldehyde, ethanol and carbon dioxide in the case of bacampicillin. These three latter compounds are natural metabolites in the human body. This may explain the better tolerance of bacampicillin compared to pivampicillin.
4. The serum levels attained following oral administration of bacampicillin are similar to those obtained after intramuscular injection of an equimolecular amount of free ampicillin.
5. The clinical trials confirm the efficiency and the safety of the prodrugs. Due to their good absorption, the drugs are given at lower dosage than ampicillin: 0.8–1.0 g daily is sufficient in common infections as compared to 2.0 g daily for ampicillin.

FIGURE 36.2 Prodrugs derived from ampicillin.[12–14]

6. It has been shown, and this seems to be a rule for prodrugs, that pivampicillin and bacampicillin are inactive *per se*, the antibiotic potency appearing only in vivo after the release of free ampicillin.

C. The bioprecursor prodrug principle

As already mentioned above, *bioprecursor-prodrugs* result from a molecular modification of the active principle, generating a new compound able to be a substrate for the metabolizing enzymes, with the metabolite being the expected active principle. The *bioprecursor-prodrug* approach exemplifies the active metabolite concept. A survey of a great number of examples of active metabolites shows that they belong exclusively to the Phase I products and result from one of the reactions mentioned in Table 36.1. As such, reactions follow some general rules, and they can often be predicted. Taking into account the common metabolic pathways, one can imagine the design of a given molecule so that it will be converted *in vivo* into the desired compound by one or more of the Phase I reactions. In other words, the active metabolite concept can be used in a forward-looking way ("Metabolic Synthesis"). By analogy to the retrosynthetic reasoning usual in organic chemistry, we can imagine retrometabolic reasoning in prodrug design. Such reasoning can lead to a particular group of prodrugs for which we proposed the name *bioprecursors* or *bioprecursor prodrugs*.[8,9]

A typical example of an effective bioprecursor prodrug is given by the anti-inflammatory drug sulindac. Sulindac, *cis*-5-fluoro-2-methyl-1-[*p*-(methylsulfinyl) benzylidene] indene-3 acetic acid[16] is a non-steroidal anti-inflammatory agent having a broad spectrum of activity in animal models and in man. The two quantitatively significant biotransformations undergone by sulindac in laboratory species[17] and in man[18,19] involve only changes in the oxidation state of the sulfinyl substituent, viz., irreversible oxidation of the parent (sulindac) to sulfone and reversible reduction to sulfide (Figure 36.3), the latter being the active species.[20] In two *in vitro* models of inflammation, prostaglandin synthase inhibition and inhibition of platelet aggregation, the sulfide has activities comparable to those of indomethacin, whereas sulindac itself is devoid of activity.

Nevertheless, sulindac is the preferred compound for clinical applications: an oral dosage of this inactive bioprecursor will circumvent initial exposure of gastric and intestinal mucosa to the active drug and might thus provide a therapeutic advantage in comparison with the sulfide dosing.

TABLE 36.1 Phase I Reactions[15]

Oxidative reactions
Oxidation of alcohol, carbonyl and acid functions, hydroxylation of aliphatic carbon atoms, hydroxylation of alicyclic carbon atoms, oxidation of aromatic carbon atoms, oxidation of carbon–carbon double bonds, oxidation of nitrogen-containing functional groups, oxidation of silicon, phosphorus, arsenic and sulfur, oxidative N-dealkylation, oxidative O- and S-dealkylation, oxidative deamination, other oxidative reactions
Reductive reactions
Reduction of carbonyl groups, reduction of alcoholic groups and carbon–carbon double bonds, reduction of nitrogen-containing functional groups, other reductive reactions
Reactions without change in the state of oxidation
Hydrolysis of esters and ethers, hydrolytic cleavage of C—N single bonds, hydrolytic cleavage of nonaromatic heterocycles, hydration and dehydration at multiple bonds, new atomic linkages resulting from dehydration reactions, hydrolytic dehalogenation: removal of hydrogen halide molecules, various reactions.

D. Other categories of prodrugs

A certain number of satellite technologies of prodrug design were proposed by various authors. Among them are the design of soft drugs, codrugs, pro-prodrugs, cascade prodrugs, site-specific chemical delivery systems, macromolecular prodrugs and drug–antibody conjugates. These often very sophisticated approaches will be discussed in Section V.

E. Practical applications of prodrug design

The domain of application of the prodrug approach is illustrated in Figure 36.4. In practice, prodrugs achieve usually one of the five following goals: increased lipophilicity, increased duration of pharmacological effects, increased site-specificity, decreased toxicity and adverse reactions, and improvement in drug formulation (stability, water solubility, suppression of an undesirable organoleptic or physicochemical property). For applications in the field of insecticides see Drabek and Neumann.[21]

FIGURE 36.3 Reductive bioactivation of sulindac.[20]

III. CARRIER PRODRUGS: APPLICATION EXAMPLES

Bioactive compounds and drugs usually bear a limited number of polar functional groups suitable for carrier prodrug synthesis. Among these, the most frequent are the alcoholic and the phenolic hydroxyls, the amino group, and the carboxylic function. The aim of the following paragraphs is to illustrate how such groups can be used to prepare carrier prodrugs with improved pharmacokinetic properties.

A. Improvement of the bioavailability and the biomembrane passage

The biomembrane passage of a drug depends primarily on its physicochemical properties and especially on its partition coefficient (Chapters 22 and 34). Thus, the transient attachment of a lipophilic carrier group to an active principle can provide a better bioavailability, mostly by facilitating cell membrane crossing by passive diffusion. Peroral absorption, as well as rectal absorption, ocular drug delivery and dermal drug delivery, are dependent on passive diffusion. Finally, lipophilic carriers can sometimes be useful to reduce first-pass metabolism.[23]

1. *Derivatization of drugs containing alcoholic or phenolic hydroxy groups*

Starting from hydroxylic derivatives, high lipophilicity can simply be obtained by esterification with lipophilic carboxylic acids. *Dipivaloyl-epinephrine,* for example (Figure 36.5), crosses the cornea and is used in the treatment of

FIGURE 36.4 Shortcomings that may be overcome through chemical formulation.[1,11,22]

glaucoma.[24] The β-blocker timolol contains a secondary amino group with a pK_a of 9.2. Since this group is highly protonated at pH 7.4, the compound shows a low lipophilicity at physiological pH ($\log P = -0.04$), which in turn is unfavorable for corneal penetration. The corresponding butyryl ester has an increased lipophilicity ($\log P = 2.08$) and causes a four- to sixfold increase in the corneal absorption of timolol following topical administration to rabbits.[23]

In a similar manner, *dibenzoyl-2-amino-6,7-dihydroxytetrahydronaphthalene* (DB-ADTN) reaches the central nervous system (CNS), whereas the parent dopamine agonist ADTN does not.[25,26] For dipivaloyl-epinephrine and dibenzoyl-ADTN, the selective acylation of the phenolic hydroxyl groups was achieved in a strong acidic medium, the amino function being protected by protonation.[25,27] Acylated thymidine analogs, such as *3′-O-hexyl-5′-amino-2′-deoxythymidine*, are prodrugs for topical application against herpes simplex type 1 (HSV-1) viruses.[28] Diacetyl and dipropionyl guanine derivatives, given orally to mice, provided concentrations of the parent drug that were more than 15-fold higher than those observed after dosing the non-acylated parent drug.[29] In augmenting the lipophilicity and simultaneously destroying the crystal lattice energy, the 2′,3′-diacetate of the antiviral agent 6-methoxypurine arabinoside allowed a 5-fold increase in bioavailability and a 3-fold increase in water solubility in comparison to the non-acetylated drug.[30] As a consequence, an intravenous formulation could be developed.

Esterification of hydroxylic functions with polyacids (e.g. succinic acid, phosphoric acid) represents an excellent way to prepare water-soluble prodrugs (see Chapter 38)

2. Derivatization of aldehydes and ketones

The ethylene ketal derivative of prostaglandin E_2 (dinoprostone) possesses much improved solid-state stability (see Chapter 41). Functionalized spirothiazolidines of hydrocortisone and hydrocortisone 21-acetate (Figure 36.6), prepared with cysteine esters or related β-aminothiols, have improved topical anti-inflammatory activity. It is speculated that the Schiff base intermediate formed upon ring-opening may accumulate in the skin by binding (through its SH function) to thiol groups in the skin.[31]

Simple and substituted oximes are biostable unless intramolecular assistance is provided. This is the case for the oximes derived from oxyamino acetic acid that are possible water-soluble prodrugs of ketones and aldehydes (See Chapter 38).

3. Derivatization of drugs containing a carboxylic acid functional group

Lipophilic prodrugs can also be derived from a carboxylic functional group, the most commonly used derivatives being *carboxylic esters*. Simple esters of aliphatic alcohols are attractive, as they are cheap to prepare, chemically stable, and yield harmless hydrolysis products.[32] Typical representatives of such prodrugs are tyrosine methyl ester,[33] levodopa ethyl ester,[34] nipecotic acid ethyl ester,[35] enalaprilat ethyl ester,[36,37] trandolapril,[38] γ-aminobutyric acid (GABA) cetyl ester,[39,40] and methotrexate cetyl ester.[41]

Lipoidal prodrugs, in which the carboxyl function esterifies the free alcoholic hydroxyl of 1,2- or 1,3-diglyceride, are well absorbed and show high lymphotropism.[42,43] Due

FIGURE 36.6 Prodrug possibilities starting from aldehydes or ketones.

FIGURE 36.5 Lipophilic prodrugs of hydroxy compounds with facilitated membrane penetration.[23–26]

FIGURE 36.7 Use of anhydrides as lipophilic prodrugs of phosphonates.

FIGURE 36.8 Methoxy-imino bioisosteres of carboxylic acid anhydrides as prodrugs for carboxylic acids.

to their greasy properties and to a difficult purification, these compounds have no industrial application.

The widespread use of *acyloxymethyl esters* in antibiotic chemistry, as illustrated above for bacampicillin, was initiated by Jansen and Russel[44] at Wyeth Laboratories and successfully applied to pivampicillin,[12] talampicillin[14] and cephalosporins.[45] In each of these cases, the oral absorption of the antibiotic was improved by some two- to threefold over that of the parent compound. The acyloxymethyl derivatization was also extended to amino acids such as α-methyldopa,[46] isoguvacine[47] or tranexamic acid,[48] anti-inflammatory drugs such as niflumic acid[49] or indomethacin,[50] and quinolone antibacterials such as norfloxacin.[51]

Mixed anhydrides represent original attempts for preparing prodrugs of carboxylic or phosphonic acid. *Clodronic acid dianhydrides* (Figure 36.7), for example, were shown to be novel bioreversible prodrugs of clodronate. They are more lipophilic than the parent clodronate, stable against chemical hydrolysis, and hydrolyzed enzymatically to clodronate in human serum.[52]

Methoxy-imino bioisosteres of carboxylic acid anhydrides, which can be prepared by *O*-acylation of the corresponding hydroxamates, represent another way to prepare prodrugs of carboxylic acids (Figure 36.8). One of these derivatives, compound FOX 988, was designed as a prodrug of 4-(2,2-dimethyl-1-hydroxypropyl) benzoic acid. In the liver this prodrug is metabolized at a rate sufficient enough to possess hypoglycemic potency (an ED_{50} of 65 μmol/kg, 28 mg/kg/day, for glucose lowering).[53] Yet by avoiding significant escape of the released active metabolite to the systemic circulation, it avoids the testicular toxicity at doses up to 1,500 μmol/kg/day. The mechanism of action of 4-(2,2-dimethyl-1-hydroxypropyl) benzoic acid consists in the sequestration of coenzyme A (CoA) in the mitochondria, thus inhibiting medium-chain acyltransferase and, as a consequence causing hepatic neoglucogenesis.[54]

Primary amides of carboxylic acids are easily converted in humans to the corresponding acid (e.g. depamide, progabide) and can thus be used in prodrug design. Amides of ketoprofen-derived arylacetic acids possess a therapeutic index one order of magnitude greater than that of indomethacin.[55]

4. Derivatization of amines

Due to the slow *in vivo* cleavage rate of the *N*-substituted amides, acylation of amines is generally not recommended. Better possibilities are offered by activated amides, peptides, imines and soft quaternary ammonium salts. However, the use of simple *N*-acyl derivatives must not systematically be discarded. The *N*-benzoyl- or *N*-pivaloyl derivatives of the inhibitory neurotransmitter GABA are examples of compounds able to penetrate the blood-barrier and to abolish pentetrazole- and bicuculline-induced convulsions. It was also demonstrated in rats that, following subcutaneous injection, rat-brain homogenates liberate free GABA from these amides.[56] There is even some biochemical and pharmacological evidence suggesting that *N*-pivaloyltaurine crosses the blood-brain barrier.[57]

Imines[58] and enamines,[59,60] stabilized through hydrogen bonds, can also be effective prodrugs of primary amines (Figure 36.9).

Small peptides constitute an alternative way of derivatizing amines. The hypotensive drug milodrine, for example (Figure 36.9), is the well-absorbed prodrug from which 2-(2,5-dimethoxyphenyl)-2-hydroxyethylamine (ST-1059) is liberated by enzymic cleavage of the glycine residue.[61] Orally given to fasted Wistar rats, the *N*-(*Z*-alanyl) amide of the hypoglycemic sulfonylurea carbutamide demonstrated a 4–6 times higher potency than the parent sulfonamide. The compound is well tolerated and is metabolized to the parent drug; the amino acid moiety just modifies the bioavailability.[62] The L-serine amide of the tubulin polymerization inhibitor analog AC-7739 is a potent combretastatin A4 analog with clearly improved bioavailability and reduced toxicity.[63]

Among the numerous variations made to α-methyldopa, acylation with a glycyl–glycyl residue was claimed

FIGURE 36.9 Imine, enamine and peptide prodrugs derived from amino functions.

FIGURE 36.10 Carbamates as amidine prodrugs.

to improve the oral bioavailability.[64] For a series of anticandidal di- and tripeptides containing *m*-fluorophenylalanine (*m*-FPhe), competitive antagonism studies supported peptide transport-mediated entry of the warhead *m*-FPhe inside the cell.[65] Dipeptides derived from α-methyldopa (Figure 36.9) show a ten- to twentyfold better penetration of the intestinal wall than α-methyldopa itself.[66]

5. *Derivatization of amidines*

Improved oral activity for amidines resulted from the formation of carbamates of 2,5-bis(4-amidinophenyl)furan, an effective drug against *Pneumocystis carinii* pneumonia (Figure 36.10).[67]

6. *Prodrugs for compounds with acidic* NH *functions*

Prodrugs obtained by *N*-alkoxycarbonyloxymethylation of 5-fluorouracil show improved delivery properties. Both 1- and 3-alkoxycarbonyloxymethyl derivatives are hydrolyzed quantitatively to 5-fluorouracil, but the 3-substituted derivatives show a greater promise as prodrugs since they combine adequate stability in aqueous solution with a high susceptibility

FIGURE 36.11 Prodrugs of acidic NH functions.

to undergo hydrolysis in plasma.[68] Sulfonamides, as well as carboxamides, carbamates and other NH-acidic compounds (Figure 36.11) can be acylated with various groups[69] or converted into phthalidyl derivatives.[70]

FIGURE 36.12 Cyclic protections of two neighboring functional groups.

TABLE 36.2 Effect of the Acyl Group in Isopropyl Hydrazide on Selective Transport[74]

Isopropyl hydrazide	Monoamine increase (%)		Ratios
	Cardiac catecholamines	Cerebral 5-HTP	
	100	100	1.0
	75	250	3.3
	145	60	0.4

Hetacillin[71,72] and droxicam[73] are examples of simultaneous cyclic protections of an acidic NH function and an amino or a hydroxylic groups located in the vicinity (Figure 36.12).

B. Site-specific delivery

Many hopes were put in the prodrug approach as a means to achieve the targeting of drugs for a specific site in the body. Actually, only a few convincing examples are found in the literature, and we are becoming somewhat disillusioned about the real possibilities of the approach. In principle, two targeting possibilities can be considered:[23] First, one can design a prodrug that affords an increased or selective transport of the parent drug to the site of action (*site-directed* drug delivery). Second, one can design a derivative that goes everywhere, but undergoes bioactivation only inside the target organ (*site-specific* drug release).

1. *Site-directed drug delivery*

Most of the successes in achieving site-directed drug delivery through prodrugs have been through *localized delivery* of lipophilic prodrugs (eyes, skin) with increased permeability characteristics. Systemic site-directed delivery, that is, to a specific internal site or organ through a selective transport, is very difficult to achieve. Nevertheless, some possibilities of local enrichment or of privileged entry to a given organ or into the central nervous system are found in the literature. Thus, the L-glutamic analog of iproniazide presents a preferential monoamine oxidase inhibition in the brain,[74] whereas the palmitoyl isopropylhydrazide demonstrates a clear cardiac selectivity (Table 36.2).

The propensity of fatty chains to concentrate in cardiac tissue is also illustrated by the findings made in the field of myocardial imaging agents. An iodine and tellurium containing fatty acid (Figure 36.13) has a high heart uptake,

FIGURE 36.13 Myocardial imaging agent.

FIGURE 36.14 Bile acids for liver specific targeting.[76]

and heart/blood ratios remained high for several hours: 13:1 after 1 h and 9:1 after 4 h.[75]

Coupling of drugs to modified bile acids was recently proposed for liver specific targeting.[76] The rationale is based on the recognition of bile acid-linked drugs by the endogenous bile acid transport system. Chlorambucil, an alkylating cytostatic agent, HR-780, an inhibitor of hydroxymethylglutaryl-CoA reductase (HMG-CoA) reductase, and an oxaproline peptide, an inhibitor of prolyl-4-hydroxylase, were chosen for conjugation to bile acids (Figure 36.14).

Cytotoxic radicals derived from daunosamine, linked through a glutaric acid spacer to the amino terminal D-Phe residue of somatostatin, are claimed to selectively target prostatic, renal, ovarian or pancreatic cancers.[77]

The 2,3-dichlorophenoxyacetic moiety of ethacrynic acid was claimed to have a high affinity for renal tissue,[78] and 2-thiouracil and 6-propylthiouracil exhibit marked affinities for melatonin-producing tissues.[79,80] They were, therefore, (unsuccessfully) tested for the treatment of malignant melanoma.[81] Many other examples of selective conducting moieties are found in X-ray contrast media and radioisotope imaging agents.[82] Similar efforts were made to find selective cancer chemotherapeutics, and a variety of drugs that are known to accumulate selectively in particular tissues have also been tried. Mustard derivatives of amino acids, steroid hormones, tetracyclines, quinacrine and uracil are examples.[83] Site-directed cancer chemotherapy can involve drugs bound to specific antibodies. Daunomycin, conjugated via an oxidized dextran bridge with anti-B-cell lymphoma 38C-13 cell-surface IgM antibodies, given to 38C-tumor bearing mice, gave increased life span and even complete cures.[84] Some other attempts to achieve delivery to the CNS will be found in Chapter 35.

FIGURE 36.15 Selective renal vasodilatation with γ-glutamyldopa.[86]

2. *Site-specific drug release*

The whole strategy of site-specific release of a given drug lies in the discovery of an enzyme present in high concentrations in the target organ and quasi-absent elsewhere. An appropriate prodrug can then be designed using the selective cleavage possibility offered by the enzyme.

A selective renal vasodilation, for example, is produced by administration of γ-glutamyldopa. It is well known that L-dopa is a precursor of the neurotransmitter dopamine, which plays an important role in the CNS and in the kidneys. The association of L-dopa with a peripheral dopa decarboxylase inhibitor results in preferential dopamine production in the brain and can be considered, at present, the best therapeutic possibility for Parkinson's disease.

On the renal side, a prodrug of L-dopa, γ-glutamyl-L-3,4-dihydroxy-phenylalanine (γ-glutamyldopa), produces a specific vasodilation of the renal tissue. Indeed, the γ-glutamyl derivatives of amino acids and peptides accumulate in the kidneys, where they undergo a selective metabolic process (for a review see Magnan *et al.*[85]). The successive actions of two enzymes present in high concentration in the kidney, γ-glutamyl transpeptidase and L-aromatic amino acid decarboxylase, release dopamine locally from γ-glutamyldopa (Figure 36.15).

In mice, renal levels of dopamine, after γ-glutamyldopa, are 5 times higher than after an equimolar administration of L-dopa. A perfusion of 10 μM/g/30 min of γ-glutamyldopa in rats produces a 60% increase in the renal plasmatic flux.[86] The same dose of L-dopa induces no vasodilation. Massive administration of γ-glutamyldopa (20 times the preceding dose) produces only a weak pressor effect, demonstrating that the systemic effects of the prodrug are low. The same principle was used for the synthesis of γ-glutamyl derivatives of dopamine itself and diacyldopamines.[87,88]

Similarly, it is possible to obtain a kidney-selective accumulation of sulfamethoxazole by administering the drug in the form of *N*-acetyl-γ-glutamate.[89] The regeneration of the free

FIGURE 36.16 Kidney-selective release of sulfamethoxazole.[89]

sulfamide requires the initial deacylation of the glutamic moiety, thanks to an *N*-acylamino acid deacylase, which is also present in the kidneys in high concentrations (Figure 36.16).

The γ-glutamyl strategy for confining drug action to the kidney and the urinary tract requires that the prodrug under consideration can function as a substrate for γ-glutamyl transpeptidase and, eventually, for *N*-acylamino acid deacylase.

The unique glucosidase activity of the colonic microflora has been utilized to deliver selectively steroid prodrugs useful in treating inflammatory bowel disease.[90] Dexamethasone 21-β-D-glucoside appeared to be a good candidate, as nearly 60% of an oral dose of the prodrug reached the cecum in the form of free steroid. Given orally, the parent dexamethasone is absorbed almost exclusively from the small intestine, and less than 1% reached the cecum.[91] In various tumor tissues, the activity of the enzyme uridine phosphorylase is markedly higher than in the surrounding normal tissues. This observation prompted the synthesis of 5-fluorouracil prodrugs. Among them, 5′-deoxy-5-fluorouracil shows high antitumor activity and less host toxicity compared to fluorouracil. This favorable therapeutic index is attributed to a preferential bioactivation by uridine phosphorylase in tumor cells.[92,93]

Paclitaxel 2′-carbamates and 2′-carbonates were prepared to fulfill the requirements of an ideal prodrug, that is, low cytotoxicity for healthy tissue, high stability against unspecific enzymes, and fast hydrolysis by the tumor-associated enzyme. These water-soluble prodrugs are specifically activated by plasmin in the tumor cells and show, on average, a decrease in cytotoxicity of more than 8,000-fold in comparison with the parent drug.[94]

C. Prolonged duration of action

Unless orally administrated drugs are accumulated in fatty tissues, they are not expected to act much longer than their transit period in the gastrointestinal tract (12–48 h). For drugs rapidly cleared from the body, the duration of activity is even shorter, and a frequent dosing within a 24 h period is required to maintain adequate plasma concentrations. This frequent dosing of short half-life drugs results in sharp peak-valley plasma concentration–time profiles, and, consequently, patient compliance is often poor.[95] However, for therapeutic, epidemiological, sociological or political reasons, durations of action prolonged over weeks or months are desired. The most easy administration route is then represented by intramuscular injection of depot preparations. The most successful applications are found in the domain of hormonal steroids, of antipsychotic drugs and, to a lesser extent, of antibiotics. The general strategy consists of preparing lipophilic prodrugs, dissolved or suspended in oily vehicles, and administering them by deep intramuscular injection.

1. *Contraceptive steroids*

Progestagens such as norethysterone enanthate and medroxyprogesterone acetate (MPA) have long durations of activity (3 months), due primarily to slow release from the injection site and storage in the fatty tissues.[96] Progestagen–estrogen combinations such as dihydroxyprogesterone acetophenide/estradiol enanthate or MPA/estradiol cypionate (= cyclopentylpropionate) are administered on a monthly basis.

2. *Treatment of menopause*

The symptomatic treatment of menopause (sweating, hot flushes, depression) has been successfully accomplished by the use of 200 mg of dehydroepiandrosterone-3-heptanoate and 4 mg of estradiol-17-β-valerate in a suspension of a castor oil/benzyl benzoate vehicle.[97]

In the estradiol prodrug estradiol 3-benzoate 17-β-cyclooctenyl ether (EBCO), the phenolic hydroxyl group is

FIGURE 36.17 Enol ethers as long-lasting steroid prodrugs.[98, 99]

FIGURE 36.18 Cascade latentiation. (a): 2-Acyloxymethylbenzoic acids provide amides the lability of esters.[100,101] (b): Substituted vinyl esters as lipophilic cascade carrier for carboxylic acid containing drugs.[46, 102]

masked as a benzoyl ester and the alcoholic 17-β-hydroxyl as an enol ether derived from cyclooctanone (Figure 36.17). Given orally to rats as a suspension in sesame oil, this derivative was active for 1–2 weeks because it was stored in body fat.[98]

Enol ethers were also used for the synthesis of other long-lasting steroidal drugs, such as penmestrol or pentagestrone.[99]

3. *Antipsychotics*

Clinically, depot neuroleptics possess several advantages over the short-acting oral forms. Among these, the main advantages are (a) ease of administration, (b) reliable therapeutic effect with no increase in tolerance, (c) enhanced patient compliance, (d) reduced relapse and rehospitalization rate and (f) enhanced rate and incidence of "normal life" reintegration and resocialization.[96]

IV. PARTICULAR ASPECTS OF CARRIER PRODRUG DESIGN

A. Use of cascade prodrugs

Classical carrier-linked prodrugs may sometimes be ineffective because the prodrug linkage is too stable (amides, non-activated esters). In such cases, a β-assistance, provided by an easily *in vivo* generated nucleophile, can represent an interesting solution. The release of the active molecule from the prodrug proceeds through a two-step trigger mechanism for which the name "cascade latentiation" was coined by Cain in 1975.[100,101]

The concept, also called distal hydrolysis[32] or the double prodrug concept,[23,103] is illustrated by the use of 2-acyloxymethylbenzoic acids as amine protective functions, providing amides with the lability of esters (Figure 36.18a) and by the use of substituted vinyl esters [= (2-oxo-1,3-dioxol-yl)methyl esters] as lipophilic cascade carriers for carboxylic acid-containing drugs such as ampicillin[102] or α-methyldopa[46] or various cephalosporins[6,13,28,104] (Figure 36.18b).

1. *Water-soluble taxol prodrugs*

Taxol is a potent microtubule-stabilizing agent that has been approved for cancer treatment. Despite taxol's therapeutic promise, its aqueous insolubility (<0.004 mg/mL) hampers its clinical application. Nicolaou *et al.*[105,106] report the design, synthesis and biological activity of prodrugs designed to improve water solubility and which can also be considered as cascade prodrugs (Figure 36.19).

The mechanistic rationale for the design of the two first protaxols lies in the spontaneous decomposition of the carbonate ester after the abstraction of one of the activated

FIGURE 36.19 Water-soluble protaxols.[105,106]

FIGURE 36.20 Taxol release mechanisms from protaxols.[105,106]

protons or of an acidic proton (Figure 36.20). The release of taxol from the pyridinium prodrug (taxol-2′-methylpyridinium acetate; taxol-2′-MPA) is presumed to be the result of a nucleophilic attack by water or another nucleophile at the 2′-position of the pyridinium moiety[106] (Figure 36.20).

2. Bioactivation of an antibacterial prodrug

Although amino acid **1** is a potent inhibitor of CMP–KDO synthase (Figure 36.21), a key enzyme in the biosynthesis of the lipopolysaccharide of Gram-negative bacteria, it is unable to reach its cytoplasmic target and is therefore inactive as an antibacterial agent. Simple lipophilic esters are not useful to enhance the delivery of amino acid **1** since they are not cleaved by the bacteria. On the other hand, double prodrug **3** has been found to solve the problem.[107]

Upon entry into bacterial cells, the disulfide bond in compound **3** is reduced by sulfydryl compounds present in the intracellular milieu, resulting in the formation of thiol **2**. This is highly unstable, and the active amino acid **1** is formed by a rapid, intramolecular displacement.

3. Double prodrugs derived from pilocarpine

Monoesters of pilocarpic acid are potentially useful prodrug forms for ocular administration and enable efficient penetration through the corneal membrane. Unfortunately, they suffer from poor solution stability, because in aqueous solution they cyclize spontaneously to pilocarpine.[108] However, double esters derived from pilocarpic acid (Figure 36.22) possess a high stability in aqueous solution (shelf lives of more than 5 years at 20°C were estimated). At the same time, they are readily converted to pilocarpine under conditions simulating those occurring *in vivo* through a sequential process involving enzymatic hydrolysis of the *O*-acyl bond followed by spontaneous lactonization of the intermediate pilocarpic monoester.[109]

4. Double prodrugs for peptides

Amsberry and Borchardt[110,111] have applied Cain's cascade concept to prepare lipophilic polypeptide prodrugs. The amine functionality of the polypeptide is coupled to 2′-acylated derivatives of 3-(2′,5′-dihydroxy-4′,6′-dimethylphenyl)-3,3-dimethylpropionic acid (Figure 36.23). Under simulated physiological conditions the parent amine is regenerated in a two-step process: enzymatic hydrolysis of the phenolic ester, followed by a non-enzymatic intramolecular cyclization, leading to the release of the free amine (polypeptide) and a lactone.

The lactonization step is highly favored because of the steric pressure created by the three methyl groups ("trimethyl lock" concept). An alternative to the hydrolytic first step involves a bioreductive generation of the intermediate phenolic amide (Figure 36.23).

FIGURE 36.21 Bioactivation of the antibacterial prodrug of an impermeant inhibitor of 3-deoxy-D-manno-2-octulosonate cytidylyltransferase.[107]

FIGURE 36.22 Double esters derived from pilocarpic acid are readily converted to pilocarpine under conditions simulating those occurring *in vivo*.[109]

FIGURE 36.23 Proposed conversion of esterase-sensitive and redox-sensitive double prodrugs of peptides.[110,111]

B. Codrugs

Codrugs are also named mutual prodrugs. Their design consists of the linking, in a unique molecule, of at least two different synergistic drugs that are released *in vivo* at the desired site of action. An example is found in the association of L-dopa to the catechol *O*-methyltransferase (COMT) inhibitor entacapone (Figure 36.24).[112]

C. Soft drugs

The "soft" quaternary ammonium salts developed by Bodor[113–116] are vulnerable derivatives of their "hard" analogs. In general, they show the same type of activity but with a much shorter half-life, as exemplified by the soft analog of cetylpyridinium chloride (Figure 36.25). Both compounds have the same hydrophobic chain length and thus similar surface-active and antimicrobial properties. However, the soft analog is about 40 times less toxic than its hard analog in terms of LD_{50}.[109] This is because the soft analog undergoes a fast and easy hydrolytic deactivation, resulting in the simultaneous destruction of the positive quaternary head and the surface-active properties.

In a similar way, the tetradecyloxymethyl quaternary salt of pilocarpine allows enhanced penetration through the cornea followed by a facile hydrolytic cleavage to pilocarpine (Figure 36.26). The corresponding hard analog (*N*-cetyl-pilocarpine) is unable to regenerate the parent drug and lacks any activity.[117]

Soft drugs can play an important role as short-acting systemic drugs as illustrated for esmolol in surgery or remifentanyl in anesthesia.[118,119]

FIGURE 36.24 The amino functional group of L-dopa is linked to the phenolic function of entacapone by means of a carbonyl group thus a carbamate function links the two active agents.

D. Carrier prodrugs: conclusion

The carrier prodrug approach is particularly successful in the antibiotics field and in the improvement of some pharmacokinetic parameters. Other prodrug examples are less convincing; they have nevertheless been included in this chapter to illustrate the "state of the art.[120]" Probably most of them have never been tested in man or in the laboratory. The design of carrier prodrugs represents in medicinal chemistry the counterpart of the design of protective groups in organic chemistry. Both approaches have a lot in common; in both of them imagination has no limits and reigns

FIGURE 36.25 The soft analog of cetylpyridinium chloride.[114]

FIGURE 36.26 The soft quaternary derivative of pilocarpine allows enhanced penetration through the cornea.[117]

as master. However, among the enormous number of candidates, only very few attain real success and celebrity.

V. BIOPRECURSOR PRODRUGS: APPLICATION EXAMPLES

The following examples illustrate the bioprecursor-prodrug approach, although the intentional use of bioprecursor design is relatively recent and, in some cases, there are some doubts about the prospective or the retrospective character of the design. The first examples relate to oxidative bioactivations; they are followed by examples of reductive bioactivations and finally by non-redox reactions. Often, however, the active species results from a cascade of metabolic reactions involving oxidative as well as reductive processes, complicated by hydrolytic reactions or hydration–dehydration sequences.

A. Oxidative bioactivations

1. *Dexpanthenol and 3-pyridine-methanol as pro-vitamins*

A simple example of a bioprecursor prodrug is found in dexpanthenol and 3-pyridine-methanol (Figure 36.27). These primary alcohols are the reduced forms of the vitamins pantothenic acid and nicotinic acid, respectively. Dexpanthenol has the advantage over the parent drug of being more stable, especially toward racemization.

2. *Oxidative bioactivation of losartan*

Losartan is a non-peptide angiotensin II receptor antagonist used as an antihypertensive medication.[121] It can also be considered as a bioprecursor prodrug insofar that, *in vivo,* the primary alcohol is oxidized to a carboxylic acid (Figure 36.28), which represents the actual active principle.[122]

3. *Methylenedioxyl derivatives as bioprecursors of catechols*

Various substituted and unsubstituted methylenedioxyl derivatives of apomorphine and *N*-*n*-propyl-norapomorphine have been studied by Baldessarini *et al.*,[123] and one of these, 10,11-methylenedioxy-*N*-*n*-propyl-norapomorphine, was found to be both a long-acting and an orally efficient prodrug (Figure 36.29a). The oral activity of the compound can be ascribed to the protection of the catechol system from first-pass effects by the methylenedioxyl group. The conversion to the free catechol is possible, thanks to the hepatic microsomal enzymes (see Chapters 33 and 34 on drug metabolism).

FIGURE 36.27 Dexpanthenol and 3-pyridine-methanol are pro-vitamins yielding again the parent molecules after *in vivo* oxidation.

FIGURE 36.28 Oxidative bioactivation of losartan.

4. *Conjugated cyclohexenenones as bioprecursors of catecholamines*

Venhuis *et al.* observed a particularly original oxidative bioactivation mechanism by which an α, β-unsaturated cyclic ketone is converted to the corresponding catechol and delivered enantioselectively into the CNS (Figure 36.29b). This concept can be generalized and has the potential to lead to new anti-Parkinson treatments.[124]

5. *Oxidative bioactivation of clopidogrel*

The antiaggregating agent clopidogrel undergoes extensive hydrolysis in humans (*ca.* 85% of a dose) to the inactive carbocylic acid.[125] A smaller proportion of the dose is activated by cytochrome P450 3 A to 2-oxoclopidogrel, which spontaneously and rapidly hydrolyzes to a highly unstable thiol *a*[126] (Figure 36.30).

FIGURE 36.29 (a) and (b): Oxidative bioactivations leading to catechols.[123,124]

FIGURE 36.30 Metabolism of clopidogrel and dihydro derivatives of 2-PAM.[125,127]

6. *Site-specific delivery of the acetylcholine-esterase reactivator 2-pam to the brain*

N-methylpyridinium-2-carbaldoxime (2-PAM = *a*; Figure 36.30b) constitutes the most potent reactivator of acetylcholinesterase poisoned by organophosphorus acylation. However, due to its quaternary nitrogen, 2-PAM penetrates the biological membranes poorly and does not appreciably cross the blood-brain barrier. For this compound Bodor *et al.*[127] designed an ingenious dihydropyridine–pyridinium salt type of redox delivery system. The active drug is administered as its 5,6-dihydropyridine derivative (Pro-2-PAM = *b*), which exists as a stable immonium salt *c*. The lipoidal *b* ($pK_a = 6.32$) easily penetrates the blood-brain barrier where it is oxidized to the active *a*.

A dramatic increase in the brain delivery of 2-PAM by the use of Pro-2-PAM is thus achieved, resulting in a re-activation of phosphorylated brain acetyl-cholinesterase *in vivo*.[128,129]

7. *6-Deoxyacyclovir as a bioprecursor of acyclovir*

The antiherpetic agent acyclovir suffers from poor oral bioavailability; only 10–20% of an oral dose is absorbed in humans. This can be essentially ascribed to low water solubility due to strong interaction forces in the crystal lattice. The corresponding deoxo derivative (6-deoxyacyclovir) was shown by Krenitsky[130] to be 18 times more water soluble and to be rapidly oxidized *in vivo* by xanthine oxidase to the parent drug (Figure 36.31). Studies in rats and in human volunteers showed that orally administered 6-deoxyacyclovir has a 5–6 times greater bioavailability than has acyclovir.[130,131]

8. *Bioactivation of cyclophosphamide*

Cyclophosphamide is a cytotoxic (cytostatic), cell cycle non-specific, antiproliferative agent, which is used in such diverse medical problems as neoplasia, tissue transplantation, and inflammatory diseases.[132] Chemically, it is an inert bioprecursor for a potent nitrogen mustard alkylation agent (Figure 36.32).

Cyclophosphamide was synthesized by Arnold *et al.*[135–137] in the hope that it would be inert until activated by an enzyme present in the body, especially in the tumor. The activation mechanism is believed to require an initial oxidative dealkylation, followed by a spontaneous or phosphoramidase-catalyzed hydrolysis to the parent nitrogen mustard.[133,134]

B. Reductive bioactivations

1. *Reductive bioactivation of nitrogen mustards*

Many conventional anticancer drugs display relatively poor selectivity for neoplastic cells, and solid tumors are particularly resistant both to radiation and to chemotherapy.

FIGURE 36.31 6-Deoxyacyclovir as a bioprecursor of acyclovir.[130,131]

FIGURE 36.32 Bioactivation of cyclophosphamide.[133,134]

However, in solid tumors there are a few unique and important microenvironmental properties such as localized hypoxia, nutrient deprivation and low pH.[138] On the other hand, as shown above for sulindac, sulfoxides can undergo two major biotransformations: reversible reduction to the sulfide and irreversible oxidation to the sulfone. The oxidation to the sulfone is the dominant process under normal physiological conditions, but the reduction to the sulfide becomes significant under anaerobiotic conditions.[139] Taking advantage of these findings, Kwon *et al.*[140] devised a hypoxia-selective alkylating bioprecursor prodrug (Figure 36.33a).

2. *Reductive bioactivation of nitroimidazolylmethyluracils*

Thymidine phosphorylase (TP) is an angiogenic growth factor and a target for anticancer drug design (Figure 36.33b). Docking studies of the modeled TP predicted that the binding of aminoimidazolylmethyluracils was energically more favored than that of the corresponding nitro counterparts.[141] Effectively, the passage from the nitro to the amino analog was accompanied by a 1,000-fold increase in TP inhibition.

3. *Reductive bioactivation of omeprazole*

Omeprazole effectively inhibits gastric secretion by inhibiting the gastric H^+, K^+-ATPase.[142] This enzyme is responsible for gastric acid production and is located in the secretory membranes of parietal cells. Thus, omeprazole is proposed as an anti-ulcerative drug, specially in the treatment of Zollinger-Ellison syndrome.[143]

In vivo omeprazole is transformed into the active inhibitor, a cyclic sulfenamide (Figure 36.34), which forms disulfide bridges with the thiol groups of the enzyme and thus inactivates it.[144,145] The high specificity in the action of omeprazole ($pK_a = 4.0$) is due to its preferential concentration in the rather acidic parietal cells where it is activated. In neutral regions of the body, omeprazole is rather stable and is only partially converted to the active species.

C. Mixed bioactivation mechanisms

Certain bioactivation mechanisms involve several chemical sequences, some of them being oxidative, others being reductive.

FIGURE 36.33 (a) and (b): Hypoxia-selective nitrogen mustard[140] and bioactivated TP inhibitors.[141]

FIGURE 36.34 Reductive bioactivation of omeprazole.

1. *Mixed oxidative/reductive bioactivation of dioxolanes*

The prodrug SAH 51-641 (**1**) (Figure 36.35) is a potent hypoglycemic agent, which acts by inhibiting hepatic gluconeogenesis via inhibition of fatty acid oxidation.[54] This compound is metabolized by a sequential oxidation/reduction to the corresponding keto-acid (2) and the hydroxy-acid (3). Compound (3) is a substrate for medium-chain fatty acyl CoA ligase and represents the actual active agent.[53]

2. *Arylacetic acids from aroylpropionic precursors*

The metabolic pathways involved in the biotransformation of nicotine and haloperidol involve the initial formation of an aroylpropionic acid [3-nicotinoylpropionic and 3-(*p*-fluorobenzoyl)propionic acid, respectively]. These aroylpropionic acids undergo afterwards a progressive degradation of the oxobutyric side-chain and finally arylacetic acids[15] (Figure 36.36). This information was used to design bucloxic acid,[146,147] fenbufene[148,149] and furobufene,[150] which are all three bioprecursor forms of anti-inflammatory arylacetic acids (Figure 36.37). For all of these compounds the bioactivation takes place through a multistep process implying reductive, oxidative and hydration–dehydration sequences.

More recently, arylhexenoic acids were shown to undergo a similar metabolic degradation to arylacetic acids.[151] The hexenoic analog of indomethacin (Figure 36.38) acts as a prodrug of indomethacin and provides sustained analgesia at oral dosings of 30 mg/kg to mice (phenylquinone writhing test) or to rats (yeast-induced hyperalgesia test).

FIGURE 36.35 Mixed oxidative/reductive bioactivation of dioxolanes.

FIGURE 36.36 Progressive metabolic degradation of β-aroylpropionic acids into arylacetic acids.[15]

FIGURE 36.37 Anti-inflammatory agents representing the aroylpropionic structure.[146,148,150]

FIGURE 36.38 The hexenoic analog of indomethacin as a bioprecursor prodrug of indomethacin.

D. Reactions without change in the state of oxidation

As a rule, within the bioprecursor category of prodrugs, non-redox reactions are infrequent. An example is found in the *in vivo* generation of L-cysteine from its cyclic thiocarbamate.

1. *L-2-oxothiazolidine-4-carboxylate: a cysteine delivery system*

The enzyme 5-oxo-L-prolinase, which catalyses the conversion of 5-oxo-L-proline to L-glutamate coupled to the consumption of adenosine triphosphate (ATP) (Figure 36.39), was shown by Williamson and Meister[152] to act also on a synthetic substrate, L-2-oxothiazolidine-4-carboxylate, which is an analog of 5-oxoproline with the 4-methylene group replaced by sulfur.

The enzyme, which exhibits a similar affinity for the analog and the natural substrate, is inhibited by the analog *in vitro* and *in vivo*. L-3-oxothiazolidine-4-carboxylate thus serves as a potent inhibitor of the γ-glutamyl cycle at the step of 5-oxoprolinase. Administration of L-2-oxothiazolidine-4-carboxylate to mice deprived of hepatic glutathione led to restoration of normal hepatic glutathione levels. Since L-2-oxothiazoline-4-carboxylate is an excellent substrate of the enzyme, it may serve as an intracellular delivery system for cysteine and thus a potential as a therapeutic agent for conditions in which there is depletion of hepatic glutathione.

VI. DISCUSSION

A. Bioprecursors versus carrier prodrugs

A comparative balance-sheet established for the two prodrug approaches led us to the following conclusions (Table 36.3):

- The *bioavailability* of carrier prodrugs is modulated by using a transient transport moiety; such a linkage is not implied for bioprecursors, which result from a molecular modification of the active principle itself.
- The *lipophilicity* is generally the subject of a profound alteration of the parent molecule in the case of carrier prodrugs, whereas it remains practically unchanged for bioprecursors.
- The *bioactivation* process is exclusively hydrolytic for carrier prodrugs; it involves mostly redox systems for bioprecursors.
- The *catalysis* leading to the active principle is hydrolytic (either through general catalysis or through extra-hepatic enzymes) for carrier prodrugs. For bioprecursors, it seems largely restricted to Phase I metabolizing enzymes.

B. Existence of mixed-type prodrugs

In some cases, the design of mixed-type prodrugs can be advantageous, as illustrated in the following three examples.

1. *Disulfide thiamine prodrugs*

The thiamin (vitamin B1) molecule contains a quaternary ammonium functionality and is thus badly absorbed. In healthy patients the necessary amounts of thiamin are absorbed, thanks to an active transport mechanism coupled with ATP consumption. However, these mechanisms are rapidly saturable and easily inhibited, especially by

TABLE 36.3 Bioprecursors Versus Carrier-Prodrugs

	Prodrugs	
	Carrier prodrugs	Bioprecursors
Constitution	Active principle + carrier group	No carrier group
Lipophilicity	Strongly modified	Usually slightly modified
Bioactivation	Hydrolytic	Mostly oxidative or reductive
Catalysis	Chemical or enzymic	Only enzymic

FIGURE 36.39 L-2-Oxothiazolidine-4-carboxylate: an intracellular cysteine delivery system.[152]

chronic alcoholic consumption. As a consequence of the insufficient absorption of thiamin, alcoholism often entails Wernicke's encephalopathy (neurological disorders such as nystagmus, ocular motor nerve paralysis, memory losses, disorientation). The design of lipophilic prodrugs able to reach the CNS by passive diffusion was then undertaken: compounds like (a) and (b) result from lipophilic disulfide derivation of the open ring thiolate anion corresponding to thiamin (Figure 36.40).[153] Such compounds can also be considered as carrier prodrugs, insofar as the thiolate is linked to an *n*-propylthio (a) or a tetrahydrofuranylmethylthio (b) transport moiety, or as bioprecursors, insofar as a bioreductive cleavage in the thiolate anion is needed to generate thiamin; the thiolate anion then functions as a less polar (no quaternary ammonium functional group) precursor form of thiamin. After oral administration, higher thiamin blood levels were observed in healthy volunteers, as well as in cirrhotic patients with the prodrug, than with thiamin hydrochloride.[154]

2. *Trigonelline esters and amines*

Generalizing the dihydropyridine⇔pyridinium salt redox delivery system successfully applied to 2-PAM, Bodor *et al.* proposed an astute sustained release methodology for brain delivery, based on the mixed-type prodrugs.[155] The biologically active compound is linked to a lipoidal dihydropyridine carrier that easily penetrates the blood-brain barrier (Figure 36.41).

Enzymatic oxidation *in vivo* by the NAD ↔ NADH system of the carrier part to the ionic pyridinium salt prevents its elimination from the brain, while elimination from the general circulation is accelerated. Subsequent cleavage of the quaternary carrier-drug species results in sustained delivery of the drug in the brain and a facile elimination of the non-toxic carrier part (trigonelline or its *N*-benzyl analog).

VII. DIFFICULTIES AND LIMITATIONS

The introduction of prodrugs in human therapy gave successful results in overcoming undesirable properties such as poor absorption, rapid biodegradation, or formulation problems. It can be expected that an increasing number of medicinal chemists will be tempted by this approach. However, they must keep in mind that prodrug design can also give rise to a large number of new difficulties, especially in the assessment of pharmacological, pharmacokinetic, toxicological and clinical properties.

FIGURE 36.40 Disulfide thiamin prodrugs as example of mixed-type prodrugs.[153]

FIGURE 36.41 Trigonelline amides (or esters) as examples of mixed-type prodrugs.[155]

At the *pharmacological level*, for example, because bioactivation is necessary to create the active species, these compounds cannot be submitted to preliminary *in vitro* screening tests, namely, binding studies, neurotransmitter re-uptake, measurements of enzymatic inhibition and activity on isolated organs.

The measurements of *pharmacokinetic parameters* can lead to numerous misinterpretations. Thus, pivampicillin has a half-life of 103 min in a buffered aqueous solution at 37°C, but it falls to less than 1 min after addition of only 1% of mouse or rat serum. In the presence of human serum (10%), however, it is 50 min, whereas in whole human blood it is only 5 min. These results exemplify the care required to avoid incorrect conclusions. In addition, when a prodrug and the parent molecule are compared, one must take into account the differences in their respective time courses of action. The maximum activity can appear later for the prodrug than for the parent compound, and often the comparison of the area under the curve could constitute a better criterion.

At the *toxicological level*, even when prodrugs derive from well-known active principles, they have to be regarded as new entities. Undesirable side effects can appear that are directly related to the prodrug (allergy to bucloxic acid) or derived from the bioactivation process (formation of unwanted or unexpected metabolites) or which can be attributed to the temporary transport moiety (digestive intolerance to pivampicillin, antivitamin-PP activity of nicafenine). This latter case is particularly illustrative: an apparently innocent carrier group such as *N*-hydroxyethylnicotinamide appeared as a promising candidate for improving the absorption of acidic anti-inflammatory drugs or clofibric acid.[156–158] However, during the clinical studies, side effects similar to vitamin PP deficiency appeared, suggesting that *N*-hydroxyethylnicotinamide could function as a nicotinamide antimetabolite. The compounds then had to be withdrawn (H. Cousse, Pierre Fabre & Co, personal communication).

In a review of potential hazards of the prodrug approach, Gorrod[159] cites four toxicity mechanisms:

1. Formation of a toxic metabolite of the total prodrug, which is not produced by the parent drug.
2. Consumption of a vital constituent (e.g. glutathione) during the prodrug activation process. As L-cysteine is needed for the biosynthesis of glutathione, a supply with L-cysteine prodrugs can eventually confer some protection of the hepatic cells.[160]
3. Generation of a toxic derivative from a transport moiety supposed to be "inert."
4. Release of a pharmacokinetic modifier (causing enzymatic induction, displacing protein-bound molecules, altering drug excretion, etc.).

Eventually, at the *clinical stage*, the predictive value of animal experiments is also questionable. Thus, for two prodrugs derived from α-methyldopa, the active doses in the rat were identical; nonetheless, they turned out to be very different during clinical investigations. One compound was just as potent as α-methyldopa, whereas the other one was 3–4 times more potent.[161,162]

An application file for a new prodrug should take into account all these aspects and can in no way be regarded just as a complement to the main file.

VIII. CONCLUSION

In the future it would be preferable to distinguish the carrier prodrug and the bioprecursor approaches. The first one, consisting in the attachment of a temporary carrier group to an active principle, largely proved its utility in the design of orally active antibiotics and, more generally, every time high bioavailability in plasma or peripheral organs is required. The CNS delivery of drugs using carrier prodrugs is less convincing insofar as usually high dosages are needed to ascertain clinical efficiency (1–2 g progabide per day, for example).

The design of bioprecursors, which represents a creative application of the active metabolite concept in the forward-looking way, seems *a priori* more adequate for CNS delivery, but its clinical usefulness still needs to be proven.

Mixed-type approaches gave good results for thiamine and appear to be an interesting alternative when each individual approach fails.

REFERENCES

1. Stella, V. Pro-drugs: an overview and definition. In *Pro-drugs as a Novel Drug Delivery System* (Higuchi, T., Stella, V., Eds), Vol. 14. American Chemical Society: Washington, DC, 1975, pp. 1–115.
2. Sinkula, A. A. Prodrugs, protective groups and the medicinal chemist. In *Medicinal Chemistry* (Mathieu, J., Ed.), Vol. 5. Elsevier: Amsterdam, 1977, pp. 125–133.
3. Wermuth, C. G. Designing prodrugs and bioprecursors. In *The Practice of Medicinal Chemistry* (Wermuth, C. G., Ed.). Academic press: London, 2003, pp. 561–585.
4. Ettmayer, P., Amidon, G. L., Clement, B., Testa, B. Lessons learned from marketed and investigational prodrugs. *J. Med. Chem.* 2004, *47*, 2393–2403.
5. Testa, B. Prodrug objectives and design. In *Comprehensive Medicinal Chemistry II* (Taylor, J. B., Triggle, D. J., Eds), Vol. 5. Elsevier: Amsterdam, 2007, pp. 1009–1041.
6. Albert, A. Chemical aspects of selective toxicity. *Nature (London)* 1958, *182*, 421–423.
7. Harper, N. J. Drug latentiation. *J. Med. Pharm. Chem.* 1959, *1*, 467–500.
8. Wermuth, C. G. Bioprécurseurs contre prodrogues. In *Drug Metabolism and Drug Design: Quo Vadis?* (Briot, M., Cautreels, W., Roncucci, R., Eds). Sanofi-Clin-Midy: Montpellier, 1983, pp. 253–271.
9. Wermuth, C. G. Designing prodrugs and bioprecursors. In *Drug Design: Fact or Fantasy?* (Jollès, G., Wooldrige, K. R. H., Eds). Academic Press: London, 1984, pp. 47–72.

10. Wermuth, C. G. Les prodrogues, des médicaments plus sûrs et plus maniables. *Bull. Soc. Pharm. Bordeaux.* 1980, *119*, 107–129.
11. Wermuth, C. G. Modulation of natural substances in order to improve their pharmacokinetic properties. In *Natural Products as Medicinal Agents* (Beal, J. L., Reinhard, E., Eds). Hippokrates Verlag: Stuttgart, 1981, pp. 185–216.
12. Daehne, W. V., Frederiksen, E., Gundersen, E., Lund, F., March, P., Petersen, H. J., Roholt, K., Tybring, L., Godtfredsen, W. O. Acyloxymethyl esters of ampicillin. *J. Med. Chem.* 1970, *13*, 607–612.
13. Bodin, N. O., Ekström, B., Forsgren, U., Jalar, L. P., Magni, L., Ramsey, C. H., Sjöberg, B. Bacampicillin: a new orally well-absorbed derivative of ampicillin. *Antimicrob. Agents Chemother.* 1975, *8*, 518–525.
14. Clayton, J. P., Cole, M., Elson, S. W., Ferres, H., Hanson, J. C., Mizen, L. W., Sutherland, R. Preparation, hydrolysis, and oral absorption of lactonyl esters of penicillins. *J. Med. Chem.* 1976, *19*(12), 1385–1391.
15. Testa, B., Jenner, P. *Drug Metabolism, Chemical and Biochemical Aspects*. Marcel Dekker: New York, 1976.
16. Shen, T. I., Witzel, B. E., Jones, H., Linn, B. O., McPherson, J., Greenwald, R., Fordice, M., Jacobs, A. Synthesis of a new anti-inflammatory agent, *cis* -5-fluoro-2-methyl-1-[*p*-(methylsulfinyl) benzylidenyl]-indene-3-acetic acid. *Fed. Proc.* 1972, *31*, 577.
17. Hucker, H. B., Stauffer, S. C., White, S. D., Rhodes, R. A., Arrison, B. H., Umbenhauer, E. R., Bower, R. J., McMahon, F. G. Physiological disposition and metabolic fate of a new anti-inflammatory agent, *cis* -5-fluoro-2-methyl-1-[*p*-(methylsulfinyl) benzylidenyl]-indene-3-acetic acid in the rat, dog, rhesus monkey, and man. *Drug Metab. Dispos.* 1973, *1*, 721–736.
18. Duggan, D. E., Hare, L. E., Ditzler, C. A., Lei, B. W., Kwan, K. C. The disposition of sulindac. *Clin. Pharmacol. Ther.* 1977, *21*, 326–335.
19. Duggan, D. E., Hooke, K. F., Noll, R. M., Hucker, H. B., Van Arman, C. G. Comparative biodisposition of sulindac and metabolites in five species. *Biochem. Pharmacol.* 1978, *27*, 2311–2320.
20. Duggan, D. E., Hooke, K. F., Risley, E. A., Shen, T. Y., Van Arman, C. G. Identification of the biologically active form of sulindac. *J. Pharmacol. Exp. Ther.* 1977, *201*(1), 8–13.
21. Drabek, J., Neumann, R. Proinsecticides. In *Progress in Pesticide Biochemistry and Toxicology* (Hutson, D. H., Roberts, T. R., Eds), Vol. 5. John Wiley & Sons, Ltd.: Chichester, 1985, pp. 35–86.
22. Higuchi, T., Stella, V. *Pro-drugs as novel drug delivery systems*, Vol. 14. American Chemical Society: Washington, DC, 1975.
23. Bundgaard, H. Design and application of prodrugs. In *A Textbook of Drug Design and Development* (Krogsgaard-Larsen, P., Bundgaard, H., Eds), 1st Edition. Harwood Academic Publishers: Chur, 1991, pp. 113–191.
24. McClure, D. The effect of a pro-drug of epinephrine (dipivaloyl-epinephrine) in glaucoma-general pharmacology, toxicology and clinical experience. In *Pro-Drugs as a Novel Drug Delivery Systems* (Higuchi, T., Stella, V., Eds), Vol. 14. American Chemical Society: Washington, DC, 1975, pp. 224–235.
25. Horn, A. S. Pro-drugs of dopaminergic agonists. *Chemistry and Industry* 1980, 441–444.
26. Westerink, B. H. C., Dijkstra, D., Feenstra, M. G. P., Grol, C. J., Horn, A. S., Rollema, H., Wirix, E. Dopaminergic prodrugs: brain concentration and neurochemical effects of 5,6- and 6,7-ADTN after administration as dibenzoyl esters. *Eur. J. Pharmacol.* 1980, *61*, 7–15.
27. Tullar, B. F., Minatoya, H., Lorenz, R. R. Esters of *N-tert*-butylarterenol. Long-acting new bronchodilators with reduced cardiac effects. *J. Med. Chem.* 1976, *19*, 834–838.
28. Lin, T. S. Synthesis and *in vitro* antiviral activity of 3′-*O*-acyl derivatives of 5′-amino-2′-deoxy thymidine: potential prodrugs for topical application. *J. Pharm. Sci.* 1984, *73*, 1568–1571.
29. Harnden, M. R., Jarvest, R. L., Boyd, M. R., Sutton, D., Vere hodge, R. A. Prodrugs of the selective antiherpesvirus agent 9-[4-hydroxy-3-(hydroxymethyl)but-1-yl]guanine (BRL 39123) with improved gastrointestinal absorption properties. *J. Med. Chem.* 1989, *32*, 1738–1743.
30. Jones, L. A., Moorman, A. R., Chamberlain, S. D., Miranda, P.d., Reynolds, D. J., Burns, C. L., Krenitsky, T. A. Di- and triester prodrugs of the varicella-zoster antiviral agent 6-methoxypurine arabinoside. *J. Med. Chem.* 1992, *35*, 56–63.
31. Bodor, N., Sloan, K. B., Little, R. J., Selk, S. H., Caldwell, L. Soft drugs 4. 3-Spirothiazolidines of hydrocortisone and its derivatives. *Int. J. Pharm.* 1982, *10*, 307–321.
32. Collis, A. J. Drug access and prodrugs. In *Medicinal chemistry – the role of organic chemistry in drug research* (Ganellin, C. R., Roberts, S. M., Eds), 2nd Edition. Academic Press: London, 1993, pp. 61–82.
33. Anden, N. E., Corrodi, H., Dahlström, A., Fuxe, K., Högfelt, T. Effects of tyrosine hydroxylase inhibition on the amine levels of central monoamine neurons. *Life Sci.* 1966, *5*, 561–568.
34. Anonymous. Etilevodopa. *Drugs Fut.* 2001, *26*, 219–223.
35. Frey, H. H., Popp, C., Löscher, W. Influence of inhibitors of the high affinity GABA uptake on seizure threshold in mice. *Neuropharmacology* 1979, *18*, 581–590.
36. Ulm, E. H., Hichens, M., Gomez, H. J., Till, A. E., Hand, E., Vassil, T. C., Biollaz, J., Brunner, H. R., Schelling, J. L. Enalapril maleate and a lysine analogue (MK-521): disposition in man. *Br. J. Pharmacol.* 1982, *14*, 357–362.
37. Swanson, B. L., Vlasses, P. H., Ferguson, R. K., Berquist, P. A., Till, A. E., Irvin, J. D., Harris, K. Influence of food on the bioavailability of enalapril. *J. Pharm. Sci.* 1984, *73*, 1655–1657.
38. Zannad, F. Trandolapril: how does it differ from other angiotensin converting enzyme inhibitors?. *Drugs* 1993, *46*, 172–183.
39. Tsybina, N., Ostrovskaya, R. U., Protopova, T. V., Parin, V. V., Selezneva, N. I., Skolainov, A. P. Synthesis and pharmacological activity of gamma-aminobutyric acid derivatives. *Khim. Pharm. Zh.* 1974, *17*, 10–13.
40. Ostrovskaya, R. U., Parin, V. V., Tsybina, N. M. The comparative neurotropic potency of gamma-aminobutyric acid and its cetyl ester. *Byul. Eksp. Biol. Med.* 1972, *73*, 51–55.
41. Beardsley, G. P., Rosowsky, A. Effect of methotrexate γ-monohexadecyl ester (γ-MHxMTX) on nucleoside uptake by human leukemic cells.In *Proc. Am. Assoc. Cancer Res. 21*, 71 Meeting, 1980, p. 264.
42. Wermuth, C. G., Gaignault, J.-C., Marchandeau, C. Designing prodrugs and bioprecursors I: carrier prodrugs. In *The Practice of Medicinal Chemistry* (Wermuth, C. G., Ed.). Academic Press: London, 1996, pp. 671–696.
43. Jones, G. Lipoidal pro-drug analogues of various anti-inflammatory agents. *Chem. Ind. (London)* 1980, 452–456.
44. Jansen, A. B. A., Russel, T. J. Some novel penicillin derivatives. *J. Chem. Soc.* 1965, 2127–2132.
45. Binderup, E., Godtfredsen, W. O., Roholt, K. Orally active cephaloglycin esters. *J. Antibiot.* 1971, *24*, 767–773.
46. Saari, W. S., Halczenko, W., Cochran, D. W., Dobrinska, M. R., Vincek, W. C., Titus, D. C., Gaul, S. L., Sweet, C. S. 3-Hydroxy-α-methyltyrosine progenitors: synthesis and evaluation of some (2-oxo-1,3-dioxol-4-yl)methyl esters. *J. Med. Chem.* 1984, *27*, 713–717.
47. Falch, E., Krogsgaard-Larsen, P., Christensen, A. V. Esters of isoguvacine as potential prodrugs. *J. Med. Chem.* 1981, *24*, 285–289.
48. Svahn, C. M., Merenyi, F., Karlson, L., Widlund, L., Grälls, M. Tranexamic acid derivatives with enhanced absorption. *J. Med. Chem.* 1986, *29*, 448–453.
49. Torriani, H. Talniflumate. *Drugs Fut.* 1979, *4*, 448–450.
50. Torriani, H. Talmetacin. *Drugs Fut.* 1982, *7*, 823–824.
51. Alexander, J., Fromtling, R. A., Bland, J. A., Pelak, B. A., Gilfillan, E. C. (Acyloxy)alkyl carbamate prodrugs of norfloxacin. *J. Med. Chem.* 1991, *34*, 78–81.
52. Ahlmark, M., Vepsaelaeinen, J., Taipale, H., Niemi, R., Jaervinen, T. Bisphophonate prodrugs: synthesis and *in vitro* evaluation of novel clodronic acid dianhydrides as bioreversible prodrugs of clodronate. *J. Med. Chem.* 1999, *42*, 1473–1476.

53. Aicher, T. D., Bebernitz, G. R., Bell, P. A., Brand, L. J., Dain, J. G., Deems, R., Fillers, W. S., Foley, J. E., Knorr, D. C., Nadelson, J., Otero, D. A., Simpson, R., Strohschein, R. J., Young, D. A. Hypoglycemic prodrugs of 4-(2,2-dimethyl-1-oxopropyl)benzoic acid. *J. Med. Chem.* 1999, *42*, 153–163.
54. Young, D. A., Ho, R. S., Bell, P. A., Cohen, D. K., McIntosh, R. H., Nadelson, J., Foley, J. E. Inhibition of hepatic glucose production by SDZ 51641. *Diabetes* 1990, *39*, 1408–1413.
55. Walsh, D. A., Moran, H. W., Shamblee, D. A., Welstead, W. J., Jr., Nolan, J. C., Sancilio, L. F., Graff, G. Antiinflammatory agents. 4. Syntheses and biological evaluation of potential prodrugs of 2-amino-3-benzoylbenzeneacetic acid and 2-amino-3-(4-chlorobenzoyl)benzen eacetic acid. *J. Med. Chem.* 1990, *33*, 2296–2304.
56. Galzinga, L., Garbin, L., Bianchi, M., Marzotto, A. Properties of two derivatives of γ-aminobutyric acid (GABA) capable of abolishing cardiazol- and bicuculline-induced convulsions in the rat. *Arch. Int. Pharmacodyn.* 1978, *235*, 73–85.
57. Ahtee, L., Halmekoski, J., Heinonen, H., Koskimies, A. Comparison of the central nervous system actions of taurine and *N*-pivaloyltaurine. *Br. J. Pharmacol.* 1979, *66*(3), 480 pages.
58. Kaplan, J. P., Raizon, B., Desarmenien, M., Feltz, P., Headley, P. M., Worms, P., Lloyd, K. G., Bartholini, G. New anticonvulsants: Schiff bases of γ-aminobutyric acid and γ-aminobutyramide. *J. Med. Chem.* 1980, *23*, 702–704.
59. Jensen, N. P., Friedman, J. J., Kropp, H., Kahan, F. M. Use of acetylacetone to prepare a prodrug of cycloserine. *J. Med. Chem.* 1980, *23*, 6–8.
60. Bodor, N. S., Sloan, K. B., Hussain, A. A., Novel transient pro-drug forms of L-DOPA, *US Patent 3,891,696* (June 24, 1975; Inter'X Res. Corpn.) 1975.
61. Koch, H. ST-1059. *Drugs Fut.* 1981, *6*, 244–246.
62. Vicentini, C. B., Guarneri, M., Sarto, G. Hypoglycemic compounds. Sulfonylurea derivatives containing amino acids and dipeptides. *Farmaco, Ed. Sci.* 1983, *38*, 595–608.
63. Rubenstein, S. M., Baichwal, V., Beckmann, H., Clark, D. L., Frankmoelle, W., Roche, D., Santha, E., Schwender, S., Thoolen, M., Ye, Q., Jaen, J. C. Hydrophilic, pro-drug analogues of T138067 are efficacious in controlling tumor growth *in vivo* and show a decreased ability to cross the blood brain barrier. *J. Med. Chem.* 2001, *44*, 3599–3605.
64. Boehringer, Dérivés de la L-(3,4-dihydroxy-phényl)-2-méthyl alanine et leur préparation, *Belg. Pat. 839,362* (March 9, 1976 to Boehringer Mannheim GMBH), 1976.
65. Kingsbury, W. D., Boehm, J. C., Mehta, R. J., Grappel, S. F. Transport of antimicrobial agents using peptide carrier systems: anticandidal activity of *m*-fluorophenylalanine peptide conjugates. *J. Med. Chem.* 1983, *26*, 1725–1729.
66. Hu, M., Subramanian, P., Mosberg, H. I., Amidon, G. L. Use of the peptide carrier system to improve the intestinal absorption of L-α-methyldopa: carrier kinetics, intestinal permeabilities, and *in vitro* hydrolysis of dipeptidyl derivatives of L-α-methyldopa. *Pharm. Res.* 1989, *6*, 66–70.
67. Rahmathullah, S. M., Hall, J. E., Bender, B. C., McCurdy, D. R., Tidwell, R. R., Boykin, D. W. Prodrugs for amidines: synthesis and anti-*pneumocystis carinii* activity of carbamates of 2,5-bis(4-amidinophenyl)furan. *J. Med. Chem.* 1999, *42*, 3994–4000.
68. Buur, A., Bundgaard, H., Falch, E. Prodrugs of 5-fluorouracil. VII. Hydrolysis kinetics and physicochemical properties of *N*-ethoxy- and *N*-phenoxycarbonylmethyl derivatives of 5-fluorouracil. *Acta Pharm. Suec.* 1986, *23*, 205–216.
69. Larsen, J. D., Bundgaard, H. Prodrug forms for the sulfonamide group. I. Evaluation of *N*-acyl derivatives, *N*-sulfonylamidines, *N*-sulfonyl-sulfilimines and sulfonylureas as possible prodrug derivatives. *Int. J. Pharm.* 1987, *37*, 87–95.
70. Bundgaard, H., Buur, A., Hansen, K. T., Larsen, J. D., Moss, J., Olsen, L. Prodrugs as drug delivery systems. 77. Phthalidyl derivatives as prodrug forms for amides, sulfonamides, carbamates and other NH-acidic compounds. *Int. J. Pharm.* 1988, *45*, 47–57.
71. Hardcastle, G. A., Johnson, D. A., Panetta, C. A., Scott, A. I., Sutherland, S. A. The preparation and structure of hetacillin. *J. Org. Chem.* 1966, *31*, 897–899.
72. Bundgaard, H. Design of prodrugs: bioreversible derivatives for various functional groups and chemical entities. In *Design of Prodrugs* (Bundgaard, H., Ed.). Elsevier: Amsterdam, 1985, pp. 1–92.
73. Anonymous. Droxicam. *Drugs Fut.* 1986, *11*, 835–836.
74. Zeller, P., Pletscher, A., Gey, K. F., Gutmann, H., Hegedus, B., Staub, O. Amino acid and fatty acid hydrazides: chemistry and action on monoamine oxidase. *Ann. NY Acad. Sci.* 1959, *80*, 555–567.
75. Knapp, F. F. J., Goodman, M. M., Callahan, A. P., Ferren, L. A., Kabalka, G. W., Sastry, K. A. R. New myocardial imaging agents: stabilization of radioiodine as a terminal vinyl iodide moiety on tellurium fatty acids. *J. Med. Chem.* 1983, *26*, 1293–1300.
76. Wess, G., Kramer, W., Schubert, G., Bickel, M., Hoffmann, A., Baringhaus, K. H. Coupling of drugs to modified bile acids for liver specific targeting. *Abst. Pap.205th Meeting of the American Chemical Society.* 1993, *16*. Pt. 1, MEDI 152.
77. Nagy, A., Schally, A. V. Targeted cytotoxic somatostatin analogs: a modern approach to the therapy of various cancers. *Drugs Fut.* 2001, *26*, 261–270.
78. Biel, J. H., Martin, Y. C. Organic synthesis as a source of new drugs. In *Drug Discovery-Science and Development in a Changing Society* (Society, A. C., Ed.). American Chemical Society: Washington, DC, 1971, Vol. 108, pp. 81–111.
79. Whittaker, J. R. Biosynthesis of a thiouracil pheomelanin in embryonic pigment cells exposed to thiouracil. *J. Biol. Chem.* 1971, *246*, 6217–6226.
80. Dencker, L., Larsson, B., Olander, K., Ullberg, S., Yokota, M. False precursors of melanin as selective melanoma seekers. *Br. J. Cancer* 1979, *39*, 449–452.
81. Wätjen, F., Buchardt, O., Langvad, E. Affinity therapeutics. 1. Selective incorporation of 2-thiouracil derivatives in murine melanomas. Cytostatic activity of 2-thiouracil arotinoids, 2-thiouracil retinoids, arotinoids and retinoids. *J. Med. Chem.* 1982, *25*, 956–960.
82. Ariëns, E. J. Modulation of pharmacokinetics by molecular manipulation. In *Drug Des.* (Ariëns, E. J., Ed.), Vol. II. Academic Press: New York, 1971, pp. 1–127.
83. Ariëns, E. J. Pharmacological basis of cancer therapy, In *Twenty-Seventh Annual Symposium on Fundamental Cancer Research*. The University of Texas M.D. Anderson Hospital and Tumor Institute at Houston, 1974; The Williams and Wilkins Company, Baltimore, MD, 1975.
84. Hurwitx, E., Kashi, R., Burowsky, D., Arnon, R., Haimovitch, J. Site-directed chemotherapy with a drug bound to anti-idiotypic antibody to a lymphoma cell-surface IgM. *Int. J. Cancer* 1983, *31*, 745–748.
85. Magnan, S. D. J., Shirota, F. N., Nagasawa, H. T. Drug latentiation by γ-glutamyl transpeptidase. *J. Med. Chem.* 1982, *25*, 1018–1021.
86. Wilk, S., Mizoguchi, H., Orlowski, M. Gamma-glutamyl-DOPA: a kidney specific dopamine precursor. *J. Pharmacol. Exp. Ther.* 1978, *206*, 227–232.
87. Kyncl, J. J., Minard, F. N., Jones, P. H. Peripheral dopamine receptors. In *Symposium on Peripheral Dopaminergic Receptors-Strasbourg, July 1978* (Imbs, J. L., Schwartz, J., Eds). Pergamon Press: Oxford, 1979, pp. 369–380.
88. Jones, P. H., Kyncl, J., Ours, C. W., Somani, P. Esters of γ-glutamyl amide of dopamine. *US Patent. 4,017,636* (12 April 1977, to Abott Laboratories) 1977.
89. Orlowski, M., Mizoguchi, H., Wilk, S. *N*-acyl-γ-glutamyl derivatives of sulfamethoxazole as models of kidney-selective prodrugs. *J. Pharmacol. Exp. Ther.* 1979, *212*, 167–172.
90. Friend, D. R., Chang, G. W. Drug glycosides: potential prodrugs for colon-specific drug delivery. *J. Med. Chem.* 1985, *28*, 51–57.

91. Friend, D. R., Chang, G. W. A colon-specific drug-delivery system based on drug glycosides and the glycosidases of colonic bacteria. *J. Med. Chem.* 1984, *27*, 261–266.
92. Cook, A. F., Holman, M. J., Kramer, M. J., Trown, P. W. Florinated pyrimidine nucleosides. 3. Synthesis and antitumor activity of a series of 5′-deoxy-5fluoropyrimidine nucleosides. *J. Med. Chem.* 1979, *22*, 1330–1335.
93. Au, J. L.-S., Walker, J. S., Rustum, Y. Pharmacokinetic studies of 5-fluorouracil and 5′-deoxy-5-fluorouridine in rats. *J. Pharmacol. Exp. Ther.* 1983, *227*, 174–180.
94. de Groot, F. M. H., van Berkom, L. W. A., Scheeren, H. W. Synthesis and biological evaluation of 2′-carbamate-linked and 3′-carbonate-linked prodrugs of paclitaxel: selective activation by the tumor-associated protease plasmin. *J. Med. Chem.* 2000, *43*, 3093–3102.
95. Stella, V. J., Charman, W. N. A., Naringrekar, V. H. Prodrugs do they have advantages in clinical practice?. *Drugs* 1985, *29*, 455–473.
96. Sinkula, A. A. Sustained drug action accomplished by the prodrug approach. In *Design of Prodrugs* (Bundgaard, H., Ed.). Elsevier: Amsterdam, 1985, pp. 157–176.
97. Dusterberg, B., Wendt, H. Plasma levels of dehydroepiandrosterone and 17β-estradiol after intramuscular administration of Gynodian-Depot® in 3 women. *Hormone Res.* 1983, *17*, 84–89.
98. Falconi, G., Galetti, F., Celasco, G., Gardi, R. Oral long-lasting estrogenic activity of estradiol 3-benzoate 17-cyclooctenyl ether. *Steroids* 1972, *20*, 627–632.
99. Ercoli, A., Gardi, R. D4-3-keto steroidal ethers. Paradoxical dependency of their effectiveness on the administration route. *J. Am. Chem. Soc.* 1960, *82*, 746–748.
100. Cain, B. F. The role of structure–activity studies in the design of antitumor agents. *Cancer Chemother. Rep.* 1975, *59*(4), 679–683.
101. Cain, B. F. 2-Acyloxymethylbenzoic acids. Novel amine protective functions providing amides with the lability of esters. *J. Org. Chem.* 1976, *41*(11), 2029–2031.
102. Sakamoto, F., Ikeda, S., Tsukamoto, G. Studies on prodrugs. II. Preparation and characterization of (5-substituted 2-oxo-1,3-dioxolen-4-yl)methyl esters of ampicillin. *Chem. Pharm. Bull.* 1984, *32*, 2241–2248.
103. Bundgaard, H. Novel chemical approaches in prodrug design. *Drugs Fut.* 1991, *16*, 443–458.
104. Borgman, R. J., Smith, R. V., Keiser, J. E. The acetylation of apomorphine. An improved method for the selective preparation of diacetylapomorphine utilizing trifluoroacetic acid/acetyl bromide. *Synthesis* 1975, 249–250.
105. Nicolaou, K. C., Riemer, C., Kerr, M. A., Rideout, D., Wrasidlo, W. Design, synthesis and biological activity of protaxols. *Nature (London)* 1993, *364*, 464–466.
106. Nicolaou, K. C., Guy, R. K., Pitsinos, E. N., Wrasidlo, W. A water-soluble prodrug of taxol with self-assembling properties. *Angew. Chem. Int. Ed.* 1994. (in press)
107. Norbeck, D. W., Rosenbrook, W., Kramer, J. B., Grampovnik, D. J., Lartey, P. A. A novel prodrug of an impermeant inhibitor of 3-deoxy-D-manno-2-octulosonate cytidylyl-transferase has antibacterial activity. *J. Med. Chem.* 1989, *32*, 625–629.
108. Bundgaard, H., Falch, E., Larsen, C., Mikkelson, T. J. Pilocarpine prodrugs Synthesis, stability, bioconversion, and physicochemical properties of sequentially labile pilocarpine acid diesters. *J. Pharm. Sci.* 1986, *75*, 36–44.
109. Bundgaard, H., Falch, E., Larsen, C., Mosher, G. L., Mikkelson, T. J. Pilocarpine prodrugs II. Synthesis, stability, bioconversion and physicochemical properties of sequentially labile pilocarpine acid diesters. *J. Pharm. Sci.* 1986, *75*, 775–783.
110. Amsberry, K. L., Borchardt, R. T. Amine prodrugs which utilize hydroxy-amide lactonization. I. A potential redox-sensitive amide prodrug. *Pharm. Res.* 1991, *8*, 323–330.
111. Amsberry, K. L., Gerstenberger, A. E., Borchardt, R. T. Amine prodrugs which utilize hydroxy-amide lactonization. II. A potential esterase-sensitive prodrug. *Pharm. Res.* 1991, *8*, 455–461.
112. Leppänen, J., Huuskonen, J., Nevalainen, T., Gynther, J., Taipale, H., Järvinen, T. Design and synthesis of a novel L-dopa-entacapone codrug. *J. Med. Chem* 2002, *45*, 1379–1382.
113. Bodor, N. S. Novel approaches for the design of membrane transport properties of drugs. In *Design of biopharmaceutical properties through prodrugs and analogs* (Roche, E. B., Ed.). American Pharmaceutical Association: Washington, DC, 1977, pp. 98–135.
114. Bodor, N. S., Kaminski, J. J., Selk, S. Soft drugs. 1. Labile quaternary ammonium salts as soft antimicrobials. *J. Med. Chem.* 1980, *23*, 469–474.
115. Bodor, N. S., Kaminski, J. J. Soft drugs. 2. Soft alkylating compounds as potential antitumor agents. *J. Med. Chem.* 1980, *232*, 566–569.
116. Bodor, N. S., Woods, R., Raper, C., Kearney, P., Kaminski, J. J. Soft drugs. 3. A new class of anticholinergic agents. *J. Med. Chem.* 1980, *23*, 474–480.
117. Bodor, N. Prodrugs versus soft drugs. In *Design of Prodrugs* (Bundgaard, H., Ed.). Elsevier: Amsterdam, 1985, pp. 333–354.
118. Peacock, J. E., Francis, G. Remifentanyl pharmacokinetics and metabolism. *Drugs Today* 1997, *33*, 611–618.
119. Benfield, P., Sorkin, E. M. Esmolol. A preliminary review of its pharmacodynamic and pharmacokinetic properties, and therapeutic efficacy. *Drugs* 1987, 392–412.
120. Pitman, I. H. Pro-drugs of amides, imides, and amines. In *Med. Res. Rev.* (deStevens, G., Ed.), Vol. 1. John Wiley & Sons: New York, 1981, pp. 189–214.
121. Carini, D. J., Duncia, J. V., Aldrich, P. E., Chiu, A. T., Johnson, A. U., Piera, M. E., Price, W. A., Santella, J. B.r., Wells, G. J., Wexler, R. R., Wong, P. C., Yoo, S.-E., Timmerman, P. B. Non-peptide angiotensin II antagonists: the discovery of a series of *N*-(biphenylmethyl) imidazoles as potent, orally active antihypertensives. *J. Med. Chem.* 1991, *34*, 2525–2547.
122. Wong, P. C., Price, W. A. J., Chiu, A. T., Duncia, J. V., Carini, D. J., Wexler, R. R., Johnson, A. L., Timmerman, P. B. Non-peptide angiotensin II receptor antagonists. XI. Pharmacology of Exp 3174: an active metabolite of DuP 753, an orally activeantihypertensive agent. *J. Pharmacol. Exp. Ther.* 1990, *255*, 211–217.
123. Baldessarini, R. J., Neumeyer, J. L., Campbell, A., Sperk, G., Ram, V. J., Arana, G. W., Kula, N. S. An orally effective, long-acting dopaminergic prodrug: (–)-10,11-methylenedioxy-*N*-propylnorapomorphine. *Eur. J. Pharmacol.* 1982, *77*, 87–88.
124. Venhuis, B. J., Wikström, H. V., Rodenhuis, N., Sundell, S., Dijkstra, D. A new type of prodrug of catecholamines: an opportunity to improve the treatment of Parkinson's disease. *J. Med. Chem.* 2002, *45*, 2349–2351.
125. Savi, P., Pereillo, J. M., Uzabiaga, M. F., Combalbert, J., Picard, C., Maffrand, J. P., Pascal, M., Herbert, J. M. Identification and biological activity of the active metabolite of clopidogrel. *Thromb. Haemostasis* 2000, *84*, 891–896.
126. Clarke, T. A., Waskell, L. A. The mechanism of clopidogrel is catalyzed by human cytochrome P450 3Aand is inhibited by atorvastatin. *Drug Metab. Dispos.* 2003, *31*, 53–59.
127. Bodor, N., Shek, E., Higuchi, T. Improved delivery through biological membranes. 1. Synthesis and properties of 1-methyl-1,6-dihydropyridine-2-carbaldoxime, a pro-drug of *N*-methylpyridinium-2-carbaldoxime chloride. *J. Med. Chem.* 1976, *19*, 102–107.
128. Shek, E., Higuchi, T., Bodor, N. Improved delivery through biological membranes. 2. Distribution, excretion, and metabolism of *N*-methyl-1,6-dihydropyridine-2-carbaldoxime hydrochloride, a pro-drug of *N*-methylpyridinium-2-carbaldoxime chloride. *J. Med. Chem.* 1976, *19*, 108–112.

129. Shek, E., Higuchi, T., Bodor, N. Improved delivery through biological membranes. 3. Delivery of *N*-methylpyridinium-2-carbaldoxime chloride through the blood-brain-barrier in its dihydropyridine prodrug form. *J. Med. Chem.* 1976, *19*, 113–117.
130. Krenitsky, T. A., Hall, W. W., de Miranda, P., Beauchamp, L. M., Schaeffer, H. J., Whiteman, P. D. Deoxyacyclovir: a xanthine oxidase-acitvated prodrug of acyclovir. *Proc. Natl. Acad. Sci. USA* 1984, *31*, 3209–3213.
131. Whiteman, P. D., Bye, A., Fowle, A. S. E., Jeal, S., Land, G., Posner, J. Tolerance and pharmacokinetics of A515U, an acyclovir analogue in healthy volunteers. *Eur. J. Clin. Pharmacol.* 1984, *27*, 471–475.
132. Gershwin, M. E., Goetzl, E. J., Steinberg, A. D. Cyclophosphamide: use in practice. *Ann. Int. Med.* 1974, *80*, 531–540.
133. Silverman, R. B. *The Organic Chemistry of Drug Design and Drug Action*. Academic Press: San Diego, CA, 1992. p. 380
134. Zon, G. Cyclophosphamide analogues. In *Progress in Medicinal Chemistry* (Ellis, G. P., West, G. B., Eds). Elsevier: Amsterdam, 1982, pp. 205–246.
135. Arnold, H., Bourseaux, F., Brock, N. NeuartigeKrebs-Chemotherapeutika aus der Gruppe der zyklischen *N*-Lost-phosphamidester. *Naturwissenschaften* 1958, *45*, 64–66.
136. Arnold, H., Bourseaux, F., Brock, N. Chemotherapeutic action of a cyclic nitrogen mustard phosphamide ester (B518-ASTA) in experimental tumors of the rat. *Nature* 1958, *181*, 931.
137. Arnold, H. Ueber die Chemie neuer zytostatisch wirksamer *N*-Chloroaethyl-phosphorsäureesterdiamide. In *Proceedings of the Fifth International Congress on Chemotherapy* (Spitzy, K., Haschek, H., Eds). Verlag der Wiener Medizinischen Akademie: Wien, 1967, pp. 751–754.
138. Kennedy, K. A., Teicher, B. A., Rockwell, S., Sartorelli, A. C. The hypoxic tumor cell: a target for selective cancer chemotherapy. *Biochem. Pharmacol.* 1980, *29*, 1–8.
139. Davis, P. J., Guenthner, L. E. Sulindac oxidation/reduction by microbial cultures; microbial models for mammalian metabolism. *Xenobiotica* 1985, *15*, 845–857.
140. Kwon, C.-H., Blanco, D. R., Baturay, N. *p*-(Methylsulfinyl)phenyl nitrogen mustard as a novel bioreductive prodrug selective against hypoxic tumors. *J. Med. Chem.* 1992, *35*, 2137–2139.
141. Cole, C., Reigan, P., Gbaj, A., Edwards, P. N., Douglas, K. T., Stratford, I. J., Freeman, S., Jaffar, M. Potential tumor-selective nitroimidazolylmethyluracil prodrug derivatives: inhibitors of the angiogenic enzyme thymidine phosphorylase. *J. Med. Chem.* 2003, *46*, 207–209.
142. Wallmark, B., Brändström, A., Larsson, H. Evidence for acid-induced transformation of omeprazole into an active inhibitor of ($H^+ + K^+$)-ATPase within the parietal cell. *Biochim. Biophys. Acta* 1984, *778*, 549–558.
143. Lamers, C. B. H. W., Lind, T., Moberg, S., Jansen, J. B. M., OIbe, L. Omeprazole in Zollinger-Ellison Syndrome. *New Engl. J. Med.* 1984, *310*, 758–761.
144. Im, W. I., Shi, J. C., Blakeman, D. P., McGrath, J. P. Omeprazole, a specific inhibitor of gastric (H^+-K^+)-ATPase, is a H^+-activated oxidizing agent of sulfhydryl groups. *J. Biol. Chem.* 1985, *260*, 4591–4597.
145. Lindberg, P., Nordberg, P., Alminger, T., Brändström, A., Wallmark, B. The mechanism of action of the gastric acid secretion inhibitor omeprazole. *J. Med. Chem.* 1986, *33*, 1329–1685.
146. Krausz, F., Demarne, H., Vaillant, J., Brunaud, M., Navarro, J. Anti-inflammatoires non stéroidiques: Dérivés de l'acide phényl-4-butyrique et phényl-4, oxo-4, butyrique. *Arzneim.-Forsch. (Drug Res.)* 1974, *24*(9a), 1360–1364.
147. Gros, P. M., Davi, H. J., Chasseaud, L. F., Hawkins, D. R. Metabolic and pharmacokinetic study of bucloxic acid. *Arzneim.-Forsch. (Drug Res.)* 1974, *24*(9a), 1385–1390.
148. Kohler, C., Tolman, E., Wooding, W., Ellenbogen, L. A review of the effects of fenbufen and a metabolite, biphenylacetic acid, on platelet biochemistry and function. *Arzneim.-Forsch. (Drug Res.),* 1980, *30*, 702–707.
149. Chicarelli, F. S., Eisner, H. J., Van Lear, G. E. Disposition and metabolism of fenbufen in several laboratory animals. *Arzneim.-Forsch. (Drug Res.),* 1980, *30*, 707–715.
150. Martel, R. R., Rochefort, J. G., Klicius, J., Dobson, T. A. Anti-inflammatory properties of furobufen. *Can. J. Physiol. Pharmacol.* 1974, *52*, 669–673.
151. Gilard, J. W., Belanger, P. Metabolic synthesis of arylacetic acid antiinflammatory drugs from arylhexenoic acids. 2. Indomethacin. *J. Med. Chem.* 1987, *30*, 2051–2058.
152. Williamson, J. M., Meister, A. Stimulation of hepatic glutathione formation by administration of L-2-oxothiazolidine-4-carboxylate, a 5-oxo-L-prolinase substrate. *Proc. Natl. Acad. Sci. USA* 1981, *78*(2), 936–939.
153. Matsukawa, T., Yuruki, S., Oka, Y. The synthesis of *S*-acylthiamine derivatives and their stability. *Ann. NY Acad Sci.* 1962, *98*, 430–444.
154. Thomson, A. D., Frank, O., Baker, H., Leevy, C. M. Thiamine propyl disulphide: absorption and utilization. *Ann. Int. Med.* 1971, *74*, 529–534.
155. Bodor, N., Farag, H. H., Brewster, M. E., III. Site-specific, sustained release of drugs to the brain. *Science* 1981, *214*, 1370–1372.
156. Cousse, H., Casadio, S., Mouzin, G. L'hydroxy éthyl nicotinamide vecteur d'acides thérapeutiquement actifs. *Trav. Soc. Pharm. Montpellier* 1978, *38*, 71–76.
157. Casadio, S., Cousse, H., Mouzin, G. Nouvelles formes modulées de médicaments utilisant les *N*-hydroxy alcoyl pyridine carboxamides comme vecteur. *Fr. Pat.N°77.13478 (May 2, 1977, to P. Fabre S.A.)* 1977.
158. Vezin, J. C., Mouzin, G., Cousse, H., Casadio, S. Nicafenine, a new analgesic. *Arzneim.-Forsch.* 1979, *29*, 1659–1661.
159. Gorrod, J. W. Potential hazards of the pro-drug approach. *Chem. Ind.* 1980, 457–461.
160. Roberts, J. C., Nagasawa, H. T., Zera, R. T., Fricke, R. F., Goon, D. J. W. Prodrugs of L-cysteine as protective agents against acetaminophen-induced hepatotoxicity. 2-(Polyhydroxyalkyl)- and 2-(polyacetoxyalkyl)thiazolidine-4(*R*)-carboxylic acids. *J. Med. Chem.* 1987, *30*, 1891–1896.
161. Saari, W. S., Freedman, M. B., Hartman, R. D., King, S. W., Raab, A. W., Randall, W. C., Engelhart, E. L., Hirschmann, R., Rosegay, A., Ludden, C. T., Scriabine, A. Synthesis and antihypertensive activity of some ester progenitors of methyldopa. *J. Med. Chem.* 1978, *21*, 746–753.
162. Vickers, S., Duncan, C. A., White, S. D., Breault, G. O., Boyds, R. B., Shepper, d. P. J., Tempero, K. F. Evaluation of succinimidoethyl and pivaloyloxyethyl esters as progenitors of methyldopa in man, rhesus monkey, dog, and rat. *Drug Metab. Dispos.* 1978, *6*, 640–646.

Part VII

Pharmaceutical and Chemical Means to Solubility and Formulation Problems

Michael J. Bowker

Section Editor

Chapter 37

Preparation of Water-Soluble Compounds Through Salt Formation

Michael J. Bowker and P. Heinrich Stahl

I. INTRODUCTION
II. THE SOLUBILITY OF COMPOUNDS IN WATER
A. The determination and predictionof solubility
B. Ionization of drugs and the importance of pK_a
III. ACIDS AND BASES USED IN SALT FORMATION
IV. EARLY SALT FORMATION STUDIES
A. Choice of salt formers
B. Prediction of the pH of the salt in solution
C. Search for crystalline salts
V. COMPARISON OF DIFFERENT CRYSTALLINE SALTS
A. Melting point
B. Aqueous solubility
C. Common ion and indifferent electrolyte effects
D. Hygroscopicity
E. Solubility in co-solvents (water-miscible solvents)
F. Dissolution rate
G. Particle size and crystal morphology
H. Polymorphism and pseudopolymorphism
I. Chemical stability
J. Other properties
VI. IMPLICATIONS OF SALT SELECTION ON DRUG DOSAGE FORMS
A. Tablet products
B. Hard gelatine capsules
C. Parenteral solutions
D. Oral solutions
E. Suspension formulations
F. MDI products
G. DPI products
H. Soft gelatine capsule formulations
I. Emulsions, creams and ointments
VII. CONCLUSION
REFERENCES

We are continually faced by great opportunities brilliantly disguised as insoluble problems

Lee Iacocca, Chrysler Corporation (1924–)

I. INTRODUCTION

Medicinal chemists in pharmaceutical companies involved in drug discovery normally prepare many compounds each year, however, the majority of these are usually found to be inactive in cell-based screens. When a compound is found to have some interesting activity in a biological screen, it normally forms the basis of a period of intense activity. A dedicated team of chemists then prepares numerous derivatives with similar structures. After screening the set of derivatives and structural variations, one or more compounds are chosen that are considered optimum in terms of biological activity and, occasionally, ease of synthesis. At this early stage, the compound is often relatively impure (typically 90–95% pure), amorphous (or at best poorly crystalline) and is usually offered as the free base or free acid. Occasionally, these compounds are neutral, non-ionizable molecules often with severe aqueous solubility problems; the development of these will not be discussed further in this chapter. Once the compound (or compounds) is accepted as a potential candidate for further development, it is the mandate of the preformulation and process chemistry scientists, in collaboration with the medicinal chemists, to convert this medicinal chemical candidate into a pharmaceutical product.

Copyright © 2008, Elsevier Ltd
All rights reserved.

During the development of the candidate it is necessary to modify its form to improve the physicochemical and physicomechanical characteristics for optimizing the biopharmaceutical and technological properties. If the drug is ionizable, these can often be modified and improved by the formation of salts. A further step in optimization can often be done by the formation of suitable polymorphs or pseudopolymorphs (hydrates or, occasionally, other solvates). Historically, salt formation was primarily used as a means to improve aqueous solubility (hydrophilicity) and crystallinity but, it also has an effect on many other properties, including bioavailability, organoleptic properties (e.g. taste), melting point, dissolution rate, hygroscopicity, stability, processing and powder characteristics. However, the improvement of solubility still remains the major reason for preparation of salts. In rare cases there is even a need to reduce the aqueous solubility in order to enable the preparation of a modified release or suspension formulation. This can be achieved by using a high molecular weight or hydrophobic salt former (also called counterion). A salt should be easily prepared and isolated in high purity, with a consistent particle size distribution and polymorphic form, and have low irritancy and toxicity.

It is important to undertake these studies early in the development process as salt formation modifies the molecular formula and hence represents a chemical entity of its own. If the salt is isolated as a hydrate or solvate, the water or solvent of crystallization also contribute to the molecular formula. The molecular formula needs to be established as early as possible during the preclinical phase, that is, before the start of human Phase I clinical trials, since early safety (toxicology) and pharmacology studies need to be undertaken using the form intended for use in Phase I. If the salt form chosen is eventually found to be inadequate, it is then still possible to change to different salt. However, some expensive and time consuming safety and/or pharmacological studies may require repeating, inevitably delaying the development process.

An alternative process for solubility modification is the formation of a prodrug of the candidate; this involves covalently "grafting" an additional group or polymer onto the original candidate in order to change, or modify, one or more properties (such as lipid solubility or hydrophobicity) to aid drug delivery to specific tissues. These grafted groups are designed for easy removal by enzymatic hydrolysis. As this process does not involve salt formation it will not be discussed further in this chapter.

II. THE SOLUBILITY OF COMPOUNDS IN WATER

A. The determination and prediction of solubility

Being the physicochemical property of a drug substance of utmost relevance, solubility information is requested at the earliest time possible. Intentions to predict or estimate water solubility from a compound's structure or from a minimum of easily accessible data date back decades.

A predictive estimation is available through the use of the General Solubility Equation defined and developed by Yalkowsky *et al.*[1,2] This simple but effective equation for nonelectrolytes was derived using sound thermodynamic principles to establish the semi-empirical correlation:

$$\text{Log}\, S = 0.5 - 0.01(\text{Mp} - 25) - \log K_{\text{ow}} \qquad (37.1)$$

where S is the molar solubility of the solute, K_{ow} is the octanol/water partition coefficient, which can be computed from the structural formula ($C\log P$)[3] and Mp is the melting point (ºC). Based on the assumption that the non-ionized form of an electrolyte is effectively a nonelectrolyte, an extension to the general solubility equation has been proposed.[4,5]

While this approach still requires the melting temperature of the substance in question, that is, requires its existence, the full *in silico* calculation based on the structural formula is now possible. Several recent publications present computational programs for the prediction of solubility. These approaches are being used for the pre-screening of candidates in advance of their synthesis, or for the design of combinatorial libraries. In a review of the different methodologies and the quality of results obtained from the most useful procedures it is concluded that viable procedures now exist for the determination of solubility with less than 1 log unit uncertainty.[6] This uncertainty is judged to be adequate for use in the above processes. In parallel with this, combination of a newly miniaturized shake flask experimental method and a computational approach for solubility prediction has been published.[7]

It may be added here, that, as extensions to these computational predictors of solubility, initial information has been published recently on the search for models for the prediction of drug solubility and permeability,[8] bioavailability[9] and for the determination of the influence of some solid-state properties (melting point, enthalpy of melting and entropy of melting) on the intrinsic solubility of drugs.[10]

In any case, the solubility is determined experimentally as soon as sufficient substance of an active compound is available. Many experimental methods have been described and are in use. The true equilibrium solubility is the data that counts. Solubility studies are best undertaken by adding sufficient drug substance to the solvent until undissolved solid is present. This suspension is then agitated under temperature-controlled conditions until equilibrium is established, which takes normally 12–24 h. The solution is filtered or centrifuged and following appropriate dilutions, the drug content is determined using a suitable analytical technique. It is essential that the final solution pH is recorded since, if the salt contains a strong counterion, the excess of undissolved drug present will influence the pH and the solubility results obtained (refer to Table 37.5).

The quantities of material required for this study may vary from a few to several hundred milligrams depending on the substance solubility. Very useful information is obtained by establishing the pH/solubility profile of the parent drug and salt, and by determining the pH at which parent drug is precipitated from solution.

It is often the case that different results are quoted for the solubility of a compound in water as different experimental procedures are regularly used in industry. In very early screening studies nephelometry or laser nephelometry is normally used to give an approximate value for the solubility as the procedure is non-specific, yields rapid results and requires very little compound.[11] It is often referred to as the kinetic solubility. Only some time later this is followed, when more compound is available, by the determination of the true or equilibrium solubility. As described above, it is defined as the concentration where the solution and undissolved solid are at equilibrium at constant temperature (usually 25ºC). For some compounds the two results are similar, whilst for others the Kinetic Solubility may be significantly higher.[12] The equilibrium solubility as the physicochemically correct result should be the one to be quoted in submissions to regulatory authorities.

B. Ionization of drugs and the importance of pK_a

Many drug substances behave as weak acids or bases and only partially dissociate into ions in aqueous solution, whereas strong acids and bases totally dissociate to their constituent ions. The majority of currently approved drug substances are weak bases, a smaller number are weak acids, while a few are neutral molecules. Occasionally zwitterionic drugs containing both acidic and basic groups (e.g. peptides, peptidomimetics and proteins) are prepared which can usually form salts with either anions or cations. It should be noted that these zwitterionic drugs exhibit their lowest solubility at the pH where there are an equal number of positive and negative charges (i.e. the molecule is effectively neutral). This pH value, called the isoelectric point (pI), can be either calculated or determined by isoelectric focusing.

The acidity or basicity of a drug substance is defined by the dissociation constant K_a which is the equilibrium constant, more conveniently represented by its logarithmic parameter pK_a, reflecting the degree of ionization of a substance at a particular pH and described by the Henderson-Hasselbalch equations (37.2) and (37.3).[13]

For a weak base ($pK_a < 12$):

$$pH = pK_a + \log([B]/[BH^+]) \tag{37.2}$$

For a weak acid ($pK_a > 2$):

$$pH = pK_a + \log([A^-]/[HA]) \tag{37.3}$$

The pK_a value is numerically equal to the pH value at which a molecule is half dissociated and the calculation or determination of the pK_a value(s) of a potential candidate drug should be one of the first physical properties determined. The approximate pK_a of the molecule of interest can be estimated by referring to literature values for substituents and similar molecules.[14,15] A range of experimental techniques such as potentiometry, solubility determinations or ultraviolet–visible spectroscopy may be used for determination of the true pK_a value(s).

Ideally, in order to have the best opportunity to form a stable salt, there should be a difference of at least 3 units in the pK_a values of the parent drug and the proposed counterion. If the pK_a values of the parent drug and counterion are closer than 3 units, the salt may be unstable or may not form at all. Also, if the parent drug has a high molecular weight and is largely hydrophobic with several aromatic rings or extended alkyl chains it may form an unstable salt where the parent molecule precipitates rapidly after dissolution of the salt in aqueous media (including blood). The relative acidities and basicities of some functional groups commonly found in drug structures are provided in Table 37.1.

For compounds that are very weakly basic, the choice of salt former is restricted to strong acids such as hydrochloric ($pK_a = -6.1$), sulfuric ($pK_{a1} = -3.00$, $pK_{a2} = -1.96$) or methanesulfonic ($pK_a = -1.2$) to ensure protonation of the parent molecule. Compounds that are more highly basic offer greater scope for salt formation and will often form salts with weaker acids such as phosphoric ($pK_{a1} = 2.15$, $pK_{a2} = 7.2$, $pK_{a3} = 12.38$), tartaric ($pK_a = 2.93$), acetic ($pK_a = 4.76$) and benzoic ($pK_a = 4.2$). For very weakly acidic compounds, strongly basic cations such as sodium ($pK_a = 14.8$), potassium ($pK_a = 16.0$) or calcium ($pK_a = 12.9$) are required for deprotonation. Compounds that are more acidic may form satisfactory salts with weaker cations such as zinc ($pK_a = 8.96$), choline ($pK_a = 8.9$) and diethanolamine ($pK_a = 9.65$).

III. ACIDS AND BASES USED IN SALT FORMATION

Over many years, the number of different acids and bases available for consideration has gradually expanded to more than 70 acids and more than 25 bases. Most of these can be classed as "small molecules" and many are naturally occurring. For many of them, their use as salt formers commenced before there was a regulatory requirement for significant and well-designed safety studies. With the advent of regulatory authorities it is now mandatory to undertake a battery of safety and pharmacological tests on the drug substance. The majority of these tests have been formalized and agreed for worldwide use under the auspices of the International Conference on Harmonization of Technical

TABLE 37.1 Relative Acid and Base Strength of Some Functional Groups Found in Drug Substances

Functional group	Relative acid/base strength	Structure	Typical pK_a values
Sulphonic acid	Stronger acid	R—SO_3H, Ar—SO_3H	–1.2, –0.7
Carboxylic acid	↑	R—CO_2H, Ar—CO_2H	4.2,–4.7
Imide		R(CO)NH(CO)R^1	8.2
Phenol, –thiol		Ar—OH, Ar—SH, R—SH	10
Sulphonamide		R—SO_2NH_2	10–11
Amide	Weaker acid	R(CO)NHR^1	13–14
N-oxide	Non-ionizable neutral	$R_3N^+ \rightarrow O^-$	
Alcohol		R—OH	
Ethers and Thioether		R—O—R^1, R—S—R^1	
Carbonyl–		R(CO)R^1	
Sulphone/sulphoxide		R—SO_2—R^1, R—SO—R^1	
Pyridine/pyridyl–	Weaker base	C_5H_5N, RC_5H_4N	5.2
Imine		R—CH═NR^1	9.2
Arylamine		Ar$NHAr^1$	9.3
Alkylamine		R—NH_2, RNHR^1	9.8–11
Amidine	↓	R—C(NH_2)NH	12.4
Guanidine		NH_2—C(NH_2)—NH	13.7
Quaternary ammonium	Stronger base	N^+R_4	

Requirements for Registration of Pharmaceuticals for Human Use (ICH).[16] Where the drug substance is a salt, the safety of both the active component and the salt former are obviously being tested in parallel. However, if a new salt former is chosen, perhaps because all others are inadequate, or because the salt former imparts a special property to the drug substance (e.g. salmeterol xinafoate), it may be considered necessary to undertake additional safety and/or pharmacological studies on the salt former alone.

The *Electronic Orange Book*[17] is a register of drug products approved in the United States that is periodically updated by the US Food and Drug Administration (FDA). A similar register is currently being compiled for drug products approved in the EC.[18] These registers, together with the recently published *Handbook of Pharmaceutical Salts: Properties, Selection and Use*[19] can be searched for information on particular salt forming agents and their uses with different drug substances.

The pK_a values of virtually all the known acidic and basic salt forming agents currently used with drugs that are approved worldwide are given in the monograph on pharmaceutical salts.[19,20] The only known example of a counterion missing from these tabulations is *tert*-butylamine (erbumine), which has been used in the recently approved perindopril erbumine (Aceon®, Solvay Pharmaceuticals, Inc.). The monograph is a timely update of the papers by Gould[21] and Berge *et al.*[22] that were used as the "salt selection bible" for more than two decades.

Stahl[20] published an updated review of the earlier compilation by Berge *et al.*[22] of the frequency of use of the 15 most commonly used acid and base salt formers.

TABLE 37.2 Frequency of Use of the 15 Most Commonly Used Anionic Salt Formers

Salt former	Approximate frequency (%)	
	Stahl (1998)	Berge *et al.* (1977)
Hydrochloride/chloride	49	43
Sulfate	6	7.5
Hydrobromide/bromide	5	2
Tartrate	3	3.5
Mesylate	3	2
Maleate	3	3
Citrate	3	3
Phosphate	2.5	3.2
Acetate	2	1.3
Embonate (pamoate)	1.5	1
Hydroiodide/iodide	1	2
Nitrate	1	0.6
Lactate	1	0.8
Methylsulphate	1	0.9
Fumarate	1	0.3

Surprisingly, after more than 20 years and with few exceptions, the frequencies of use of the different salt formers show only minor variation (see Tables 37.2 and 37.3).

TABLE 37.3 Frequency of Use of the 15 Most Commonly Used Cationic Salt Formers

Salt former	Approximate frequency (%)	
	Stahl (1998)	Berge *et al.* (1977)
Sodium	58	62
Calcium	12	10.5
Potassium	5	10.8
Magnesium	4.5	1.3
Meglumine	2.5	2.3
Ammonium	2	<0.3
Aluminium	1.5	0.7
Zinc	1	2.9
Piperazine	1	<0.3
Tromethamine	1	<0.3
Lithium	1	1.6
Choline	0.5	0.3
Diethylamine	0.5	<0.3
4-Phenyl-cyclohexylamine	0.5	<0.3
Benzathine	0.5	0.7

TABLE 37.4 Categories of Common Pharmaceutical Salts

Salt class	Examples
Anions	
Inorganic acids	Hydrochloride, sulphate, nitrate, phosphate
Sulphonic acids	Mesylate, esylate, isethionate, tosylate, napsylate, besylate
Carboxylic acids	Acetate, propionate, maleate, benzoate, salicylate, fumarate
Anionic amino acids	Glutamate, aspartate
Hydroxyacids	Citrate, lactate, succinate, tartrate, glycollate
Fatty acids	Hexanoate, octanoate, decanoate, oleate, stearate
Insoluble salts	Pamoate (embonate), polystyrene sulphonate (resinate)
Cations	
Organic amines	Diethylamine, diethanolamine, ethylenediamine
Metallic salts	Sodium, potassium, calcium, magnesium, zinc
Cationic amino acids	Arginate, lysinate, histidinate
Insoluble salts	Procaine, benzathine

The following sections briefly describe how salt selection and optimization studies normally evolve within major pharmaceutical companies.

IV. EARLY SALT FORMATION STUDIES

A. Choice of salt formers

At an early stage during the preformulation program for the candidate, particularly for those with unfavorable physicochemical and physicomechanical properties as a free acid or free base, it is normal procedure to investigate a suitable range of potential salts. Typically 4–6 different salts are evaluated, if they can be isolated. Often the properties of the free acid (or free base) are hardly satisfactory and it is recommended[23] that it is included in the comparative exercise using a battery of tests (see Section V). If a parenteral dosage form is intended and the properties of the free acid or free base are better than any of the salts, it is often possible to prepare a salt *in situ* in the liquid vehicle.

Once an appropriate amount of data is available from these tests, a decision should then be made on whether to proceed with the salt form with the better overall range of properties or to continue with the parent drug substance. This decision should be made knowing the intended route of administration and dosage form, together with the approximate dose required. Physicochemical properties that are favorable for one dosage form may be unsuitable or irrelevant for other presentations. Since the properties of individual salts may vary considerably and, for regulatory purposes, each salt form is regarded as a separate chemical entity, it is normally possible for companies to only develop one salt form.

Although the choice of salt former is governed largely by the acidity or basicity respectively of the ionizable groups, a number of points regarding the safety of the counterion, the indication and route of administration and potential dosage form must also be considered at an early stage. Primarily, the toxicological and pharmacological implications of the selected salt former must be considered as well as the effects of the parent drug. In the Pharmaceutical Salts monograph, the author has classified the different salt formers into three classes,[20] depending on the available toxicological information. Salt formers can also be subdivided into a number of categories depending on their functionality and purpose. Some of the most commonly used and interesting examples are listed in Table 37.4. The intended dose is also an important factor in some cases, especially where the activity is low and the dose is expected to be high. Here the molecular weight of the salt former also may become an important consideration, especially for solid oral dosage forms, as it reduces the percentage content of the active

drug. In an extreme case, it may result in the unit dose being two capsules or tablets, rather than the desired one, with the likelihood of the patient compliance being reduced.

The vast majority of salt forms are developed to enhance the aqueous solubility of drug substances. In some cases, a salt may be preferred with reduced solubility for use in suspension formulations where solubility as low as possible is optimum to prevent "Ostwald ripening," for taste masking or to prepare a sustained release product. Chlorpromazine is marketed as a tablet and syrup using the hydrochloride salt but the insoluble embonate salt is used in suspension formulations to extend the duration of action and minimize taste problems. Occasionally the selection of a salt with only modest aqueous solubility may be more suitable for use in tablet products prepared by wet granulation. For example, prochlorperazine maleate is used in Stemetil® tablets rather than the highly soluble mesylate salt.

In the case of basic salt forming agents, about one-third of them are organic amines that are normally only suitable for use in parenteral or topical dosage forms as they are oils with a characteristic, unpleasant fishy odor. They are rarely considered for use in solid oral dosage forms because of their adverse organoleptic properties. One further issue with the use of several of the secondary amines was the announcements by the German Regulatory Authorities in 1984 and 1989[19] that they were capable of forming carcinogenic nitrosamines in the digestive tract and their use in pharmaceutical products should be restricted.

B. Prediction of the pH of the salt in solution

If the salt form of the drug substance is required for use as a parental dosage form, or a topical dosage form that may find use on broken skin, it is often useful to be able to predict the pH of an aqueous solution in order to not prepare a salt that cannot be used. The pH range that is acceptable for injectable dosage forms is 4–9, with outer limits of 3–10. Outside this acceptable range there is increasing pain on injection and the distinct possibility of necrosis of the tissue around the injection site. In cases where the drug solubility is high, it is usually possible to buffer the solution to within the acceptable pH range without causing precipitation of the drug as its free base or free acid. However, where the salt is only sparingly or moderately soluble, attempting to buffer the solution often results in precipitation.

If the drug candidate is considered as a weak electrolyte, there are three possible combinations for the salt: weak acid + strong base, weak base + strong acid and weak acid + weak base. It is possible to calculate the expected pH of the salt solutions at various salt concentrations at 25ºC:

$$\text{Weak acid + strong base, pH} = \frac{1}{2}(\text{p}K_{\text{a}_{\text{acid}}} + \text{p}K_{\text{w}} + \log c) \tag{37.4}$$

TABLE 37.5 Solubilities of Salts of an Experimental Antimalarial Drug

Form or salt	Melting point (°C)	Solubility (mg/mL)	pH saturated solution
Free base	215	7–8	–
Hydrochloride	331	15–32	5.8
DL-lactate	172	1,800	3.8
L-Lactate	193	900	–
2-Hydroxyethane sulphonate	250	620	2.4
Methane sulphonate	290	300	5.1
Sulphate	270	20	–

$$\text{Weak base + strong acid, pH} = \frac{1}{2}(\text{p}K_{\text{a}_{\text{base}}} - \log c) \tag{37.5}$$

$$\text{Weak acid + weak base, pH} = \frac{1}{2}(\text{p}K_{\text{a}_{\text{acid}}} + \text{p}K_{\text{a}_{\text{base}}}) \tag{37.6}$$

where, $\text{p}K_{\text{w}}$ (or $-\log K_{\text{w}}$) has the value 14 at 25ºC and c is the salt concentration.

The equations above demonstrate that a strong acid or strong base salt former yields a solution with a more extreme pH, which is dependent on concentration as it is greatly influenced by the stronger component (i.e. the counterion). For this reason, salts prepared using strong acids (e.g. hydrochloric, methanesulfonic) or bases (sodium, potassium) often give solutions that are highly acidic or alkaline. For hydrochloride salts this could result in the formation of an excessively acidic solution, which may be a limitation to the tolerance of parenteral formulations and may even corrode any stainless steel equipment used in its preparation. In contrast, for salts prepared from a weak acid and a weak base such as dextromoramide tartrate ($\text{p}K_{\text{a}}$ values 7.1 and 3.0), the overall pH of 5.1 is an average of the component $\text{p}K_{\text{a}}$ values and is independent of concentration.

Unfortunately, the predicted values for solubility and those experimentally determined do not always correlate well because the effect of pH on solubility is often complicated by other factors.[24] This was clearly illustrated during a study on the solubility of salts of a poorly soluble and weakly basic experimental antimalarial drug.[25] The solubility of the DL-lactate salt was shown to be 200 times greater than the hydrochloride (see Table 37.5) and twice that of the L-lactate, suggesting that the DL-lactate could provide a route to a parenteral formulation.

For inorganic acid salts of the bactericide chlorhexidine, the aqueous solubilities of the dihydrochloride, sulfate, dinitrate and dihydrogenphosphate salts were determined as 0.6, 0.1, 0.03 and 0.03 mg/mL respectively, demonstrating that solubility is greater with increased salt former acidity.[26]

Some of the solubility differences for the antimalarial drug described above arise from a pH effect, however, the melting point of the drug and the size and charge of the anion also play a significant role. This can be best explained by considering that, for a drug to dissolve in a solvent, three basic processes must occur:

- Firstly, it is necessary to break down the bonds forming the crystal. The lattice energy of the crystal is reflected by the melting point, or to be more precise, the heat of fusion (DH_f). The melting point is increased when there are strong interactions with small, highly charged ions such as chloride and sodium or with larger counterions, which form strong hydrogen bonds, such as the hydroxyacids. For this reason sodium, hydrochloride and hydroxyacid salts often have high melting points, but may also be quite soluble due to the counterion hydrophilicity.
- Secondly, it is necessary to generate cavities in the water structure to permit solubilization of the drug. Large hydrophilic, counterion molecules encourage cavity formation and increase solubility.
- Finally, there must be significant interaction between the solvent and solute to favor solubilization. This is related to the relative polarities of the solvent and solute (partition coefficients, dielectric constants and solubility parameters). Solubilization is improved by the formation of hydrogen bonds (e.g. hydroxyacids) or the electrostatic effects that occur with small, highly charged ions (e.g. sodium).

Counterions that are large and hydrophobic, such as napsylate, benzathine and embonate, result in a solubility decrease and may be employed when a reduction in solubility is required.[19]

Therefore, apart from the pH effect, the solubility of different drug salts can be affected by the size, charge and polarity of the counterion and its ability to form hydrogen bonds or to promote complexation of the parent drug molecule.

C. Search for crystalline salts

So far, no definitive decision mechanisms covering the choice of salt formers have been developed, so the final evaluation and selection process remains essentially empirical to a large extent. In addition to the reasons mentioned earlier that in the major pharmaceutical companies the search for crystalline salts occurs very early along the development timeline, there is further reason. Preclinical investigations are necessarily initiated at a time when the candidate drug substance is usually impure (often around 95% pure) and material is at a premium. Here the formation of a crystalline salt often provides a mechanism whereby batches of material can be purified by crystallization. So discovery chemists and also chemical process development scientists are interested in this opportunity to provide material pure enough for valid research results.

Recognition of this problem has led the scientists involved to adopt the use of microplate techniques for the screening of salts. Usually, this involves dissolving approximately 50 mg of sample in a suitable, volatile solvent and adding a fixed volume of this stock solution, containing about 0.5 mg of sample, into each microplate well. Stock solutions of each potential salt former are prepared and a few microliters of each is added in equimolar proportions, or other appropriate stoichiometric ratio, sequentially to each well. Thus, all the wells in line 1 (*x*-direction) will contain the same combination of sample and counterion 1; all the wells in line 2 contain the same combination of sample and counterion 2, etc. Different, potential crystallizing solvents can be investigated methodically in the *y*-direction. Crystallization can be promoted by evaporation of any excess solvent in some wells using a slow stream of dry nitrogen gas. The wells are inspected at regular intervals for the appearance of crystals using an inverted microscope (e.g. Leica, Model DMIRB) or other suitable technique. Further advances in the automation of the salt formation process using robotics and high-throughput procedures have been published.[27–31]

Once the combinations of counterion and solvent(s) are identified, studies at a slightly larger scale (usually 10–50 mg, occasionally up to 500 mg) can be initiated to confirm the suitability and viability of the crystalline salts produced.

Major pharmaceutical companies may find it economical to invest heavily in the automation of salt selection because of the large number of studies they undertake in an average year; this is usually not the case for smaller companies. A small number of specialized companies now offer individually tailored salt selection and optimization studies under contract; companies that undertake these studies include the following:

- Avantium Technologies BV, Amsterdam, The Netherlands
- Evotec OAI, Abingdon, UK
- PharmaDirections, Inc., Cary, NC, USA
- PharMaterials (UK) Ltd., Reading, UK
- Pharmorphix Ltd., Cambridge, UK
- SSCI, Inc., West Lafayette, IN, USA
- Symyx Technologies AG, Basel, Switzerland
- Thorn BioScience LLC, Louisville, KY, USA
- TransForm Pharmaceuticals, Inc., Waltham, MD, USA

V. COMPARISON OF DIFFERENT CRYSTALLINE SALTS

Before the final form of the candidate can be definitively selected, it is necessary to prepare and characterize a range

of potentially suitable salts in parallel with the parent compound. The extent of early preformulation studies including tests on stability, compatibility and formulation principles at this time realistically depends on the amount of compound available for study. The greater the depth of study, the stronger the database of information available for the decision-making process. Results from the full range of tests usually yields enough information to give an informed judgment on both which is the best salt form to take forward and whether the intended type of formulation will be feasible. Table 37.6 gives an outline of the critical tests that can be performed using small quantities of material; these tests are discussed in greater detail below.

After the preformulation studies have been performed and appropriate pharmacokinetic parameters considered, the most suitable salt form can be proposed and decided upon. Subsequently, when more material is available, a more comprehensive range of preformulation studies can be carried out on the selected salt to assist in formulation development.

TABLE 37.6 Typical Range of Tests for Comparison of Salt Forms and Parent Compound for Oral and Parenteral Dosage Forms

Test	Suitable techniques	Comments
Melting point	Capillary mpt, hot stage microscopy, differential scanning calorimetry	Perform on each salt and compare to parent
pH of solution		Examine pH of saturated solution if quantities permit
Aqueous solubility	Overnight equilibration at 25°C, analysis by HPLC	Perform on each salt and compare to parent
pH/solubility profile	Overnight equilibration at 25°C, analysis by HPLC	Determine at pH 2, 5, 7.5 and 9
Common ion effect on solubility	Overnight equilibration at 25°C in suitable media and analysis by UV–VIS or HPLC	Compare solubility in demineralized water with 1.2%w/v saline for salts and parent
Hygroscopicity	Expose samples to various RH values and measure weight gain after 1 week	Perform at 93 or 97%RH, 53% RH, and other values of interest. Assign hygroscopicity classification to each salt
Solubility in co-solvents	Overnight equilibration at 25°C, analysis by HPLC	Examine solubilities in ethanol, polyethylene glycol, propylene glycol and glycerol and compare to parent
Intrinsic dissolution rate	IDR apparatus	Compare dissolution rates between salts. Can provide data on wettability
Crystal morphology	SEM or optical microscopy	Compare crystal habits and levels of agglomeration
Particle size	SEM or laser diffraction	Examine particle size distributions
Polymorphism and solvates	Recrystallizations, HSM, DSC, TGA	Preliminary exploration
Powder properties	Bulk density measurement	Determine Carr's index
Stability	Various	Perform preliminary tests on parent drug substance and on appropriate salts

A. Melting point

If the parent drug has a particularly low melting point, it may be necessary to attempt preparation of a higher melting salt. Melting points below 80°C–100°C can prove problematical if the drug substance is to be subjected to mechanical stress or heat. Typical pharmaceutical processes, where heat and pressure are applied, include milling operations, drying of granules and their subsequent compression into tablets or the preparation of cream or ointment formulations where melting of the ointment base is required. If the candidate is a hydrate or solvate, the temperature at which dehydration or desolvation occurs is an important consideration; it should not be below 80°C–100°C as this temperature could be attained in some pharmaceutical processes. Unless rehydration subsequently occurs readily, this would lead to the formulation containing a different drug substance form than is intended for approval.

Increases in melting point are due to an increase in crystallinity and rigidity from hydrogen bonding (e.g. with the hydroxyacids), strong electrostatic interactions (presence of small highly charged counterions such as sodium) or steric effects. Zwitterionic compounds often have high melting points due to electrostatic or hydrogen bonding effects in the crystalline state and form an "internal salt." The influence of melting point on aqueous solubility is described in the next section.

B. Aqueous solubility

As mentioned in Section II.A., modification of drug solubility is generally the main reason for the preparation of salts. When the modification results in an increase in solubility,

there is usually a concomitant improvement in the physical and mechanical properties.

It has been proposed that, when the solubility of the parent compound is less than 1 mg/mL in its unionized form (intrinsic aqueous solubility), *in vivo* absorption may be poor.[32] Where the intrinsic solubility is between 1 mg/mL and 10 mg/mL, salt formation may sometimes be desirable to ensure good absorption. Salts are rarely required when the aqueous solubility is 10 mg/mL or greater, unless other physical properties are poor.

Much of the benefit in solubility enhancement from salt formation is attributable to the change in solution pH caused by the presence of the counterion. This occurs because the ionization and solubility of acidic drugs (such as barbiturates and non-steroidal anti-inflammatory drugs) increases in basic conditions but decreases in acidic conditions. This behavior is exemplified by derivations of the Henderson-Hasselbalch equations (37.2) and (37.3). The opposite situation occurs for basic drugs such as chlorpromazine, morphine and codeine, which are more soluble in acidic conditions.

For a weak base:

$$\% \text{ Ionization} = 100/(1 + \text{antilog } (\text{pH} - \text{p}K_a) \quad (37.7)$$

$$C_S = C_o[1 + \text{antilog } (\text{p}K_a - \text{pH})] \quad (37.8)$$

For a weak acid:

$$\% \text{ Ionization} = 100/(1 + \text{antilog } (\text{p}K_a - \text{pH})) \quad (37.9)$$

$$C_S = C_o[1 + \text{antilog } (\text{pH} - \text{p}K_a)] \quad (37.10)$$

where, C_S = solubility at the selected pH and C_0 = solubility of unionized drug (intrinsic solubility)

Concerning the different salt categories given in Table 37.2, some generalizations can be made regarding drug solubility:

- For inorganic acid salt formers (the most frequently used anions), the degree of solubility enhancement is strongly influenced by the counterion charge and size as well as the pH effect. The inorganic acid salts of higher molecular weight bases are often prone to salting out and precipitation of the free base on dilution. Hydrochlorides often form an anhydrate plus one or more hydrate salts; hydrates are invariably less soluble than anhydrates.
- For sulfonic and carboxylic acid counterions, solubility is also affected by the size and polarity of the counterion. In the sulfonic acid series, large counterions such as besylate and napsylate have lower solubilities than the smaller mesylate, esylate and isethionate (2-hydroxyethanesulfonate) ions. Addition of a hydroxyl group to sulfonic acids, as in isethionate, further improves solubility.
- Hydroxyacids such as citrate and tartrate may increase or decrease solubility and melting point depending on the parent drug structure. Generally, hydroxyacids give high melting points but high solubilities as a consequence of their increased number of hydroxyl groups and potential for hydrogen bonding. However, the parent drug must have a sufficient basicity to form the salt and there is less of a pH solubilizing effect compared to other acids.
- Fatty acids are often used as counterions to produce drugs with low melting points and to increase the hydrophobicity of the drug substance, that is more suitable for use in oil-based formulations (e.g. topical preparations or intramuscular injections). Oral applications of these salts include the formulation of soft gelatin capsules, particularly when the capsule vehicle is semi-solid or an oil.
- Insoluble salts may be prepared using high molecular weight counterions to reduce the solubility of the drug for formulation of a suspension, to improve chemical stability, provide improvements in solid-state stability, reduce acid lability and permit formulation of tasteless or controlled release products. Erythromycin stearate, which is poorly soluble in acidic media (an "enteric" salt), is less prone to decomposition in gastric fluids than the more acid-soluble free base.
- Metallic ions are the most commonly used counterion for weak acids since they provide salts with high melting points but good solubility due to their high counterion hydrophilicity. Alkali metal salts especially provide a strong pH solubilizing effect due to their high charge to size ratio. Sodium salts are often hygroscopic and occasionally form a gel on standing in solution. This gelling occurs frequently with solutions of fatty acid sodium salts.
- Organic amines may be used as salt formers for acidic drugs but generally yield low melting salts. Some amines can form carcinogenic nitrosamines in the gastrointestinal tract and are not recommended for use in oral products. Also, some salts prepared from low molecular weight amines retain the "fishy" smell of the parent. The use of most amines as salt formers is thus largely restricted to the formation of salts intended for use in parenteral or topical products.
- Insoluble counterions such as benzathine and procaine may be used for some acidic drugs such as penicillins where their main function is to improve stability but they also provide a longer period of action due to their reduced solubility.
- Amino acids, although not very commonly used, can provide an excellent choice of salt for weakly acidic drugs (using lysine, arginine and histidine) and basic drugs (using aspartic acid and glutamic acid). There is often a good solubility enhancement due to the hydrophilic nature of the amino acids.

Solubility studies are best undertaken by adding drug substance to the solvent until undissolved solid is present. This suspension is then agitated under temperature-controlled conditions overnight or for 24 h until equilibrated. The solution

is filtered or centrifuged and following appropriate dilutions, the drug content is determined using a suitable analytical technique. It is essential that the final solution pH is recorded since, if the salt contains a strong counterion, the excess of undissolved drug present will influence the pH and the solubility results obtained (refer to Table 37.5). The quantities of material required for this study may vary from a few to several hundred milligrams depending on the drug substance solubility. It is also useful to examine the pH/solubility profile of the parent drug and salt and determine the pH at which parent drug is precipitated from solution.

C. Common ion and indifferent electrolyte effects

Occasionally salts may demonstrate lower aqueous solubilities and dissolution rates than expected due to the common ion effect. For example, the dissolution rate of doxycycline hydrochloride dihydrate salt was determined as four times greater in water than in 0.1 M HCl solution.[33] Changing the dissolution medium from 0.1 M HCl to 0.1 M methanesulfonic acid resulted in an increase in rate, clearly demonstrating specificity for chloride ions. This effect is due to the presence of chloride ions in the dissolution medium, which disturb the position of equilibrium for the drug solubilization process:

$$BH^{+}Cl^{-}_{(solid)} \leftrightarrow BH^{+}_{(aq)} + Cl^{-}_{(aq)}$$

The common ion effect, which is not limited to chloride ions, can result in a reduction of the solubility in gastric juice and may reduce drug absorption. It is possible to determine whether the effect is likely to occur by comparing the intrinsic dissolution rate (IDR) or solubility of the drug in distilled water and in sodium chloride solution, 1.2% w/v.

Conversely, the addition of some other ions can promote solubility by the indifferent electrolyte effect. The use of hydrophilic molecules such as the hydroxyacids (e.g. citrate, tartaric) or aromatic carboxylic acids (e.g. benzoic) can create cavities in the water structure thereby promoting solubilization. Many salt formers increase drug solubility by this type of mechanism. Citrate buffers and sodium benzoate, the latter often used in formulations as an antimicrobial preservative, are known to enhance the solubility of a number of drugs.

D. Hygroscopicity

A small percentage of salts of drug substances are hygroscopic and take up atmospheric moisture; occasionally the uptake is so large that the salt dissolves and is classed as deliquescent. The degree of uptake depends on the hygroscopicity of the drug and the relative humidity (RH) of the environment. Hygroscopicity often occurs when the aqueous solubility is high; salts prepared using strongly acidic or basic counterions (e.g. hydrochloride, mesylate and sodium) and hydroxyacid salts (e.g. citrate and tartrate) often demonstrate a high degree of moisture uptake.

Hygroscopic drug substances should be avoided, where possible, as processing and storage of bulk drug and formulations under low or controlled, humidity conditions may be required. Expensive moisture-proof packaging may be necessary for protection the finished product. Many companies have established decision rules covering hygroscopicity, especially when oral tablets are required as highly hygroscopic drugs often give tablet weight fluctuations that lead to tablet and/or film-coat cracking.

Typically, the decision rules may be of the following form:

1. Candidates with a water uptake (hygroscopicity value) less than 2% w/w after storage at 90% RH for 24 h should present no difficulties for oral solid dose formulation.
2. Candidates with a water uptake (hygroscopicity value) more than 2% w/w but less than 7% w/w may present unpredictable difficulties for oral solid dose formulation. Formulations should be closely monitored for problems.
3. Candidates with a water uptake (hygroscopicity value) more than 7% w/w should not be considered for oral solid dose formulation.

For weakly acidic drugs, producing a form with reduced solubility, such as a calcium salt[34] rather than sodium may be beneficial and yield a non-hygroscopic drug form. Atorvastatin is formulated as the trihydrate hemi-calcium salt in Lipitor® tablets, rather than as the hygroscopic sodium salt, for this reason. Avoidance of strongly acidic counterions for potentially hygroscopic drugs is often desirable since these may be hygroscopic or even deliquescent. Where the parent drug has a very low degree of hygroscopicity, it is unlikely that the hygroscopicity of the majority of its salts will be problematical.

In order to compare the moisture uptake of salts, small samples of powder can be placed in hygrostats and exposed to the atmosphere above a range of saturated inorganic salt solutions and the weight gain determined after a suitable period (usually 1 week).[35] Alternatively, moisture uptake information can be obtained using an automated microbalance apparatus (Dynamic Vapor Sorption) in which scans of water uptake versus humidity are possible using smaller quantities of material. Comparison of the uptake of moisture at several RH values will help identify suitably stable salts, especially when the humidity is ramped up and down as a hysteresis cycle. The use of a Powder X-ray apparatus with an environmental chamber often gives even more useful information as it enables the identification of polymorphism and pseudopolymorphism changes induced by increasing, or decreasing, relative humidities.

FIGURE 37.1 Diagrammatic representation of the drug dissolution and absorption process.

E. Solubility in co-solvents (water-miscible solvents)

Information on the solubility of salts in several co-solvents may be useful, particularly when the candidate is poorly soluble and it is necessary to administer intravenous formulations during the preclinical development program. Although co-solvent solubility information may be less important when it is intended to develop oral formulations, co-solvents are occasionally used in oral liquid preparations and soft gelatine capsules. The degree of solubilization achieved by co-solvents will depend on the comparative polarity of the drug and solvent and in some cases the parent drug may exhibit a substantially greater co-solvent solubility than any salts.

The solubility of drugs in co-solvents can be determined using a similar procedure to that described previously for aqueous media. Suitable commonly used pharmaceutical co-solvents include, ethanol, propylene glycol, glycerol and polyethylene glycol. Other solvents such as fixed oils and esters may find occasional use in soft gelatine capsules or intramuscular injections for incorporating hydrophobic or oily drugs and salts. Solubilization by co-solvents is covered in greater detail in chapter 39.

F. Dissolution rate

The dissolution rate of a drug substance is related to its aqueous solubility. For neutral drugs there is a direct correlation between the solubility and dissolution rate as demonstrated by the Noyes and Whitney equation (37.11). However, depending on the pH of the dissolution medium salts demonstrate a higher dissolution rate than may be expected from the solubility of the parent drug or salt as a desired consequence of salt formation.

During the dissolution of the drug particle a boundary layer of highly concentrated dissolved drug ions and counterions forms around the particle.[36] This layer has a pH close to that of a saturated solution of the salt but quite different to that of the bulk medium (see Figure 37.1). For instance, with the sodium salt of a weak acid, the boundary layer has an alkaline pH and dissolution of the ion of weak acid from the drug particle is accelerated. The dissolution process proceeds independently of the bulk medium pH and under the control of the microenvironment around the particle. Similarly, use of a strongly acidic salt former accelerates the dissolution of a weakly basic drug. This increase in dissolution rate often leads to enhanced bioavailability for poorly soluble drugs, where solubility is the rate-limiting step for drug absorption.[37] Examples that demonstrate these differences in dissolution rate are reproduced in Table 37.7.

Combining the Noyes–Whitney equation (37.11) for the dissolution of particles[38] and the derivations of the Henderson-Hasselbalch equations (37.2) and (37.3), demonstrates that this dissolution enhancing effect is very pronounced and increases exponentially with pH (Equation (37.12)).

$$\frac{dc}{dt} = KC_s \tag{37.11}$$

For a weak base:

$$\frac{dc}{dt} = KC_0[1 + \text{antilog}(pK_a - pH)] \tag{37.12}$$

TABLE 37.7 Dissolution Rates of Weak Acids and their Sodium Salts

Drug or salt form	pK_a	pH at C_S	Intrinsic dissolution rate mg/min/cm^2	
			0.1 M HCl pH 1.5	Phosphate buffer pH 7.5
Salicylic acid	3.0	2.40	1.7	27
Sodium salicylate		8.78	1870	2500
Sulphathiazole	7.3	4.97	<0.1550	0.5
Sodium sulphathiazole		10.75		810

Source: Data from Ref. [37]. C_S = saturation solubility.

where dc/dt = dissolution rate under sink conditions, C_s = concentration of the saturated solution, $K = DS/h$, D is a diffusion coefficient, S is the surface area of drug particles and h is the diffusion layer thickness.

During preformulation studies on candidate drug substances, it is useful to compare the dissolution rate of various salt forms in media of physiological pH (e.g. solutions between pH 2 and pH 7.4). However, because the different samples of salts under investigation are unlikely to have the same particle size distribution and surface area, meaningful comparisons cannot be made using normal dissolution testing procedures. A more satisfactory method of comparing the dissolution rate of salts is to measure the IDR.[39]

The improved apparatus, devised by Wood, Syarto and Letterman, is used to form a flat pellet in a die. The die and pellet is then submersed in a vessel of dissolution medium and rotated at 50 or 100 rpm. Since the surface area of the pellet is virtually identical between samples and is dependent only on the diameter of the die, the dissolution rate is no longer affected by particle size. The dissolution medium is sampled at intervals and the concentration of drug determined by HPLC then plotted against time. The initial, linear portion of the graph is used to calculate the IDR in mg/min/cm^2. Kaplan proposed that drugs with an IDR below 0.1 mg/min/cm^2 could exhibit poor bioavailability,[32] while the bioavailability of drugs with a rate greater than 1 mg/min/cm^2 should be high since absorption would not be limited by dissolution rate.

The measurement of IDR forms an important part of preformulation studies as they permit direct solubility and dissolution rate comparisons between different drug candidates, salts, polymorphic and pseudopolymorphic forms. The measurements are relatively straightforward to carry out but require a minimum of about 200 mg of sample per determination. Production of the pellet can occasionally result in a pressure-induced polymorphism change; the new polymorph may have different dissolution properties. Therefore, the presence or absence of a new polymorph should be verified by IR spectroscopy or by Differential Scanning Calorimetry (DSC).

G. Particle size and crystal morphology

Particle size and shape are important attributes of drug substances that can affect the dissolution rate, bioavailability, and the processing of oral solid dosage forms and the physical stability of semi-solid dosage forms. During manufacture of the bulk drug substance, the recrystallization conditions required to prepare a particular salt will inevitably influence the particle size and crystal habit; conditions should be established early in the development program that routinely yield material with consistent characteristics.

Particle size analysis should be carried out on all batches of salt candidates to establish suitable recrystallization procedures. A rapid assessment of both the particle size and crystal habit can be carried out by Scanning Electron Microscopy (SEM). Laser diffraction techniques can provide a rapid assessment of the particle size distribution using less than 10 mg of material. It is also useful to retain a photomicrograph for each batch.

If it was shown, for example, that a particular salt had an unacceptably large particle size or the material could only be prepared in a highly aggregated form and that neither could be effectively milled or size reduced, it may be necessary to consider alternative salts.

H. Polymorphism and pseudopolymorphism

A large proportion of drug substances, whether neutral molecules, free acids, free bases or salts, are capable of exhibiting polymorphism or pseudopolymorphism (hydrate or solvate formation). It has been reported that 70% of barbiturates, 60% of sulfonamides and 23% of steroids exhibit polymorphism.[40] Polymorphism often influences a range of physicochemical properties such as solubility, dissolution rate, stability and powder properties as well as bioavailability. Usually, it is possible to determine the most stable polymorph and discover recrystallization solvents that uniquely produce this form and improve the physicochemical and physicomechanical properties and chemical stability of the drug.

A brief investigation of the polymorphic form of each candidate salt should be undertaken and as far as possible, the existence of polymorphism should be confirmed or ruled out. Polymorphs can be produced by simple recrystallization from a range of solvents, or solvent mixtures, of different polarities and dielectric constants and their existence most easily confirmed by a combination of hot stage microscopy (HSM), DSC or infrared or Raman spectroscopy and by powder X-ray techniques.[41]

The formation of solvates and hydrates (pseudopolymorphs) and the potential for conversion between different forms may be an area of concern for some drugs. Hydrochloride salts are prone to the formation of hydrates, often exhibiting high degrees of hygroscopicity and changes in hydrate stoichiometry on storage or processing. Drugs that exhibit these properties should be avoided, where possible, and a stable hydrate or anhydrate selected. Hydrates are invariably less soluble than their anhydrous counterparts. The behavior of solvates and hydrates may be investigated by a combination of a number of experimental techniques including powder X-ray techniques, HSM, DSC and thermogravimetric analysis (TGA).

For the very first clinical studies it is often possible to find a polymorph with sufficient physical stability to permit completion of Phase I trials. The most stable polymorph usually exhibits the highest density, the lowest solubility (and often the lowest bioavailability). Very detailed evaluations of polymorphism and the inter-relationships between the different polymorphs and pseudopolymorphs usually need to be delayed until the candidate salt has been chosen and larger quantities of drug from several batches are available. These studies are often started around the initiation of Phase II clinical trials.

I. Chemical stability

An early determination of the stability profile of a candidate drug substance is important since it can influence the choice of dosage form, the manufacturing process, the choice of packaging materials and, importantly, the efficacy and safety of the product. A preliminary assessment of the stability of drug substances in both the solid and solution state at an early stage is always extremely valuable. Because of time constraints and availability of material, it may not be possible to complete all tests quickly enough on a range of salts. It may be necessary to only carry out accelerated stress tests initially and make predictions based on the stability of the parent compound (if known) and the physical characteristics of the salt such as hygroscopicity and solution pH. Here it is not intended to provide a detailed discussion of stability except the key issues relating to the stability of salts concerning hydrolysis and solid-state compatibility.

In solution, the major route of drug decomposition is usually by hydrolysis. Since this is normally a pH dependent mechanism, a suitable method of improving stability is to control the pH of the solution. As described previously, the pH of a drug salt in an unbuffered solution is affected by the solubility and the pK_a of the parent drug and counterions. Use of strongly acidic or basic counterions (such as chloride, mesylate or potassium and sodium) may promote hydrolytic degradation. A second important degradation route is oxidation, either directly by peroxides or indirectly, resulting from the formation of peroxides by the action of light on dissolved oxygen.

Alternatively, the use of counterions such as organic carboxylic acids and hydroxyacids, which produce solutions with a more neutral pH, may be successful for preventing or reducing acid or base catalyzed hydrolysis. To fully understand degradation in solution it is necessary to produce a pH-stability profile for the parent drug or its salts.

For liquid formulations, stability may be improved by the preparation of a suspension containing a sparingly soluble salt, thereby reducing the amount of drug in solution and the degree of degradation. For example, penicillin G solutions are not stable for more than 2 weeks, even when refrigerated, but the use of the less soluble, but chemically stable, procaine or benzathine salts allows preparation of ready to use suspensions with much longer shelf lives.

Hydrolytic degradation is not confined to liquid formulations; solids that are hygroscopic may be susceptible to degradation in the solid state. Where drugs are hygroscopic and surface moisture films that form around particles are particularly acidic or alkaline due to the presence of strong counterions, degradation may be surprisingly rapid. The moisture may be gained from the air, from excipients (such as dicalcium phosphate dihydrate) that have loosely bound water on their surface or from the gelatine in capsule shells (which contain 12–16% w/w loosely bound water). Similarly, drugs that have very poor hydrolytic stability may be prone to decomposition during tablet granulation or even in the gastrointestinal tract following administration.

J. Other properties

Some other characteristics such as organoleptic and powder properties may influence the acceptability of a particular salt. Determining the taste of early stage candidates, where little toxicological data is available, would be highly undesirable and therefore it is not generally possible to use this as a pivotal criterion for salt selection. Although taste represents no significant problem for tablet and capsule formulations, syrups and chewable tablets may be rendered unpalatable by an unpleasant tasting drug. An improvement in the taste profile can be most readily achieved by the preparation of a less soluble or insoluble salt (e.g. embonate or napsylate) but the pK_a and acidity (and hence the tendency for the salt to dissociate) of the counterion will also affect the taste profile. Other approaches to taste

masking have included forming salts with ion exchange resins or the preparation of water-soluble pleasant tasting salts using cyclamate, aspartamate or saccharinate or film coating the tablet.

Powder properties may be important for the processing characteristics of solid dosage forms and if sufficient material is available, the bulk powder density and flow properties of salt candidates should be determined. However, it should be noted that powder properties could change quite significantly by the time the recrystallization process for the bulk drug is optimized.

In recent years, the concept of "developability" has been applied to these early studies. The scientists undertaking the studies described above and assessing the results obtained can often reach a conclusion on whether the candidate can be successfully developed pharmaceutically.[42,43]

VI. IMPLICATIONS OF SALT SELECTION ON DRUG DOSAGE FORMS

It is important to select drug candidates rapidly and to propose a suitable salt for development very early in the drug discovery program. This decision may need to be made using small batches of drug substance and well in advance of major formulation studies. It may be some time before the wisdom of the salt selection process is confirmed. For this reason, it is important to carry out preformulation studies as comprehensive as possible on the salt candidates and to clearly understand the impact of drug properties on dosage form design from the preclinical stages through to the marketplace.

Since the ideal attributes of drug substances differ markedly for different dosage forms, properties appropriate to one dosage form may be inappropriate for another. For this reason, the drug substance attributes should suit the primary mode of delivery (i.e. oral, parenteral, inhalation, or other routes, if appropriate) and the dose (or dose range) required. If other dosage forms are developed, such as "line extensions," at a future date, it may be necessary to counteract any adverse properties of the selected salt using suitable formulation techniques.

The majority of pharmaceutical marketed products are tablets, capsules or injections where moderate to high aqueous solubility, dissolution properties and stability of the drug substance are necessary. Even if the final dosage form intended for marketing is a tablet or capsule, parenteral solutions, particularly intravenous solutions, are inevitably required in order to undertake a range of preclinical studies including the determination of absolute bioavailability. A reasonably high degree of solubility is thus a distinct advantage for the drug substance. However, for some dosage forms, such as liquid filled capsules, suspensions, inhalation products, controlled release forms or creams and ointments, a high aqueous solubility may not be as critical for obtaining good drug delivery and could even be detrimental.

A. Tablet products

Low hygroscopicity and good powder flow is required for the efficient production of tablet products. For direct compression tablets, which are prepared by compression of a powder blend (rather than granules), good powder flow is essential. Normally careful attention is paid to the crystal habit of drugs used in these types of products, and ideally the drug substance should have a moderately large particle size (median 70–140 mm), narrow size distribution and a relatively isometric crystal habit. Poor flow properties severely limit the low concentration end of drug amount that can be incorporated into direct compression tablet products, otherwise content uniformity problems will arise.

Where tablets are prepared by a wet granulation process the drug and most of the excipients are converted to a damp paste, passed through a sieve, then the resulting granules are dried. Adverse drug properties can often be masked by judicious choice of the excipients and it is often possible to incorporate quite high levels of relatively poorly flowing drugs. However, in these granulated formulations, the solubility of the drug substance in water often has a profound effect on granule quality. Generally, moderately soluble salts rather than highly soluble salts are most suitable for tablets prepared by wet granulation. Exceptionally, solvents such as ethanol may be substituted for water during granulation but the process requires fire and explosion proofing measures, adding to the cost of production.

Although several excipients with different functions are always added to the powder mixtures for tablet products to facilitate processing and yield good tablet properties, there are some drugs that, at the doses required, cannot be compressed into a satisfactory tablet. In these instances, selection of an alternative salt may be a solution; or it may be necessary to investigate whether a capsule formulation can be produced which is more expensive in production. Since the granules are exposed to compression forces during tablet manufacture, the compression properties of the drug substance are important. Drugs that are hygroscopic often stick to tooling on compression. The deformation (plastic, brittle or elastic) of the drug substance under a compression load will vary for different salts and may have a significant effect on tablet quality.

As already pointed out for drug substances with low activity and correspondingly high unit doses, a further factor to be considered is the molecular weight of the counterion as a large counterion effectively reduces the drug proportion per mass unit.

Other factors to consider are the powder properties of the material. Poor flow can be attributed to a small particle size, wide particle size distribution, typically by needle-shaped crystal habit, and with materials with a high propensity of flow-induced electrostatic charge build-up causing lumping and demixing. Good powder flow is generally required to ensure satisfactory processing properties.

To a certain extent, powder flow can be improved by good formulation and the addition of glidants (flow-aids).

B. Hard gelatine capsules

Simple capsule formulations are frequently prepared for early clinical studies, prior to development of tablet formulations. However, in a significant proportion of cases, the marketed product may be required as a capsule, when for one reason or another, a tablet product is shown to be unsuitable. As long as the drug has an adequate solubility profile to satisfy pharmacopoeial dissolution requirements, salt solubility is probably less critical for capsules than for many other products. One property of critical importance however, is hygroscopicity. Since capsule shells and many excipients contain loosely bound water, uptake of this water by hygroscopic drug salts can occur. This could result in instability of the product, changes in hydration state and possibly changes in the dissolution rate with potential changes to the drug delivery. For solid dosage forms, it is essential that a salt with a low degree of hygroscopicity is selected.

Most of the drug powder properties considered for tablets in the previous section also apply to capsule fill mixtures, especially where flow properties are concerned. Limitations for a poorly flowing drug/powder mixture are seen with respect to the amount of mass that can be accommodated and the constancy of fill weight for each capsule.

C. Parenteral solutions

For parenteral products it is necessary to ensure that the salt former is physiologically compatible and has a satisfactory toxicological profile. Salt formers that are suitable for use in oral products may not be satisfactory for parenteral use and *vice versa*.

The most important requirement is that the salt possesses sufficient solubility at physiologically compatible pH values to permit incorporation into the dosage form. Buffering the solution to an appropriate pH can often enhance solubility. Salts may also be prepared *in situ* in the formulation. This is particularly useful when the main route of administration utilizes the parent drug form. Where the aqueous solubility of the salt is not sufficiently high, co-solvents may need to be added to enhance solubility (e.g. propylene glycol is used as the vehicle in phenobarbitone sodium injection). Parenteral solutions based on co-solvent vehicles normally cannot be directly injected intravenously because there is the risk of precipitation at the injection site. Therefore, such products are diluted with isotonic saline or 5%w/v dextrose solution to produce a lower concentration that remains soluble and can be safely administered by infusion. Alternative delivery routes are by subcutaneous or intramuscular administration by which, in addition to aqueous solutions, suspensions or oily solutions can also be delivered.

Another important factor that should be considered is the pH of the dissolved salt in aqueous media and its likely tolerance. Wherever possible, parenteral solutions should have a pH range between 4 and 9 (certainly within pH 3 and 10.0). On some occasions it may not be possible to buffer a highly acidic or alkaline salt solution to a suitable pH because of precipitation of parent drug. The chemical stability profile of the drug at the chosen pH value must be considered.

Finally, it may be necessary to ensure compatibility with the different tonicity adjusters used in formulations and infusion fluids. Where precipitation is likely to occur with a hydrochloride salt, selection of another drug salt or use of an alternative tonicity adjuster (e.g. glucose) must be considered.

D. Oral solutions

Many of the criteria described for parenteral solutions should be considered for other solution formulations; however, the pH values of the solutions are not as critical as for the parenteral route. If it is not possible to achieve the required solubility, oral liquids may simply be formulated as a suspension, though solution forms are preferable to ensure good product homogeneity and patient compliance.

Another factor to consider for oral solution products is the taste of the drug. Where drugs have an unpleasant taste, using a less soluble salt in a suspension form, or using flavors to disguise the taste or even using an aspartamate or saccharin salt may be advisable.

E. Suspension formulations

An essential property required to obtain a physically stable suspension formulation is that the drug substance has a low solubility in the suspension vehicle. This applies to all types of suspension whether intended for oral, parenteral or inhalation delivery. Since most suspensions are aqueous, a low solubility in water is required to prevent drug dissolution and crystal growth on storage. Some drugs intended for intramuscular use are formulated as suspensions in oil (e.g. penicillin injections are formulated in sesame oil) to improve chemical stability via a reduction in solubility.

It is essential to prepare a stable polymorphic form of the selected salt, since transitions from metastable to more stable, and often less soluble and less bioavailable forms, can occur over time in suspensions. Salts with complex polymorphic profiles may be unsuitable for suspension formulations.

In heterogeneous systems, the characteristics of the drug particles become important. To reduce sedimentation rates and to prevent caking of the sediment, it is necessary to ensure that the drug has an appropriate particle size

distribution and crystal habit. Occasionally, it is also possible to closely match the densities of the drug particles and the vehicle. Although a reduction in particle size can be achieved by milling, it is more usual and cost effective to crystallize the drug substance with appropriate properties at the outset. The particle size distribution and crystal morphology that can be achieved with particular salts will then become key factors.

As described previously, unpleasant tasting drugs for administration as oral liquids are often presented as suspensions of insoluble salts to improve patient acceptability. An example of this is the bitter tasting analgesic drug, propoxyphene, which is presented as the hydrochloride salt in tablets and as the napsylate salt in oral suspension formulations.

F. MDI products

MDIs are presented as either solution or suspension formulations in hydrofluoroalkane (HFA) propellants; previously chlorofluorocarbon (CFC) propellants were used until their use was proscribed. HFA propellants are essentially lipophilic. Hydrophilic drugs with good aqueous solubility usually demonstrate poor solubility in propellant systems unless a co-solvent such as ethanol is added. To formulate a satisfactory solution MDI, it is essential that the drug is sufficiently soluble in the propellant or propellant/ethanol mixture to form a solution. Conversely, to formulate a suspension MDI, the drug must have negligible solubility in the propellant. Since hydrophobic drugs are more soluble in propellants, it could be advantageous to select the parent drug rather than a salt if a solution MDI is required.

In suspension MDIs, it is essential that the solubility of the drug in the propellant is negligible to prevent crystal growth as this can affect the physical stability of the formulation, the aerodynamic properties of the drug particles and the efficiency of drug delivery. For suspensions, use of a more hydrophilic salt, which is insoluble in the propellant, could be advantageous. Examples of relatively hydrophilic drugs presented as MDI suspensions include salbutamol sulfate and sodium cromoglycate. In Serevant® inhalers, the long-acting bronchodilator, salmeterol is presented as the less hydrophilic xinafoate (1-hydroxy-2-naphthoate) salt to increase the duration of action.

Other factors that may need to be considered are the irritancy of the salt to the airways and the potential risk of corrosion or attack on the MDI containers when the moisture content of the propellant is not carefully controlled or when ethanol is included in the formulation. For this reason, aggressive salt formers such as hydrochloride may be undesirable, if they result in particularly acidic formulations.

G. DPI products

Drug substances in DPI systems are normally aerosolized and delivered from capsules, blisters or device reservoirs containing a blend of drug substance and a simple excipient carrier such as lactose. It is essential that the drug substance is not hygroscopic and the drug/lactose mixture has a low moisture content to ensure reproducible and efficacious drug delivery to the airways.

In common with suspension MDI formulations, it is essential to reduce the particle size of the drug to a median diameter of 5–10 mm, to ensure delivery to the lower airways. If it is not possible to micronize the salt candidate, it may be necessary to select a different salt that can be micronized. A further characteristic of the drug that influences micronization is the melting point. Drugs or salts with low melting points (<100°C) are often difficult to micronize and may be better formulated as a solution MDI, whereas high melting crystalline drugs are often readily micronized and incorporated into a DPI or solution MDI.

H. Soft gelatine capsule formulations

Soft gelatine capsules are one-piece capsules that are formed, filled and sealed in a unique synchronous step. The shells are composed of gelatine, with glycerol added as a plasticizer; the gelatine also contains a small quantity of water. The liquid fill is non-aqueous or substantially non-aqueous and generally consists of a fatty oil or a water-miscible liquid such as polyethylene glycol. The consistency of the fill can vary between a slightly viscous liquid and a thick paste and can be either a solution or a suspension. Soft gelatine capsules can satisfy special needs such as the solubilization of drugs that are poorly soluble; having the drug pre-dissolved often gives an improvement in bioavailability.

In deciding which type of formulation to opt for, it is necessary to assess the solubility of the drug in several non-aqueous solvents. When a solution is required, it should be decided whether the fill will be semi-polar (polyethylene glycol based) or non-polar (oil based). If presented as a suspension, again low solubility is normally required to prevent crystal growth.

An alternative to the soft gelatine capsule is the same two-piece hard gelatine capsule that is normally used for standard capsule formulations. The shells can also be filled with a thick non-aqueous paste or a thermosetting matrix, which is warmed to allow pumping into the capsule bottom shell; if necessary a banderole is fitted to cover the join to prevent leakage.

I. Emulsions, creams and ointments

Where it is necessary to incorporate the drug into the oil phase of a dispersed o/w system, a hydrophobic drug is required. The drug substance should have a low aqueous solubility and it may be necessary to select a more lipophilic, low melting salt (e.g. a fatty acid ester salt such as stearate or an amine salt such as diethanolamine) or

the parent compound, if it is more lipophilic than the salt. However, the salt should not be too lipophilic since this could impair partitioning from the oil phase into the aqueous phase and adversely affect drug delivery.

In topical formulations, where the drug is occasionally formulated as a suspension of solid particles, a relatively high melting point and absence of polymorphic transitions of the salt are essential requirements. This is important because the majority of creams and ointments are prepared by warming the oil phase or ointment base during incorporation of the drug substance into the formulation.

VII. CONCLUSION

Seen from a distance, to decide on the physicochemical form of a newly discovered active substance *en route* to development might appear a trivial act. However, from this chapter, it should be evident that the selection of the optimum solid state and salt form of a pharmacologically active substance is a multidisciplinary task that cannot be accomplished by the discovery chemist alone as it had been in past times. Too many factors now need to be taken into account, many requirements satisfied, limitations respected, and a considerable amount of experimental effort expended in order to convert the drug candidate into a producible, stable, therapeutically useful drug product. So it may be concluded that this task requires a highly interactive team representing the many disciplines involved: drug discovery, process development, physical pharmacy, biopharmacy, dosage form development, analytics, toxicology, pharmacokinetics, and regulatory affairs.

REFERENCES

1. Yalkowsky, S. H. Solubility and solubilization of nonelectrolytes. In *Techniques of Solubilization of Drugs* (Yalkowsky, S. H., Ed.). Marcel Dekker: New York, 1985, pp. 1–14.
2. Jain, N., Yalkowsky, S. H. Estimation of the aqueous solubility, I: application to organic nonelectrolytes. *J. Pharm. Sci.* 2001, *90*, 234–252.
3. CLogP. BioByte Corp. and Pomona College: Clairmont, CA, USA, 2001.
4. Yang, G., Jain, N., Yalkowsky, S. H. Estimation of the aqueous solubility, III: application to weak electrolyte drugs in buffered solutions, *AAPS 2001 Annual Meeting and Exposition*, Denver, CO, USA, *AAPS PharmSci. 3*(3) (Suppl.), poster No. W4446.
5. Ran, Y., Jain, N., Yalkowsky, S. H. Prediction of aqueous solubility of organic compounds by the general solubility equation. *J. Chem. Inf. Comput. Sci.* 2001, *41*(5), 1208–1217.
6. Jorgensen, W. L., Duffy, E. M. Prediction of drug solubility from structure. *Adv. Drug Deliv. Rev.* 2002, *54*(3), 355–366.
7. Bergstrom, C. A., Norinder, U., Luthman, K., Artursson, P. Experimental and computational screening models for prediction of aqueous drug solubility. *Pharm. Res.* 2002, *19*(2), 182–188.
8. Bergstrom, C. A., Strafford, M., Lazorova, L., Avdeef, A., Luthman, K., Artursson, P. Absorption classification of oral drugs based on molecular surface properties. *J. Med. Chem.* 2003, *46*(4), 558–570.
9. Turner, J. V., Maddalena, D. J., Agatonovic-Kustrin, S. Bioavailability prediction based on molecular structure for a diverse series of drugs. *Pharm. Res.* 2004, *21*(1), 68–82.
10. Wassvik, C. M., Holmen, A. G., Bergstrom, C. A., Zamora, I., Artursson, P. Contribution of solid-state properties to the aqueous solubility of drugs. *Eur. J. Pharm. Sci.* 2006, *29*(3–4), 294–305.
11. Bevan, C. D., Lloyd, R. S. A high-throughput screening method for the determination of aqueous drug solubility using laser nephelometry in microtiter plates. *Anal. Chem.* 2000, *72*(8), 1781–1787.
12. http://www.documentarea.com/qsar/richard_lloyd.pdf.
13. de Levie, R. The Henderson-Hasselbalch equation: its history and limitations. *J. Chem. Educ.* 2003, *80*, 146.
14. Gennaro, A. R. (Ed.) *Remington's Pharmaceutical Sciences,* 17th Edition. Mack Publishing Co., Easton, PA, 1985, p. 391.
15. Albert, A. A., Serjeant, E. P. *Ionisation Constants of Acids and Bases*. Methuen: London, 1962.
16. http://www.ich.org.
17. *The Electronic Orange Book: Approved Drug Products with Therapeutic Equivalence Evaluations*. US Department of Health and Human Services, Public Health Service, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Office of Information Technology, Division of Data Management and Services. http://www.fda.gov/cder/ob/default.htm
18. http://eudrapharm.eu.
19. Stahl, P. H., Wermuth, C. G. Monographs on acids and bases. In *Handbook of Pharmaceutical Salts: Properties, Selection, and Use* (Stahl, P. H., Wermuth, C. G., Eds). Wiley-VCH: Zurich, 2002, pp. 265–327.
20. Stahl, P. H. Appendix. In *Handbook of Pharmaceutical Salts: Properties, Selection, and Use* (Stahl, P. H., Wermuth, C. G., Eds). Wiley-VCH: Zurich, 2002, pp. 329–350.
21. Gould, P. L. Salt selection for basic drugs. *Int. J. Pharm.* 1986, *33*, 201–217.
22. Berge, S. M., Bighley, L. D., Monkhouse, D. C. Pharmaceutical salts. *J. Pharm. Sci.* 1977, *66*, 1–19.
23. Bowker, M. J. A procedure for salt selection and optimization. In *Handbook of Pharmaceutical Salts: Properties, Selection, and Use* (Stahl, P. H., Wermuth, C. G., Eds). Wiley-VCH: Zurich, 2002, pp. 161–189.
24. Senior, N. Some observations on the formulation and properties of chlorhexidine. *J. Soc. Cosmet. Chem.* 1973, *24*, 259–278.
25. Argharkar, S., Lindenbaum, S., Higuchi, T. Enhancement of solubility of drug salts by hydrophilic counterions: properties of organic salts of an antimalarial drug. *J. Pharm. Sci.* 1976, *65*(5), 747–749.
26. Martin, A. *Physical Pharmacy – Physical Chemical Principles in the Pharmaceutical Sciences*, 4th Edition. Lea & Febiger: Philadelphia, PA, 1993. p. 153
27. Tong, W. Q., Whitesell, G. *In situ* salt screening – a useful technique for discovery support and preformulation studies. *Pharm. Dev. Technol.* 1998, *3*(2), 215–232.
28. Morissette, S. L., Almarsson, O., Peterson, M. L., Remenar, J. F., Read, M. J., Lemmo, A. V., Ellis, S., Cima, M. J., Gardner, C. R. High-throughput crystallisation: polymorphs, salts, co-crystals and solvates of pharmaceutical solids. *Adv. Drug Deliv. Rev.* 2004, *56*(3), 275–300.
29. Remenar, J. F., MacPhee, J. M., Larson, B., Tyagi, V. A., Ho, J. H., McIlroy, D. A., Hickey, M. B., Shaw, P. B., Almarsson, O. Salt selection and simultaneous polymorphism assessment via high-throughput crystallisation: the case of sertraline. *Org. Proc. Res. Dev.* 2003, *7*(6), 990–996.
30. Ware, E. C., Lu, D. R. An automated approach to salt selection for new unique trazodone salts. *Pharm. Res.* 2004, *21*(1), 177–184.
31. Kojima, T., Onoue, S., Murase, N., Katoh, F., Mano, T., Matsuda, Y. Crystalline form information from multiwell plate salt screening by use of Raman microscopy. *Pharm. Res.* 2006, *23*(4), 806–812.

32. Kaplan, S. A. Biological implications of *in vitro* dissolution testing. In *Dissolution Technology* (Leeson, L., Carstensen, J., Eds). Academy of Pharmaceutical Sciences: Washington, DC, 1974, p. 167.
33. Bogardus, J. B., Blackwood, R. K., Jr. Dissolution rates of doxycycline free base and hydrochloride salts. *J. Pharm. Sci.* 1979, *68*(9), 1183–1184.
34. Hirsch, C. A., Messenger, R. J., Brannon, J. L. Fenoprofen drug form selection and preformulation stability studies. *J. Pharm. Sci.* 1978, *67*(2), 231–236.
35. Callahan, J. C., Cleary, G. W., Elefant, M. et al. Equilibrium moisture content of pharmaceutical excipients. *Drug Dev. Ind. Pharm.* 1982, *8*, 355–369.
36. Serajuddin, A. T. M., Jarowski, C. I. Effect of diffusion layer pH and solubility on the dissolution rate of pharmaceutical bases and their hydrochloride salts I: phenazopyridine. *J. Pharm. Sci.* 1984, *74*(2), 142–147.
37. Nelson, E. Comparative dissolution rates of weak acids and their sodium salts. *J. Am. Pharm. Assoc. Sci. Ed.* 1958, *47*, 297–299.
38. Noyes, A. A., Whitney, W. R. The rate of solution of solid substances in their own solutions. *J. Am. Chem. Soc.* 1897, *19*, 930–934.
39. Wood, J., Syarto, J., Letterman, H. Improved holder for dissolution rate studies. *J. Pharm. Sci.* 1965, *54*(7), 1068.
40. Chawla, G., Bansal, A. K. Challenges in polymorphism of pharmaceuticals. *CRIPS* 2004, *5*(1), 9–14.
41. Threlfall, T. L. Analysis of organic polymorphs. *Analyst* 1995, *120*, 2435–2460.
42. Huang, L. F., Tong, W. Q. Impact of solid state properties on developability assessment of drug candidates. *Adv. Drug Deliv. Rev.* 2004, *56*(3), 321–324.
43. Gardner, C. R., Walsh, C. T., Almarsson, O. Drugs as materials: valuing physical form in drug discovery. *Nat. Rev. Drug Discov.* 2004, *3*, 926–934.

Chapter 38

Preparation of Water-Soluble Compounds by Covalent Attachment of Solubilizing Moieties

Camille G. Wermuth

I. INTRODUCTION
II. SOLUBILIZATION STRATEGIES
A. How will the solubilizing moiety be grafted?
B. Where will it be grafted?
C. What kind of solubilizing chain will be utilized?
III. ACIDIC SOLUBILIZING CHAINS
A. Direct introduction of acidic functions
B. Alkylation of OH and NH functions with acidic chains
C. Acylation of OH and NH functions with acidic chains
IV. BASIC SOLUBILIZING CHAINS
A. Direct attachment of a basic residue
B. Bioisosteric exchange of a basic functionality
C. Development of a water-soluble prodrug of diazepam
D. Attachment of the solubilizing moiety to an alcoholic hydroxyl
E. Attachment of the solubilizing moiety to an acidic NH function
F. Attachment of the solubilizing moiety to a basic NH_2 function
G. Attachment of the solubilizing moiety to carboxylic acid functionalities
V. NON-IONIZABLE SIDE CHAINS
A. Glycolyl and glyceryl side chains
B. Polyethylene glycol derivatives
C. Glucosides and related compounds
VI. CONCLUDING REMARKS
REFERENCES

"Ajouter à sa queue, ôter à ses oreilles"
(Add to her tail, remove from her ears)

Jean de LA FONTAINE, La Besace

I. INTRODUCTION

The strategy described in here aims to convert a water-insoluble drug into a water soluble one by attaching *covalently* an appropriate solubilizing side chain. Surprisingly few reviews covering this subject are found in the literature.[1–5] Seen from the chemical side, the solubilizing moiety can be a neutral hydrophilic group or an ionizable organic base or acid. With the exception of possible crystallization problems, no major difficulties are expected in the synthesis of such compounds. One problematic aspect of the solubilization approach to be taken is to decide if the solubilizing moiety has to be fixed in a reversible manner, generating a prodrug, or in an irreversible manner, yielding a new chemical entity. In the latter case, the solubilizing procedure may exact a cost in terms of the recognition mechanisms, the soluble analog being less potent or even showing a different pharmacological profile. In some instances, changes in one part of the molecule have to be compensated by changes in another part. As Jean de La Fontaine said of the elephant: "Add to her tail, remove

Copyright © 2008, Elsevier Ltd
All rights reserved.

TABLE 38.1 Successful Examples of Drug Solubilization by a Chemical Means

Solubilizing side chain	Therapeutic class	Compound
Phosphoric ester	Steroids	Betamethazone
Phosphoric ester	Vitamins	Menadione
Hemisuccinate	Cardiotonics	Benfurodil
Hemisuccinate	Antibiotics	Chloramphenicol
Hemisuccinate	Steroides	Prednisolone
Hemisuccinate	Benzodiazepines	Oxazepam
Acidic side chain	Theophylline	Etaphylline
Acidic side chain	Antisyphilitic	Solusalvarsan
Neutral side chain	Analgesic	Glafenine
Neutral side chain	Bronchodilator	Dyphylline
Neutral side chain	Antibacterial	Sulfapyridine N-glucoside
Basic side chain	Antibiotics	Rolitetracycline
Basic side chain	Flavonoids	Solurutine
Basic side chain	Morphine	Pholcodine

from her ears." In addition, the solubilized analog of an already approved drug is considered by the governmental drug agencies as a totally new chemical entity, demanding a completely new development process. The financial investment that is then necessary can only be justified if enough sales of the solubilized form are expected. As a consequence the attachment of solubilizing moieties has to be considered very early in the drug discovery process or else limited to drugs with sizeable markets and undertaken when all other solubilizing stratagems fail. Despite these difficulties, many examples are found in therapy of successful drug solubilization by means of a chemical transformation of a parent drug (Table 38.1).

Chemically solubilized active principles render possible the preparation of parenteral, and especially intravenous forms appreciated in the clinical practice. But even at the preclinical level, the use of water-soluble molecules is recommended as they are effectively much easier to study by *in vitro* tests, in cell or microorganism cultures and on isolated organs. The inconveniences are that chemically modified structures may show modified pharmacological, pharmacokinetic and toxicological properties.

II. SOLUBILIZATION STRATEGIES

Three points are decisive in terms of the solubilizing strategies: How will the solubilizing moiety be grafted? Where will it be grafted? What kind of side chain will be utilized?

A. How will the solubilizing moiety be grafted?

The solubilizing chain can be *reversibly* or *irreversibly* grafted to the parent molecule. In the case of reversible linkages we are dealing in fact with prodrugs. Reversible linkages are usually provided by esters, peptides or glucosides.

Irreversible attachment of side chains is achieved by *O*- and *N*-alkylation and creation of carbon–carbon bonds. The grafted side chains can be basic (dimethylaminoethyl or morpholinoethyl chains), acidic (carboxylic, sulfonic, etc.) or neutral (glyceryl).

Intermediate situations are found for enol and phenol phosphates as well as for some amides. For these compounds only partial reversibility is observed *in vivo*.

B. Where will it be grafted?

First of all a careful examination of the parent molecule must be undertaken in order to identify the parts of the molecule that present adequate chemical reactivity and are suitable as attachement points for the solubilizing chain. Functions such as OH, SH, NH, acidic CH or CO_2H are reactive sites that furnish nucleophilic or basic entities. Conversely, aromatic double bonds are sensitive to electrophilic attack whereas carbonyl groups and conjugated carbon–carbon double bonds are sensitive to nucleophilic attack. The second criterion that has to be considered is of a biological nature: the solubilizing chain can only be grafted to those parts of the molecule that are not involved in the drug–receptor interaction. Fixed at the wrong place, the solubilizing chain can totally inactivate the molecule.

C. What kind of solubilizing chain will be utilized?

The *size* of the solubilizing chain is one of the selection parameters. The chains can be limited to the strict minimum and simply represent functional groups, or they can be made from larger residues containing several atoms (Table 38.2).

The *nature* of the side chains is the next selection parameter. It has to be decided if they can be ionizable (acidic or basic) or non-ionizable.

Acidic ionizable moieties (e.g. carboxylic acids) yield readily crystalline compounds and often do not alter the pharmacologic profile of the parent molecule. However, owing to their amphiphilic nature, they can show hemolytic properties. In addition, only a limited number of cations can be used to neutralize them. Traditional inorganic cations such as sodium, potassium or magnesium can induce mineral surcharges and are no longer recommended. They are advantageously replaced by organic bases such as tromethamine, lysine or *N*-methylglucamine (see Chapter 37).

TABLE 38.2 Small and Large Solubilizing Moieties

Small groups or simple functionalities	Larger solubilizing moieties
—CO_2H	R—OH→R—O—CH_2—CH_2—CO_2H
—SO_3H, —OSO_3H	R—NH_2→R—NH—CH_2—CH_2—CH_2—SO_3H
—PO_3H_2, —OPO_3H_2	$(R)_2C{=}O$ → $(R)_2C{=}N$—O—CH_2—CO_2H
—NH_2, —NHR, —NR_2	R—OH → O-morpholinylethyl
N-Oxides	R—OH → O-glucoside
S-Oxides	R—OH → O—CO—CH_2—CH_2—CO_2H
Sulfones	R—OH → *m*—O—C_6H_4—SO_3H

Basic ionizable moieties (e.g. substituted amines) can be neutralized by a very large number of organic and inorganic acids (see Chapter 37). The salts obtained are also readily crystallized and usually show less surface-active properties than salts from acidic chains. Their main disadvantage, which somewhat limits their utility, is their tendency to interfere with biogenic amines and neurotransmitters. In other words, attaching a basic amine functional group can seriously modify the pharmacological activity with regard to the parent drug. Drugs with acidic side chains cannot be mixed with drugs having basic side chains, as it is likely that a salt formed between the two drugs might precipitate. *Non-ionizable moieties* (e.g. polyhydroxylated chains) do not present this disadvantage and are compatible with other drug preparations. As they can be delivered at pH values close to 7, they do not produce painful injections. The main problems encountered with non-ionizable solubilizing moities is their lesser propensity to crystallize. In addition, increased cost can be expected from the necessity of added protection/deprotection steps during their synthesis.

III. ACIDIC SOLUBILIZING CHAINS

When planning the solubilization by means of a carboxylic acid side chain, one has to take into account the therapeutic properties peculiar to the carboxylic group. Thus, all arylacetic acids show more or less potent anti-inflammatory activities and many α-functionalized carboxylic acids are chelating agents. Among them we find the chelating α-amino acids[6] and antibacterial nalidixic acid-derived quinolones[7] and probably kynurenic acid analogs acting as antagonists at the glycine site.[8]

A. Direct introduction of acidic functions

Direct introduction of a solubilizing function can be achieved by carboxylation and by sulfonation. The historical example of carboxylation is the Kolbe synthesis of salicylic acid. Sulfonation was employed to solubilize guaiacol, camphor and 7-chloro-8-hydroxyquinoline (Figure 38.1).

FIGURE 38.1 Sulfonic acid solubilization.

FIGURE 38.2 Carboxylic acid and heterocyclic bioisosteres as solubilizing groups.

In a recent example, aryl-carboxylic solubilization was used to solubilize core-modified porphyrins.[9] Solubilizing aryl carboxylic functions were also replaced by their tetrazole or 1,2,4-oxadiazolone bioisosteres in the design of second generation, benzodiazepine-derived, CCK-B antagonists (Figure 38.2).[10]

In a series of adenosine A_1 receptor antagonists,[11] simple homologation of the carbon chain bearing the carboxylic group improved the bioavailability, the selectivity and the water solubility (Figure 38.3).

B. Alkylation of OH and NH functions with acidic chains

This procedure alkylates the hydroxy and the amino groups already present on the molecule with reactive intermediates bearing acidic functional residues (Table 38.3).

FIGURE 38.3 Homologation of the carboxylic solubilizing group provides improved bioavailability.

TABLE 38.3 Alkylation of OH and NH Functions with Acidic Chains

Starting derivative	Solubilized analog	Example	Reference
Ar—OH	Ar—O—CH_2—CO_2H	Solusalvarsan	3
Ar—OH	Ar—O—CH_2—CO_2H	Flavodic acid	12
Ar—NH_2	Ar—NH—CH_2—CO_2H	Acediasulfone	13
Ar—NH_2	Ar—NH—CH_2—CO_2H	Iodopyracet3	14
Ar—NH_2	Ar—NH—CH_2—SO_2H	Sulfoxone sodium	15

These compounds are prepared starting from chloroacetic acid or its ethyl ester. For chains longer than acetic, cyanoethylation and hydrolysis of the nitrile obtained leads to the propionic chain, alkylation with ethyl 4-bromobutyrate and saponification leads to the butyric chain. The propanesulfonic chains are particularly accessible by means of ring opening of propane-sultone.

1. *Dihydroartemisinin ethers*

A water-soluble derivative of artemisinin, the sodium salt of artesunic acid (the succinic half-ester derivative of dihydroartemisinin; Figure 38.4), can be administered by intravenous injection, a property that makes it especially useful in the treatment of advanced and potentially letal cases of *Plasmodium falciparum.* Sodium artesunate is capable of rapidly reversing parasitaemia and causing the restoration to conciousness of the comatose cerebral malaria patient. The utility of sodium artesunate, however, is impaired by its poor stability due to the facile hydrolysis of the ester linkage. To overcome the ease of hydrolysis of the ester function in sodium artesunate, Lin *et al.*[16] prepared a series of analogs in which the solubilizing moiety is joined to dihydroartemisinin by a ether rather than an ester linkage. One of the compounds prepared, artelinic acid (Figure 38.4) is both soluble and stable in 2.5% K_2CO_3 solution and possesses superior *in vivo* activity against *Plasmodium berghei* in comparison to artermisinin or artesunic acid.[16]

Continuing the search for water-soluble dihydroartemisin derivatives with higher efficacy and longer plasma half-life than artesunic or artelinic acid, Lin *et al*. prepared a series of dihydroartemisinoxy-butyric acids bearing an aryl substituent at the 4-position of the butyric side chain (Figure 38.4) The *p*-chlorophenyl and the *p*-bromophenyl derivatives showed a five- to eightfold increase in *in vitro* antimalarial activity against D-6 and W-2 clones of *Plasmodium falciparum* than artemisin or artelinic acid. They also showed *in vivo* oral antimalarial activity superior to that of artelinic acid.[17] Other ether-linked artemisinin-solubilizing chains containing asymmetric centers did not show activities superior to that of artelinic acid.[18]

B. Acylation of OH and NH functions with acidic chains

The acylation of OH and NH functions with acidic chains is probably the most popular mode of acidic solubilization. Alcohols and phenols are converted not only into half-esters such as hemisuccinates, hemiglutarates, hemiphtalates[19] *meta*-benzenesulfonates[20] but also into phosphates or even sulfates (Figure 38.5). All these derivatives can give water-soluble sodium or amine salts. Similar acylation possibilities exist for amines, but peptide-like derivatives are often preferred because the enzymatic regeneration of the parent molecule *in vivo* is easier.

Carboxylic half-esters (e.g. hemisuccinates) of phenols are easily hydrolyzed in aqueous solution and are therefore not recommended for the solubilization of phenolic compounds. Even hemisuccinates of alcohols suffer somewhat from stability problems and must be supplied as lyophilized (freeze-dried) powders for reconstitution in water and used within 48 h (see, for example, the monographies chloramphenicol sodium succinate or hydrocortisone sodium succinate in *The Handbook on Injectable Drugs*,[21] see also Anderson *et al*.[22,23].

FIGURE 38.4 Solubilized forms of artemisin.[16,17]

FIGURE 38.5 Acylation of OH and NH functions with acidic chains.

An additional difficulty occurring with hemisuccinates was discovered by Sandman *et al.*[24] These authors, in studying the stability of chloramphenicol succinate, found an unusual partial acyl transfer reaction of the succinyl group to give a cyclic hemi-orthoester (Figure 38.6).

Apparently hemiglutarates of phenolic drugs are more stable than hemisuccinates, an example is provided by the water-soluble diglutaryl-probucol which prevents cell-induced low-density lipoproteins (LDL) oxidation (Figure 38.7).[25]

In the search for an improvement of solution stability, i.e., in minimizing the ester hydrolysis and decreasing the acyl migration, Anderson *et al.*[26] synthesized a series of more stable water-soluble methylprednisolone esters. Several of the analogs were shown to have shelf lives in solution of greater than 2 years at room temperature. Ester hydrolysis studies of these compounds in human and monkey serum indicated that derivatives having anionic solubilizing residues such as carboxylate

FIGURE 38.6 Formation of cyclic hemi-orthoesters from a hemisuccinate.[4,24]

FIGURE 38.7 Structure of the diglutaryl ester of probucol.[25]

or sulfonate are more slowly hydrolyzed by serum esterases then compounds with a cationic solubilizing moiety (tertiary amine).[26]

Phosphoric esters (Figure 38.8) are generally more stable. They have been used in the steroid[27,28] and in the vitamin field (vitamin C,[29] vitamin B_1,[30] benfotiamine,[31,32] riboflavine,[33] dihydrovitamin K_1.[34] Riboflavine-5′-phosphoric acid dihydrate, for example, has a solubility in water of 112 g/L at pH 6.9, compared to 0.06–0.33 g/L for riboflavine itself. Phosphoric esters of trichloroethanol, diphenylhydantoin (open form) and clindamycin are discussed in.

A large number of reported peptidomimetic compounds possess very low aqueous solubility at physiological pH owing to the high lipophilicity inherent in these structures. Phosphorylation can yield improved biological activities for such compounds. This is at least the case for the phosphorylated neurokinin-1 receptor antagonist[35] and the human immunodeficiency virus (HIV) protease inhibitor[36] of Figure 38.8 described by scientists from Merck and Upjohn respectively. Clean mono-phosphorylation methods are available: some of them are shown in Figure 38.9.

Formation of sulfate esters is one of the metabolic conjugation reactions (Phase II reactions, see Chapter 34). Sulfates of estradiol,[47] glucose,[47] menadiol,[48] and hydroxyethyl-theophylline[49] have been prepared (Figure 38.10). As a rule sulfuric acid esters, compared to their phosphoric analogs, are resistant to enzymatic hydrolysis *in vivo*[50,51] and their conversion to the parent drug is questionable.

Sulfonic acids can be prepared by direct sulfonation (see above, Figure 38.1). Compounds containing conjugated double bonds have been solubilized by addition of sodium bisulfite. Treatment of menadione (vitamin K_3) with sodium bisulfite leads to two addition compounds (Figure 38.11). Mild warming of the reactants for a short time predominantly affords adduct (a) which arises from attack of bisulfite ion at carbon 2. Heating at reflux for an extended period yields adduct (b) from addition of bisulfite ion to carbon in position 3[52].

In a similar way, the treatment of N^4-cinnamylidenesulfanilamide (prepared from cinnamic aldehyde and sulfanylamide) with sodium bisulfite affords noprylsulfamide (Figure 38.11), according to the "soluseptazine principle" (noprylsulfamide is also called soluseptazine).[53] Noprylsulfamide is freely soluble in water (200 g/L), and breaks down in the body with the liberation of free sulfanilamide. Treatment of 6-chloropurine riboside with *p*-aminobenzenesulfonic acid leads to the highly water-soluble N^6-(*p*-sulfophenyl)adenosine (solubility >1.5 g/mL, <3 M).

FIGURE 38.8 Phosphate esters.

This compound (Figure 38.12) is a potent A_1 adenosine agonist in receptor binding, in inhibitory electrophysiological effects in hippocampal slices, and in inhibition of lipolysis *in vivo*.[54]

Treatment of primary and secondary amines with formaldehyde and sodium bisulfite generates stable methanesulfonates that can also act as solubilizing groups. The first example of this reaction found in the literature is the conversion of *p*-phenetidine into the corresponding methanesulfonate.[55] The compound obtained (Figure 38.13) still possesses antipyretic properties and is much less toxic than *p*-phenetidine. Nevertheless, it did not break into a new market.[3]

Applied to noraminopyrine, the same solubilization strategy led to dipyrone,[56,57] a water-soluble (1 g/1.5 mL), injectable form of aminopyrine (Pyramidon) used worldwide. Replacement of formaldehyde by acetaldehyde to yield ethanesulfonates has been claimed to lead to compounds with faster hydrolysis kinetics *in vivo*.[58] Replacement of formaldehyde by glucose afforded glucosulfone sodium, a soluble preparation of the leprostatic 4,4′-diaminodiphenylsulfone.[59,60] Replacement of

FIGURE 38.9 Useful syntheses of monophosphate esters (a) Ref. [37] (b) Ref. [38,39], (c) Ref. [40], (d) Ref. [41,42], (e) Ref. [43–45],

FIGURE 38.10 Sulfuric esters.

FIGURE 38.11 Bisulfite adducts.

formaldehyde-bisulfite by formaldehyde sulfoxylate is also claimed in some references.[3,61]

Ketones can be solubilized as carboxymethoximes (Figure 38.14). Menadoxime is freely water soluble, can be sterilized by autoclaving and, like menadione itself, shows high antihemorrhagic activity.[62] The carboxymethoxime of griseofulvin, on the other hand, is devoid of activity *in vitro* as well as *in vivo*.[63] This may be due either to an absence of conversion to the parent molecule, or to rapid renal elimination. Unsubstituted oximes can be converted enzymatically to their corresponding ketones.[64] For carboxymethoximes intramolecular assistance should even facilitate the hydrolysis to the initial carbonyl function. This was shown to be the case for the carboxymethoxime of naloxone.[65] Other substituted oximes seem to be rather resistant as apparent from the metabolic stability of noxiptylin.[66]

FIGURE 38.12 N^6-(*p*-sulfophenyl)adenosine, a freely water-soluble adenosine A1 receptor agonist.[54]

IV. BASIC SOLUBILIZING CHAINS

Solubilization with basic side chains involves two essential strategies: either direct binding of the amine function on a carbon atom of the parent molecule, or linking it to a function already present: alcoholic or phenolic hydroxyl, carboxylic acid, amine or amide.

A. Direct attachment of a basic residue

Simple tertiary amines can be grafted to a carbon skeleton, either by exchange reactions or by Mannich reactions. The hydrochloride salt of the camptothecin derivative (a) (Figure 38.15) is soluble in water at concentrations up to 1 mg/mL; the comparable value for camptothecin itself is 0.0025 mg/mL.[67] Similar results were obtained in solubilizing the benzodiazepine (b)[68] and the quinazolinone (c).[69] The adenosine A_1 antagonist KW-3902 (d) was solubilized in an original manner by converting it to an amidinic and cyclized bioisostere (e).[70]

FIGURE 38.13 Sulfonates.

FIGURE 38.14 Carboxymethoximes as (reversible?) solubilizing chains for carbonyl containing molecules.

B. Bioisosteric exchange of a basic functionality

The K channel opener **8a** (Figure 38.16) suffers from poor aqueous solubility; enhanced solubility was achieved by appending heterocycles to the **8a** scaffold (**14b** and **15a**) but the most surprising observation was that simple amino groups (**8j**) yielded excellent *in vitro* maxi-K ion channel opening activity and enhanced brain-to-plasma partitioning compared to the appended heterocycles.[71]

C. Development of a water-soluble prodrug of diazepam

In the search for water-soluble derivatives of the tranquillizer and anticonvulsant diazepam, scientists from Roche UK[72] imagined an original solubilization approach, based on the attachment of a basic side chain to the ring-opened diazepam molecule (Figure 38.17).

Avizafone, the peptide between the open form of diazepam (Valium®) and L-lysine bears two primary amino functions which can be salified (by HCl) and thus render the compound freely water soluble. *In vivo* generation of

FIGURE 38.15 Solubilization by means of basic side chains.

free diazepam takes only some minutes after the injection, thanks to an enzymatic cleavage of the peptidic bond followed by a spontaneous cyclization at the pH 7.4 of the bloodstream. The peptide with D-lysine is not substrate for the esterases (Box 38.1).

D. Attachment of the solubilizing moiety to an alcoholic hydroxyl

Esterification of an alcoholic hydroxyl with dialkylglycine or its analogs is a very popular mode of solubilization of alcohols. It is illustrated (Figure 38.18) by soluble esters of forskolin,[73] of allopurinol,[74] and of metronidazole.[75] This mode of solubilization can also be applied to phenols such as paracetamol (see Chapter 43).

Many α-amino acid esters or related short-chained aliphatic amino esters show satisfying hydrolysis kinetics in plasma, but exhibit poor stability in aqueous solution.[75] This poor stability is predominantly due to electron withdrawal by the positively charged amino group, but may also involve intramolecular catalysis or assistance of the ester hydrolysis by the neighboring amino group. The replacement of the glycine unit by its benzologue, as shown for

FIGURE 38.16 Solubility and bioavailability are enhanced by basic substituents on an aromatic ring.

FIGURE 38.17 Avizafone, the peptide between diazepam and L-lysine, is an original water-soluble prodrug.

BOX 38.1 Avizafone, an Essential Component of the Nerve Agent Antidote Systems

Nerve agents such as tabun, sarin and soman, are a class of *organofluorophosphates* that act in blocking the enzyme *acetylcholinesterase*. As a result they produce an excessive accumulation of acetylcholine in the cholinergic synapses. Poisoning by a nerve agent leads to contraction of pupils, profuse salivation, convulsions, involuntary urination and defecation, and eventual death by asphyxiation as control is lost over respiratory muscles. Some nerve agents are readily vaporized or aerosolized and the primary portal of entry into the body is the respiratory system. Nerve agents can also be absorbed through the skin, requiring that those likely to be subjected to such agents wear a full body suit in addition to a *respirator*.

It has been known for many years that benzodiazepine compounds effectively antagonize seizures induced by organofluorophosphates. In the event of nerve gas poisoning a combination of three drugs is commonly used: an anticholinergic drug (e.g. atropine) an acetylcholinesterase reactivator (e.g. pralidoxime = 2-pyridine aldoxime), and anticonvulsant (i.e. benzodiazepine). Atropine and most of the oximes are freely soluble in water whereas many benzodiazepines are not. However, the diazepam prodrug avizafone is freely water soluble and chemically stable. Thus, a typical self-administrable emergency treatment containing 2 mg of atropine, 350 mg of pralidoxime and 20 mg of avizafone could be developed. The French Army Auto Injector, made by the Sedat Company is displayed below.

FIGURE 38.18 Dialkylglycines and related aminoesters.

FIGURE 38.19 Basic chains on acidic NH functions.

the metronidazole derivative of Figure 38.18, prevents the hydrolysis-facilitating effect of the amino group. Alternative solutions place the amino group more distant from the ester linkage in using 6-aminocaproic acid esters and sebacic acid-derived spacer groups[76] or render the ester function more resistant by replacing it by a carbonate or a carbamate function.[77]

E. Attachment of the solubilizing moiety to an acidic NH function

The NH group in theophylline is reactive and can easily be alkylated, yielding among others, etamiphyllin (Figure 38.19) which is rendered water soluble as hydrochloride or as camphosulfonate.[78]

Rolitetracycline, the Mannich base derived from the carboxamide function of tetracycline, formaldehyde and pyrrolidine,[79] is surprisingly stable and is used as an injectable form of tetracycline.

F. Attachment of the solubilizing moiety to a basic NH_2 function

An appropriate manner of solubilizing basic NH_2 functions is to form peptides with common amino acids as shown above for the L-lysine peptide of the ring-opened form of diazepam. Similarly acylation of the 3-amino group of the pyrrolidine ring in a series of quinolone antibacterial agents yielded interesting compounds (Figure 38.20).

FIGURE 38.20 *S*-alanine-derived peptide of a quinolone antibacterial agent confers high water-solubility and is sensitive to enzymatic cleavage *in vivo*.[80]

The amino acid analogs were less active *in vitro*, but had equal or increased efficacy *in vivo*. Indeed, it was shown that these compounds, which were stable to acid and base under the reaction conditions for their preparation, were rapidly cleaved in serum to give the parent quinolones. The amino acid derivatives showed a 3–70 times improved solubility when compared to the parent compounds.[80]

G. Attachment of the solubilizing moiety to carboxylic acid functionalities

As carboxylic esters of aminoalcohols are usually too sensitive to hydrolysis, amides with aminoalkylamines are preferred (Figure 38.21).

The water-soluble E-lactone ring modified 7-ethyl-camptothecin analog bearing a dimethylaminoethyl amidic chain compares favorably with the sodium salt resulting simply from the lactone ring opening.[81] Introduction of basic substituents into modified hydroxyethylene dipeptide isosteres gave inhibitors with improved solubility as well as improved potency against human plasma rennin.[82]

V. NON-IONIZABLE SIDE CHAINS

The most frequently employed solubilization approaches using non-ionizable moieties involve hydroxylated and poly-oxymethylenic side chains or glucosides and their analogs.

A. Glycolyl and glyceryl side chains

These chains are present in some classical drugs such as the muscle relaxant mephenesin[83] and the bronchodilator dyphylline (diprophylline);[84] (Figure 38.22). More recent applications are found in the venotropic troxerutin[85] and

FIGURE 38.21 Basic amides of carboxylic acids.

FIGURE 38.22 Glycolyl and glyceryl side chains.

in the analgesic-anti-inflammatory drugs glafenin[20,86] and etofenamate.[87]

B. Polyethylene glycol derivatives

Only a few examples of solubilization involving esters or ethers of poly(ethylene glycol) are found in the literature and it is not always clear whether the main purpose of the polyoxyethylenic chain grafting was to increase the aqueous solubility or to produce another improvement such as sustained release. One of the most representative examples is the antitussive benzonatate[88] (Figure 38.23). In a similar way, Nagakawa *et al.* prepared soluble forms of vitamins A and E, as well as various steroids (prednisolone, testosterone, hydrocortisone, gitoxine).[89,90]

The symmetrical attachment of the local anesthetic procaine to poly(ethylene glycol) increases the duration of action[91] and probably improves water-solubility. The hypnotic etodroxizine bears a three-ethylene-oxy-unit chain but is nevertheless used as a dimaleate.[92] In roxithromycin the oxygenated side chain is attached to the oxygen atom of the oxime of the antibiotic erythromycin.[93]

C. Glucosides and related compounds

Despite their ubiquitous distribution in plants, man-made glucosidic derivatives of alcohols or phenols are rarely prepared in medicinal chemistry. An old-timer is the sedative-hypnotic α-chloralose (Figure 38.24) which is presently used only as a surgical anesthetic for laboratory animals.

FIGURE 38.23 Polyethylene glycol derivatives.

FIGURE 38.24 Glycosidic derivatives of alcoholic functions.

Other examples are deoxycorticosterone β-maltoside[94,95] and menthol β-glucoside. Deoxycorticosterone glycosides show various solubilities depending on the sugar conjugate (Table 38.4).[96]

Menthol β-glucoside is a water-soluble, non-irritating prodrug of menthol that can be used, like glucovanillin, the β-D-glucoside of vanillin, as a pharmaceutic flavor adjuvant.[97] The use of sugar moieties as drug carriers has been reviewed by Chavis and Imbach.[98]

Attachment of the sugar moiety to nitrogen atoms of amine, amide or hydrazine functions is much more frequently encountered (Figure 38.25). Prontoglucal is the N^4-β-D-glucoside of sulfanylamide,[99] the tuberculostatic glyconiazide is the isonicotinoylhydrazone of D-glucuronic acid lactone,[100] and glucometacin results from the amidification of indomethacin with D-glucosamine.[101] Many highly water-soluble radiological contrast agents are solubilized as sugar conjugates. This is the case for metrizamide [102] and compound P297[103].

TABLE 38.4 Solubilities of Deoxycorticosterone Glycosides[96]

Deoxycorticosterone glycoside	Solubility in water
Glucoside	1.2‰
Galactoside	2.2‰
Lactoside	3.4‰
Lactosidoglucoside	Unlimited

VI. CONCLUDING REMARKS

The different chemical solutions to solubilizing problems discussed in this chapter reveal that in many cases the chemical transformation used also improves the activity profile of the parent molecule. This can be due to purely *pharmacokinetic* factors such as a better resorption from the organism and faster transport and diffusion. These factors explain also why solubilized drugs are generally faster eliminated and therefore show fewer symptoms of toxicity. But the *pharmacological* profile also can be affected. Chlorpromazine, for example (Figure 38.26), has neuroleptic properties, whereas the parent phenothiazine possesses anthelminthic properties. In this example, the attachment of the basic moiety has totally modified the pharmacological profile. However, the replacement of the basic moiety by its carboxylic counterpart yielded a compound totally inactive as neuroleptic (C. G. Wermuth, unpublished result).

Conversely, in the tricyclic antidepressant series, the passage from the basic imipramine to the acidic amineptine conserved the antidepressant properties. In other words, there are no general rules available for the selection of the most appropriate solubilizing moiety. It is therefore recommended, that for each new solubilization problem, acidic, basic and neutral solubilized versions of the parent molecule be prepared.

FIGURE 38.25 Glycosidic derivatives of amine, amide and hydrazine functions.

FIGURE 38.26 Acidic and basic side chains on tricyclic skeletons.

REFERENCES

1. Li, M., Wu, R. S., Tsai, J. S. C. Water-soluble derivatives of lipophilic drugs. *US 7,029,918 B2*, April 18, 2006.
2. Marini-Bettolo, G. B. Metodos Modernos de Solubilizacion de Medicamentos Organicos. *Ann. Asoc. Quim. Pharm. Urugay* 1948, *50*, 3–17.
3. Büchi, J. *Grundlagen der Arzneimittelforschung und der Synthetischen Arzneimittel*. Birkhaüser Verlag: Basel, 1963, pp. 220–235.
4. Stella, V. Prodrugs: an overview and definition. In *Prodrugs as Novel Drug Delivery Systems* (Higuchi, T., Stella, V., Eds), Vol. 14. American Chemical Society: Washington, DC, 1975, pp. 1–115.
5. Silverman, R. B. *The Organic Chemistry of Drug Design and Drug Action*. Academic Press: San Diego, CA, 1992, pp. 358–360.
6. Albert, A. *Selective Toxicity*. Chapman and Hall: London, 1979, pp. 403–440.
7. Hammond, M. L. Recent advances in anti-infective agents. In *Annu. Rep. Med. Chem.* (Bristol, J. A., Ed.), Vol. 28. Academic Press: San Diego, CA, 1993, pp. 119–130.
8. Leeson, P. D., Carling, R. W., Kulagowski, J. J., Mawer, I. M., Moore, K. W., Moseley, A. M., Rowley, M., Smith, J. D., Stevenson, G. I., Williams, B. J., Baker, R., Foster, A. C., Kemp, J. A., Tricklebank, M. D. Drugs interacting with the glycine binding site. In *Perspectives in Medicinal Chemistry* (Testa, B., Kyburz, E., Fuhrer, W., Giger, R., Eds). Verlag Helvetica Chemica Acta/VHC: Basel and Weinheim, 1993, pp. 239–257.
9. Stilts, C. E., Nelen, M. I., Hilmey, D. G., Davies, S., Gollnick, S. O., Oseroff, A. R., Gibson, S. L., Hilf, R., Detty, M. R. Water-soluble, core-modified porphyrins as novel, longer-wavelength-absorbing sensitizers for photodynamic therapy. *J. Med. Chem.* 2000, *43*, 2403–2410.
10. Bock, M. G., DiPardo, R. M., Mellin, E. C., Newton, R. C., Veber, D. F., Freedman, S. B. Second-generation benzodiazepine CCK-B antagonists. Development of subnanomolar analogs with selectivity and water solubility. *J. Med. Chem.* 1994, *37*, 722–724.
11. Kiesman, W. F., Zhao, J., Conlon, P. R., Dowling, J. E., Petter, R. C., Lutterodt, F., Jin, X., Smits, G., Fure, M., Jayarag, A. Potent and orally bioavailable 8-bicyclo[2.2.2]octylxanthines as adenosine A1 receptor antagonists,. *J. Med. Chem.* 2006, *49*, 7119–7131.
12. Mesanguy, M. Procédé de préparation du flavone 5,7-dioxyacétate de sodium. *French Patent 2.076.753*, January 27, 1970.
13. Jackson, E. L. Certain *N*-alkyl, *N*-carboxyalkyl and *N*-hydroxyalkyl derivatives of 4,4′-diaminodiphenyl sulfone. *J. Am. Chem. Soc.* 1948, *70*, 680–684.
14. Reitmann, J. Aliphatic amine salts of halogenated pyridones containing an acid group. *US Patent 1,993,039* (March 5, 1935, to Winthrop Chemical Company, Inc., New York), 1935.
15. Bauer, H. Organic compounds in chemotherapy. II. The preparation of formaldehyde sulfoxylate derivatives of sulfanilamide and of amino compounds. *J. Am. Chem. Soc.* 1939, *61*, 617–618.
16. Lin, A. J., Klayman, D. L., Milhous, W. Antimalarial activity of new water-soluble dihydroartemisin derivatives. *J. Med. Chem.* 1987, *30*, 2147–2150.
17. Lin, A. J., Sikry, A. B., Kyle, D. E. Antimalarial activity of new dihydroartemisin derivatives. 7. 4-(*p*-substituted phenyl)-4(*R* or *S*)-[10 (α or β)-dihydroartemisinoxy]butyric acids. *J. Med. Chem.* 1997, *40*, 1396–1400.
18. Lin, A. J., Lee, M., Klayman, D. L. Antimalarial activity of new water-soluble dihydroartemisin derivatives. 2. Stereospecificity of the ether side chain. *J. Med. Chem.* 1989, *32*, 1249–1252.
19. Coker, J. D., Elks, J., May, P. J., Nice, F. A., Phillips, G. H., Wall, W. F. Action of some steroids on the central nervous system of the mouse. I. Synthetic methods. *J. Med. Chem.* 1965, *8*, 417–425.
20. Allais, A., Rousseau, G., Girault, P., Mathieu, J., Peterfalvi, M., Branceni, D., Azadian-Boulanger, G., Chifflot, L., Jequier, R. Sur l'activité analgésique et antiinflammatoire des 4-(2′-alcoxycarbonyl phénylamino) quinoléïnes. *Eur. J. Med. Chem.* 1966, 65–70.
21. Trissel, L. A. *Handbook on Injectable Drugs*, 14th Edition. American Society of Hospital Pharmacists: Bethesda, MD, 2007.
22. Anderson, B. D., Taphouse, V. Initial rate studies of hydrolysis and acyl migration in methylprednisolone 21-hemisuccinate and 17-hemisuccinate. *J. Pharm. Sci.* 1981, *70*, 181–186.
23. Anderson, B. D., Conradi, R. A., Lambert, J. W. Carboxyl group catalysis of acyl transfer reactions in corticosteroid 17- and 21-monoesters. *J. Pharm. Sci.* 1984, *73*, 604–611.
24. Sandman, B., Szulczewski, D., Winheuser, J., Higuchi, T. Rearrangement of chloramphenicol-3-monosuccinate. *J. Pharm. Sci.* 1970, *59*, 427–429.
25. Parthasarathy, S. Evidence for an additional intracellular site of action of probucol in the prevention of oxidative modification of low density lipoprotein. Use of a new water-soluble probucol derivative. *J. Clin. Invest.* 1992, *89*, 1618–1621.
26. Anderson, B. D., Conradi, R. A., Spilman, C. H., Forbes, A. D. Strategies in the design of solution-stable, water-soluble prodrugs III: Influence of the pro-moiety on the bioconversion of 21-esters of corticosteroids. *J. Pharm. Sci.* 1985, *74*, 382–387.
27. Flynn, G. L., Lamb, D. J. Factors influencing solvolysis of corticosteroid-21-phosphate esters. *J. Pharm. Sci.* 1970, *59*, 1433–1438.
28. Melby, J. C., St Cyr, M. Comparative studies on absorption and metabolic disposal of water-soluble corticosteroid esters. *Metabolism* 1961, *10*, 75–82.
29. Cutolo, E., Larizza, A. Synthesis of 3-phosphoric ester of L-ascorbic acid. *Gazz. Chim. Ital.* 1961, *91*, 964–972.
30. Wenz, A., Göttmann, G., Koop, H. Verfahren zur Herstellung der freien Base und der Salze des Aneurin-orthophosphor-säureesters. *German Patent 1 110 646* (July 13, 1961; E. Merck A.G. Darmstadt), 1961.
31. Ito, A., Hamanaka, W., Takagi, H., Wada, T., Kawada, T. Verfahren zur Herstellung eines Acylierungsproduktes von Vitamin-B_1-orthophosphorsäureester und Salzen davon. *German Patent 1,130,811* (April 14, 1960 to Sankyo Kabushiki Kaisha, Tokyo), 1960.
32. Wada, T., Tagaki, H., Minakami, H., Hamanaka, W., Okamoto, K., Ito, A., Sahashi, Y. A new thiamine derivative, *S*-benzoylthiamine *O*-monophosphate. *Science* 1961, *134*, 195–196.
33. Viscontini, M., Ebnoether, C., Karrer, P. Einfache Synthese kristallisierter Lactoflavin-5′-phosphorsäure (Coferment des Flavinenzyms). *Helv. Chim. Acta* 1952, *35*, 457–459.

34. Fieser, L. F. Antihemorrhagic esters and methods for Producing the Same. *US Patent 2,407,823* (September 17, 1946 to Research Corporation, New York), 1946.
35. Hale, J. J., Mills, S. G., MacCoss, M., Dorn, C. P., Finke, P. E., Budhu, R. J., Reamer, R. A., Huskey, S. W., Luffer-Atlas, D., Dean, B. J., McGowan, E. M., Feeney, W. P., Lee Chiu, S. H., Cascieri, M. A., Chicchi, G. G., Kurtz, M. M., Sadowski, S., Ber, E., Tattersall, F. D., Rupniak, N. M. J., Williams, A. R., Rycroft, W., Hargreaves, R., Metzger, J. M., MacIntyre, D. E. Phosphorylated morpholine acetal human neurokinin-1 receptor antagonists as water-soluble prodrugs. *J. Med. Chem.* 2000, *43*, 1234–1241.
36. Chong, K.-T., Ruwart, M. J., Hinshaw, R. R., Wilkinson, K. F., Rush, B. D., Yancey, M. F., Strohbach, J. W., Thaisrivongs, S. Peptidomimetic HIV protease inhibitors: phosphate prodrugs with improved biological activities. *J. Med. Chem.* 1993, *36*, 2575–2577.
37. Fieser, L. F., Fieser, M. *Reagents for organic synthesis*. John Wiley & Sons: New York, 1967, p. 198.
38. Khawaja, T. A., Reese, C. B. *o*-Phenylene phosphochloridate. A convenient phosphorylating agent. *J. Am. Chem. Soc.* 1966, *88*, 3446–3447.
39. Khawaja, T. A., Reese, C. B., Stewart, J. C. M. A Convenient general procedure for the conversion of alcohols into their monophosphate esters. *J. Chem. Soc. (C)* 1970, 2090–2100.
40. Taguchi, Y., Mushika, Y. 2-(*N*,*N*-diethylamino)-4-nitrophenyl phosphate and its use in the selective phosphorylation of unprotected nucleosides. *Tetrahedron Lett.* 1975, *24*, 1913–1916.
41. Gajda, T., Zwierzak, A. Phase-transfer-catalysed halogenation of di-*t*-butyl phosphite. preparation of di-*t*-butyl phosphorohalidates. *Synthesis* 1976, 243–244.
42. Gajda, T., Zwierzak, A. Di-*t*-butyl phosphorobromidate. A new selective phosphorylating agent containing acid-labile protecting groups. *Synthesis* 1977, 623–625.
43. Tener, G. M. 2-Cyanoethyl phosphate and its use in the synthesis of phosphate esters. *J. Am. Chem. Soc.* 1961, *83*, 159–168.
44. Moffatt, J. G. The synthesis of orotidine-5′ phosphate. *J. Am. Chem. Soc.* 1963, *85*, 1118–1123.
45. Brownfield, R. B., Shultz, W. A direct method for the preparation of steroid-21-phosphates. *Steroids* 1963, *2*, 597–603.
46. Montgomery, H. A. C., Turnbull, J. H. Phosphoramidic halides; phosphorylating agents derived from morpholine. *J. Chem. Soc.* 1958, 1963–1967.
47. Fex, H., Lundvall, K. E., Olsson, A. Hydrogen sulfates of natural estrogens. *Acta Chem. Scand.* 1968, *22*, 254–264.
48. Fieser, L. F., Fry, E. M. Water-soluble antihemorrhagic esters. *J. Am. Chem. Soc.* 1940, *62*, 228–229.
49. Stieglitz, E.; Matz, M. Verfahren zur Herstellung von Xanthinalkylschwefelsäuren oder deren Salze. *German Auslegeschrift 1,090,669* (January 11, 1958 to Arzneimittelfabrik Krewel-Leuffen G.m.b. H., Eitorf/Sieg), 1958.
50. Miyabo, S., Nakamura, T., Kuwazima, S., Kishida, S. A comparison of the bioavailability and potency of dexamethazone phosphate and sulphate in man. *Eur. J. Clin. Pharmacol.* 1981, *20*, 277–282.
51. Williams, D. B., Varia, S. A., Stella, V. J., Pitman, I. H. Evaluation of the prodrug potential of the sulfate ester of acetaminophen and 3-hydroxymethyl-phenytoin. *Int. J. Pharm.* 1983, *14*, 113–120.
52. Greenberg, F. H., Leung, K. K., Leung, M. The reaction of vitamin K_3 with sodium bisulfite. *J. Chem. Ed.* 1971, *48*, 632–634.
53. Despois, R. L. Procédé de préparation de composés aminés aromatiques solubles possédant une valeur thérapeutique, French Patent 831,366 (June 7, 1938; Rhône-Poulenc), 1938.
54. Jacobson, K. A., Nikodijevic, O., Ji, X.-D., Berkich, D. A., Eveleth, D., Dean, R. L., Hiramatsu, K.-I., Kassel, N. F., van Galen, P. J. M., Lee, K. S., Bartus, R. T., Daly, J. D., LaNoue, K. F., Maillard, M. Synthesis and biological activity of N^6-(*p*-sulfophenyl)alkyl and N^6-sulfoalkyl derivatives of adenosine: water-soluble and peripherally selective adenosine agonists. *J. Med. Chem.* 1992, *35*, 4143–4149.
55. Anonymous. Verfahren zur Darstellung von p-äthoxyphenylaminomethylschwefligsauren Salzen. *German Patent 209,695* (8 May 1907, to Roberto Lepetit, Garessio, Italy), 1907.
56. Anonymous. Verfahren zur Darstellung von ω-methylschwefligsauren Salzen aminosubstituierter Arylpyrazolone. *Ger. Patent 259,577* (October 26, 1911; to Farbwerke vorm. Meister Lucius & Brüning in Hoechst am Main), 1911.
57. Anonymous. Verfahren zur Darstellung von ω-methylschwefligsauren Salzen aminosubstituierter Arylpyrazolone. *Ger. Patent 254,711* (July 21, 1911; to Farbwerke vorm. Meister Lucius & Brüning in Hoechst am Main), 1911.
58. Mutch, N. A new sulphonamide (sulfonamide E.O.S.). *Br. Med. J.* 1941, *2*, 503–507.
59. Jain, B. C., Iyer, B. H., Guha, P. C. The preparation of promin. *Sci. Cult.* 1946, *11*, 568–569. (C.A. **40**, 4687[7], 1946).
60. Anonymous. Soluble diphenylsulfone. *Swiss Patent 234,108* (December 1, 1944; Aktien-Gesellschaft vorm. B. Siegfried), 1949.
61. Bockmühl, M., Krohs, W., Racke, F., Windisch, K. *N*-methylsulphites and *N*-methanesulphinic acid salts of 1-aryl-2,3-dialkyl-4-alkylaminopyrazolones. *US Patent 2,193,788* (March 19, 1940; Winthrop Chemical Company, Inc.), 1940.
62. Holland, D. O. Preparation of 2-alkyl-1,4-naphtaquinone-4-carboxyalkoximes. *British Patent 621,934* (April 22, 1949; to Glaxo Laboratories Ltd), 1949.
63. Fischer, L. J., Riegelman, S. Absorption and activity of some derivatives of griseofulvin. *J. Pharm. Sci.* 1967, *56*, 469–476.
64. Tatsumi, K., Ishigai, M. Oxime-metabolizing activity of liver aldehyde oxidase. *Arch. Biochem. Biophys.* 1987, *253*, 413–418.
65. Negwer, M. *Organic-chemical drugs and their synonyms. Entry No 5760: Codoxime*, Vol. II. Akademie Verlag: Berlin, 1987.
66. Aichinger, G., Behner, O., Hoffmeister, F., Schütz, S. Basische tricyclische Oximinoäther und ihre pharmakologischen Eigenschaften. *Arzneim.-Forsch.* 1969, *19*, 838–845.
67. Kingsbury, W. D., Boehm, J. C., Jakas, D. R., Holden, K. G., Hecht, S. M., Gallagher, G., Caranfa, M. J., McCabe, F. L., Faucette, L. F., Johnson, R. K., Hertzberg, R. P. Synthesis of water-soluble (aminoalkyl)camptothecin analogues: inhibition of topoisomerase I and antitumor activity. *J. Med. Chem.* 1991, *34*, 98–107.
68. Showell, G. A., Bourrain, S., Neduvelil, J. G., Fletcher, S. R., Freedman, S. B., Kemp, J. A., Marshall, G. R., Patel, S., Smith, A. J., Matassa, V. G. High-affinity and potent, water-soluble 5-amino-1,4-benzodiazepine CCKB/gastrin receptor antagonists containing a cationic solubilizing group. *J. Med. Chem.* 1994, *37*, 719–721.
69. Ishikawa, F., Saegusa, J., Inamura, K., Sakuma, K., Ashida, S.-I. Cyclic guanidines. 17. Novel (*N*-substituted amino)imidazo[2,1-*b*]quinazolin-2-ones: water-soluble platelet aggregation inhibitors. *J. Med. Chem.* 1985, *28*, 1387–1393.
70. Suzuki, F., Shimada, J., Nonaka, H., Ishii, A., Shiozaki, S., Ichikawa, S., Ono, E. 7,8-Dihydro-8-ethyl-2-(3-noradamantyl)-4-propyl-1*H*-imidazo[2,1-*i*]purin-5(4*H*)-one: a potent and water-soluble adenosine A1 antagonist. *J. Med. Chem.* 1992, *35*, 3572–3581.
71. Romine, J. L., Martin, S. W., Meanwell, S. A., Gribkoff, V. K., Boissard, C. G., Dworetzky, S. I., Natale, J., Moon, S., Ortiz, A., Yeleswaram, S., Pajor, L., Gao, Q., Starrett, J. E., Jr. 3-[(5-Chloro-2-hydroxyphenyl)methyl]-5-[4-(trifluoromethyl)phenyl]-1,3,4-oxadiazol-2(3*H*)-one, BMS-191011: opener of large-conductance Ca^{2+}-activated potassium (maxi-K) channels, identification, solubility, and SAR. *J. Med. Chem.* 2007, *50*, 528–542.
72. Hassall, C. H., Holmes, S. W., Johnson, W. H., Kröhn, A., Smithen, C. E., Thomas, W. A. Peptido-aminobenzophenones-novel latentiated benzo-1,4-diazepines. *Experientia* 1977, *33*, 1492–1493.
73. Khandelwal, Y., Rajeshwari, K., Rajagopalan, R., Swamy, L., Dohadwalla, A. N., de Souza, N. J., Rupp, R. H. Cardiovascular effects of new water-soluble derivatives of forskolin. *J. Med. Chem.* 1988, *31*, 1872–1879.

74. Bundgaard, H., Falch, E. Improved rectal and parenteral delivery of allopurinol using the prodrug approach. *Arch. Pharm. Chem. Sci. Ed.* 1985, *13*, 39–48.
75. Bundgaard, H., Falch, E., Jensen, E. A novel solution-stable, water-soluble prodrug type for drugs containing a hydroxyl or an NH-acidic group. *J. Med. Chem.* 1989, *32*, 2503–2507.
76. Anderson, B. D., Conradi, R. A., Knuth, K. E. Strategies in the design of solution-stable, water-soluble prodrugs I: a physical-organic approach to pro-moiety selection for 21-esters of corticosteroids. *J. Pharm. Sci.* 1985, *74*, 365–374.
77. Balkovec, J. M., Black, R. M., Hammond, M. L., Heck, J. V., Zambias, R. A., Abruzzo, G., Bartizal, K., Kropp, H., Trainor, C., Schwartz, R. E., McFadden, D. C., Nollstadt, K. H., Pittarelli, L. A., Powles, M. A., Schmatz, D. M. Synthesis, stability, and biological evaluation of water-soluble prodrugs of a new echinocandin lipopeptide. Discovery of a potential clinical agent for the treatment of systemic candidiasis and *Pneumocystis carinii* pneumonia (PCP). *J. Med. Chem.* 1992, *35*, 194–198.
78. Klosa, J. Beitrag zur Reaktionsfähigheit der 7-Stellung des Theophyllins. 2.Mitt. Über Synthesen in der Theophyllinreihe. *Arch. Pharm.* 1955, *288*, 301–303.
79. Gottstein, W. J., Minor, W. F., Cheney, L. C. Carboxamido derivatives of tetracyclines. *J. Am Chem. Soc.* 1959, *81*, 1198–1201.
80. Sanchez, J. P., Domagala, J. M., Heifetz, C. L., Priebe, S. R., Sesnie, J. A., Trehan, A. K. Quinolone antibacterial agents. Synthesis and structure–activity relationships of a series of amino acid prodrugs of racemic and chiral 7-(3-amino-1-pyrrolidinyl)quinolones. Highly soluble quinolone prodrugs with *in vivo* pseudomonas activity. *J. Med. Chem.* 1992, *35*, 1764–1773.
81. Sawada, S., Yaegashi, T., Furuta, T., Yokokura, T., Miyasaka, T. Chemical modification of an antitumor alkaloid, 20(*S*)-camptothecin: E-lactone ring modified water-soluble derivatives of 7-ethylcamptothecin. *Chem. Pharm. Bull.* 1993, *41*, 310–313.
82. Boyd, S. A., Fung, A. K. L., Baker, W. R., Mantei, R. A., Armiger, Y.-L., Stein, H. H., Cohen, J., Egan, D. A., Barlow, J. L., Klinghofer, V., Verburg, K. M., Martin, D. L., Young, G. A., Polakowski, J. S., Hoffman, D. J., Garren, K. W., Perun, T. J., Kleinert, H. D. C-Terminal modifications of nonpeptide renin inhibitors: improved oral bioavailability via modification of physicochemical properties. *J. Med. Chem.* 1992, *35*, 1735–1746.
83. Morch, P. Glykresin. *Arch. Pharm. Chem.* 1947, *54*, 327–332.
84. Roth, H. J. Zur Darstellung von β-Oxyalkyl-dimethyl-purinen; Umsetzung von Theophyllin und Theobromin mit 1,2-Epoxyden. *Arch. Pharm.* 1959, *292*, 234–238.
85. Courbat, P. J. Qercetin and Quercetin Gycoside. *US Patent3,420,815* (January 7, 1969; to Zyma S.A.), 1969.
86. Mouzin, G., Cousse, H., Autin, J. M. A new, convenient synthesis of glafenine and floctafenine. *Synthesis* 1980, 54–55.
87. Boltze, K. H., Kreisfeld, H. Zur Chemie von Etofenamat, einem Antiphlogisticum aus der Klasse der *N*-Arylanthranylsäurederivate. *Arzneimit.-Forsch.-Drug Des.* 1977, *27*, 1300–1312.
88. Matter, M. Polyethoxy ethers of isocyclic organic carboxylic acids. *US Patent 2,714,608* (August 2, 1955; to Ciba Pharmaceutical Products, Inc., Summit, N.J.), 1955.
89. Nagakawa, T., Muneyuki, T., Mori, Y. Water solubilization by means of polyethylene glycol derivatives. *Japan Patent7455 CA (1961) 55, 5879b*, 1961.
90. Nagakawa, T., Muneyuki, T., Mori, Y. Water solubilization of drugs. *Japan Patent 17006*, November 17, 1960.
91. Weiner, B.-Z., Zilkha, A. Polyethylene glycol derivatives of procaine. *J. Med. Chem.* 1973, *16*, 573–574.
92. Morren, H. New derivatives of *N*-mono-benzhydril-piperazine and process for the preparation thereof. *British Patent 817,231* (July 29, 1959), 1959.
93. Gouin d'Ambrières, S., Lutz, A., Gasc, J. C. Novel erythromycin a derivatives. *US Patent 4,349,545* (September 14, 1982; to Roussel-UCLAF), 1982.
94. Miescher, K., Fischer, W. H., Meystre, C. 6. Über Steroide. (33. Mitteilung). Über Glukoside des Desoxy-corticosterons. *Helv. Chim. Acta* 1942, *25*, 40–42.
95. Miescher, K., Meystre, C. 26. Über Steroide. 34. Mitteilung. Über Sacharide des Desoxy-corticosterons II. *Helv. Chim. Acta* 1943, *26*, 224–233.
96. Meystre, C., Miescher, K. Über Steroide. (35. Mitteilung). Zur Darstellung von Saccharidderivaten der Steroide. *Helv. Chim. Acta* 1944, *27*, 231–236.
97. Higashiyama, T., Sakata, I., Mentholglykoside, Verfahren zur ihrer Herstellung, ihre Verwendung zur Entwicklung des Pfefferminzgeschmack sowie diese Verbindungen enthaltende Arzneimittel. *German Offen. 2.242.237* (August 28, 1972; to Toyo Hakka Kogyo Kabushiki Kaishy), 1972.
98. Chavis, C., Imbach, J.-L. Sur une méthode de pharmacomodulation: l'utilisation du vecteur sucre. In *Actualités de Chimie Thérapeutique* (Combet-Farnoux, C., Ed.), Vol. 5. Société de Chimie Thérapeutique: Châtenay-Malabry, 1977, pp. 3–28.
99. Bognar, R., Nanasi, P. *N*-substituted glycosylamines derived from sulphanilamide and *p*-aminosalicylic acid. *J. Chem. Soc.* 1953, 1703–1708.
100. Sah, P. D-Glucuronolactone isonicotinyl hydrazone. *J. Am. Chem. Soc.* 1953, *75*, 2512–2513.
101. Demetrio, A., Ganzina, F., Magi, M., Serino, E., Paroli, E., Samueli, F., Antiinflammatory di-(+)glucosamide of 1-(4-chlorobenzoyl)-2-methyl-5-methoxyindole-3-acetic acid. *German Patent 2,223,051* (December 13, 1973; to SIR Lab. Chimico Biologici SpA); CA 80, 83529e, 1974.
102. Almen, T. H. O., Haavaldsen, J., Nordal, V. *N*-(2,4,6-triiodobenzoyl)-sugar amines. *US Patent 3,701,771* (October 31,1972; to Nyegaard & Co, A/S), 1972.
103. Hilae, S. K., Dauth, G. W., Hess, K. H., Gilman, S. Development and evaluation of a new water-soluble iodinated myelographic contrast medium with markedly reduced convulsive effects. *Radiology* 1978, *126*, 417–422.

Chapter 39

Drug Solubilization with Organic Solvents, or Using Micellar Solutions or Other Colloidal Dispersed Systems

Michael J. Bowker and P. Heinrich Stahl

I. INTRODUCTION
II. FACTORS CONTROLLING SOLUBILITY AND ABSORPTION
A. The nature of drug substances
B. The polarity of the solvent system
III. WATER–COSOLVENT SYSTEMS
IV. SOLUBILIZATION MEDIATED BY SURFACTANTS
V. SOLUBILIZATION BY LIPID VEHICLES
A. Emulsions and microemulsions
B. Liposomes
VI. NANOPARTICLES AND OTHER NANOCOLLOIDAL TECHNOLOGIES
VII. DRUG DELIVERY AND CLEARANCE MECHANISMS OF NANOCOLLOIDS
VIII. DRUG DELIVERY AND ACCUMULATION USING COLLOIDAL SYSTEMS FOR THE TREATMENT OF CANCER
A. Liposome formulations
B. Formulations based on nanoparticles, microparticles and conjugated systems
IX. MODIFICATION OF DRUG TOXICITY BY NANOCOLLOIDAL DRUG DELIVERY SYSTEMS
REFERENCES

There's plenty of room at the bottom.

Richard Feynman (1959)

I. INTRODUCTION

Modern methods for the synthesis of drug candidates normally involve combinatorial chemistry techniques and the preparation of libraries of compounds that often have related structures. These libraries are usually dissolved in dimethylsulphoxide (DMSO) solution and archived, from where they are used in high-throughput, *in vitro* screening procedures for activity against enzyme- or receptor-based assay systems. DMSO is used because it dissolves a wide variety of different compound types, including compounds that are practically insoluble in water; it exhibits relatively low toxicity to most types of cells and has low chemical reactivity. When activity in a compound mixture is detected, the active compound is isolated and the structure is optimized by chemical modification. Ideally, the development scientist expects water-soluble candidates from this process, but in contrast, there is a general impression within the drug development community that the extensive use of DMSO in early screening and pharmacology studies is gradually skewing the properties of active compounds toward more hydrophobic (lipophilic) structures, rather than toward hydrophilic ones.

For most drug substances the preferred delivery route is oral but a significant percentage are required to be injected. However, even when the drug is intended for oral

Copyright © 2008, Elsevier Ltd
All rights reserved.

administration, an injectable dosage form is still required at several points along the development pathway, for example, for use in safety studies or for the determination of absolute bioavailability in humans. In most cases, the oral route can cope with a lower aqueous solubility as long as the physicochemical and mechanical properties of the drug substance are adequate for formulation purposes. For a drug intended for parenteral administration, with water solubility that is lower than required, the chemist and formulator must work together to explore the possibilities around increasingly "exotic" formulation approaches. Typically, these studies would involve the following–in increasing order of complexity:

1. If the drug is ionizable, a search for salts that are adequately soluble in water should be undertaken (see Chapter 37). For typical drug substances, target solubilities would be expected to be around 10 mg/mL for parenteral products and around 1 mg/mL for oral products for drugs with moderate activity.
2. An investigation of the solubility of the drug in an acceptable cosolvent, or if necessary cosolvent combinations, perhaps including a surfactant or lipid. The range of cosolvents requiring evaluation would differ, depending on the intended route of administration.
3. An investigation into the formation of water-soluble complexes with cyclodextrins; this solubilization approach is covered in Chapter 40.
4. Modification of the drug structure by the covalent attachment of a water-solubilizing group through a biodegradable linker (e.g. poly(ethylene glycol) (PEG),[1] poly(ethylene imine)[2,3] (PEI), a water-soluble polymethacrylate derivative,[4,5] a poly(amino acid) such as poly (l-glutamic acid) (poliglumex),[6] etc.). Alternatively, the drug could be made totally lipophilic by the attachment, again through a biodegradable linker, of a hydrophobic group (e.g. a long chain fatty acid such as docosahexaenoic acid (DHA; recommended INN: doconexent),[7] cholesterol,[8] an adamantine derivative,[9] etc.). This latter approach leads to a drug that may be more suitable for incorporation into liposomes or lipophilic nanoparticles. This process is often termed "chemical drug delivery" and is the subject of Chapters 34 and 35.
5. Investigations into the possibilities for the formulation of the drug substance as a more complex drug delivery system such as an emulsion, liposomal, micro- or nanoparticle, or other colloidal product.

If it is found necessary to investigate the more complex systems such as (4) and (5) above, the anticipated route of administration must be considered as most of these chemical modifications or formulation types can be degraded to some extent in the gastrointestinal tract. They are primarily intended for parenteral delivery.

Thus, there is a wide range of techniques available to the formulation scientist to improve and achieve the required aqueous solubility using different types of formulations. These formulations can exist as dispersed droplets, globules or particles, etc. with different size ranges, stabilities and properties; an overview of the different types and their approximate size ranges is given in Figure 39.1.

In certain cases, depending on particle size and the nature of the external surface coating, the drug-loaded colloids can be used to either passively or actively target specific cells, organs or disease sites.

FIGURE 39.1 Approximate size ranges for the commonly used colloidal drug delivery systems. SUV = small unilamellar vesicle; LUV = large unilamellar vesicle; MLV = large multilamellar vesicle; MVV = multivesicular vesicle.

II. FACTORS CONTROLLING SOLUBILITY AND ABSORPTION

A. The nature of drug substances

Aqueous solubility is a characteristic of drug substances that can vary considerably, depending on the balance of hydrophilic and hydrophobic domains or functional groups within the molecule. If the drug is polar, and contains ionizable functional groups, the aqueous solubility also usually varies with solution pH. Considering the structure as a whole, the drug molecule could be classed anywhere in the range between polar (hydrophilic) and non-polar (hydrophobic or lipophilic). If the molecule has both polar and non-polar domains, it is often termed as amphiphilic.

When taken orally, drug substances and other molecules are absorbed from the gastrointestinal tract when they are in their unionized form; highly ionized or zwitterionic forms are, at best, only poorly absorbed. After evaluating a large data set (2,245 compounds) of marketed oral drugs, Lipinski *et al.*[10] at Pfizer reached a series of conclusions on the optimum properties of drugs intended for oral administration. In order to achieve efficient oral absorption and cell permeability, drug candidates should have:

- Less than five hydrogen bond donors (sum of OH and NH groups)
- A molecular weight less than 500 Da
- A log P less than 5
- Less than 10 hydrogen bond acceptors (sum of O and N atoms in the structure).

This rapidly became known as "Lipinski's Rule of Five" as all the cut-off points are multiples of five. It should be noted that substrates for biological transporters and natural products (e.g. antibiotics, fungicides, vitamins, cardiac glycosides) form exceptions to this rule.

In a similar, but later, analysis of data from 1,791 synthetic, oral drug substances marketed since 1937, Proudfoot[11] concluded that the median molecular weight of drug molecules has increased substantially over the past 60+ years from around 300 Da to around 500 Da. It was also found that less than 5% of these oral drugs have more than four hydrogen bond donors and only 2% have a molecular weight greater than 500 and more than three hydrogen bond donors. Other approaches, such as the use of polar surface area calculations to predict absorption and drug transport[12–14] have elicited a decision rule that the maximum polar surface area should not exceed about 120 Å^2 for compounds that are intended to be absorbed by passive diffusion.

Also concerning the relationship between solid oral dosage form dissolution (i.e. drug solubility and absorption), Amidon *et al.* published a seminal paper[15] on the biopharmaceutics classification system (BCS) that concerns the evaluation of the bioequivalence of immediate-release solid oral dosage forms. Four different categories of drugs were defined, based on their solubility and oral absorption characteristics:

1. Class I: High solubility, high permeability (i.e. amphiphilic drugs)
2. Class II: Low solubility, high permeability (i.e. lipophilic drugs)
3. Class III: High solubility, low permeability (i.e. hydrophilic drugs)
4. Class IV: Low solubility, low permeability (i.e. hydrophobic drugs)

The basis of this classification system was taken up by the US Food and Drug Administration (FDA) and issued as one of their guidance for industry series of documents.[16]

There is one final concept, related to solubility and absorption, that often gives useful information on the oral absorption of drug candidates; this is the maximum absorbable dose[17] (MAD) (equation (39.1):

$$\mathrm{MAD} = \mathrm{S} \times K_a \times \mathrm{SWWW} \times \mathrm{SITT} \qquad (39.1)$$

where S = drug solubility at pH 6.5 (reflects typical pH of small intestine), K_a = absorption rate constant (can be determined from perfusion experiment using rat small intestine), SWWW = water volume in small intestine that is available for drug dissolution (generally accepted to be 250 mL), SITT = small intestine transit (generally accepted to be approximately 270 min).

Whilst the equation was developed for K_a, values experimentally determined using a portion of rat small intestine, it can be used to give an approximate range for the minimum acceptable solubility required for a given target dose. It has been reported[18] that the majority of drug substances have human K_a values in the range 0.001–0.05 min^{-1}. If these values are used in equation (39.2), together with a value for the MAD (or target dose), an approximate range for the target solubility required is calculated as:

$$\text{Minimum acceptable solubility} = \frac{\text{Target dose}}{\left(\dfrac{\mathrm{S} \times K_a}{\mathrm{MAD}}\right)} \qquad (39.2)$$

Thus, for a target dose of 70 mg, and entering the range of the human K_a, minimum solubilities between 20 μg/mL and 1 mg/mL are obtained.

B. The polarity of the solvent system

Several parameters are suitable as measures of polarity: the dielectric constant ε, the Hildebrand solubility parameter δ and its extension, the three-component (or three-dimensional) Hansen solubility parameters, which will not be discussed further. However, the most practical and widely used parameter is P_{octanol} because it is the most easily determined

TABLE 39.1 Example for the Property "Lipophilicity"

Substance **1**: R = H

Substance **2**: R = —CH_3

Parameter	Substance 1 R = H	Substance 2 R = —CH_3
Solubility in water S_w (mg/100 mL) at 25°C	210	40
$\log P_{octanol}$ (pH = 7.4) $P_{octanol}$	1.81 64.4	2.26 182
Calculated solubility in octanol (mg/100 mL) $S_{octanol} = P_{octanol}S_w$	13.5	7.28
Ratio $S_{octanol}$ (Subst. 1/ Subst. 2)	1	0.54

TABLE 39.2 Categorization of Solvents based on their Hemolytic Effect on Human Blood After Storage for 90 min at 37°C

Hemolytic effect	Concentration ratio (solvent: blood)	Solvents
Slight	>10:90	Propylene glycol Dimethyl isosorbide Diglyme
Moderate	<10:90	Tetrahydrofurfuryl alcohol N-methyl-2-pyrrolidone Glycerol formal Ethanol Solketal (glycerol dimethylketal) Glycofurol
High	<5:95	PEG 200 Acetone
Very high	<1:99	Ethyl lactate Dimethylsulphoxide

Source: Adapted from Doelker.[19]

or calculated parameter, and many useful correlations for its utilization in a number of relationships have been established. It is most commonly used in its logarithmic form, $\log P$ or $C\log P$ (the calculated value).

It is common to view $\log P_{octanol}$ as a measure of lipophilicity. This perception leads to the erroneous expectation that a compound with a high $\log P$ value should dissolve easily in lipids. However, this is not necessarily the case as the example Table 39.1 demonstrates. Substance 2, which is the methyl homolog of Substance 1, has a higher partition coefficient. Nevertheless, the calculated solubility in octanol is just about half of the lower homolog with its significantly lower partition coefficient. This is due to the much lower solubility of Substance 2 in water (only a quarter of that of Substance 1). Hence, it is more appropriate to regard $\log P_{octanol}$ as a measure of hydrophobicity, that is, the tendency to escape an aqueous environment.

III. WATER–COSOLVENT SYSTEMS

During the discovery phase, when availability of compound is at a premium, the formulation of a simple injectable product for use in animal studies can be problematical. If a drug candidate required for formulation as a parenteral product is non-ionizable, lipophilic or does not form a water-soluble salt with acceptable properties, solubilization by use of cosolvents should be considered as the next option. In cases where the candidate is ionizable, or where it is zwitterionic, a combination of use of cosolvent(s) together with optimization of the apparent pH of solution should be investigated. However, the selection of the cosolvent should be made with care; the majority cannot be injected undiluted by the intravenous (i.v.) route as they cause hemolysis (rupture of red blood cells) and pain at the injection site. In a recent study Doelker *et al.*[19] determined the hemolytic effect of 13 individual solvents on human blood and found that the solvents could be grouped into four categories, depending on the severity of hemolysis; their results are reproduced in Table 39.2. As a parallel exercise, they reviewed the toxicity data available for the same 13 solvents.[20]

Cosolvency, the addition of water miscible solvents to an aqueous system, is one of the oldest, most powerful and most popular means of solubilization, rendering thermodynamically stable homogeneous solutions. Cosolvents are used in 13% of FDA-approved parenteral products.[21] Pharmaceutically acceptable solvents are listed in Table 39.3 and their structures are given in Figure 39.2. Of these, the ethanol, propylene glycol, glycerol and polyethylene glycol are used in approximately 66% of those products. Generally, the objective is to use, in a mixture with water, a minimum proportion of cosolvent necessary for increasing the solubility of a substance enough to accommodate the dose to be administered, a solution volume of 1 mL being the ideal for an intramuscular (i.m.) injection. They are also used in oral liquid dosage forms such as syrups or drops.

TABLE 39.3 Pharmaceutically Acceptable Hydrophilic Solvents

Solvent	Remarks on use
Ethanol 96%	Moderately hemolytic cosolvent in injectables, up to 25%; used in oral dosage forms
Propylene glycol (1,2-propanediol)	In injectables up to 70%, low hemolytic effect Used in oral dosage forms
Glycerol (glycerine)	In injectables up to 30%; used in oral dosage forms
Polyethylene glycol (PEG 200, PEG 300, PEG 400, PEG 600)	Used in injectable and in oral dosage forms Can be used undiluted even for subcutaneous administration without tissue damage
Isopropanol	For topical (external) use only
Butylene glycol (1,3-butanediol)	Miscible with water or ethanol
Benzyl alcohol (phenylmethanol)	Used in injectables up to 4%; used for its preservative and local anesthetic activity Limited miscibility with water
Tetrahydrofurfuryl alcohol	Moderately hemolytic cosolvent for injectables
Glycofurol	In injectables up to 50%; moderately hemolytic
Solketal (glycerol dimethylketal)	Moderately hemolytic cosolvent for injectables
Glycerol formal	Moderate hemolytic effect. Mixture of isomers
Dimethyl-isosorbide	Low hemolytic effect
Diglyme (diethylene glycol dimethyl ether)	Low hemolytic effect
N,N-dimethylacetamide	Rarely used for injectables
N-methyl-2-pyrrolidone (pharmasolve)	Rarely used for injectables
2-Pyrrolidone	Occasionally in experimental use
Dimethylsulfoxide (DMSO)	Acceptable for investigational purpose only Although in U.S.P. XX for topical use, definitively not recommended for formulation of dosage forms intended for use in humans.

The most frequently used solvents are in bold print.

For estimating the solubility of a given solute in cosolvent-water mixtures, a simple relationship between suitable parameters of the solute and the cosolvent would be helpful. A straightforward and reliable approach for selecting cosolvents and predicting their solubilization effects on drugs that requires little or no experimental data, and thus minimal time and drug, is the log-linear model proposed by Yalkowsky *et al.*[22–24] This model describes an exponential increase in the solubility of a non-polar drug with a linear increase in cosolvent concentration. The relationship is described by:

$$\log S_{tot} = \log S_w + \sigma f_c \qquad (39.3)$$

where S_{tot} is the total solute solubility in the cosolvent-water mixture, S_w is its water solubility, σ is the cosolvent solubilization power for the particular cosolvent–solute system, and f_c is the volume fraction of the cosolvent in the aqueous mixture. The σ term can be obtained from the slope of the $\log(S_{tot}/S_w)$ versus cosolvent volume fraction (f_c) profile of each selected drug and cosolvent. This implies the need for experimental data for each cosolvent–solute system, a time and material consuming prospect. However, it was demonstrated that a linear relationship also exists between σ and the logarithm of the solute's $\log P_{octanol}$ ($\log P_{ow}$ in the equations).[25] Therefore, σ can be related to the solute and the cosolvent by the following simple relationship:

$$\sigma = S \log P_{ow} + T \qquad (39.4)$$

where S and T are constants dependent only on the particular cosolvent and $\log P_{ow}$ is the solute's partition coefficient. In other words, the only *ab initio* data required to predict the solubility of a solute is the compound's octanol–water partition coefficient. Fortunately, this value can be determined experimentally or accurately predicted by a number of computational methods ($C\log P$®, for example). A useful form of the log-linear equation is obtained by substituting from equation (39.4) into equation (39.3) to give:

$$\log S_{tot} = \log S_w + f_c(S \log P_{ow} + T) \qquad (39.5)$$

which expresses the total solubility in a mixed solvent system solely in terms of the pure components: water, cosolvent and solute – obviating the need for any individual solute–cosolvent experiments. In Figure 39.3, the solubility is plotted against the composition (per cent of volume) of the water–cosolvent mixtures for three drug substances representative for different polarities: tioconazole **1** ($\log P_{octanol} = 4.52$), caffeine **2** ($\log P_{octanol} = 0.57$), and oxfenicine **3** ($\log P_{octanol} = -2.5$).

Several features can be derived from the graphs:

1. As the polarity of a compound increases, the solubilization by cosolvents decreases, eventually turning to desolvation in case of polar substances. The non-polar

FIGURE 39.2 Structures of commonly used hydrophilic cosolvents.

FIGURE 39.3 Solubility of tioconazole, caffeine and oxfenicine in cosolvent–water mixtures. Cosolvents are ethanol (triangles), propylene glycol (black circles) and PEG 400 (hollow circles).
Source: Redrawn from.[26]

tioconazole exhibits a logarithmic increase of solubility as the proportion of cosolvent increases, whereas a polar compound's solubility decreases, following the model represented by equation (39.5).

2. For non-polar solutes, the solubilization power of the cosolvents increases in the series glycerine < PEG 400 < propylene glycol < ethanol. The lower the polarity of the cosolvent, the higher is its solubilization power for non-polar solutes. As can be seen above, this correlates well with the $\log P_{ow}$ of the cosolvents.
3. The intermediate case of caffeine does not follow this log-linear relationship, instead generally, a parabolic course is observed with semi-polar compounds. The deviation from linearity is related to the non-ideal nature of water–cosolvent interactions. To account for such deviations, equation (39.5) can be empirically modified by the addition of a quadratic term to provide a more accurate description of the data although no theoretical foundation can be given:

$$\log S_{tot} = \log S_w + \sigma' f_c - \tau f_c \qquad (39.6)$$

where σ' and τ are constants for each solvent. For the propylene glycol–water system, the coefficients $\sigma' = \log P_{ow}$ of the solute and $\tau = 0.5$ work reasonably well. In consequence, for semi-polar substances the maximum solubility must be sought experimentally. But at least the rank order of the degree of solubilization is reflected by the relative magnitude of σ values that may be calculated from the drug's $\log P_{octanol}$ value and the solvent coefficients in Table 39.4.

As exemplified by the case of oxfenicine, the solubilization approach by cosolvent mixtures appears not to be useful for polar drug substances simply because they are less soluble in non-polar solvents than they are in water. Electrolytes are at least partly ionized, and therefore they exhibit pH-dependent solubility in aqueous media, while the fraction of the less water-soluble, non-ionized species may benefit from cosolvent solubilization as shown in the example Figure 39.4. Although the solubility of dexoxadrol hydrochloride in water exceeds that of the free base by more than two orders of magnitude, a significant increase can be achieved by additional cosolvent (PEG 300), although the effect is much less pronounced than with the free base.

Finally, it should be noted that the log-linear model holds for mixed cosolvent systems in the following form:

$$\log S_{mix} = \log S_w + \sum(\sigma f_c) \qquad (39.7)$$

where the individual solubilization powers and volume fractions of each cosolvent are linearly summed.

Homogeneous solutions are the preferred formulation systems for parenteral administration because they can be easily visually inspected for the absence of particulate matter. For this reason, cosolvent solubilization is the first choice for parenteral products once purely aqueous systems provide insufficient solvency. The compositions of three commercial, injectable products are given in Table 39.5. The first product (1) has a low percentage of cosolvent in the separate solvent ampoule. The drug substance is provided as a dry powder because of its limited stability in solution. The second one (2) is solubilized with two cosolvents amounting to 50% of the total volume, whereas in the third product the drug dose is dissolved in a water-free mixture of "cosolvents." This draws the attention to a further point to consider when cosolvents are employed in formulations. The formulation has to be devised such that the effect of dilution of

TABLE 39.4 Coefficients of Cosolvents for Calculating the Solubilization Power σ According to equation (39.5)

Cosolvent	Cosolvent $\log P_{octanol}$	S	T
Ethanol	−0.31	0.93	0.38
Propylene glycol	−0.92	0.76	0.57
PEG 400	−0.88	0.73	1.19
Glycerin	−1.96	0.34	0.29

Taken from Ref. [27]

FIGURE 39.4 Solubility of the free base and the hydrochloride of dexoxadrol at 23°C in a water–cosolvent system, as a function of the percentage of the cosolvent, PEG 300.
Source: Redrawn from Ref. [28].

the drug solution by the aqueous body fluids (blood, tissue interstitial fluids, etc.) during administration is considered in order to avoid precipitation upon injection. Solubilization needs to be maintained until the drug substance is diluted to its final concentration in the circulating bloodstream or in an infusion fluid, if the product is added to an –i.v. infusion.[29] Otherwise, there is the risk that precipitated drug particles could mechanically or chemically irritate tissue, or even clog blood vessels.

A further part of the cosolvent selection process for a drug that is to be injected as a bolus involves the determination of the optimum concentration of solvent. This is a balance between the percentage that is required to maintain the drug in solution during manufacture and storage and the maximum concentration that can be safely injected. Similarly for drugs that are infused i.v. the selection process should involve a dilution study using normal saline or 5% dextrose solution with the final concentration of the cosolvent at or close to its maximum accepted concentration to check for precipitation. Generally, the optimum concentration of each cosolvent should be determined empirically, as the solvation properties of drugs differ, but data from existing marketed products can give a lead to approximate values. Some data from marketed products has been collated by Strickley[30] and Bertholom *et al.*[31]

A decision tree approach to parenteral formulation for use at the discovery stage of development has been published[32] covering compounds that are strong or weak acids, strong or weak bases or neutral entities. In addition to this paper, several useful reviews have also been published recently that specifically cover the requirement for early formulations.[33–36]

TABLE 39.5 Compositions of Cosolvent-Solubilized Drug Products

(1) Librium® for i.m. injection	Component	Amount in 2 mL
In 5-mL dry filled ampoule:	Chlordiazepoxide	100 mg
In 2-mL solvent ampoule:	HCl	
Buffering agent	Maleic acid	1.6%
	NaOH	To adjust pH
Solvents		
– Solvent and surfactant	Polysorbate 80	1.5%
	Benzyl alcohol	4%
– Solvent and preservative	Propylene glycol	20%
	Water	To volume
(2) Bactrim® solution for i.v. infusion		**in 5 mL**
Active substances	Trimethoprim	80 mg
	Sulfamethoxazole	400 mg
Preservative	Benzyl alcohol	1%
Preservative	Ethanolamine	0.3%
Antioxidant	Sodium metabisulfite	0.1%
Solvents	Propylene glycol	40%
	Ethanol	10%
	water	To volume
(3) Ativan® for i.m. injection		**in 1 mL**
Active substance	Lorazepam	2 or 4 mg
Solvent	PEG 400	0.18 ml
Solvent and preservative	Benzyl alcohol	2%
Solvent	Propylene glycol	To volume

IV. SOLUBILIZATION MEDIATED BY SURFACTANTS

Cosolvent solubilization is usually successful for drug substances with relatively low molecular weights. It is often less successful with medium to high-molecular weight compounds for which other solubilization strategies need to be evaluated. The most common of these is solubilization by non-ionic or anionic surfactants whereas cationic surfactants are rarely used as they tend to be more toxic. The preferred surfactants are non-ionic with a mix of polar and non-polar domains in their structures; these are termed amphiphilic (or amphipathic) surfactants. In aqueous solution they are not randomly distributed throughout the solution; instead, they are concentrated, and adopt a specific orientation, at interfaces according to the nature of the polarity of the molecules and that of the interface. At the air–water interface their polar region is oriented toward the aqueous surface while the non-polar region is pointing outwards into the low-polar gas phase (Figure 39.5a). When hydrophobic drug substance particles are stirred into and suspended in the solution, the non-polar region of the surfactant molecule attaches to the surface of these particles while the polar group maintains intimate

FIGURE 39.5 Formation of surfactant micelles and micellar solubilization.

contact with the aqueous bulk phase. At low concentrations, surfactants are therefore used as wetting agents. As the concentration in aqueous solution is increased, it exceeds a critical value, the critical micelle concentration (CMC), and the surfactant molecules assemble cooperatively to form soluble, structured aggregates (Figure 39.5b). The CMC value is a characteristic of a particular surfactant and depends on environmental factors such as temperature, pH, ionic strength and others. Surfactant micelles are in a dynamic equilibrium with the surfactant monomers. Therefore, the number of molecules constituting a micelle is distributed around an average value over time and over a population of micelles. For example, the average aggregation number $\bar{n}$ in micelles of sodium laurylsulfate is 64 at 25°C; for other surfactants this number varies over several orders of magnitude from 2–8 (bile salts) to many thousands. With a value of $\bar{n}$ as that of sodium laurylsulfate, the shape of micelles is likely to be spherical; as $\bar{n}$ increases the shapes tend to be disks, rods, sheets or even vesicles. Surfactant micelles can accommodate solute molecules in a number of orientations, arranged according to their polarity and hydrophobicity: (i) at the interface between the micelle surface and the aqueous bulk phase; (ii) between the polar hydrophilic head groups; (iii) within the hydrophobic part (the palisades layer) of the micelle shell (Figure 39.5c) or (iv) in the core (Figure 39.5d). This is the mechanism of micellar solubilization.

The totality of micelles represents a colloidal phase, into which a substance is "dissolved" in the aqueous phase partitions. The capacity of the micellar phase to solubilize a solute can therefore be expressed as a partition coefficient K_m. Hence, a linear relationship can be expected between the concentration S_{mic} of substance solubilized by micelles and the concentration of the surfactant C_{sfc} in the system. Because only micelles contribute to the solubilizing effect but not the monomeric surfactant molecules, the critical micelle concentration C_{cmc} must be subtracted from the total of the surfactant concentration. The resulting total concentration of solute in the micellar solution is then:

$$S_{tot} = S_w + S_{mic} \tag{39.8}$$

where

$$S_{mic} = K_m\ (C_{sfc} - C_{cmc}) \tag{39.9}$$

and

$$K_m = \frac{S_{mic}}{S_w} \tag{39.10}$$

The solubilization of chloramphenicol with one member each of four classes of surfactants is shown in Figure 39.6 while Table 39.6 gives the key parameters of the surfactants.

The typically linear relationship in micellar solubilization is obvious in the graph Figure 39.6. In Table 39.6, the respective data reflecting the solubilization power of the surfactants are presented.

FIGURE 39.6 Solubilization of chloramphenicol by surfactants. Squares, cetrimide; black circles, sodium laurylsulfate; empty circles, polysorbate 80; triangles, PEG 40 monostearate.
Source: Compiled and redrawn from Ref. [37].

TABLE 39.6 Parameters of the Solubilization of Chloramphenicol by Surfactants.[37] Solubilization of chloramphenicol by surfactants

Surfactant	Solubility (g/g solubilizer)		K_m (g/g)	
	25°C	35°C	25°C	35°C
Cetrimide	0.3386	0.393	96	82
SDS	0.135	0.2263	40	50
Polysorbate 80	0.0875	0.1224	24	25
Myrij 52	0.0634	0.082	19	19

The hydrophile/lipophile balance (HLB) system for the characterization of surfactants, devised by Griffin,[38] was originally developed for selecting non-ionic surfactants as emulsifiers for stabilizing oil-in-water (o/w) emulsions. According to the ratio of polar and non-polar groups in the molecule, a surfactant is assigned a number, the HLB value, on a scale originally ranging from 1 to 20 with the most lipophilic surfactants having the lowest number,

FIGURE 39.7 Use and behavior in water of surfactants, according to HLB ranges.

and the more hydrophilic surfactants having higher numbers. Figure 39.7 gives an overview of the functions of surfactants within the different ranges in the HLB classification system and also the visual impression of the suspensions or solutions of the surfactants in water. Later inclusion of ionic surfactants in the system required an extension of the scale beyond 20. The HLB values correlate with polarity-indicating parameters such as the $\log P_{\text{octanol}}$ and the dielectric constant. The surfactants that are most frequently used in pharmaceutical formulations are listed in Table 39.7 by their HLB values, and their structures shown in Figure 39.8.

To illustrate how these surfactants are used in typical formulations two different compositions of the lipid soluble vitamin K compound phytomenadione are given in Table 39.8. In (4) a non-ionic surfactant is used, in (5) lecithin–bile acid mixed micelles are the solubilizing principle.

Surfactants must be chosen with due care, as their strong interfacial activity can elicit adverse reactions in biological systems. Cell membrane constituents (phospholipids, cholesterol, etc.) can be solubilized at supermicellar concentrations, impairing the integrity of membranes. This disruption of the membrane enhances the permeability of drug substances and other substances present in the extracellular fluid. At low concentrations such alterations are reversible and membranes recover rapidly, but higher concentrations

TABLE 39.7 Surfactants Most Frequently Used in Pharmaceutical Formulations

Surfactant	[b]	Chemical name or other description	Type[c]	HLB	CMC	Use[e]
Glycerol monostearate	10		N	3.8		po, par
Span 80	11	Sorbitan mono-oleate	N	4.3		po, ex
Lecithin	10	From soybeans or egg yolk	Z	7–10		po, par, ex
DPPC		Dipalmitoyl phosphatidylcholine			0.47 nM	
Cremophor EL		Polyethylene glycol 35 castor oil	N	12–14	0.02 wt.%	po, par
Cremophor RH 40		Polyethylene glycol 40 hydrogenated castor oil	N	14–16		po, par
Solutol HS 15	10	Polyethylene glycol 660 hydroxystearate, mono- and di-esters	N	14–16	0.005–0.02 wt.%	po, par
Vitamin E TPGS	10	D-α-tocopheryl polyethylene glycol-1000-succinate	N	15–19	0.02 wt.%	po, par
Tween 80[a]	11	Polysorbate 80, polyethylene glycol 20 sorbitan mono-oleate	N	15	20 mM	po, par
Myrij 52[a]	11	Polyethylene glycol 40 monostearate	N	16.9		po, ext

(Continued)

TABLE 39.7 (Continued)

Surfactant	[b]	Chemical name or other description	Type[c]	HLB	CMC	Use[e]
Brij 35[a]	11	Polyethylene glycol 23 lauryl ether	N	16.9	0.09 mM[d]	po, ext
Bile salts	10 10	Deoxycholate Na Taurocholate Na Taurodeoxycholate Na Glycocholate	A	20–25	[d] 1.5 mM 3.3 mM 2.7 mM 7.1 mM	po, par
Pluronic F-68	11	Polyethylene glycol-polypropylene glycol block copolymer	N	29	190 mg/L	po, par
Sodium laurylsulfate	11	Dodecylsulfate sodium	A	40	8.2 mM	po, ext
Docusate sodium	11	Dioctylsulfosuccinate sodium	A	40	2.5 mM	po, ext
Cetrimonium bromide	11	Cetyltrimethylammonium bromide	C	40	[d]1 mM	ext

[a]Typical members of large families of surfactants with the number of polyoxyethylene units and the ester acyl moieties varied.
[b]Structural formula in Tables 10 or 11, respectively.
[c]A, anionic; C, cationic; N, non-ionic; Z, zwitterionic.
[d]50 mM Na^+.
[e]ext, topical; par, parenteral, po, oral applications.

FIGURE 39.8 Structures of frequently used surfactants. *Remark:* The "Tweens" are derivatives of the same sorbitol anhydride ring structures including isosorbide (not shown) as shown for the Span 80.

FIGURE 39.8 (*Continued*)

can cause irreversible damage to mucosa and other tissues. Some of the polyglycolized surfactants used as solubilizers in parenteral and oral applications, in particular Cremophor EL®, have been reported to occasionally elicit anaphylactic reactions.[39,40] Cremophor EL® is used in conjunction with ethanol as a solvent system for paclitaxel in Taxol®, a valuable and very effective anticancer drug. Severe reactions to Cremophor EL® in a significant percentage of patients[41] has resulted in patients having to be pre-dosed with steroids and antihistamines prior to chemotherapy.

Cationic surfactants are bactericidal agents and are used as antiseptics and disinfectants in topical and other external preparations and are never used in parenterals. Lecithins and other natural phospholipids, as one of the main constituents

TABLE 39.8 Compositions of Surfactant-Solubilized Drug Products Containing Phytomenadione

(4) AquaMephyton® solution for s.c./i.m. injection		in 1 mL
Active substance	Phytomenadione	2 or 10 mg
Solubilizer	Polyoxyethylated fatty acid derivative	70 mg
Preservative	Benzyl alcohol	0.9%
To adjust isotonicity	Dextrose	37.5 mg
Solvent	Water	To volume in 1 mL
(5) Konakion® for i.m. injection		
Active substance	Phytomenadione	2 mg
Solubilizer	Soy lecithin glycocholic acid	
NaOH/HCl		To adjust pH
Solvent	Water	To volume

of all biological membranes, are the surfactants and solubilizers best tolerated in parenteral and oral applications.

In order to keep the use concentrations low, the solubilizing power of one surfactant is rarely utilized alone. Thus, combinations of two or more surfactants are used to optimize the solubilization effect by the formation of mixed micelles, especially with lecithin and bile acids. For injectable solutions mostly several solubilization principles are combined as can be seen in the following examples:

1. By employing pH adjustment and 5% sodium dodecyl sulfate, the solubility of the investigational cancer drug pyroxamide was increased by a factor of nearly half a million over its intrinsic solubility.[42] Although the micellar solubilization can be demonstrated *in vitro* there is no need for a corresponding enhancement of bioavailability by the oral route in any case. It should be noted that with particular combinations of drug substance and surfactant, reduced instead of improved bioavailability has been experienced. Such effects depend on the concentration of the surfactant and may have various reasons such as shift of equilibrium in favor of micellar inclusion or formation of other types of drug–surfactant aggregates.
2. The presence of hydrophilic and hydrophobic regions in the same molecule is not a monopoly of surfactants; many drug molecules exhibit hydrophilic and hydrophobic regions as well, and this is indeed the structural requirement for their ability to permeate cell membranes and pass through absorption barriers. The amphiphilic properties are obvious in some drug classes, for example, the local anesthetics.[43]
3. On some individual occasions a drug substance may exceed the CMC, whereupon micellar "self-solubilization" can occur. A recent example has been described by Hussain *et al.* who found that the hydrochloride of the basic analgesic DuP 747 is surface active; the saturated solution (3 mg/mL at 22°C) lowered the surface tension of water to 50 dyne/cm. It was hoped that a different salt might be soluble enough to reach higher concentrations exceeding a CMC. This was indeed the case: the CMC of the methanesulfonate was found to be *ca.* 4 mg/mL and a micellar solubility of 60 mg/mL could be achieved.[44]

V. SOLUBILIZATION BY LIPID VEHICLES

Lipid soluble drug substances like the fat soluble vitamins and steroids can be readily dissolved in lipid vehicles such as natural triglycerides (i.e. purified fatty oils), semi-synthetic triglycerides such as the medium-chain triglycerides (e.g. Miglyol 812, consisting of the C_8–C_{12} fatty acid triglycerides) and synthetic fatty acid esters (e.g. isopropyl myristate) for oral administration. Administration by the parenteral route is restricted to subcutaneous (s.c.) and i.m. injection. When given i.v. oils would form large droplets leading to fatal fat embolism; therefore homogeneous lipid phases must never be injected by the intravascular route. However, this problem can be circumvented if the oil is finely dispersed to a stable droplet size that can pass through narrow capillaries.

A. Emulsions and microemulsions

An emulsion is defined as a dispersion of two immiscible liquids, one of which is finely subdivided and uniformly distributed as droplets (the dispersed phase) throughout the other (the continuous phase). A third component (or multiple additional components), the emulsifying agent(s), is necessary to help stabilize the emulsion. The emulsifying agent(s) coats the droplets and prevents droplet coalescence by either reducing the interfacial tension or by creating a physical repulsion between the droplets.[45] The dispersed phase is occasionally also defined as the internal phase; the continuous phase is occasionally also defined as the external phase or dispersion medium. Virtually all emulsions are inherently physically unstable.

There are three main types of emulsion:

1. Oil-in-water emulsions (o/w), where the continuous phase is aqueous. These are normally used for oral or parenteral products.
2. Water-in-oil emulsions (w/o), where the continuous phase is an oil. These are normally used for topical products.

3. Multiple emulsions that can be either w/o/w or o/w/o and comprise droplets of one phase inside droplets of the second phase, with the continuous phase the same as phase one. These are often used for delayed release products.

Emulsions are usually used as a means of administering drugs that are insoluble, or relatively insoluble, in aqueous media by dissolution of the drug within the oil phase.[46–49] Emulsion systems play their most prominent role in products for topical applications (e.g. creams and ointments); they are also prepared for oral and parenteral administration. They may also have the advantage of preventing hydrolysis of the drug or uptake of the drug by infusion sets.[50] Most emulsions have droplet sizes in the range 0.1–100 μm, though microemulsions with droplet sizes below 0.05 μm (50 nm) can be produced with relative ease; 100 μm represents a relatively coarse emulsion. It is recommended that emulsions developed for the i.v. route of administration should have a sub-micron droplet size, although emulsions with a 10 mm droplet size have been used parenterally.[51,52]

For parenteral nutrition drug-free emulsions are used. A typical composition of a nutritional fat emulsion for i.v. administration is 200 g soybean oil or safflower oil, 12 g egg phospholipids and 25 g glycerine per liter of emulsion with the pH adjusted to within the range 5.5–8. It is most important that the fat droplet particle size that does not exceed 0.5 μm. For experimental purposes, such fat emulsions (e.g. Abbolipid®, Intralipid®, Lipofundin®) are well suited and used as early dosing vehicles for lipophilic drug candidates.

Depending on the drug properties, emulsion formulations require significant optimization effort in tailoring the oily phase and emulsifiers. As an example, for the i.v. administration of the poorly soluble antimitotic agent paclitaxel in Taxol®, large volumes of a micellar and cosolvent-mediated dispersion consisting of Cremophor EL + ethanol + isotonic saline 5 + 5 + 90 have to be infused to reach the therapeutic dose of 30 mg because the solubility of the drug is only 0.6 mg/mL in that vehicle. Intralipid as an infusion vehicle was ruled out (solubility of paclitaxel in soybean oil is 0.3 mg/mL). Therefore, an emulsion was developed based on triacetin as the oil phase (solubility of paclitaxel in triacetin is 75 mg/mL). An optimized "Taxol emulsion" formulation consists of 1.5% soy lecithin, 1.5% poloxamer 188, 2% ethyl oleate, 50% triacetin, and 10 mg taxol/mL.[46]

Emulsions may also be used for drug targeting. Formulations with a droplet size between 100–200 μm usually result in a high liver uptake after i.v. injection.[53,54] Emulsified drugs may be diverted away from the liver upon i.v. injection by incorporating poly (oxyethylene) surfactants as part or all of the emulsification system. Diversion from the liver prolongs the half-life of the drug, allowing passive targeting to the lung, kidneys and areas of inflammation. Active targeting may be achieved by the conjugation of antibodies to the distal ends of the poly (oxyethylene) chains, provided that the emulsion droplets have a sub-micron droplet size.[55] Often, drug toxicity may be reduced by formulation as an emulsion. The use of a w/o emulsion of amphotericin deoxycholate, as opposed to a simple solution, was shown to reduce the incidence and severity of renal impairment and chills in patients, whilst still maintaining the antifungal efficacy of the drug.[56]

In contrast to ordinary emulsions, microemulsions are homogeneous, transparent and thermodynamically stable; they can be considered as swollen micelles. Microemulsions usually form spontaneously when the components are mixed at their optimized ratios. It is essential that, for optimum ease of formation and subsequent stability, an interfacial tension as low as possible is achieved and maintained between the phases. Generally, this is not accomplished with just a single surfactant; a second amphiphile, typically a medium-chain-length alcohol, as the so-called cosurfactant needs to be added. Microemulsions have several advantages over coarse emulsions, especially in respect of transparency and stability, however, they require significantly higher percentages of surfactants for formulation, thus restricting the choice of acceptable components.[57]

The pharmaceutical significance of microemulsions lies in the fact that they are not merely solvents that solubilize drug substances that are difficult to dissolve in other solvents, but serve as systems that impart high penetration power to a drug product. For this reason, they have gained increasing interest. Limiting factors for their use on a broad scale is the need for well-tolerated components, particularly surfactants and cosurfactants, whose concentrations are higher than in ordinary emulsions. Furthermore, thermodynamic stability must be maintained over a temperature range between 4°C and 40°C. Drug carrying microemulsions so far have been used in the topical and transdermal field, but with the natural phospholipids the principle can be applied even to parenteral administered products. The breakthrough in oral drug delivery was achieved with the successful attempts of Kissel and Posanski to increase the bioavailability of cyclosporine-A by the microemulsion principle.[58] A solution formulation for an oral microemulsion with a drug content of 10% cyclosporine-A is composed of corn oil mono-, di-, and triglycerides, PEG-40 hydrogenated castor oil, propylene glycol and anhydrous ethanol. This product, replacing the former self emulsifying drug delivery system (SEDDS) (see below), is actually a concentrate, which upon dilution with an aqueous phase generates the effective microemulsion. This principle of anhydrous concentrates was named "SMEDDS" (self micro-emulsifying drug delivery system). Such formulations lack the aqueous phase. On dilution, a SMEDDS spontaneously converts to an optically clear, thermodynamically stable microemulsion that contains the drug in molecular dispersion. The same principle of a water-free concentrate when leading to a "macro"-emulsion

TABLE 39.9 Relationship Between Droplet Size and Appearance

Droplet Size, μm	Appearance
<0.05	Transparent
0.05–0.1	Grey, translucent
0.1–1	Blue-white (Tyndall effect)
>1	Milky

is called SEDDS. A review on self-dispersing lipid-based systems has been published recently.[59]

As stated above, emulsions are essentially swollen micellar systems. The differences between a micelle, containing solubilized oil, and an emulsion, comprising an oil droplet surrounded by an interfacial layer composed largely of surfactant, are difficult to assess. The droplet sizes of the dispersed phase of an emulsion can be estimated approximately by its appearance as seen in Table 39.9.

Also, liposomes (see later) are very similar but are formed by the self-assembly of phospholipid molecules in an aqueous environment. The amphiphilic phospholipid molecules form a closed spherical bilayer in an attempt to shield their hydrophobic groups from the aqueous environment, whilst still maintaining contact with the aqueous phase via the hydrophilic head group. The resulting closed sphere may encapsulate water-soluble drugs within the inner aqueous compartment or may encapsulate lipid soluble drugs within the bilayer membrane. Alternatively, lipid soluble drugs may be complexed with a hydrophilic cyclodextrin and then encapsulated within the liposome aqueous compartment.[60]

1. *The choice and classification of emulsifying agents*

Emulsification agents for parenteral emulsions are often chosen on the basis of their low toxicity and irritancy, in addition to their emulsification ability. Whilst there is no list of emulsifying agents specifically approved for use in pharmaceutical products,[61] there is one of approved emulsifiers for use as food additives within the European Union. It is widely assumed that emulsifiers from this list are suitable for pharmaceutical products intended for internal use. The list comprises naturally occurring materials, together with some of their semi-synthetic derivatives, such as polysaccharides, glycerol esters, cellulose ethers, sorbitan esters and polysorbates. These are largely non-ionic and are less irritant and less toxic than their anionic derivatives; cationic derivatives are usually significantly more toxic than the corresponding anionic derivatives.

Typical surface active agents are as follows:[57]

1. Non-ionic agents
   - Polyoxyethylene sorbitan monooleate (polysorbate 80)
   - Sorbitan esters (e.g. the monooleate)
   - Glycerol esters (e.g. the monostearate)
   - Poloxalkols (polyoxyethylene/polyoxypropylene copolymers)
   - Higher fatty alcohols (e.g. cetyl and stearyl alcohols)
   - Hexaoxyethylene monohexadecyl ether
2. Amphoteric (ampholytic) agents
   - Lecithin
   - *N*-dodecyl alanine
3. Anionic agents
   - Sodium dodecanoate
   - Sodium dodecyl sulfate
   - Sodium dioctyl sulphosuccinate
   - Calcium oleate
   - Amine soaps (based on triethanolamine, for example, triethanolamine stearate)
4. Cationic agents
   - Hexadecyl trimethyl ammonium bromide (cetrimide)
   - Dodecyl pyridinium iodide

As already mentioned, cationic surfactants are too toxic for parenteral administration and are restricted for use only in dermal antiseptic creams.[61]

2. *Other formulation additives*

Other formulation additives may be added when the emulsion is intended for topical use; these may include humectants (such as glycerol, PEG or propylene glycol), sweetening agents (sucrose, sorbitol or glycerol), or flavors, coloring agents and perfumes. These are unlikely to be used in parenteral formulations. Occasionally, parenteral emulsions may require the use of buffers, microbiological preservatives, antioxidants or density modifiers with adequate caution.[61] It may be necessary to add a buffer system to improve chemical stability, control tonicity or to ensure physiological compatibility. However, as buffers are electrolytes, they will generally significantly reduce the physical stability of the emulsion.

Microbiological preservatives and/or antioxidants may be necessary, especially if a multi-dose product is envisaged, or the oil used in the formulation is contaminated with peroxide. Care must be taken with both types of additive, as it is often the case that it partitions between the two phases and some of the required activity is lost. Activity may be also lost if there is incompatibility with other excipients, or there is adsorption onto the container/closure system. The choice of either component must therefore be checked thoroughly by effectiveness testing of the final product during the pharmaceutical development process.

Sodium metabisulfite and EDTA have been used individually as microbiological preservatives in different approved formulations of propofol injectable emulsion. The choice of antioxidant must be made with care as some are restricted by use and/or concentration in different countries. BHA is widely used in fixed oils and fats at concentrations up to 0.02% but is permitted up to 0.1% in some essential oils. Its close relative, BHT, is recommended as an alternative to tocopherol at concentrations up to 10 ppm for the stabilization of liquid paraffin. Other antioxidants widely used in emulsion formulations include the propyl, octyl and dodecyl esters of gallic acid at concentrations up to 0.001% (fixed oils and fats) and up to 0.1% (essential oils).[61]

Density modifiers may be used to adjust the density of either phase to reduce "creaming." Sucrose, dextrose, glycerol or propylene glycol can be used with care for minor adjustments to the density of the aqueous phase. However, due to differences in the coefficient of expansion, they can only be used over a small temperature range.

3. *Stability of emulsions*

a. Physical stability

A stable emulsion is one where the globules retain their initial character and remain uniformly distributed throughout the continuous phase. The addition of emulsifying agents results in the formation of an interfacial film around each of the dispersed droplets; the physical nature of this film creates a barrier that controls eventually the coalescence of the droplets that approach one another. If this film is electrically charged through the use of charged surfactants, repulsion occurs before contact is made and improvements in stability are normally seen.

i. Cracking or breaking This is the separation of an emulsion back to its two constituting coherent phases and is usually caused by agents that destroy the interfacial film. Some of the factors causing cracking are:

1. The addition of agents with charges opposite to those of the film; these can include surfactants, electrolytes or iodized drugs with the opposite charge.
2. Bacterial growth
3. Wide temperature changes, including freeze–thaw cycles.

ii. Flocculation This is an intermediate phenomenon, not to be confused with creaming, that is likely to occur in most pharmaceutical emulsions if there are weaknesses in the interfacial film. The droplets aggregate due to a shift in the interaction between attractive and repulsive forces, without fully coalescing; shaking can normally re-disperse them.

iii. Creaming Creaming is primarily due to density differences between the two phases and many emulsions cream on prolonged standing. The disperse phase (normally oily) is less dense and rises to the top and forms a layer of more concentrated emulsion. Flocculation may occur in addition to creaming. Droplets of the creamed layer do not normally coalesce and can usually be re-dispersed by gentle shaking. Although it is not such a serious problem as cracking, it often leads to the formation of a product that is less elegant with the added danger of incorrect dosing.

iv. Phase inversion Normally the relative volumes of each phase contribute to the type of emulsion formed. It is normal to have less than 50% of the disperse phase but there are numerous examples where this volume exceeds 50%. Occasionally, attempts to significantly increase the amount of disperse phase result in cracking or phase inversion (an o/w emulsion switches over to a w/o emulsion).

b. Chemical stability

As stated earlier, ionic emulsifying agents are incompatible with materials or agents with the opposite charge. The presence of electrolytes can markedly affect the stability by either reducing the energies involved in the interactions between adjacent globules or by promoting salting out. High concentrations of electrolytes can strip the hydration layers from emulsifying agents, causing their precipitation. Changes in pH may also exert a significant effect by converting an ionised surfactant, for example, a sodium salt of an anionic surfactant, to its free acid or base form.

Emulsions by their nature usually contain high concentrations of oils or fats of natural (animal or vegetal) or semi-synthetic origin. These may contain peroxide or may be susceptible to the production of peroxides by the action of atmospheric oxygen or microorganisms unless due care is taken. Products can be formulated with antioxidants and/or microbiological preservatives, if considered necessary (see below). Obviously, if the active ingredient has been shown during the preformulation evaluation to be susceptible to peroxide oxidation, consideration should be given to the use of excipients with minimal propensity of peroxide formation, or the formulation should be adequately protected by antioxidants.

c. Microbiological contamination and stability

The mixture of excipients that constitute an emulsion forms an excellent substrate for the growth of bacteria and fungi. Degradation due to microbial contamination can markedly affect both the physical and chemical stability resulting in color and odor changes (rancidification), the production of gases (mainly carbon dioxide), hydrolysis of any ester

groups in the fats and oils, pH changes in the aqueous phase (due to the production of organic acids) and a general breakdown of the emulsion.

4. Drug release from emulsion products

Emulsions are commonly used to deliver drugs by the oral, rectal and topical routes. Drug delivery rates are usually high due the high surface area of the droplets of the dispersed phase.[50,62–68] They are also a convenient formulation for use in i.v. feeding even though the choice of emulsifying agent is rather limited and the average globule size must be no greater than 4 μm.[61] However, to be on the safe side it is generally recognized that for drug delivery by the i.v. route, the average droplet size should not be greater than 1 μm as particles greater than 5 μm cause capillary embolism.[52]

It has been demonstrated that the gastrointestinal absorption of highly lipophilic substances can be enhanced impressively when given in lipid vehicles by entering the lymphatic route of fat absorption "piggy-back" where bile acids are the potent natural solubilizers. Charman and Stella have suggested that drugs require a $\log P_{\text{octanol}}$ in excess of 5 and a triglyceride solubility of at least 50 mg/mL before mesenteric lymphatic transport is likely to become a major contributor to bioavailability.[69,70] Simple solutions of the antimalarial halofantrine hydrochloride (solubility in water: ~1 μg/mL, calculated $\log P = 8.5$) in triglycerides increased the lymphatic uptake of the drug compared to an aqueous suspension. The long chain glycerides (peanut oil) were most effective and surpassed by far the medium-chain triglycerides (fractionated coconut oil) and the still less effective tributyrin.[71]

Emulsions can also be used as sustained release products by reversing the nature of the emulsion from o/w to w/o. The i.m. injection of w/o emulsions of some water-soluble vaccines results in a slow release of the antigen, a greater antibody response and hence an immunity with longer duration.[61]

Emulsions may also be used to reduce the toxicity of the drug substance and the use of a w/o emulsion, rather than an aqueous solution can reduce the incidence and severity toxic effects.

B. Liposomes

Liposomes are formed by the self-assembly of amphiphilic phospholipids, possessing a hydrophilic, ionised phosphate group and a long hydrophobic fatty acid chain, in aqueous solutions. In solution the phospholipid molecules form a closed bilayer sphere in an attempt to shield their hydrophobic groups from the polar aqueous environment while still maintaining contact with the aqueous phase via the hydrophilic phosphate group. The resulting sphere may encapsulate water-soluble drugs within the central aqueous compartment; alternatively, lipid soluble drugs may be encapsulated within the hydrophobic bilayer membrane. If the drug is toxic, there is a benefit as they are screened off from tissues until the liposome is destroyed or disintegrates. The encapsulation of these drug molecules within liposomes can often alter the pharmacokinetic behavior of the drug; this often gives the formulator the ability to achieve targeted therapies. The surface of liposomes can be modified in order to optimize drug targeting to specific tissues. Such alterations have led to the formation and current availability of four different classes of liposomes.

There is a large number of naturally occurring lipids but not all are available commercially. The Lipid Library[72] is a useful source of general information on most classes of lipids, fatty acids and eicosanoids. In addition to this, a more specific database, the Lipid Metabolites and Pathways Strategy[73] (LipidMAPS), provides a classification system for all known lipids and gives information on the structures of these lipids.

1. Conventional liposomes

These are defined as liposomes that are typically only composed of neutral or negatively charged phospholipids and a cholesterol-like molecule. They may vary in size, composition, surface charge, etc. but are characterized by their relatively short lifetime in the circulatory system because various cells take them up, removing them from circulation. The use of liposomes as anticancer drug delivery systems was originally hampered because they are rapidly cleared from the systemic circulation and largely taken up by the liver macrophages. Liposome macrophage uptake by the liver and spleen (the reticuloendothelial system, RES) restricted the use of liposomes as efficient drug delivery systems for over 20 years.

2. Long circulating liposomes

Two pivotal discoveries enabled the full potential of liposomes to be realized. The first was the discovery in the late 1980s and early 1990s that the presence of additional molecules such as poly(oxyethylene) bonded onto the liposome surface decreased their clearance by partially preventing liver and spleen uptake of i.v. injected liposomes. These are now often referred to as "stealth liposomes" as this effect enables them to escape recognition by the liver and spleen with the benefit of long circulation times. This observation was made almost in parallel with the discovery that polystyrene nanoparticles coated with a poly(oxyethylene) polymer also showed reduced liver and spleen uptake. It is likely that a reduction in the coating of these liposomes by plasma proteins (opsonization), and their reduced aggregation in the blood are responsible for the increased circulation time of

these stealth liposomes. Liposome size has been shown to affect biodistribution. A size of between 70 and 200 nm is necessary to achieve prolonged circulation times with stealth liposomes. Stealth liposomes of about 100 nm in size, on i.v. administration passively target solid tumors by extravasation, due to the disorganised tumor vasculature, into their extracellular space.

The second discovery that enabled liposomes to be developed into efficacious anticancer drug delivery systems was the development of an efficient loading procedure that resulted in liposomes with high drug concentrations. This in turn permitted a reduction in the amount of phospholipid to be administered concomitantly with a given drug dose. This advanced technique was developed during formulation studies on doxorubicin that eventually led to a liposomal formulation that is now licensed as Caelyx® for the treatment of Kaposi's sarcoma and ovarian cancer.

3. *Immunoliposomes*

The production of stealth liposomes, which are not rapidly cleared by the liver and spleen, made the active and specific targeting of drugs a real possibility. In an attempt to increase the specificity, various antibodies have been conjugated to the surface of stealth liposomes to produce immunoliposomes for active targeting, as opposed to the passively targeted species discussed above. These antibodies can selectively bind to specific sites, where there is a preponderance of antibody receptors and hence can be designed to avoid healthy tissue. However, the effectiveness of the antibody in helping to target the liposome appears to be directly related to the concentration and molecular weight of the poly(oxyethylene) on the surface. A high poly(oxyethylene) density and poly(oxyethylene) with a molecular weight in excess of 2,000 reduces the directing ability of the antibody by shielding the antibody from its target.

4. *Cationic liposomes for gene delivery*

The use of cationic liposomes as gene delivery systems was pioneered in the late 1980s when *in vitro* studies demonstrated that the complexation of genes with liposomes promoted gene uptake by cells *in vitro*. Such liposomes are prepared from phospholipids with a hydrophilic amine head group (e.g. phosphatidylcholines, phosphatidylethanolamines or phosphatidylserines). The amines may be either quaternary ammonium, tertiary, secondary or primary and the liposomes prepared in this way are commonly referred to as cationic liposomes, because they possess a positive surface charge at physiological pH. These bind effectively to the negatively charged DNA. Since this time, cationic liposomes of varying description have been used to promote the cellular uptake of DNA with resultant therapeutic protein expression by various organs *in vivo*.

However, following i.v. injection, cationic liposomes are rapidly cleared from the blood. The knowledge gained in the development of anticancer liposomes, such as the prolongation of liposome circulation using poly(oxyethylene) lipids, may be applied in the development of liposomes for gene delivery in order to produce efficient carriers. This strategy should aid liposome targeting in gene delivery. Thus far, targeting efforts with cationic liposomal gene formulations have been few and far between. Clinical trials with liposomal gene delivery have been initiated in cystic fibrosis patients but they have concentrated on administration via non-parenteral routes for the obvious reason that the disease manifests itself in the alveolar epithelium. Although the experimental data indicate that cationic liposomes are able to facilitate the transfer of DNA into live mammalian cells, there are still major problems that need to be overcome in order to approach the ideal. In the future, it may be possible to treat this condition by the i.v. administration of the replacement gene–liposome complex.

VI. NANOPARTICLES AND OTHER NANOCOLLOIDAL TECHNOLOGIES

Huge advances have been made over the last 15 years in the research, development, application, commercialization and understanding of the mechanisms of action of nanoscale drug delivery systems. Many of these advances have had a spectacular impact on the prevention, diagnosis and treatment of a wide range of diseases. This is especially true in the treatment of many cancers, and it is in the oncology field that the majority of the formulation advances have been made. Pharmaceutically, the most important nanotechnologies are exemplified and discussed below (see Table 39.10 for definitions and descriptions).

Injectable nanocrystals, dendrimers, aquasomes, polyplexes and lipopolyplexes are not discussed in detail as their use is not yet as common as the main drug delivery systems. An excellent, comprehensive review of the properties and use of these colloidal systems, together with other similar technologies, (collectively often referred to as Nanomedicines) for use in medical imaging, has been published recently.[74]

1. *Polymeric micelles*

These are stable micelles that are formed with polymeric surfactants. Amphiphilic block copolymers such as the pluronics (polyoxyethylene–polyoxypropylene block copolymers) are able to self-assemble into polymeric micelles and hydrophobic drugs may be solubilized within the core of the micelle or, alternatively, conjugated to the micelle-forming polymer. Although micelles are rather dynamic systems that continuously exchange units between the micelle structure and the free units in solution, those composed of polyoxyethylene – poly(aspartic acid) have been found sufficiently

TABLE 39.10 Micro- and Nano Drug Delivery Systems

Liposomes	
Vesicles consisting of a phospholipids bilayers and aqueous core. Spontaneous formation from hydrated phospholipids at temperatures above transition temperature. Three main types:	
1. Multilamellar vesicles several lipid bilayers separated from each other by aqueous layers	Size range: few hundred nm – several μm
2. Small unilamellar vesicles (SUV) single bilayer	Size: <100 nm
3. Large unilamellar vesicles (LUV) single bilayer	Size: >100 nm
Hydrophobic drug molecules trapped/dissolved within lipid bilayer; hydrophilic drug molecules or dissolved/suspended aqueous core phase Lipid surface can be modified by covalently bonded targeting ligands or polymers, yielding "stealth" variants	
Nanoparticles (nanosuspensions, injectable nanocrystals)	
Solid particles	Size range: tens – hundreds nm
Two main types, consisting of	
1. Downsized extremely sparingly soluble synthetic drug substances: Nano-milled drug substance particles surface coated with polymer(s) for stabilization.	
2. Synthetic natural polymers (e.g. albumin or collagen) Drug substance dissolved, trapped, encapsulated in or attached to the polymeric matrix Drug release rate may be controlled Surface may be modified by covalently bonded biological materials, specific targeting ligands or polymers for directing to specific body sites, organs or cells.	
Polymeric micelles	
Micelles	Size: <50 nm
Spheroidal structures spontaneously formed in solution from amphiphilic (AB type) copolymers or block copolymers (ABA type), where A = hydrophobic, B = hydrophilic Hydrophobic core shielded from water by a layer of hydrophilic groups. Hydrophobic drugs/diagnostic contrast agents physically trapped in core or covalently linked to polymer	
Dendrimers	
Macromolecules	Size: a few nm
Highly tree-like branched structures emanating from a central core Polymer growth is initiated at the core molecule; diameter controlled by growth control. Voids in the structure and folding of the branches provide cavities and channels trapping drug molecules/particles Drugs and diagnostic agents can be covalently attached to surface groups.	
Aquasomes	
Inorganic solid particles	Size: 60–300 nm
Spherical shape, consisting of nanocrystalline calcium phosphate or ceramic diamond covered with a polyhydroxylated oligomeric film. Drugs and antigens are adsorbed onto the particle surface.	
Polyplexes and Lipopolyplexes	
Aggregates	Size: variable
Consisting of nucleic acids and polycations form spontaneously between nucleic acids and polycations, cationic liposomes, or polycations conjugated to targeting ligands or hydrophilic polymers. Used for transfection. Shape, size distribution and transfection capability of these complexes depend on composition and charge ratio of nucleic acid to that of the cationic lipid/polymer. Effective polycations: poly-L-lysine, linear and branched poly(ethyleneimmine), poly(amidoamine), poly-β-aminoesters and cationic cyclodextrin.	

Source: Abstracted from [74].

stable in the blood to alter the pharmacokinetics of the solubilized drug. They are able to circulate for prolonged periods and to deliver more of the drug to tumor tissue when compared with administration of the drug in solution.

2. *Nanoparticles*

Nanoparticles are defined as particulate dispersions having a particle size of between 30 and 500 nm (microparticles

are larger than 0.5 μm). They are prepared from polymers and have a solid polymer matrix that forms the core. They are usually prepared by the controlled precipitation of polymers solubilized in one of the phases of an emulsion. Precipitation of the polymer from the solvent and simultaneous drug loading takes place on solvent evaporation, leaving particles of the polymer suspended in the remaining solvent. The presence of surfactants in the aqueous phase may be used to ensure a small particle size. Nanoparticles can be prepared by the polymerization of a monomer in a good solvent for the monomer but a poor solvent for the polymer or by high-pressure homogenization of the mixture. The loading of water-soluble active ingredients into the hydrophobic polymer particle matrix is still difficult but may be enhanced by the salting out of the active ingredient with soluble inorganic salts, by the formation of lipophilic ion pairs of soluble amine drugs with monoalkyl phosphate esters, or by exploiting the ionic attraction between drugs with a basic amine group and the carboxylic groups of polylactic acid/glycolic acid copolymer.

Many of the early investigations into solid particulate drug delivery used non-biodegradable polymers such as polystyrene. However, recent studies have quite rightly focused on the development of biodegradable or at least bioerodible particles. Polylactic acid, polyglycolic acid and poly β-hydroxybutyrate are used to make the biodegradable particles, while the alkyl cyanoacrylates produce bioerodible ones. These solid nanoparticles may be used to prepare sustained release parenteral formulations or to achieve drug targeting. The particle size of an injected dispersion has a significant impact on its biodistribution and pharmacodynamics; the drug delivery applications of nanoparticles are different to the closely allied microparticles, which are prepared by similar processes.

Because of their small size, nanoparticles may be injected i.v. and used to target drugs to particular organs. As with all i.v. injected and non-polymer coated colloidal particulates, they are cleared from the circulation by the liver and spleen. To facilitate drug targeting to tumor tissue a RES avoidance facility, a "stealth" mask in the same fashion as described earlier for liposomes, may be incorporated using polyoxyethylene block copolymers. With these copolymers the hydrophobic portion of the molecule (e.g. polylactic acid) forms the nanoparticle matrix while the water-soluble polyoxyethylene block forms a hydrophilic coating on the particle. Stealth nanoparticles may also be prepared by coating with soluble polyoxyethylene or by using dialkyl-polyoxyethylenes and phospholipids to make nanoparticles. The hydrophobic portions of these amphiphilic molecules, in the latter case, comprise the particle matrix.

Nanoparticles are not actually taken up by a tumor. However, stealth nanoparticles accumulate in its vicinity thereby forming a depot of anticancer drug, which provides a high concentration and tumoricidal drug activity.

3. *Microparticles*

Microparticles are very similar to nanoparticles, except that they have a significantly larger size (particle sizes in excess of 0.5 μm, compared with the particle size of between 30 and 500 nm for nanoparticles). They may be prepared by processes similar to those described for nanoparticles and can also be prepared by chemical cross-linking of soluble polymers.

Because of their larger size, microparticles are generally injected either intraperitoneally, intramuscularly, s.c. or directly to the target organ and are used to provide a sustained release depot of the drug, which is gradually released on erosion or by diffusion from the particles. The rate of release may be increased by decreasing the polymer molecular weight and particle size and also by controls on the nature of the polymer/copolymer.

Microparticle drug delivery has been evaluated in experimental tumors. However, due to the relatively large particle sizes the administration of such formulations is restricted to the application routes given above, thus ultimately limiting their potential. Hence, antitumor microparticles are either administered intra-arterially, and hence straight into the organ, or into body cavities such as the peritoneum. Apart from the intra-arterial and intraperitoneal routes, microparticles may also be injected directly into tumors. Direct injection of microparticles into solid tumors creates a depot, which increases the tumoricidal activity of the drug.

4. *Hydrogels*

Hydrogels are solid or liquid polymeric masses that are generally hydrophilic and compatible with the surrounding aqueous environment. Historically, they have been investigated mainly for potential use as a solid, drug-loaded implant, capable of acting as a depot that is inserted under the skin by surgery. They are used as depot formulations either to provide high concentrations of drug substance only to the immediate area surrounding the disease site or to provide sustained drug release for systemic therapy (e.g. hormone replacement therapy or for contraception).

Recently, forms of hydrogels have been developed that produce highly viscous liquids or semi-solid formulations, both of which may be injected with a needle, despite their high viscosity. Some hydrogels are liquids that congeal *in situ*; gel formation occurs when the polymer solubilizing solvent diffuses away from the injection site, leaving the polymer in contact with an aqueous environment *in vivo*. Other types are thermosensitive and either congeal on cooling after being injected at a temperature above, or upon warming after injection at a temperature below body temperature. These gels that form *in situ* may be used to prepare sustained release, depot formulations following injection close to, or into, the disease site. Recently, it has been demonstrated that they can be injected directly into

tumors, giving high local drug concentrations without any systemic effects.

Hydrogels can be prepared from a variety of biodegradable polymers, such as polysaccharides and poly (lactic acid) -glycolic acid copolymer; some polymethacrylate derivatives will form hydrogels but are non-biodegradable and require removal by surgery. These non-biodegradable materials do provide therapeutic levels of some highly active drug substance for up to 1 year *in vivo* and can be used for the treatment of diseases that are difficult to target by other means.

VII. DRUG DELIVERY AND CLEARANCE MECHANISMS OF NANOCOLLOIDS

Having discussed the properties of some of the more important lipidic and colloidal solubilizing systems used in pharmaceuticals, it will be useful to describe how these systems are able to deliver their drug payload. Under normal circumstances, when colloidal drug-loaded delivery systems are injected or infused into the systemic circulation, they are rapidly cleared by the liver (Kupfer cells) and the spleen (marginal zone and red pulp) macrophages. The reticuloendothelial clearance system (RES, also referred to as the mononuclear phagocyte system, MPS) is the body's normal defense against "foreign bodies" and is often aided by opsonization, which involves the surface deposition of blood opsonic factors such as fibronectin, immunoglobulins and complement proteins onto the surface of the colloidal particles. Surface deposition of these factors aids particle and "foreign body" recognition by the macrophages[75] and hence their destruction.

The particle size and the surface characteristics of nanoparticles have an important role to play in the opsonization process and the overall kinetics of the clearance process.[75–77] Whatever the drug delivery system, larger particles (⩾200 nm) activate the human complement system more effectively and are therefore cleared faster and more effectively from the blood by Kupfer cells than are particles smaller than 200 nm. The binding of blood proteins and opsonins to nanoparticles differs markedly in both amount and pattern, depending on the surface characteristics such as the type of functional groups present and the local surface charge density.[75–77] As a result, differential opsonization may occur and may account for differences seen in clearance rates and macrophage sequestration of nanoparticles. Small nanoparticles may aggregate to form larger agglomerates that are more easily recognized and cleared.

There may be an interaction between certain lipid-based nanocolloids and lipoproteins that occur naturally in the blood. This may result in dramatic size changes of the colloidal particles. It has been reported that some hydrophobic components of Cremophor EL micelles are rapidly incorporated into both high-density lipoproteins (HDL) and low-density lipoproteins (LDL), leaving the remaining components to form larger lipid structures in the plasma. These larger structures assemble C3 convertases and activate the complement system.[78] This effect may be responsible, at least in part, for the hypersensitivity reactions seen during and after the administration of the anticancer drug Taxol®.

There are many examples given in the literature of nanocolloids enrobed with polymers based on PEG. These produce "macromolecular hairs" which significantly reduce recognition of the underlying particle or droplet by macrophages. It is believed that there is an array of mechanisms that help achieve reduced protein adsorption and surface opsonization. The efficiency of the process is dependent on polymer type, the surface stability of the polymer, surface reactivity and physics (e.g. surface density and conformation).[75,79,80] This stealth like behavior is similar to that used by pathogenic microorganisms to minimize their detection by the body's immune system. Suppression of opsonization is the key step in enhancing passive retention of long circulating, surface manipulated nanoparticles at sites in contact with the blood.

Recent advances in polymer chemistry have yielded better control over the architecture and polydispersivity of simple polymers, polymer-conjugates and block copolymers such that they can be efficiently produced and can spontaneously form sterically stabilized, nanoscale, self-assembling structures with macrophage evading properties.[81,82]

The escape of particles from the vasculature is normally restricted to sites where the endothelial cells in capillaries have a more open ("leaky" or fenestrated) structure, as in the sinus endothelium of the liver, or where the integrity of the endothelial barrier is damaged by inflammatory processes or by tumor growth (as a result of unregulated angiogenesis).[75,83,84] In the liver, the size of the fenestrae ("windows") in the sinus endothelium can be up to 150 nm, whilst in tumor capillaries (angiogenic blood vessels) there is a wide variation up to 600–800 nm[85] for the size of the gaps between adjacent endothelial cells. If confinement within the vascular system is necessary, the predominant filtration processes at the inter-endothelial slits in the spleen must be considered. This is particularly true for rigid particles, or those that cannot be deformed, whose size exceeds the width (200–250 nm) of the cell slits.[76]

If the target for the drug is a particular lymph node, or a group of nodes in a small area, injection of nanoparticles or other colloidal drug delivery systems into the interstitial space surrounding the nodes is often successful.[86,87] It has been reported[86] that many of the overlapping endothelial cells in the lymphatic capillaries lack adhesion and they are separated after the injection. This separation opens channels through which the colloidal particles are conveyed. Whilst in the lymphatic system, the fate of the particles is controlled by their size and surface characteristics,[86–88] the size of the nanoparticles must be large enough to prevent their rapid leakage into blood capillaries; a size range of

30–100 nm has been found to satisfy this criterion. Particles larger than 100 nm move very slowly and are susceptible to clearance by interstitial macrophages. Conversely, very small nanoparticles (1–20 nm), especially those that are long circulating, can slowly extravasate from the vasculature into the interstitial spaces prior to their transport into the lymph nodes through the lymphatic vessels.[86]

Multi-drug resistance (MDR) is a further important clearance mechanism that can severely limit the effectiveness of some forms of cancer chemotherapy.[89] The most commonly seen form of MDR occurs through a membrane-spanning transporter molecule, ATPase, that is located in the apical membranes of intestinal absorptive cells and is responsible for the efflux of positively charged amphipathic drugs from the cell.[90–92] Overexpression of P-glycoprotein (P-gp170) in tumor cells can lead to a significant decrease in the sensitivity of the cell to some drugs such as doxorubicin. Delivery of doxorubicin entrapped in liposomes, however, has resulted in an effective treatment for a number of cancers that previously were ineffectively treated.[93–96] This has raised some questions whether liposomal doxorubicin is able to sensitize tumor cells and thus partially reverse the MDR effect.

In the last few years it has been found that many drugs and drug families are substrates for P-gp and this results in a reduction in the bioavailability of the drugs affected. Thus, there is considerable interest in finding means of inhibiting P-gp.[97] Two strategies are gradually evolving to accomplish this inhibition: (i) incorporation of a non-specific lipid and/or polymer into the formulation and (ii) coadministration of a second P-gp substrate or specific inhibitor. Several lipids and polymers are emerging as efficient inhibitors, with TPGS (α-tocopheryl-polyethylene glycol-1000-succinate) being the one that is currently the most frequently used. The second strategy involves addition of a second drug substance to the formulation, which adds significant complications to development in terms of regulatory constraints. Examples of drug substances that are efficient P-gp inhibitors are verapamil and cyclosporine.

VIII. DRUG DELIVERY AND ACCUMULATION USING COLLOIDAL SYSTEMS FOR THE TREATMENT OF CANCER

Currently, significant R&D effort and real clinical advances are being seen in the delivery of therapeutic agents to cancers and cancer cells. Huge advances have been made over the last 10 years using liposomes, nanoparticles, microparticles and conjugates[98–100] as drug delivery systems for pharmaceuticals. The first liposomal anticancer agent to be approved was PEGylated (stealth-) liposomal doxorubicin (Doxil®/Caelix®); it was approved in the United States in 1995. In 2006, the first nanoparticle anticancer drug delivery system was approved in the United States; this comprises albumin bound paclitaxel suspension (Abraxane®, previously ABI007) for the treatment of MBC. Both of these dosage forms demonstrate clear advantages over the earlier, simple formulations in the following respects:

- increased drug solubility
- prolonged duration of exposure
- selective delivery of drug to cancer cells
- improved therapeutic index
- partial overcoming of resistance associated with the earlier, simple formulations[98,99]

Even more recently, in early January 2007, the first nanoparticle based drug delivery system produced outside the United States was approved in India; it contains paclitaxel and is to be marketed as Nanoxel®. It is polymer-based and does not contain Cremophor® EL.

A. Liposome formulations

The pharmacokinetic disposition of therapeutic liposomes is primarily dependent on the carrier and not the drug,[101–103] that is, until the drug is released from the formulation. As stated earlier, the pharmacokinetic parameters of drug-loaded liposomes are highly dependent on their physicochemical characteristics (size, surface charge, membrane lipid packing and orientation, steric stabilization, drug content and route of administration).[103] The primary sites of accumulation of conventional liposomes are in the tumor, liver and spleen. In comparison, incorporation of PEGylated lipids into the liposomes results in their accumulation primarily in the systemic circulation, where they may have a half-life of >40 h (results quoted for Doxil® indicate that the half-life may be close to 56 h). About 10–15% may accumulate in the liver (compare to 80–90% in the liver for conventional liposomes).[104–107] The use of PEGylated lipids in the liposome does not prevent uptake by the RES, they significantly reduce the rate of uptake.

Current opinion concerning the mechanism of delivery of liposomal therapeutics to tumors is that, once in the tumor, standard liposomes are localized in the extra-cellular fluid that surrounds the tumor cell but do not enter it.[108–110] Therefore, for delivery of the therapeutic agent, the drug must first be released into the extra-cellular fluid, from where it must then diffuse into the cell.[111] There is the evidence[106] that doxorubicin is released efficiently from Doxil® liposomes, to be taken up by the cell in ovarian cancer and Kaposi's sarcoma. Concentrations of doxorubicin achieved in cells have been estimated as up to 13 times higher from Stealth liposomes than with a simple doxorubicin injection in patients undergoing treatment of Kaposi's sarcoma.[112] In contrast, evidence from the use of cisplatin seems less definitive. In studies comparing Stealth liposomal cisplatin (SPI-077) with standard cisplatin in

mice tumor models, the total exposure to platinum was only 4 times higher but more prolonged with SPI-077.[107] Total platinum was determined in tumor extracts and it was impossible to differentiate between liposomal-platinum, protein bound platinum or unbound platinum. Also, even though there was a 4-fold higher concentration, this did not give a higher clinical response. One possible explanation offered[113] for the higher tumor exposure but low efficacy was that the cisplatin is not released efficiently from the liposome into the extra-cellular fluid of the tumor cell.

B. Formulations based on nanoparticles, microparticles and conjugated systems

The first protein stabilized nanoparticle formulation to be approved by the FDA is Abraxane® (previously called ABI-007). This albumin-stabilized nanoparticle containing paclitaxel was developed in order to reduce the high toxicity and solubility issues previously seen with the micellar Cremophor® formulation, Taxol®. It has been demonstrated that Abraxane® enables a more rapid distribution of paclitaxel from the vascular system as well as providing tumor targeting.[100,114] The approved dosing schedule for Abraxane® is 260 mg/m^2 infused over 30 min every 3 weeks, compared with 135–200 mg/m^2 infused over 3 h for Taxol®. Taxol® infusions generally require premedication with steroids in order to suppress toxicity during the infusion process; premedication has not been found necessary with Abraxane®.

A lower incidence of myelosuppression has been observed and reported[113] during studies on Abraxane® but other toxic effects (sensory neuropathy, mucositis) are similar to those seen with Taxol given at high doses. Abraxane® has been reported to produce keratopathy, which is a toxic effect rarely seen with drugs.[114] Thus, as with the liposomal formulations described earlier (Section VIII.A.), the administration of nanoparticle based formulations can dramatically alter the pharmacokinetics, the distribution of the drug in both tissues and tumors, and the toxicity profile. Also, similar to what has been found with liposomes, the mechanism(s) by which nanoparticles release their drug payload is not well understood as yet.

IX. MODIFICATION OF DRUG TOXICITY BY NANOCOLLOIDAL DRUG DELIVERY SYSTEMS

The incorporation of drugs into liposomal formulations often results in a significant alteration of the toxicity profile for some drug substances and, surprisingly, new toxicities may be introduced.[115–117] This has been found especially so in the case of doxorubicin and other related anthracycline drugs where cardiotoxicity associated with cumulative doses has limited their use in simple formulations, despite them being very effective against many types of cancer.[118] Preclinical studies undertaken on doxorubicin demonstrated that, if it was incorporated into liposomes, the antitumor activity was retained but the incidence and severity of the dose-related cardiotoxicity was reduced.[117] Subsequently, doxorubicin has been formulated as a conventional, but extemporaneously prepared, liposomal product (Myocet®, The Liposome Company/Elan Corporation) and also as a stealth liposome (Doxil®/Caelix®, Alza Corporation/Schering Plough). The two liposomal formulations have significantly different lipid compositions, surface properties and drug loading processes, which lead to unique pharmacokinetic and toxicity profiles, each differing significantly from the unencapsulated form. Liposome-encapsulated drug was found to accumulate in the liver, spleen and tumors and not in other major organs, such as the heart.[119]

This evidence has also been supported by other findings from histological examination of cardiac biopsies from patients who received cumulative doses of Doxil® from 440 to 840 mg/m^2, with no prior exposure to anthracyclines, which revealed significantly less cardiotoxicity than in matched controls that had been given only unencapsulated doxorubicin ($P < 0.001$).[120]

It has been found that administration of Doxil®/Caelix® stealth liposomes gives rise to a significant incidence of palmar plantar erythrodysesthia syndrome (PPE), often called hand and foot syndrome and stomatitis. The occurrence of this syndrome appears to be highly dose dependent and both peak drug concentration and total cumulative dose determine its appearance.[121] It was concluded that the use of liposomal doxorubicin permits a greater lifetime cumulative dose to be administered, but the acute maximum tolerated doses (MTD) differ significantly. The MTD of Myocet® is essentially equivalent to that of unencapsulated doxorubicin, whilst higher doses of Doxil®/Caelix® may be administered.

In vitro studies demonstrated that liposomal encapsulation of doxorubicin can also overcome excessive drug efflux due to MDR-1 overexpression.[122,123] A similar phenomenon has been investigated in murine models and demonstrated that escalation of the dose of Doxil® from 2.5 to 20 mg/kg results in a saturation of drug clearance and a disproportional increase in the amount of drug accumulating in the tumor.[124] Results from experiments with radiolabeled, highly negatively charged liposomes injected into mice that had been previously treated with Doxil® were consistent with a partial blockade of the RES, with a relative reduction of uptake by the liver and a greater prolongation of the liposome circulation time.

Liposomal formulations with different lipid compositions have been used in order to determine whether drug release rate had a significant effect upon the accumulation of liposomal lipid and doxorubicin into tumors and in the cutaneous tissues (skin and paws) of mice.[125] Liposomes

with different leakage rates were prepared using lipids of different fluidity, achieved by changing the fatty acyl chain length and/or degree of saturation of the phosphatidylcholine component of the liposome. This resulted in liposome formulations with fast, intermediate and slow release rates. These were administered to mice (orthotopic 4T1 murine carcinoma model). The plasma pharmacokinetics of the liposomal lipid proved to be similar for all formulations but the plasma pharmacokinetics of the doxorubicin was dependent on the formulation. Liposomes with the slowest rate of release of doxorubicin produced the highest doxorubicin concentrations in the tumor and cutaneous tissue. Lipid accumulated to similar extent in the tumor and tissue independent of the formulation. Liposomes with the fastest release rate gave low doxorubicin levels in the tumor and tissue. The intermediate formulations produced unexpected toxicity, not related to the lipid content of the formulation. It was concluded that slow release rates gave the best therapeutic activity.

Batist *et al.*[126] reported results from an early trial using Myocet® for the treatment of MBC and concluded that "Myocet® improves the therapeutic index of doxorubicin by significantly reducing cardiotoxicity and Grade 4 neutropenia and provides comparable antitumor efficacy, when used in combination with cyclophosphamide for the first line treatment of MBC." However, these conclusions were not accepted by FDA as they concluded that the trial had insufficient patients, hence insufficient power, to infer this based on statistical analysis.[127]

REFERENCES

1. Greenwald, R. B. PEG drugs: an overview. *J. Control Release* 2001, *74*, 159–171.
2. Bannerjee, P., Weissleder, R., Bogdanov, A., Jr. Linear polyethyleneimine grafted to a hyperbranched poly(ethylene glycol)-like core: a copolymer for gene delivery. *Bioconjug. Chem.* 2006, *17*(1), 125–131.
3. Glodde, M., Sirsi, S. R., Lutz, G. J. Physiochemical properties of low and high molecular weight poly(ethylene glycol)-grafted poly(ethylene imine) copolymers and their complexes with oligonucleotides. *Biomacromolecules* 2006, *7*(1), 347–356.
4. Kissel, M., Peschke, P., Subr, V., Ulbrich, K., Strunz, A. M., Kuehnlein, R., Debus, J., Friedrich, E. Detection and cellular localisation of the synthetic soluble macromolecular drug carrier pHPMA. *Eur. J. Nucl. Med. Mol. Imaging* 2002, *29*(8), 1055–1062.
5. Zamai, M., vandeVen, M., Farao, M., Gratton, E., Ghiglieri, A., Castelli, M. G., Fontana, E., d'Argy, R., Fiorino, A., Pesenti, E., Suarato, A., Caiolfa, V. R. Camptothecin poly[*N*-(2-hydroxypropyl) methacrylamide] copolymers in antitopoisomerase-I tumor therapy: intratumor release and antitumor efficacy. *Mol. Cancer Ther.* 2003, *2*, 29–40.
6. Hennenfent, K. L., Govindan, R. Novel formulations of taxanes: old wine in a new bottle?. *Ann. Oncol.* 2006, *17*(5), 735–749.
7. Bradley, M. O., Webb, N. L., Anthony, F. H., Devanescan, P., Witman, P. A., Hemamalini, S., Chander, M. C., Baker, S. D., He, L., Horwitz, S. B. Swindell, tumor targeting by covalent conjugation of a natural fatty acid to paclitaxel. *Clin. Cancer Res.* 2001, *7*, 3229–3238.
8. Stevens, P. J., Sekido, M., Lee, R. J. A folate receptor-targeted lipid nanoparticle formulation for a lipophilic paclitaxel prodrug. *Pharm. Res.* 2004, *21*(12), 2153–2157.
9. Hwang Pun, S., Davis, M. E. Development of a nonviral gene delivery vehicle for systemic application. *Bioconjug. Chem.* 2002, *13*(3), 630–639.
10. Lipinski, C. A., Lombardo, F., Dominy, B. W., Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Del. Rev.* 1997, *46*, 3–25.
11. Proudfoot, J. R. The evolution of synthetic oral drug properties. *Bioorg. Med. Chem. Lett.* 2005, *15*(4), 1087–1090.
12. Palm, K., Sternberg, P., Luthman, K., Artursson, P. Polar molecular surface area properties predict the intestinal absorption of drugs in humans. *Pharm. Res.* 1997, *14*(5), 568–571.
13. Palm, K., Luthman, K., Ungell, A. L., Strandlund, G., Artursson, P. Correlation of drug absorption and molecular surface properties. *J. Pharm. Sci.* 1996, *85*, 32–39.
14. Kelder, J., Grootenhuis, P. D., Bayada, D. M., Delbressine, L. P., Ploemen, J. P. Polar surface area as a dominating determinant for oral absorption and brain penetration of drugs. *Pharm. Res.* 1999, *16*, 1514–1519.
15. Amidon, G. L., Lennernäs, H., Shah, V. P., Crison, J. R. A theoretical basis for a biopharmaceutic drug classification: the correlation of *in vitro* drug product dissolution and *in vivo* bioavailability. *Pharm. Res.* 1995, *12*, 413–420.
16. US Department of Health and Human Services, FDA Center for Drug Evaluation and Research (CDER). *Guidance for Industry. Waiver of In Vivo Bioavailability and Bioequivalence Studies for Immediate-Release Solid Oral Dosage Forms Based on a Biopharmaceutics Classification System*. CDER: Rockville, MD (August 31, 2000).
17. Johnson, K. C., Swindell, A. C. Guidance in the setting of drug particle size specification to minimise variability in absorption. *Pharm. Res.* 1996, *13*(12), 1795–1798.
18. Curatolo, W. Physical chemical properties of oral drug candidates in the discovery and exploratory development settings. *PSTT* 1988, *1*, 387–393.
19. Mottu, F., Stelling, M.-J., Ruefenacht, D. A., Doelker, E. Comparative hemolytic activity of undiluted organic water-miscible solvents for intravenous and intra-arterial injection. *PDA J. Pharm. Sci. Technol.* 2001, *55*(1), 16–23.
20. Mottu, F., Laurent, A., Ruefenacht, D. A., Doelker, E. Organic solvents for pharmaceutical parenterals and embolic liquids: a review of toxicity data. *PDA J. Pharm. Sci. Technol.* 2000, *54*(6), 456–469.
21. Nema, S., Washkuhn, R., Brendel, R. Excipients and their use in injectable products. *PDA J. Pharm. Sci. Technol.* 1997, *51*, 161–171.
22. Yalkowsky, S. H., Roseman, T. J. Solubilization of drugs by cosolvents. In *Techniques of Solubilization of Drugs* (Yalkowsky, S. H., Ed.). Marcel Dekker, Inc.: New York and Basel, 1981, pp. 91–134.
23. Yalkowsky, S. H., Flynn, G., Amidon, G. L. Solubility of nonelectrolytes in polar solvents. *J. Pharm. Sci.* 1972, *61*, 983–984.
24. Yalkowsky, S. H., Valvani, S. C., Amidon, G. L. Solubility of non electrolytes in polar solvents IV: nonpolar drugs in mixed solvents. *J. Pharm. Sci.* 1976, *65*, 1488–1494.
25. Valvani, S. C., Yalkowsky, S. H., Roseman, T. J. Solubility and partitioning IV: aqueous solubility and octanol–water partition coefficients of liquid nonelectrolytes. *J. Pharm. Sci.* 1981, *70*, 502–507.
26. Gould, P. L., Goodman, M., Hanson, P. A. Investigation of the solubility relationships of polar, semi-polar and non-polar drugs in mixed co-solvent systems. *Int. J. Pharm.* 1984, *19*, 149–159.
27. Alvarez-Nunez, F. A., Pinsuwan, S., Lerkpulsawad, S., Yalkowsky, S. H. Effect of the most common co-solvents upon the extent of solubilization of some drugs: log-linear model. *AAPS Annual Meeting*, San Francisco, CA, USA, Vol. 1 (4), 1998, Poster 2433.
28. Kramer, S. F., Flynn, G. L. Solubility of organic hydrochlorides. *J. Pharm. Sci.* 1972, *61*, 1896–1904.
29. Yalkowsky, S. H., Valvani, S. C. Precipitation of solubilized drugs due to injection or dilution. *Drug Intel. Clin. Pharm.* 1977, *11*, 417–419.

30. Strickley, R. G. Review article: solubilizing excipients in oral and injectable formulations. *Pharm. Res.* 2004, *21*(2), 201–230.
31. Bertholom, C., Le Ray, A. M., Olivier, E., Merle, C. Critères de choix des solvents injectables non aqueux miscibles à l'eau. *STP Pharma Pratiques* 2000, *10*(3), 137–143.
32. Lee, Y.-C., Zocharski, P. D., Samas, B. An intravenous formulation decision tree for discovery compound formulation development. *Int. J. Pharm.* 2003, *253*, 111–119.
33. Bittner, B., Mountfield, R. J. Intravenous administration of poorly soluble new drug entities in early drug discovery: the potential impact of formulation on pharmacokinetic parameters. *Curr. Opin. Drug Disc. Dev.* 2002, *5*(1), 59–71.
34. Bittner, B., Mountfield, R. J. Formulation and related activities for the oral administration of poorly water-soluble compounds in early discovery animal studies: Part 1. *Pharm. Ind.* 2002, *64*(8), 800–807.
35. Bittner, B., Mountfield, R. J. Formulation and related activities for the oral administration of poorly water-soluble compounds in early discovery animal studies: Part 2. *Pharm. Ind.* 2002, *64*(9), 985–991.
36. Neervannan, S. Preclinical formulations for discovery and toxicology: physicochemical challenges. *Expert Opin. Drug Metab. Toxicol.* 2006, *2*(5), 715–731.
37. Aboutaleb, A. E., Abdelzaher, A. A study on the solubilization of chloramphenicol. *Pharmazie* 1981, *36*, 102–106.
38. Griffin, W. C. Classification of surface active agents by HLB. *J. Soc. Cosmet. Chem.* 1949, *1*, 311.
39. Volcheck, G. W., Van Dellen, R. G. Anaphylaxis to intravenous cyclosporine and tolerance to oral cyclosporine: case report and review. *Ann. Allergy Asthma Immunol.* 1998, *80*, 159–163.
40. Ebo, D., Piel, G. C., Conraads, V., Stevens, W. J. IgE-mediated anaphylaxis after first intravenous infusion of cyclosporine. *Ann. Allergy Asthma Immunol* 2001, *87*, 243–245.
41. Szebeni, J., Muggia, F. M., Alving, C. R. Complement activation by Cremophor EL as a possible contributor to hypersensitivity to paclitaxel: an *in vitro* study. *J. Natl. Cancer Inst.* 1998, *90*(4), 300–306.
42. Jain, N., Yang, G., Yalkowsky, S. H. Solubilization of NSC-639829 by combined effects of pH with cosolvents, surfactants and complexants. *Int. J. Pharm.* 2001, *225*, 41–47.
43. Thoma, K., Albert, K. Amphiphilic drugs. 2. Relation between colloidal properties and pharmaceutic-technological, pharmacokinetic and pharmacodynamic behavior. *Pharmazie* 1983, *38*, 807–817.
44. Hussain, M. A., Wu, L. S., Koval, C., Hurwitz, A. R. Parenteral formulation of the kappa agonist analgesic, DuP 747, via micellar solubilization. *Pharm. Res.* 1992, *9*, 750–752.
45. Florence, A. T., Attwood, D. *Physicochemical principles of Pharmacy*, 3rd Edition. McMillan Press: Hampshire, 1998.
46. Tarr, B. D., Sambandan, T. G., Yalkowsky, S. H. A new parenteral emulsion for the administration of Taxol. *Pharm. Res.* 1987, *4*, 162–165.
47. Strickley, R. G., Anderson, B. D. Solubilisation and stabilisation of an anti- HIV thiocarbamate, NSC 629243, for parenteral delivery, using extemporaneous emulsions. *Pharm. Res.* 1993, *10*, 1076–1082.
48. Lundberg, B. Preparation of drug-carrier emulsions stabilised with phosphatidylcholine surfactant mixtures. *J. Pharm. Sci.* 1994, *83*, 72–75.
49. Lundberg, B. A submicron lipid emulsion coated with amphipathic polyethylene glycol for parenteral administration of paclitaxel (Taxol). *J. Pharm. Pharmacol.* 1997, *49*, 16–21.
50. Prankerd, R. J., Stella, V. J. The use of oil-in-water emulsions as a vehicle for parenteral drug administration. *J. Parent. Sci. Technol.* 1990, *44*, 139–149.
51. Floyd, A. G. Top ten considerations in the development of parenteral emulsions. *PSTT* 1999, *2*, 134–143.
52. Krishna, G., Sheth, B. B. A novel self-emulsifying parenteral drug delivery system. *PDA J. Pharm. Sci. Tech.* 1999, *53*(4), 168–176.
53. Lee, M. J., Lee, M. H., Shim, C. K. Inverse targeting of drugs to the reticuloendothelial system – rich organs by microemulsions emulsified with Poloxamer – 338. *Int. J. Pharm* 1995, *113*, 175–187.
54. Liu, F., Liu, D. Long circulating emulsions (oil-in-water) as carriers for lipophilic drugs. *Pharm. Res.* 1995, *12*, 1060–1064.
55. Song, Y. K., Liu, D. X., Maruyama, K., Takizawa, T. Antibody mediated lung targeting of long circulating emulsions. *PDA J. Pharm. Sci. Tech.* 1996, *50*, 372–377.
56. Chavanet, P. Y., Garry, I., Charlier, N., Caillot, D., Kisterman, J. P., D'Athis, M. Trial of glucose versus fat emulsion in preparation of amphotericin for use in HIV infected patients with candidiasis. *BMJ* 1992, *305*, 921–925.
57. Aulton, M. E. (Ed.). *Pharmaceutics: The Science of Dosage Form Design*, 2nd Edition. Churchill Livingstone, 2001.Chapter 6, pp. 87–94,
58. Drewe, J., Meier, R., Vonderscher, J. D. K., Posanski, U., Kissel, T., Gyr, K. Enhancement of the oral absorption of cyclosporin in man. *Br. J. Clin. Pharmacol* 1992, *34*, 60–64.
59. Gershanik, T., Benita, S. Self dispersing lipid formulations for improving oral absorption of lipophilic drugs. *Eur. J. Pharm. Biopharm.* 2000, *50*, 179–188.
60. McCormack, B., Gregoriadis, G. Drugs in cyclodextrins in liposomes – a novel concept in drug delivery. *Int. J. Pharm.* 1994, *112*, 249–258.
61. Aulton, M. E. (Ed.) *Pharmaceutics: The Science of Dosage Form Design*, 2nd Edition. Churchill Livingstone, 2001. Chapter 6, pp. 334–359.
62. Hasrani, P. K., Davis, S. S., Groves, M. J. The preparation and properties of sterile intravenous emulsions. *J. Parenter. Sci. Technol.* 1983, *37*, 145–150.
63. Davis, S. S., Hadgraft, J., Palin, K. J. Medical and pharmaceutical applications of emulsions. In *Encyclopedia of Emulsion Technology* (Becher, P., Ed.). Dekker: New York, 1985, pp. 159–238.
64. Davis, S. S., Washington, C., West, P., Illum, L., Liversedge, G., Sternson, L., Kirsh, R. Lipid emulsions as drug delivery systems. *Ann. NY Acad. Sci.* 1987, *507*, 75–88.
65. Singh, M., Ravin, J. L. Parenteral emulsions as drug carrier systems. *J. Parenter. Sci. Technol.* 1986, *40*, 34–41.
66. Boyett, J. B., Davis, C. W. Injectable emulsions and suspensions. In *Pharmaceutical Dosage Forms: Disperse Systems* (Liberman, H. A., Rieger, M. M., Banker, G. S., Eds). Dekker: New York, 1989, pp. 379–416.
67. Collins-Gold, L. C., Lyons, R. T., Barlow, L. C. *Adv. Drug. Del. Rev.* 1990, *5*, 189–208.
68. Collins-Gold, L., Feichtinger, N., Waernheim, T. Are lipid emulsions the drug delivery solution?. *Mod. Drug. Disc.* 2000, *3*(3), 44–48.
69. Charman, W. N. Lipid vehicle and formulation effects on intestinal lymphatic drug transport. In *Lymphatic Transport of Drugs* (Charman, W. N., Stella, V. L., Eds). CRC Press: Boca Raton, FL, 1992, pp. 113–179.
70. Charman, W. N. Lipids, lipophilic drugs, and oral drug delivery – some emerging concepts (mini-review). *J. Pharm. Sci.* 2000, *89*, 967–978.
71. Charman, W. N. Lipids, lymph and lipidic formulations. *Bull. Tech. Gattefossé.* 1997, *90*, 27–33.
72. The Lipid Library-http://www.lipidlibrary.co.uk/lipids.html
73. Lipid Metabolites and Pathways Strategy – http://www.lipidmaps.org/data/classification/fa.html
74. Moghimi, S. M., Hunter, A. C., Murray, J. C. Nanomedicine: current status and future prospects. *FASEB J.* 2005, *19*, 311–330.
75. Moghimi, S. M., Hunter, A. C., Murray, J. C. Long-circulating and target-specific nanoparticles: theory to practice. *Pharmacol. Rev.* 2001, *53*, 283–318.
76. Moghimi, S. M., Szebeni, J. Stealth liposomes and long-circulating nanoparticles: critical issues in pharmacokinetics opsonisation and protein-binding properties. *Prog. Lipid Res.* 2003, *42*, 463–478.
77. Moghimi, S. M., Hunter, A. C. Recognition by macrophages and liver cells of opsonised phospholipid vesicles and phospholipid headgroups. *Pharm. Res.* 2001, *18*, 1–8.
78. Szebeni, J., Alving, C. R., Savay, S., Barenholz, Y., Priev, A., Danino, D., Talmon, Y. Formation of complement activating particles in aqueous solutions of Taxol: possible role in hypersensitivity reactions. *Int. Immunopharmacol.* 2001, *1*, 721–735.

79. Moghimi, S. M. Chemical camouflage of nanospheres with a poorly reactive surface: towards development of stealth and target-specific nanocarriers. *Biochim. Biophys. Acta* 2002, *1590*, 131–139.
80. Ghadamosi, J. K., Hunter, S. C., Moghimi, S. M. PEGylation of microspheres generates a heterogeneous population of particles with differential surface characteristics and biological performance. *FEBS Lett.* 2002, *532*, 338–344.
81. Adams, M. L., Lavasanifar, A., Kwon, G. S. Amphiphilic block copolymers for drug delivery. *J. Pharm. Sci.* 2003, *92*, 1343–1355.
82. Gao, Z. G., Lukyanov, A. N., Singhal, A., Torchilin, V. P. Diacyl-lipid polymer micelles as nanocarriers for poorly soluble anticancer drugs. *Nano Lett.* 2002, *2*, 979–982.
83. Braet, F., Dezanger, R., Baekeland, M., Crabbe, E., van der Smissen, P., Wisse, E. Structure and dynamics of the fenestrae-associated cytoskeleton of rat liver sinusoidal endothelial cells. *Hepatology* 1995, *21*, 180–189.
84. Mann, L. L. Aberrant vascular architecture in tumours and its importance in drug-based therapies. *Drug Discov. Today* 2003, *8*, 396–403.
85. Allen, T. M., Cullis, P. R. Drug delivery systems: entering the mainstream. *Science* 2004, *303*, 1818–1822.
86. Moghimi, S. M., Bonnemain, B. Subcutaneous and intravenous delivery of diagnostic agents to the lymphatic system: applications in lymphoscintigraphy and indirect lymphography. *Adv. Drug Del. Rev.* 1999, *37*, 295–312.
87. Philips, W. T., Klipper, R., Goins, B. Novel method of greatly enhanced delivery of liposomes to lymph nodes. *J. Pharmacol. Exp. Ther.* 2000, *295*, 309–313.
88. Moghimi, S. M. Modulation of lymphatic distribution of subcutaneously injected poloxamer 407 coated nanospheres: the effect of the ethylene oxide chain configuration. *FEBS Lett.* 2003, *540*, 241–244.
89. Murray, J. C., Carmichael, J. Targeting solid tumours: challenges, disappointments and opportunities. *Adv. Drug Del. Rev.* 1995, *17*, 117–127.
90. Endicott, J. A., Ling, V. The biochemistry of P-glycoprotein mediated multi-drug resistance. *Annu. Rev. Biochem.* 1989, *58*, 137–171.
91. Pastan, T., Gottesman, M. M. Multi-drug resistance. *Annu. Rev. Med.* 1991, *42*, 277–286.
92. Gottesman, M. M., Pastan, T. Biochemistry of multi-drug resistance mediated by the multi-drug transporter. *Annu. Rev. Biochem.* 1993, *62*, 385–427.
93. Treat, J., Greenspan, A., Forst, D., Sanchez, J. A., Ferrans, V. J., Potkul, L. A., Wooley, P. V., Rahman, A. Antitumor activity of liposome-encapsulated doxorubicin in advanced breast cancer: phase II study. *J. Natl. Cancer Inst.* 1990, *82*, 1706–1710.
94. Vaage, J., Donovan, D., Loftus, T., Abra, R., Working, P., Huang, A. Chemoprevention and therapy of mouse mammary carcinomas with doxorubicin encapsulated in sterically stabilised liposomes. *Cancer* 1994, *73*, 2366–2371.
95. Muggia, F. M., Hainsworth, J. D., Jeffers, S., Miller, P., Groshan, S., Tan, M., Roman, L., Uziely, B., Muderspach, L., Garcia, A., Burnett, A., Greco, F. A. Phase II study of liposomal doxorubicin unrefractory ovarian cancer: anti-tumor activity and toxicity modification by liposomal encapsulation. *J. Clin. Oncol.* 1997, *15*, 987–993.
96. Northfelt, D. W., Dezube, B. J., Thommes, J. A., Levine, R., Von Roenn, J. H., Dozik, G. M., Rios, A., Krown, S. E., DuMond, C., Mamaelok, R. D. Efficacy of pegylated liposomal doxorubicin in the treatment of AIDS-related Kaposi's sarcoma after failure of standard chemotherapy. *J. Clin. Oncol.* 1997, *15*, 653–659.
97. Constatinides, P. P., Wasan, K. M. Lipid formulation strategies for enhancing intestinal transport and absorption of P-glycoprotein (P-gp) substrate drugs: *In vitro*/*In vivo* case studies. *J. Pharm. Sci.* 2007, *96*, 235–248.
98. Drummond, D. C., Meer, O., Hong, K., Kirpotin, D. B., Papahdjopoulos, D. Optimising liposomes for delivery to solid tumors. *Pharmacol. Rev.* 1999, *51*, 691–743.
99. Papahdjopoulos, D., Allen, T. M., Gabizon, A., Mayhew, E., Matthay, K., Huang, S. K., Lee, K. D., Lasic, D. D., Redemann, C. Sterically stabilised liposomes: improvements in pharmacokinetics and antitumor therapeutic efficacy. *Proc. Natl. Acad. Sci. USA* 1991, *88*, 11460–11464.
100. Adis International Limited, ABI 007, Drugs R&D, Vol. 5. Adis International Limited: , 2004.,,pp. 155–159.
101. Laginha, K., Mumbengegwi, D., Allen, T. Liposomes targeted via two different antibodies: assay, B-cell binding and cytotoxicity. *Biochim. Biophys. Acta* 2005, *1711*, 25–32.
102. Lasic, D. D., Frederik, P. M., Stuart, M. C., Barenholz, Y., McIntosh, T. J. Gelation of liposome interior: a novel method for drug encapsulation. *FEBS Lett.* 1992, *312*, 255–258.
103. Barenholz, Y. Relevancy of drug loading to liposomal formulation therapeutic efficacy. *J. Liposome Res.* 2003, *13*(1), 1–8.
104. Allen, T. M., Hansen, C. Pharmacokinetics of stealth versus conventional liposomes: effect of dose. *Biochim. Biophys. Acta* 1991, *1068*, 133–141.
105. Allen, T. M., Stuart, D. D. Liposomal pharmacokinetics – classical, sterically-stabilised, cationic liposomes and immunoliposomes. In *Liposomes: rational design* (Janoff, A. S., Ed.), Marcel Dekker, Inc.: New York, 2005, pp. 63–87.
106. Working, P. K., Newman, M. S., Stuart, Y. Pharmacokinetics, bio distribution and therapeutic efficacy of doxorubicin encapsulated in STEALTH liposomes. *J. Liposome Res.* 1994, *46*, 667–687.
107. Newman, M. S., Colbern, G. T., Working, P. K., Engbers, C., Amantee, M. A. Comparative pharmacokinetics, tissue distribution and therapeutic effectiveness of cisplatin encapsulated in long circulating PEGylated liposomes (SPI-077) in tumor-bearing mice. *Cancer Chemother. Pharmacol.* 1999, *43*, 1–7.
108. Harrington, K. J., Mohammadtaghi, S., Uster, P. S., Glass, D. M., Peters, A. M., Vile, R. G., Stewart, J. S. W. Effective targeting of tumours in patients with locally advanced cancers by radiolabeled PEGylated liposomes. *Clin. Cancer Res.* 2001, *7*, 243–254.
109. Harrington, K. J., Rowlinson-Busza, G., Syrigos, K. N., Uster, P. S., Abra, R. M., Stewart, J. S. Bio-distribution and pharmacokinetics of (111)In-DPTA-labeled pegylated liposomes in a human tumour xenograft model: implications for novel targeting strategies. *Br. J. Cancer* 2000, *83*, 232–238.
110. Harrington, K. J., Rowlinson-Busza, G., Syrigos, K. N., Abra, R. M., Uster, P. S., Stewart, J. S. Influence of tumour size on uptake of (111)In-DPTA-labeled pegylated liposomes in a human tumour xenograft model. *Br. J. Cancer* 2000, *83*, 684–688.
111. Zamboni, W. C., Gervais, A. C., Egorin, M. J., Schellens, J. H., Zuhowsky, E. G., Pluim, D., Joseph, E., Hamburger, D. R., Working, P. K., Colbern, G., Tonda, M. E., Potter, D. M., Eiseman, J. L. Systemic and tumor disposition of platinum after administration of cisplatin or STEALTH liposomal cisplatin formulations (SPI-077 and SPI-077 B103) in a preclinical tumor model of melanoma. *Cancer Chemother. Pharmacol.* 2004, *53*, 329–336.
112. Requier, D. Targeting and cancer treatment. *The Drug Delivery Companies Report* 2001, PharmaVentures.
113. Zamboni, W. C. Liposomal, nanoparticle and conjugated formulations of anticancer drugs. *Clin. Cancer Res.* 2005, *11*(23), 8230–8234.
114. Ibrahim, N. K., Desai, N., Legha, S., Soon-Shiong, P., Theriault, R. L., Rivera, E., Esmaeli, B., Ring, S. E., Bedikian, A. Y., Hortobagyi, G. N., Ellerhorst, J. A. Phase I and pharmacokinetic study of ABI-007, a Cremophor free, protein stabilised, nanoparticle formulation of paclitaxel. *Clin. Cancer Res.* 2002, *8*, 1038–1044.
115. Rose, P. G. Pegylated liposomal doxorubicin: optimising the dosing schedule in ovarian cancer. *Oncologist* 2005, *10*, 205–214.
116. Cattel, L., Ceruti, M., Dosio, F. From conventional to stealth liposomes: a new frontier in cancer chemotherapy. *J. Chemother.* 2004, *16*(Suppl. 4), 94–97.
117. Veil, D. M., Amantee, M. A., Colbert, G. T., Martin, F. J., Hilger, R. A., Working, P. K. Pegylated liposomal doxorubicin: proof of principle

using preclinical animal models and pharmacokinetic studies. *Semin. Oncol.* 2004, *31*, 16–35.

118. Ewer, M. S., Martin, F. J., Henderson, C., Shapiro, C. L., Benjamin, R. S., Gabizon, A. A. Cardiac safety of liposomal anthracyclines. *Semin. Oncol.* 2004, *31*, 161–181.
119. Waterhouse, D. N., Tardi, P. G., Mayer, L. D., Bally, M. B. A comparison of liposomal formulations of doxorubicin with drug administered in free form: changing toxicity profiles. *Drug Safety* 2001, *24*(12), 903–920.
120. Berry, G., Billingham, M., Alderman, E., Richardson, P., Torti, F., Lum, B., Patek, A., Martin, F. J. The use of cardiac biopsy to demonstrate reduced cardiotoxicity in AIDS Kaposi's sarcoma patients treated with PEGylated liposomal doxorubicin. *Ann. Oncol.* 1998, *9*, 711–716.
121. Nagore, E., Insa, A., Sanmartin, O. Antineoplastic therapy-induced palmar plantar erythrodyesthia (hand-foot) syndrome: incidence, recognition and management. *Am. J. Clin. Dermatol.* 2000, *1*(4), 225–234.
122. Rahman, A., Husain, S. R., Siddiqui, J., Verma, M., Agresti, M., Center, M., Safa, A. R., Glazer, R. I. Liposome mediated modulation of multi-drug resistance in human HL-60 cells. *J. Natl. Cancer. Inst.* 1992, *84*, 1909–1915.
123. Thierry, A. R., Vige, D., Coughlin, S. S., Belli, J. A., Dritscilo, A., Rahman, A. Modulation of doxorubicin resistance in multi-drug resistant cells by liposomes. *FASEB J.* 1993, *7*, 572–579.
124. Gabizon, A., Tzemach, D., Mak, L., Bronstein, M., Horowitz, A. T. Dose dependency of pharmacokinetics and therapeutic efficacy of pegylated liposomal doxorubicin (Doxil) in murine models. *J. Drug Target* 2002, *10*(7), 539–548.
125. Charrois, G. J. R., Allen, T. M. Drug release rate influences the pharmacokinetics, biodistribution, therapeutic activity and toxicity of PEGylated liposomal doxorubicin formulations in murine breast cancer. *Biochim. Biophys. Acta* 2004, *163*, 167–177.
126. Batist, G., Ramakrishnan, G., Sekhar, R., Rao, C. S., Chandrasekharan, A., Gutheil, J., Guthrie, T., Shah, P., Khojasteh, A., Nair, M., Hoelzer, K. K., Tkaczuk, K., Park, Y. C. L., Lee, W. Reduced cardiotoxicity and preserved antitumor efficacy of liposome-encapsulated doxorubicin and cyclophosphamide compared with conventional doxorubicin and cyclophosphamide in a randomised multicenter trial of metastatic breast cancer. *J. Clin. Oncol.* 2001, *19*, 1444–1454.
127. Williams, G., Cortazar, P., Pazdur, R. Developing drugs to decrease the toxicity of chemotherapy. *J. Clin. Oncol.* 2001, *19*, 3439–3441.

Chapter 40

Improvement of Drug Properties by Cyclodextrins

Kaneto Uekama and Fumitoshi Hirayama

I. INTRODUCTION
II. PHARMACEUTICALLY USEFUL CyDs
A. Physicochemical profiles of CyDs
B. Biological profiles of CyDs
III. IMPROVEMENT OF DRUG PROPERTIES
A. Solubilization
B. Stabilization in solution
C. Control of solid properties
D. Release control
E. Enhancement of drug absorption
F. Reduction of side-effects
G. Use in peptide and protein drugs
H. Combined use of CyDs with additives
IV. CyD-BASED SITE-SPECIFIC DRUG DELIVERY
A. Colon targeting
B. Cell targeting
C. Brain targeting
V. CONCLUSION
REFERENCES

I. INTRODUCTION

Cyclodextrins (CyDs) were first isolated in 1891 as degradation products of starch and were characterized as cyclic oligosaccharides (Figure 40.1).[1] The α-, β- and γ-CyDs are the most common natural CyDs, consisting of six, seven and eight glucose units, respectively (Table 40.1). Because of their different internal cavity diameters, each CyD shows a different degree of molecular encapsulation with different-sized guest molecules.[2,3] These CyDs have therefore been utilized for the modification of physical, chemical or biological properties of guest molecules. In the pharmaceutical field, CyDs have been recognized as potent candidates to overcome the undesirable properties of drug molecules through the formation of inclusion complexes.[4–6] Recently, a number of new dosage forms have been developed as drug delivery systems (DDS). For the design of such advanced dosage forms, various kinds of CyD derivatives have been prepared to extend the physicochemical properties and inclusion capacity of natural CyDs as novel drug carriers.[7,8] This chapter deals with some aspects of the utilization of chemically modified CyDs in pharmaceutical formulations, and will discuss some fundamental characteristics of CyDs which should be considered in the development of advanced dosage forms.

II. PHARMACEUTICALLY USEFUL CYDS

To obtain drug carrier properties better than those of natural CyDs, the hydroxyl groups of CyDs are available as starting points for structural modification, and various functional groups have been incorporated in the CyD molecule (Table 40.2). These chemically modified CyDs can be classified into three types: hydrophilic, hydrophobic and ionizable derivatives.[8] Hydrophilic derivatives such as methylated CyDs,[9,10] hydroxyalkylated CyDs,[11–15] and branched CyDs[16,17] deserve special attention because their solubility in water is high, suggesting their use as solubilizers for poorly water-soluble drugs rather than the use of surface-active agents. In contrast, hydrophobic CyDs are useful as sustained-release drug carriers of water-soluble

Copyright © 2008, Elsevier Ltd
All rights reserved.

FIGURE 40.1 Structure of β-cyclodextrin.

TABLE 40.1 Some Characteristics of Natural CyDs

CyD	Number of glucose units	Molecular weight	Cavity[a] diameter(Å)	Solubility[b] (g/dL)
α-CyD	6	973	5	15
β-CyD	7	1,135	6	1.85
γ-CyD	8	1,297	8	23

[a] *Estimated by the Corey–Pauling–Koltun model.*
[b] *In water at* 25°C.

drugs[18,19] and of peptides,[20] since they have the ability to decrease the solubility of guest molecules. On the other hand, the ionizable CyDs can modify the release rate of drug, depending on the pH of the solution[21] and bind to the surface membranes of cells, which may alter the function of biological barriers.[22] Sulfates and sulfoalkyl ethers of CyDs have been evaluated as a new class of parenteral drug carriers with heparin-mimetic biological activities. However, new CyD derivatives must be thoroughly characterized before practical use in pharmaceutical formulations.[23] In this section, some physicochemical and biological profiles of CyD derivatives are briefly described, comparing them with those of natural CyDs.

A. Physicochemical profiles of CyDs

Widespread use of natural CyDs as hosts for drugs is restricted by their low aqueous solubility, particularly that of β-CyD.[24] Methylation or hydroxyalkylation of the hydroxyl groups of β-CyD has been used to obviate this problem. For example, hydroxyalkylated CyDs are amorphous mixtures of chemically related components with different degrees of substitution.[11,12] This multicomponent character prevents crystallization, and thus the hydroxyalkylated CyDs have higher solubility (>50%) in both water and ethanol. The solubility of natural β-CyD in water increases with increase in temperature. On the other hand, 2,6-dimethyl-β-CyD (DM-β-CyD) shows exothermic dissolution in water, so that the solubility decreases with increase in temperature, because of dehydration at elevated temperatures. Thus, methylated CyDs have clouding points, a behavior similar to that of non-ionic surfactants. The solubilities of maltosyl-β-CyD (G_2-β-CyD) and hydroxypropyl-β-CyD (HP-β-CyD) show a little temperature dependence.[25,26] Such information is particularly useful for the design of aqueous injectable CyD solutions which are to be heat-sterilized. The aqueous solubility of ionizable CyDs depends on the pH of the solution. *O*-Carboxymethyl-*O*-ethyl-β-CyD (CME-β-CyD), in which the hydroxyl groups of ethylated β-CyD are substituted by carboxymethyl groups, is slightly soluble in low pH regions but freely soluble in neutral and alkaline regions owing to the ionization of the carboxyl group (pK_a 3–4).[21] Thus, CME-β-CyD can serve as an enteric-type drug carrier similar to carboxylmethylethylcellulose, but may be of greater advantage than the cellulose derivatives for the stabilization of labile drugs owing to the inclusion ability.[27]

The glycosidic bonds of CyDs are fairly stable in alkaline solution, whereas they are hydrolytically cleaved by strong acids to give linear oligosaccharides.[28] The ring-opening rate of CyDs increases with increasing cavity size, and is accelerated when the ring is distorted. For example, permethylated γ-CyD (TM-γ-CyD) having a distorted ring conformation is most susceptible to acid-catalyzed hydrolysis.[6] It should be noted that the ring-opening rate of β-CyD is decreased by the addition of guest molecules, the decrease being marked for guests with a close fit to the β-CyD cavity.[29] This reduction of rate can be attributed to the inhibition of access of catalytic oxonium ions to the glycosidic bond because the CyD cavity is occupied by guests. The glycosidic bonds of CyDs are cleaved by some starch-degrading enzymes with the proper substrate specificity, although the reaction rate is much slower than that of linear sugars.[17] Generally, the introduction of substituents on the hydroxyl groups slows enzymatic hydrolysis of CyDs by lowering the affinity of CyDs for enzymes or owing to the intrinsic reactivity of enzymes.

B. Biological profiles of CyDs

α- and β-CyDs are resistant to metabolism in the body, whereas γ-CyD, having a large cavity, is hydrolyzed even by human salivary α-amylase.[30] On the other hand, β-CyD

TABLE 40.2 Pharmaceutically Useful β-Cyclodextrin Derivatives

Derivative	Characteristic	Possible use (dosage form)
Hydrophilic derivatives		
Methylated β-CyD	Soluble in cold water and in organic solvents, surface active, hemolytic	Oral, dermal, mucosal*
DM-β-CyD		
TM-β-CyD		
DMA-β-CyD	Soluble in water, low hemolytic	Parenteral
Hydroxyalkylated β-CyD		
2-HE-β-CyD 2-HP-β-CyD 3-HP-β-CyD 2,3-DHP-β-CyD	Amorphous mixture with different degrees of substitution, highly water-soluble (>50%), low toxicity	Oral, dermal, mucosal, parenteral (intravenous)
Branched β-CyD		
G_1-β-CyD G_2-β-CyD GUG-β-CyD	Highly water soluble (>50%) Low toxicity	Oral, mucosal, parenteral (intravenous)
Hydrophobic derivatives		
Alkylated β-CyD		
DE-β-CyD TE-β-CyD	Water insoluble, soluble in organic solvents, surface active	Oral, parenteral (subcutaneous) (slow release)
Acylated β-CyD		
TAcyl-β-CyD TValeryl-β-CyD	Water insoluble, soluble in organic solvents Film formation	Oral, dermal (slow release)
Ionizable derivatives		
Anionic β-CyD		
CME-β-CyD	$pK_a = 3–4$, soluble at pH > 4	Oral, dermal, mucosal (delayed release)
β-CyD·sulphate SBE-β-CyD β-CyD phosphate	$pK_a < 1$, water soluble	Oral, mucosal Parenteral (intravenous)
Al β-CyD·sulphate	Water insoluble	(slow release)

Abbreviations: DM, 2,6-di-*O*-methyl; TM, 2,3,6-tri-*O*-methyl; DMA, acetylated DM-β-CyD; 2-HE, 2-hydroxyethyl; 2-HP, 2-hydroxypropyl; 3-HP, 3-hydroxypropyl; 2,3-DHP, 2,3-dihydroxypropyl; G_1, glycosyl; G_2, maltosyl; GUG, Glucuronyl-glucosyl; DE, 2,6-di-*O*-ethyl;TE, 2,3,6-tri-*O*-ethyl; CME, *O*-carboxymethyl-*O*-ethyl; TAcyl, 2,3,6-tri-*O*-acyl (C_2-C_{18}); TValeryl, 2,3,6-tri-*O*-valeryl; SBE, sulfobutyl ether.
* Mucosal, nasal, sublingual, ophthalmic, pulmonary, rectal, vaginal, etc.

is hardly hydrolyzed at all in whole blood of rats, rabbits, dogs and humans, and also in rat liver homogenates.[31] In the case of G_2-β-CyD, the α-1,4 glycosidic bonds in the CyD ring and the α-1,6 bond at the junction between the CyD ring and the substituents are hydrolytically stable in human body fluids. Subacute or subchronic intravenous administration of HP-β-CyD to rats and monkeys showed no significant alteration in the morphological and clinical pathology parameters.[32] When G_2-β-CyD and HP-β-CyD were administered intravenously rats, they disappeared rapidly from the plasma[31,33] and were recovered almost completely as a form of water-soluble G_1-β-CyD and intact HP-β-CyD, respectively. Since the nephrotoxicity of natural β-CyD at higher doses was ascribed to the crystallization of less-soluble β-CyD or its cholesterol complex in renal tissues,[34] the metabolic fates of G_2-β-CyD and HP-β-CyD are suggestive of lower renal toxicity compared with the parent β-CyD. To assess tolerance *via* the parenteral administration route, various blood chemistry parameters in rats and rabbits after the multiple intravenous administrations of hydrophilic β-CyDs were compared with those of the parent β-CyD.[31] Multiple injections of β-CyD or DM-β-CyD at a total dose of 900 mg/kg in rats and 1,200 mg/kg in rabbits produced some kidney and liver failure, while those for HP-β-CyD, G_2-β-CyD and β-CyD sulfate at the same doses failed to induce any kidney or liver failure, suggesting that these hydrophilic β-CyDs can be safely used in parenteral formulations.

FIGURE 40.2 *In vivo* local tissue irritancy and *in vitro* hemolytic activity of CyDs. [a]Dose (100 mg/mL) injected into *M. vastus lateralis* of rabbits. [b]The reciprocal of the concentration (w/v%) of CyDs to induce 50% lysis of human erythrocytes. [c]Used as suspension due to the limited solubility. [d]Less than 0.2 w/v%.

The hemolytic activities of natural CyDs are reported to be in the order β- > α- > γ-CyD.[35,36] These differences are ascribed to the differential solubilization of membrane components by each CyD. When the CyD cavity is modified by chemical derivatization, its effects on cell membranes can be changed dramatically from those of parent CyDs.[13,17] When the muscle tissue damage due to the injection of hydrophilic CyDs was compared with that of mannitol and non-ionic surfactants, following a single injection (100 mg/mL) of the compounds into *M. vastus lateralis* of rabbits (Figure 40.2), α-CyD and DM-β-CyD showed a relatively high irritation reaction, the degree of which corresponded to that of Tween 80®. On the other hand, G_2-β-CyD, HP-β-CyD, SBE-β-CyD, DMA-β-CyD and CyD sulfates showed no or only slight irritation reaction, the degree of which was comparable to those of γ-CyD, mannitol and HCO-60®.

It is generally recognized that the gastrointestinal (GI) absorption of CyDs in intact form is limited because of their bulky and hydrophilic nature. Only an insignificant amount of intact β-CyD was absorbed from the GI tract in rats.[37,38]

In the case of the rectal route, when oleaginous suppositories containing β-CyD derivatives were administered to the rat rectum, large amounts of intact β-CyDs were excreted into the urine up to 24 h after administration.[39] The relatively high absorption observed for β-CyDs was ascribed to a change in the permeability of the rectal mucosa and/or an interaction between the surface-active β-CyDs and glycerides, which are principle components of the suppository bases. Similarly, the hydrophilic CyDs are supposed to be hardly absorbed through the skin at all. DM-β-CyD and HP-β-CyD, however, do permeate into the skin, particularly when they are applied under occlusive-dressing conditions and/or using vehicles containing absorption-promoting agents.[40,41] HP-β-CyD, even when it is applied as an aqueous solution under non-occlusive condition, penetrates into the skin of rats and distributes homogeneously over all epidermal and dermal structures without irritation.

III. IMPROVEMENT OF DRUG PROPERTIES

An important characteristic of CyDs is the formation of inclusion complexes in both the solution and solid states, in which each guest molecule is surrounded by the hydrophobic environment of the CyDs cavity. This can lead to the alteration of physicochemical properties of guest molecules, and can eventually have considerable pharmaceutical potential (Table 40.3).[42] In the practical application of CyDs, attention should be directed toward the dissociation equilibrium and stoichiometry of the inclusion complex. When a CyD complex is dissolved in water or introduced into body fluids, it dissociates rapidly to free components in equilibrium with the complex. The stability constant (K_c) is a useful index for estimating the binding strength of the complex and changes in the physicochemical properties of a guest in the complex. The degree of dissociation is dependent on the magnitude of K_c, and various environmental factors

TABLE 40.3 Pharmaceutical Products Containing CyDs

Component	Dosage form	Efficacy	Trade name
Prostaglandin E_1-α-CyD			
20 μg/ampoule	Injection	Chronic arteriosclerotic obstruction	Prostandin
500 μg/vial	Injection	Hypertension during surgery	Prostandin 500
Prostaglandin E_2-β-CyD	Tablet	Labor induction	Prostarmon E
Limaprost-α-CyD	Tablet	Buerger's disease	Opalmon
Benexate-β-CyD	Capsule	Anti-ulcer	Ulgut, Lonmiel
Nitroglycerin-β-CyD	Sublingual tablet	Angina pectoris	Nitropen
Cefotiam hexetil hydrochloride-α-CyD	Tablet	Antibiotic	Pansporin T
Piroxicam-β-CyD	Tablet and suppository	Analgesic, anti-inflammatory	Brexin, Cicladol
Iodine-β-CyD	Gargling	Throat disinfection	Mena-Gargle
New oral cephalosporin α-CyD	Tablet	Antibiotic	Meiact
Hydrocortisone-HP-β-CyD	Liquid	Mouth washing	Dexocort
Itraconazole-HP-β-CyD	Liquid	Esophageal candidiosis	Sporanox
Ziprasidone-SBE-β-CyD	Intramuscular injection	Antipsychotic	Geodon

TABLE 40.4 Effect of Temperature on Stability Constants (K_c) of Inclusion Complexes of Ethyl 4-Biphenylylacetate with Various β-CyDs, and Their Thermodynamic Parameters in Water

Complex	K_c (1/mol)				ΔH (kJ/mol)	ΔS (J/K/mol)
	15°C	25°C	37°C	45°C		
β-CyD	3,330	3,050	2,850	1,630	−18.2	5.7
HP-β-CyD	7,950	4,200	3,000	1,650	−37.2	−55.7
DM-β-CyD	31,500	12,500	10,000	5,700	−40.4	−57.1

such as dilution, temperature, pH and additives will affect the K_c value. Since the formation of the inclusion complex is usually exothermic (Table 40.4), the dissociation is facilitated by raising the temperature.[40] In the case of ionizable guests, change of the pH of solutions shifts the equilibrium to such an extent and in such a direction that the unionized form is predominantly included in the cavity. This is due to the favoring of the inclusion of hydrophobic guests compared with hydrophilic guests. The obvious requirement for inclusion complexation is that the hydrophobic moiety of a guest molecule must fit entirely or partially into the CyD cavity. In the case of prostaglandin E_1 (PGE_1) (Figure 40.3), α-CyD, with a smaller cavity, preferentially includes an aliphatic chain of the PGE_1 molecule, while β-CyD accommodates the five-membered ring of PGE_1.[6] In contrast, the larger γ-CyD cavity accommodates PGE_1 in such a manner that the whole PGE_1 molecule penetrates the cavity. When a guest molecule is too large to be included in one CyD cavity or the host cavity is too small to include a whole guest, more than one CyD is available for complete inclusion. However, the inclusion complexation often shows a different stoichiometry depending on the guest/host concentration employed, which consequently affects the physicochemical properties of the guest molecules.

A. Solubilization

The solubilizing effect of CyDs is related both to their ability to form inclusion complexes and to the intrinsic solubility of the host molecule in water. The former factor is related to the K_c values. Among β-CyD and its derivatives, DM-β-CyD shows the highest solubilizing effect for poorly water-soluble drugs, probably owing to the elongation of the cavity.[7,9,10] The steric effects of substituents also play a role in the solubilizing ability of the host molecules. For example, HP-β-CyD, similarly to DM-β-CyD, has higher complexing ability, but its solubilizing ability is lowered

FIGURE 40.3 Cavity-size dependency of inclusion complexation of CyDs with prostaglandin E_1.

at very high degrees of substitution.[13] Glucose and maltose units in the branched β-CyD also hinder the inclusion of drug molecules.[17]

HP-β-CyD can be recommended as a parenteral drug carrier because of its low toxicity, high toleration and excellent solubilizing abilities.[43] In the preparation of an HP-β-CyD-based pharmaceutical formulation, the solubility of some lipophilic drugs in water was synergistically increased with increasing temperature, suggesting that heating or heat-sterilization may be useful steps.[44] A detailed evaluation of HP-β-CyD as an excellent solubilizer and stabilizer for brain-targeting chemical delivery systems such as estradiol-dihydropyridine conjugate has been described.[45,46] The hydrophilic CyDs can solubilize some specific lipids from biological membranes through the rapid and reversible formation of inclusion complexes, leading to an increase in the membrane permeability.[35] This may allow the extended use of CyDs as adjuvants to improve the transmucosal absorption of poorly absorbable peptide and protein drugs,[47] which will be discussed in later sections.

The solubility of guests changes according to the stoichiometry of the complexes: for example, solubilities in water of nocloprost, a derivative of PGE_1[48] and sofalcon, an anti-ulcer agent,[49] increased on 1:1 complexation with β- and γ-CyDs, respectively, owing to the partial inclusion. On the other hand, the solubilities were decreased by 1:2 complexation at higher CyD concentrations because of the complete inclusion of the whole guest molecules within the CyD dimers, which are less hydrated.

B. Stabilization in solution

The drug must remain sufficiently stable not only during storage but also in the GI fluids, since reactions which result in a product that is pharmacologically inactive or less active will reduce the therapeutic effectiveness. CyDs are known to accelerate or decelerate various kinds of reactions depending on the nature of the complex formed.[2,4] Generally, when a drug's active centre is included in the CyD cavity there is a deceleration effect, and the reaction rate is dependent primarily on the amount of free drug concentration resulting from the dissociation of the complex. On the other hand, if the drug does not totally fit into the cavity or is only partially included, leaving the active centre sterically fixed in close proximity to the catalysts, it experiences an acceleration effect.[28]

1. *Prostanoids*

Prostaglandins (PGs) are essentially long-chain unsaturated fatty acids containing a substituted cyclopentane ring system. The β-hydroxy-keto moiety of E-type prostaglandins (PGEs) is extremely susceptible to dehydration under acidic or alkaline conditions to give A-type prostaglandins (PGAs), which are isomerized subsequently to form B-type prostaglandins (PGBs) under alkaline conditions. The biological activities of PGEs decrease with progress of these reactions. The chemical instability and the low aqueous solubilities of PGEs have limited dosage form design and presented a substantial challenge to pharmaceutical scientists. Natural CyDs have successfully been applied to PGEs, and stable and soluble complexes were first marketed in Japan as a PGE_2-β-CyD tablet and PGE_1-α-CyD injection (see Table 40.3). In aqueous solution, however, attempts to stabilize PGEs were rather disappointing because of the positive catalytic effect of natural CyDs (Figure 40.4). In most cases, the stabilizing effect of TM-β-CyD is smaller than that of DM-β-CyD, preventing deep penetration of the bulky PGE_2 molecules into the narrow TM-β-CyD cavity.[50] G_2-β-CyD has been shown to improve the undesirable properties of PGE_1 in lyophilized preparations.[51] The decomposition of lyophilized PGE_1 is significantly retarded by both β-CyD and G_2-β-CyD. The rapid-dissolution property of lyophilized PGE_1 with G_2-β-CyD is maintained during storage, while it tends to decrease with β-CyD, depending on the moisture-adsorbing and wetting properties and crystallinity changes of the additives.

Prostacyclin (PGI_2) is a potent therapeutic agent in the treatment of thrombosis and relaxation of vascular smooth muscle. However, this compound undergoes an extremely facile hydrolysis of the vinyl-ether moiety to yield 6-keto-$PGF_1\alpha$ in aqueous solution losing its activity within a few minutes. Upon binding to CyDs, the rate of hydrolysis of PGI_2 is retarded in the order β- > α- > γ-CyD. The retardation effect of CyDs seems to be at least in part a result of the inhibition of intramolecular carboxylate ion catalysis due to the decrease in acidity of the terminal carboxyl group. Methylated β-CyDs in which the hydroxyl groups are blocked show a much greater retardation effect in the hydrolysis of PGI_2 in comparison with that of parent β-CyD.[52]

FIGURE 40.4 Effects of β-CyDs on the dehydration of PGE_2 (left) and isomerization of PGA_2 (right) in phosphate buffer (pH 11.0, $\mu = 0.2$) at 60°C. ○, PG alone; △, β-CyD; ▲, DM-β-CyD; □, TM-β-CyD.

2. *Cardiac glycosides*

Digoxin, one of the potent cardiac glycosides, is susceptible to hydrolysis in an acidic medium, and therapeutic efficiency as well as oral bioavailability may decrease as a result. In the degradation pathways of digoxin, prevention of the appearance of digoxigenin might be clinically important because the cardioactivity of digoxigenin is about one-tenth of that of digoxin, but other digoxosides (mono- and bis-digoxosides) possess approximately the same activity. In the presence of CyDs, acid hydrolysis of digoxin is suppressed in the order β- >- α- > γ-CyD, where β-CyD inhibits the conversion from digoxosides to digoxigenin almost completely. ^{1}H NMR data reveal that the A-ring of digoxin is located at the entrance to the α-CyD cavity, that it could penetrate further into the β-CyD cavity, and that it is loosely bound to γ-CyD.[53] This indicates that either a smaller (α-CyD) or a larger (γ-CyD) cavity is unfavorable for preventing the hydrolysis of digoxin. Moreover, the methylated CyDs suppressed the hydrolysis of digitoxin in the order DM-β- >- α- > TM-β-CyD, where DM-β-CyD completely inhibited the conversion from digitoxosides to digitoxigenin.[54] In a dissolution study of digoxin tablets, increase in dissolution rate and decrease in acid hydrolysis were achieved by β- and γ-CyDs, where the appearance of digoxigenin was almost negligible even 60 min after the initiation of the dissolution test.[55] When tablets containing a γ-CyD complex of digoxin were administered sublingually to human volunteers, the serum levels of drug were significantly increased compared with digoxin alone, owing to the prevention of acid hydrolysis in the stomach.[56] The stoichiometric ratio is in practice responsible for the dosage form design of a drug molecule. In the case of digoxin, for example, one digoxin molecule needs four molecules of γ-CyD to form a 1:4 stable crystalline complex, resulting in an eight-fold increase in the molecular weight of digoxin.[53] This may facilitate the preparation of tablets containing small amounts of drug resulting in a better uniformity of content. In the case of high-dosage drugs, the increase in molecular weight becomes a disadvantage, since the relationship between the required dose and molecular weight determines the feasibility of oral administration in CyD complexes.

C. Control of solid properties

Many solid compounds are known to exist in different crystalline modifications such as amorphous, crystalline or glassy states; affecting solubility, dissolution rate, stability and bioavailability. In order to improve the pharmaceutical potential, therefore, it is important to control the crystallization, the polymorphic transition, and whisker generation of solid drugs.[5,6]

1. *Selective crystallization of metastable polymorphs from aqueous CyD solution*

For the development of high-quality drugs, it is important to select a proper polymorph and to be able to control its polymorphic transformations, and diverse crystal forms provide valuable intellectual property protection in drug development.[57] Traditional methods to obtain stable polymorphs of a compound are recrystallization under different conditions by varying solvent composition, temperature or concentration.[58] Recent advanced approaches control crystal surfaces at the molecular level on the basis of crystal structures, which include crystallizations with tailor-made additives using structurally related molecules or polymers, utilization of epitaxial crystal growth or polymer heteronuclei and seeding with appropriate polymorphs. Polymorphic transformations proceed generally *via* the solid–solid or the solution-mediated mechanism.[59] The solid–solid transformation is dependent on internal rearrangements or conformational changes of the molecules in crystals. On the other hand, the solution-mediated transformation is controlled

FIGURE 40.5 Possible inhibition mechanism of the solution-mediated polymorphic transition of Form IV to Form I by the complexation with DM-β-CyD.

by differences in solubility of stable and metastable forms, where a metastable form with higher solubility appears first from solution and it then dissolves, nucleates and transforms into a stable form with the lowest solubility, according to "Ostwald's Rule of Stages".[60] It is expected that the solution-mediated polymorphic transformation of compounds may be suppressed by inclusion complex formation with CyDs, because CyDs generally enhance the solubility of drugs in water, slow down diffusion rate of drugs due to increase in molecular volume and inhibit aggregation of drugs due to the masking of intermolecular interaction sites by inclusion formation. On the basis of these premises, we studied effects of CyDs on the crystallization and polymorphic transformation of an oral hypoglycemic agent, tolbutamide, in aqueous solution.[61,62] The crystallization of tolbutamide from aqueous solution was conducted by titrating alkaline solutions of the drug with HCl solutions, and the contents of different polymorphs in precipitated crystals were determined by powder X-ray diffractometry. In the absence of CyDs, tolbutamide crystallized into stable Form I. Stable Form I crystals of tolbutamide were also obtained from the solutions of parent α-, β- and γ-CyDs, HP-α-CyD, HP-β-CyD, DM-α-CyD and glucose. By sharp contrast, the solution containing DM-β-CyD yielded exclusively metastable Form IV crystals of the drug, and this selective crystallization of the metastable tolbutamide crystal was observed above 0.1 mM DM-β-CyD solution. The stability constant of the tolbutamide/DM-β-CyD complex was much higher than those of other CyD complexes, and the selective crystallization of the metastable form was disappeared by the addition of inclusion competitors depending on the stability constants of competitor/CyD complexes. These results suggest an important role of the guest/host inclusion complexation in the crystallization of the metastable form. Figure 40.5 shows a possible mechanism for the selective crystallization of tolbutamide Form IV crystals in DM-β-CyD solutions. In the absence of CyDs, the metastable Form IV crystal with higher solubility precipitates from the solution first according to Ostwald's rule. However, once Form I crystal nuclei are formed even in small amounts, the tolbutamide solution becomes supersaturated state with respect to stable Form I crystals, and therefore Form I nuclei rapidly grow and Form IV crystals dissolve. In the presence of DM-β-CyD, on the other hand, an additional equilibrium of inclusion complexation is introduced to the system, and this equilibrium may compete with the nucleation and crystal growth processes, inhibiting the crystallization of Form I. The newly introduced inclusion equilibrium may work like a buffer system that prevents nucleation and the following crystal growth processes. Furthermore, a metastable chlorpropamide polymorph and an amorphous indomethacin were also obtained from DM-β-CyD solutions. DM-β-CyD complexation is useful for stabilization and isolation of Ostwald's metastable forms that occur at an early stage of crystallization. The method described here for control of polymorphic transformation by CyD complex formation will provide an opportunity to isolate labile intermediate metastable polymorphs and will be a valuable tool for the detection and preparation of new polymorphs that are undiscovered so far.

2. *Modification of crystalline states of nifedipine*

The oral bioavailability of crystalline nifedipine, a potent calcium-channel antagonist, is very low because of its poor solubility and its slow dissolution rate in water. Various hydrophilic macromolecules such as poly(vinylpyrrolidone) (PVP) and poly(ethylene glycols) have been used to improve the dissolution characteristics of nifedipine.[63] However, amorphous nifedipine in these matrices gradually crystallizes during storage at high temperature and humidity, deteriorating its rapid dissolution characteristics. We have reported that crystalline nifedipine is converted to an amorphous state by spray-drying with HP-β-CyD, and the oral bioavailability is improved significantly.[64] HP-β-CyD was useful in preventing the crystal growth of amorphous nifedipine, maintaining a relatively fine and uniform size of crystals even under adverse storage conditions (60°C, 75% RH). Moreover, the rapidly dissolving form of the metastable state (Form B: m.p. 163°C) transiently formed at an

early stage of storage of amorphous nifedipine, and the glassy state (transition temperature at 48°C) of nifedipine could be obtained by cooling melts of stable form of nifedipine (Form A: m.p. 171°C). In the presence of HP-β-CyD, the crystallization of the glassy nifedipine and the polymorphic transition of Form B to Form A were significantly suppressed, indicating that Form B could be prepared in high yield (>75%) by heating in the amorphous HP-β-CyD matrix. Although the initial dissolution rate of nifedipine increased in the order of glassy state > Form B > Form A, the glassy state of nifedipine was readily converted to Form A of larger crystal size at higher humidity and temperature. These facts suggest that HP-β-CyD is particularly useful for the selective preparation of the fast-dissolving Form B, and will provide a rational basis for the design of formulation and storage conditions in solid dosage forms of nifedipine.

3. *Stabilization of carmofur*

Carmofur (1-hexylcarbamoyl-5-fluorouracil, HCFU) is one of the masked forms of 5-fluorouracil (5-FU) and has been widely used to treat carcinomas of breast and the GI tract. It is expected that the highly hygroscopic character of natural CyDs may significantly influence degradation of HCFU, because HCFU is extremely susceptible to base- and water-catalyzed hydrolysis to give 5-FU. In fact, the degradation rate of HCFU in the solid CyD complex was very fast under accelerated conditions (70°C, 75% RH). This problem can be solved by adding an organic acid, such as citric or tartaric acid, as a pH-controlling agent.[65] The organic acids may provide an acidic environment around the CyD complex after moisture sorption since HCFU is chemically stable under acidic conditions. The methylated CyDs are highly effective in preventing the decomposition of HCFU in the solid state for long periods under accelerated storage conditions (40°C, 75% RH), because they are less hygroscopic than natural CyDs. Since the stabilizing effect of CyDs is responsible for the bioavailability of HCFU,[66,67] the acidic property of CME-β-CyD is particularly effective in improving the oral bioavailability of HCFU, preventing the degradation of HCFU into 5-FU which irritates the GI mucosa.[27]

D. Release control

For the design of advanced oral formulations, control of drug release rate from dosage forms is of critical importance in realizing their therapeutic efficacy. Most of the slow-release preparations have been aimed at achieving zero-order release of drugs to provide a constant blood level for extended periods. For this purpose, wax-type matrices or water-soluble cellulose derivatives are generally used as slow-release carriers for water-soluble drugs. When hydroxyl groups of CyDs are substituted by ethyl, acetyl or longer alkyl groups, the solubility of these CyDs in water decreases proportionally to their degree of substitution or the length of the alkyl chains.[18,19] We have demonstrated that the ethylated β-CyDs, such as DE-β-CyD and perethylated β-CyD (TE-β-CyD), are useful as slow-release carriers of isosorbide dinitrate[68] and diltiazem hydrochloride.[69] New series of peracylated β-CyDs with different alkyl chains (acetyl to octanoyl) also were prepared by acylating all hydroxyl groups of β-CyD, and their physical properties were evaluated with anticipation of more effective slow-release carriers for water-soluble drugs.[70] The concentrated solutions of peracylated β-CyDs in organic solvents were highly viscous and sticky, and gelation took place upon evaporation of the solvents. Since these properties were thought to be particularly useful for a slow-release carrier, the solid complexes of peracylated β-CyDs with water-soluble drugs were prepared and their *in vitro* and *in vivo* release behaviors were examined. Drug release was markedly retarded by the complexation with peracylated β-CyDs in the decreasing order of the solubility of the host molecules. Since the peracylated derivatives with substituents longer than the hexanoyl moiety hardly released the water-soluble drugs, peracetyl-(TA-), perbutanoyl-(TB-) and perhexanoyl-(TH-)-β-CyDs were used in the *in vivo* studies following oral administration of the hydrophobic complexes to dogs. As shown in Figure 40.6, TB-β-CyD suppressed a peak plasma level of molsidomine, a peripheral vasodilator, and maintained a sufficient drug level for long periods, while other peracylated β-CyDs with shorter or longer chains were ineffective in controlling the *in vivo* release behavior of molsidomine. It is noteworthy that the TB-β-CyD complex significantly increased the AUC and prolonged the mean residence time (MRT) of the drug. The

FIGURE 40.6 Plasma levels of molsidomine after oral administration of capsules containing molsidomine or its peracylated β-CyD complexes (equivalent to molsidomine at 10 mg/body) in dogs. ○, Molsidomine alone (diluent: starch); ●, TA-β-CyD complex; □, TB-β-CyD complex; ■, TH-β-CyD complex. Each point represents the mean of 3–6 dogs. $^*p < 0.05$ versus molsidomine alone.

FIGURE 40.7 Effect of competing agent on drug absorption from inclusion complex following dissolution and dissociation in GI tract.

prominent retarding effect of TB-β-CyD was ascribable to its muco-adhesive property and hydrophobicity compared with other peracylated β-CyDs.[71] These facts suggest that peracylated β-CyDs are useful as novel multifunctional carriers to modify the release rate of water-soluble drugs.

CME-β-CyD is effective in modifying the release rate of water-soluble drugs. For example, the CME-β-CyD complex of diltiazem, a potent calcium-channel antagonist, releases the drug very slowly in the stomach but rapidly in the intestine, the main absorption site. The *in vitro* release behavior was clearly reflected in the plasma levels of diltiazem after oral administration to dogs, where the complex produced a twofold increase in bioavailability compared with the drug alone.[21] Such enhancement may arise from the reduction of the first pass effect – the metabolism in the liver immediately after absorption is saturated owing to the high local concentration of the drug at the intestinal tract. To investigate the delayed-release characteristics of the CME-β-CyD complex, the *in vivo* release of diltiazem was evaluated using dogs whose gastric acidity had been controlled by pretreatment with tetragastrin or omeprazole.[71] Following oral administration of the CME-β-CyD complex tablets in dogs with high controlled gastric acidity, the absorption of diltiazem was significantly retarded since the complex dissolved only in the intestinal fluid after passing through the stomach. In dogs with low gastric acidity, rapid absorption of diltiazem was observed owing to the dissolution of the complex in the gastric fluid. In this delayed-release formulation, a good correlation between the *in vitro* and *in vivo* release rates of diltiazem was obtained in both groups of dogs with controlled gastric acidity.[72]

A suitable combination of various CyD derivatives is effective in sustaining the release rate of drugs. For example, the release rate of theophylline, which has a narrow therapeutic range in the blood level, can be controlled by hybridizing its hydrophilic, hydrophobic and ionizable CyD complexes.[72] This formulation has many advantages in reducing the frequency of dosing, prolonging the drug efficacy and avoiding the toxicity associated with the administration of a simple plain tablet of theophylline.

The oily injection of buserelin acetate (BLA), a luteinizing hormone-releasing hormone (LHRH) agonist, with a sustained-release feature can be achieved using hydrophobic CyDs.[73] The release of BLA from the peanut-oil suspension into the aqueous phase was significantly retarded by the complexation with TA-CyDs. A single subcutaneous injection of the oily suspension of BLA containing TA-β-CyD and TA-γ-CyD in rats provided the retardation of plasma BLA levels, giving 25 and 30 times longer MRT, respectively, than that with BLA alone. Simultaneously, the suppression of plasma testosterone levels to induce castration, the pharmacological effect of BLA, continued for 1–2 weeks and significant weight reduction in genital organs was observed due to the anti-gonadal effect. Since TA-β-CyD and TA-γ-CyD were degraded enzymatically in rat skin homogenates, both TA-CyDs can be useful as bio-degradable sustained-release carriers for the water-soluble peptides following subcutaneous injection of oily suspensions.[74]

E. Enhancement of drug absorption

Many factors affect drug absorption and they are largely dependent on the pharmaceutical formulation and administration route. In order to enhance drug absorption from CyD complexes, the dissociation equilibrium can be controlled by adjusting environmental factors.[75] Figure 40.7 shows the overall process of drug absorption from the complex in the presence of competing agent, where k_d is the dissolution rate constant, K_1 is the stability constant of the drug–CyD complex, K_2 is the stability constant of the competing agent–CyD complex and k_a is the absorption rate constant of the drug. When the solid complex is administered orally, the drug absorbed after the complex is dissolved and dissociated since only the free form of the drug in solution is capable of penetrating the lipid barrier of the GI tract. Therefore, the absorption of the drug from the complex is mainly dependent on the magnitude of the dissolution rate (k_d) and the stability constants (K_1 and K_2) of the complexes; high dissolution rates and the relative ability

of the complexes ($K_2 > K_1$) favor a free drug which is readily available for absorption. Since the displacement of drug from the cavity by exogenous and endogenous substances in the formulation and GI tract is responsible for acceleration of the drug absorption, the equilibrium can be controlled by adding an appropriate competitor and by adjusting its amount.[76] Such competition may occur not only in the body fluids containing various biological components, such as lipids and sterols, but also in pharmaceutical dosage forms containing various excipients.

1. *Gastrointestinal absorption*

The hydrophilic CyDs are useful for improving oral bioavailability of poorly water-soluble drugs, including steroids, cardiac glycosides, non-steroidal anti-inflammatory drugs, barbiturates, anti-epileptics, benzodiazepines, anti-diabetics, vasodilators, etc.[1–6] These improvements are mainly due to the increase in solubility, dissolution rate and wettability of the drugs through formation of inclusion complexes. As described in Figure 40.7, CyDs are supposed to act only as carrier materials and help to transport the drug through an aqueous medium to the lipophilic absorption site in the GI tract. Results fully confirming this view were obtained in the enhancement of the absorption of spironolactone,[77] proscillaridin,[78] prednisolone,[79] benzodiazepines,[80] cardiac glycosides,[53] acetohexamide[81] and anti-inflammatory drugs[82] using natural CyDs. Additionally, CyDs cause some modification of the GI mucosa at high concentrations owing to the removal of membrane components such as cholesterol, phospholipids or proteins.[83,84] Therefore, free CyDs after dissociation of the complex may alter the lipid barrier of the absorption site, which consequently facilitates drug absorption. However, despite the larger stability constant, DM-β-CyD significantly enhances both the extent and rate of absorption of drugs compared with parental β-CyD, including fat-soluble vitamins,[3,85,86] anti-inflammatory drugs[87,88] and anti-tumor drugs.[89] For example, the rapidly dissolving DM-β-CyD complex of α-tocopheryl nicotinate, following oral administration to fasting dogs, resulted in the area under the plasma concentration–time curve (AUC) being about 70 times as great as that of the drug alone.[90] To explain this anomalous enhancing effect of DM-β-CyD on fatty drugs, additional factors influencing drug absorption were considered. Since highly surface-active DM-β-CyD is supposed to substitute bile in GI tract,[3] the lipophilic drugs can be finely emulsified by DM-β-CyD, which eventually provides the marked GI absorption in the fasting condition.

2. *Transdermal absorption*

Recent studies on human volunteers demonstrated that CyDs have a significant safety margin in dermal application.[91] Optimized release of the drug from the topical preparation containing its CyD complex may be obtained by using a vehicle in which the complex is barely dissociated and maintains a high thermodynamic activity. For example, the *in vitro* release rate of corticosteroids from water-containing ointments (hydrophilic, absorptive or polyacrylic base) was markedly increased by the hydrophilic CyDs, whereas in other ointments (a fatty alcohol propylene glycol (FAPG) or macrogol base) CyDs retarded the drug release.[92] The enhancement of drug release can be ascribed to the increase in solubility, diffusibility and concentration of the drug in the aqueous phase of the ointment through water-soluble complex formation. Moreover, the drug in the CyD complex may be displaced by some components of the treatment, depending on the stability constant of the complex.[93,94] CyDs also interact with some components of skin. For example, DM-β-CyD markedly extracts cholesterol from rabbit skin *in vitro* (Figure 40.8), a process which may reduce the function of skin as a barrier and may eventually contribute in part to the enhancement of drug absorption.

FIGURE 40.8 Release profiles of cholesterol from rabbit skin treated with β-CyDs (1.2 mM) in isotonic phosphate buffer (pH 7.4) at 37°C. ○, Control; □, β-CyD; ■, DM-β-CyD. Each point represents the mean ± SE of 12 experiments.

When a hydrophilic ointment containing anti-inflammatory ethyl 4-biphenylylacetate (EBA) or its CyD complexes was applied to the skin of rats, the release of EBA from the ointment into the skin was assisted by DM-β-CyD or HP-β-CyD, while β-CyD had no appreciable effect.[41] Although the entry of CyDs into the skin made the stratum corneum more resistant to the permeation of the drug, the greater release of EBA from the vehicle compensated for the negative effect on the skin permeation of the drug and delivered the drug more effectively to the site of action. The application of EBA in the complexed form resulted in a transient rise in the total amount of EBA and its active metabolite biphenylylacetic acid (BPAA) in the stratum corneum and in the viable skin. Interestingly, the fraction of active BPAA in the viable skin was increased when EBA was applied in the complexed form, indicating that β-CyDs assist the bioconversion of EBA to BPAA in the skin and consequently

facilitate the delivery of active BPAA to subcutaneous tissues, where its action is most desired. In the model of carrageenan-induced acute edema in rat paw, the inflammation was inhibited by pretreatment with ointments containing EBA-CyD complexes of DM-β-CyD or HP-β-CyD.[41,95]

3. *Rectal absorption*

The release of drug from suppository bases is one of the important factors in the transmucosal absorption of drugs, since the rectal fluid is small in volume and is viscous compared with GI fluid. Generally, the hydrophilic CyDs enhance the release of poorly water-soluble drugs from fatty bases because of the lesser interaction of the resultant complex with vehicles,[96,97] which eventually improves the rectal absorption. For example, the methylated CyDs significantly enhance the rectal absorption of hydrophobic drugs such as flurbiprofen,[40] HCFU[67] and BPAA.[26] The superior effect of methylated CyDs can be explained by the faster release of the drug together with the lowering of the affinity of the complexed drug for the oleaginous suppository base. HP-β-CyD is particularly useful for improving the rectal absorption of hydrophobic drugs from fatty bases. The most illuminating effect of HP-β-CyD was obtained for the enhancement of rectal absorption of EBA, a lipophilic prodrug of BPAA.[26,97] The relative potency of β-CyD analogs in enhancing the dissolution rate of EBA in water and reducing the binding affinity of drugs to the fatty base was DM-β-CyD complex > HP-β-CyD complex > β-CyD complex; this order was consistent with that of the magnitudes of the stability constants of the complexes. However, *in vivo* absorption of EBA was enhanced in the order HP-β-CyD complex > DM-β-CyD complex > β-CyD complex in rats after single and multiple administrations of suppositories containing the complexes. The enhancement of rectal absorption of EBA *in vivo* can be explained by the facts that HP-β-CyD increases the release rate of EBA from the vehicle and stabilizes it in the rectal lumen and that the drug is partly absorbed in the form of the complex. The rather small enhancing effect of DM-β-CyD was ascribable to the considerable dissociation of the complex in the vehicle together with increased viscosity of the suppository base, since DM-β-CyD is extremely surface-active and oil-soluble compared with HP-β-CyD. Consequently, HP-β-CyD had the highest potential to improve rectal absorption of EBA among the three β-CyDs tested.

4. *Nasal absorption*

Highly water-soluble CyD complexes of steroid hormones are well suited to nasal administration.[98,99] This kind of formulation may provide a rapid rise of drug levels in systemic circulation and avoid intestinal and hepatic first pass metabolism of the drugs. Inherently, the blood level of endogenous hormones rises a few times a day in episodes lasting approximately 1 h. Such pulsatile release of steroids can be imitated by the nasal administration of water-soluble CyD complexes, which may provide some desirable pharmacological profiles as demonstrated in sublingual administration of a rapidly dissolving complex of steroids with HP-β-CyD.[100] The effects of CyDs on the nasal epithelial membranes seem to be of minor importance for the absorption enhancement, because CyDs will lose their ability to interact with the membranes when their cavities are occupied by the steroids. Nasal preparations must be evaluated critically for their possible effect on the nasal mucociliary functions, which are known to defend the respiratory tract against dust, allergens and bacteria. When compared with other absorption-promoting agents and preservatives used in nasal formulations, DM-β-CyD exerted only a minor effect on the ciliary beat frequency of human nasal adenoid tissue *in vitro*.[101] This may represent an advantage of DM-β-CyD over the enhancers in promoting the nasal absorption of drugs, especially for short-term therapy. A nasal spray containing estradiol solubilized by DM-β-CyD was effective in the treatment of symptoms of estrogen deficiency in bilateral oophorectomized women, and the twice daily administration of this formulation over a period of 6 months was well-tolerated by the patients.[102]

F. Reduction of side-effects

The molecular entrapment of a drug into the CyD cavity may prevent direct contact of the drug with biological surfaces, and both the drug's entry into the cells of non-targeted tissues and local irritation are thus decreased. Since the CyD complex eventually dissociates into its components, there is no drastic loss of the therapeutic benefits of the drug. Therefore, CyDs act as wafer-like carriers which decrease the drug-induced local tissue damage at the administration site and then deliver the drug close to the site of its action.

1. *Reduction of local irritancy*

CyDs alleviate muscular tissue damage following the intramuscular injection of drugs.[103,104] The protective effects of CyDs may be attributed mainly to the poor affinity of the hydrophilic complexes of drugs for the sarcolemmal membranes of muscle fibres, a situation expected from the results of *in vitro* hemolysis studies. Similarly, CyDs diminished the ulcerogenic potency of several acidic anti-inflammatory drugs when these were administered orally.[105] In addition, HP-β-CyD significantly reduced the irritation of rectal mucosa in rats caused by BPAA, both for single and multiple administrations of EBA-β-CyD complexes in oleaginous suppositories (Figure 40.9).[97]

Chlorpromazine (CPZ), a typical anti-psychotic agent, frequently causes cutaneous phototoxic and photoallergic

FIGURE 40.9 Irritation effects of single or repeated administration (four times, 12-h intervals) of EBA or its β-CyD complexes as suppositories (10 mg/kg as EBA) on rectal mucosa in rats. Open bars, EBA alone; hatched bars, β-CyD complex; tinted bars, DM-β-CyD complex; solid bars, HP-β-CyD complex. Each value represents the mean ± SE of five rats.

FIGURE 40.10 Effect of β-CyD sulfate (300 mg/kg/day, i.p., 6 h post-administration) on serum, levels of urea nitrogen in rats treated with gentamicin (100 mg/kg/day, s.c.) for 14 days. ○, Saline control; △, β-CyD sulfate alone; ●, gentamicin alone; ▲, gentamicin with β-CyD sulfate. Each value represents the mean ± SE of 3–12 rats. $^{*}p < 0.05$ versus saline control.

responses in patients being treated with prolonged and high doses. These adverse effects may be mainly attributable to the toxic photoproducts of CPZ. We have reported that DM-β-CyD significantly reduced the photosensitized skin irritation caused by CPZ in guinea pigs according to gross and histological examination.[106] The inhibitory effect of DM-β-CyD was ascribed to the alteration of the photochemical reactivity of CPZ, rather than the direct interaction of the photoproducts with DM-β-CyD. When CPZ was photoirradiated with DM-β-CyD, promazine, which is less toxic than CPZ, was produced in high yield.[107] In addition, DM-β-CyD suppresses the formation of numerous oxidation and polymerization photoproducts which are responsible for CPZ-photosensitized skin irritation. β-CyD derivatives also decreased the photoinduced free-radical production from CPZ,[108] the photodecarboxylation of benoxaprofen[109] and the photodimerization of protriptyline,[110] resulting in the reduction of phototoxicity.

Systemic detoxication

The addition of β-CyD to dialysis fluids accelerated the removal of phenobarbital by peritoneal dialysis, thereby proving effective in the treatment of drug overdose.[111] HP-β-CyD is useful not only in the administration of drugs, but also in redistributing other endogenous lipophiles in the body. When HP-β-CyD was infused intravenously into a patient with familial hypervitaminosis A, retinyl ester overloading in the liver was reduced.[112] The possibility of HP-β-CyD redistributing cholesterol deposited in the vascular systems of hereditary hyperlipidaemic rabbits was reported.[113] A single intravenous administration of HP-β-CyD to rabbits slightly and temporarily decreased the level of total cholesterol in the serum. Similar results were obtained using normal rats.[114] Repeated administration of HP-β-CyD to the rabbits led to a gradual increase in total cholesterol in the circulation and eventually to a slight relief of atherosclerotic lesions in the thoracic aorta. HP-β-CyD may serve as an artificial lipid carrier to catalytically augment the establishment of equilibria in lipid distribution.[115]

Gentamicin, an aminoglycoside antibiotic, is widely used in the clinical treatment of Gram-negative infections, but its use is sometimes complicated by the development of drug-induced acute renal failure. Gentamicin is thought to interact with negatively charged phospholipids of lysosomal membranes in the proximal tubular cells, the interaction of which may lead eventually to lysosomal dysfunction, resulting in necrosis of the cells. Since some polyanions such as dextran sulfates are able to interact electrostatically with gentamicin and to reduce the drug's entry into the renal cortex, the effects of CyD sulfates on development of rat renal dysfunction induced with gentamicin were studied.[116] Daily subcutaneous infection of gentamicin (100 mg/kg, 14 days) developed nephrotoxicity in the rat, as assessed by an increase in serum urea nitrogen (Figure 40.10) and histopathological changes in the renal cortex. When CyD sulfates were given intraperitoneally at 300 mg/kg at 6-h intervals after gentamicin administration, they protected the rats against the drug-induced renal impairment, while parent CyDs were ineffective. Since the post-administration of CyD sulfates did not reduce the total amount of gentamicin accumulated in the kidney, the protection may occur through

interference with intracellular events leading from the drug accumulation to nephrotoxicity. Since CyD sulfates have anti-inflammatory activity and have high affinity for growth factors and stabilize them against proteolysis, the effect of CyD sulfates on the renal regeneration processes should be considered as another possible protective mechanism. In any case, CyD sulfates may serve as potent antidotes against renal failure associated with aminoglycoside treatment.

G. Use in peptide and protein drugs

There are considerable hurdles in the practical use of biologically active peptides and proteins because of chemical and biological instability, poor absorption through biological membranes, rapid plasma clearance, peculiar dose–response curves and immunogenicity. Many attempts have addressed these problems by chemical modification or by co-administration of adjuvants to promote absorption of peptides and proteins and protect them from proteolytic enzymes. The absorption-enhancing effects of CyDs mimic those of bile salts in regard to increased membrane permeability accompanied by inhibition of proteolysis, although they may differ somewhat in their manner of action on membranes. Thus, CyD complexation seems to be an attractive alternative to these approaches, but the field is still in its infancy.[75]

1. *Absorption enhancement*

Internasal delivery of peptide and protein drugs is severely restricted by pre-systemic elimination due to enzymatic degradation or mucociliary clearance and by the limited extent of mucosal membrane permeability. α-CyD has been shown to remove some fatty acids from nasal mucosa and to enhance the nasal absorption of leuprolide acetate in rats and dogs.[117] The utility of chemically modified CyDs as absorption enhancers for peptide drugs in rats has been demonstrated.[75] For example, DM-β-CyD was shown to be a potent enhancer of insulin absorption in rats,[118] and a minimal effective concentration of DM-β-CyD for absorption enhancement exerted only a mild effect on the *in vitro* ciliary movement.[119] The scope of interaction of insulin with CyDs is limited, because CyDs can only partially include the hydrophobic amino acid residues in peptides with small stability constants.[118] Under *in vivo* conditions, these complexes will readily dissociate into separate components, and hence the displacement by membrane lipids may further destabilize the complexes. The direct interaction of peptides with CyDs is therefore of minor importance in the enhancement of nasal absorption. Of the hydrophilic CyDs tested, DM-β-CyD had the most prominent inhibitory effect on the enzymatic degradation of both BLA and insulin in rat nasal tissue homogenates. Because of the limited interaction between peptides and CyDs, they may reduce the proteolytic activities of enzymes by preventing the formation of the enzyme–substrate complexes. This view was supported by the following observations. Leucine aminopeptidase in the nasal mucosa is known to cleave the B-chain of insulin from the N-terminal end. CyDs, especially DM-β-CyD and HP-β-CyD, reduce the activity of leucine aminopeptidase in a concentration-dependent manner.[120] The inhibition of proteolysis by these CyDs may participate in the absorption enhancement of peptides. Another potential barrier to the nasal absorption of peptide and protein drugs is the limitation in the size of hydrophilic pores through which they are thought to pass. The methylated CyDs significantly extracted membrane lipids, depending on the size and hydrophobicity of the CyD cavity in which lipids were included.[36] Therefore, lipid solubilization mediated by CyDs may result in transcellular processes, and these changes could be transmitted to the paracellular region, which is the most likely route for the transport of polypeptides.

The combined effect of β-CyD with absorption enhancers such as sodium glycocholate or Azone® on the nasal absorption of human fibroblast interferon-β in powder form in rabbits has been described.[121] HP-β-CyD was useful as a biocompatible solubilizer for lipophilic absorption enhancers involved in the nasal preparations of peptides.[122] When insulin was administered nasally to rats, simultaneous use of an oily penetration enhancer, HPE-101, (1-[2-(decylthio)-ethyl]azacyclopentane-2-one) or oleic acid solubilized in HP-β-CyD showed a marked increase in serum immunoreactive insulin levels and marked hypoglycemic (Figure 40.11). The potentiation of the enhancing effect of HPE-101 by HP-β-CyD can be explained by the facilitated transfer of HPE-101 into the nasal mucosa. Studies on the release of membrane proteins and scanning electron microscopic observations of rat nasal mucosa indicated that the local mucosal damage due to the combination with HP-β-CyD may not be serious obstacles to their safe use.

2. *Stabilization*

Initial use of HP-β-CyD in peptide formulations has been reviewed.[123] For example, HP-β-CyD prevented the heat-induced charge alterations and loss of antigen-binding capacity of lyophilized monoclonal antibody (MN12) during storage.[124] Also, G_2-β-CyD has been shown to be an effective stabilizer in preventing the degradation of lyophilized tumor necrosis factor (TNF) during storage and freeze-drying processes.[125]

Basic fibroblast growth factor (bFGF) is a potent mitgen that stimulates the proliferation of a wide variety of cells and could play a crucial role in wound healing processes. The therapeutic potential of bFGF has not been fully realized, however, because of its susceptibility to proteolytic inactivation and short duration of retention at the site of action. Sulfated oligosaccharides, including a sodium salt of

FIGURE 40.11 Serum levels of immunoreactive insulin (IRI) and glucose after nasal administration of insulin (21 U/body) with HPE-101 (1% w/v) and/or HP-β-CyD (10% w/v) to rats. ○, Insulin alone; ▲, with HP-β-CyD; △, with HPE-101; ■, with HPE-101 and HP-β-CyD. Each point represents the mean ± SE of four rats.

β-CyD sulfate (Na β-CyD sul), have high affinity for bFGF and protect it from heat, acid and proteolytic degradation. Unfortunately, the highly hydrophilic nature of Na β-CyD sul is not suited to the design of bFGF formulations with controlled-release features. A water-insoluble aluminum salt of β-CyD sulfate (Al β-CyD sul) was prepared, and its possible utility as a stabilizer and sustained-release carrier for recombinant human bFGF was evaluated. An adsorbate of bFGF with Al β-CyD sul was prepared by incubating the protein with a suspension of Al β-CyD sul in water.[126] The mitogenetic activity of bFGF released from the adsorbate, as indicated by the proliferation of kidney cells of baby hamster (BHK-21), was almost comparable with that of the intact bFGF. Interestingly, Al β-CyD sul significantly protected bFGF from proteolytic degradation by pepsin and a chymotrypsin compared with their sodium salts and other oligosaccharides (Table 40.5). The *in vivo* release of bFGF from the adsorbate was proportional to the increase in the Al β-CyD sul/protein ratio in the adsorbate. Of the bFGF preparations evaluated, the adsorbate of bFGF with Al β-CyD sul, when given subcutaneously to rats, showed the most prominent increase in the formation of granulation tissues, owing to the stabilization and slow-release of the mitogen. These results suggest that the adsorbate of bFGF with A1 β-CyD sul has potent therapeutic efficacy for wound healing, and may be applicable to oral protein formulations for the treatment of intestinal mucosal erosions.

TABLE 40.5 Effects of β-CyD Suls and Additives (25 mg/mL) on Stability of bFGF(250 μg/mL) in the Presence of Pepsin(50 μg/mL, pH 1.6) for 3 h at 37°C

CyDs	pH	Remaining bFGF(%)*
Without additives	1.67	ND
α-CyD	1.67	ND
β-CyD	1.60	ND
γ-CyD	1.66	ND
HP-β-CyD	1.65	ND
NA·α-CyD·Sul	1.74	1.6 ± 1.3
NA·β-CyD·Sul	1.74	3.3 ± 1.4
NA·γ-CyD·Sul	1.73	5.1 ± 1.0
NA·HP-β-CyD·Sul	1.63	4.5 ± 3.2
Al·α-CyD·Sul	3.81	96.4 ± 1.2
Al·β-CyD·Sul	3.88	99.2 ± 2.6
Al·γ-CyD·Sul	3.82	96.5 ± 3.1
Al·HP-β-CyD·Sul	3.86	93.6 ± 2.1
Dextran sulfate	1.84	20.2 ± 9.6
Sucralfate	3.78	59.5 ± 6.1

*Each value represents the mean ± SD of three determinations. ND: Not detectable.

3. *Prevention of self-association*

Self-association of the insulin molecule into oligomers and macromolecular aggregates leads to complications in the development of long-term insulin therapeutic systems and limits the rate of subcutaneous absorptions, a process which is too slow to mimic the physiological plasma insulin profile at the time of meal consumption. These problems are further complicated by the tendency for insulin to adsorb on to the surfaces of containers and devices, perhaps by mechanisms similar to those inducing aggregation. Thus, many attempts have been made to prevent aggregation and

FIGURE 40.12 Effects of β-CyDs (0.1 M) and EDTA (0.2 mM) on permeation of insulin (0.1 mM) through ultrafiltration membranes in phosphate buffer (pH 6.8, $\mu = 0.2$) at 25°C. (A) Insulin alone, (B) with EDTA, (C) with HP-β-CyD, (D) with G_2-β-CyD. The membranes of YM-10, YM-30 and XM-50 were rated at pore sizes 10,000; 30,000 and 50,000 nominal molecular weight cut-off, respectively.

surface adsorption in parenteral formulations.[125] Among the hydrophilic CyDs tested, HP-β-CyD and G_2-β-CyD significantly reduced the adsorption to containers and the self-association of insulin at neutral pH. Both hydrophilic β-CyDs facilitated the permeation of insulin through ultrafiltration membranes (Figure 40.12). The increased permeation of insulin was much greater than that of EDTA which is known to prevent the self-association of insulin by sequestering zinc ions from insulin oligomers. By the addition of β-CyDs, the surface tension of insulin solutions was increased, whereas it was decreased by EDTA. In the circular dichroism (CD) spectra of insulin, β-CyDs increased the negative CD intensity around 208 nm assigned to the a helix structure of insulin, while it decreased that around 275 nm assigned to the anti-parallel β-structure of insulin oligomers. These spectral changes were in good agreement with those observed when insulin aggregates are dissociated to monomer or lower-order aggregates. In ^{1}H NMR measurements, G_2-β-CyD significantly altered the aromatic regions of insulin, particularly B24(Phe) and B26(Tyr), which are involved in the association of insulin to form an anti-parallel β-sheet around the carboxyl terminal of the B-chain. The inhibition mechanism of β-CyDs seems to be different from that of EDTA: EDTA sequesters the binding zinc ions from insulin and dissociates the hexamer or higher-order aggregates to the dimer, whereas β-CyDs may interact with hydrophobic amino acid residues to prevent the direct contact of insulin molecules.

4. *Polypseudorotaxane-based sustained release of insulin*

Pegylation technology has been used to improve therapeutic efficacies of a range of molecules, from proteins both small and large, through liposomes and viruses.[127] For example, when poly(ethylene glycol) (PEG) is covalently attached to a protein, it transfers many of the polymer's favorable characteristics to the resulting conjugate, i.e. a number of benefits such as increased circulating half-life, enhanced proteolytic resistance, reduced antigenicity and immunogenicity, reduced aggregation, improved bioavailability, etc. There are many examples of pegylation of proteins such as adenosine deamidase, insulin, interferon-α2, β-lactoglobulin, α-chymotrypsin, lipase, bovine liver catalase, asparaginase, superoxide dismutase, etc. of which the first three conjugates are on the market.[128] Recently, supramolecular assemblies have attracted a great attention due to its intriguing topologies and its application in various fields such as nanodevices, sensors, molecular switches, DDS etc. Macrocyclic compounds are most often used as host molecules in supramolecular chemistry, of which CyDs have been widely applied to drug delivery system because of their good bioadaptability. Harada et al. first reported the supramolecular assemblies of PEG and α-CyD, in which a number of the cyclic molecule are spontaneously threaded onto the polymer chain.[129] These assemblies are formed by mixing both components in water giving precipitates of the supramolecular complexes. These complexes are called polypseudorotaxane because the CyDs can be dethreaded of the polymer chain when dissolved in water. This complexation shows the size-dependency, i.e. the small cavity of α-CyD forms the polypseudorotaxane with PEG while the large cavity of β-CyD with poly(propylene glycol). When both ends of the polymer chains in polypseudorotaxanes are covalently capped with bulky molecules, CyDs are trapped in and can not be dethreaded from the assembly giving so-called polyrotaxane. We will show a new polypseudorotaxane-based sustained release system of proteins using pegylated insulin.

The PEG moiety (molecular weight about 2,200) was introduced at the lysine residue (Lys^{B29}) in insulin B-chain using α-succinimidyloxysuccinyl-ω-methoxy-polyoxyethylene. Figure 40.13 shows the solutions after mixing the pegylated insulin and α-, β- and γ-CyDs solutions and standing for 12 h at 4°C.[130] The α- and γ-CyD solutions gave white precipitates, whereas the β-CyD solution gave no precipitates, indicating the formation of polypseudorotaxanes of the pegylated insulin with α- and γ-CyDs, the phenomenon was same as those observed for PEG/CyD systems reported by Harada et al.[129] ^{1}H-NMR spectroscopic studies indicated that 20 and 11 moles of α- and γ-CyDs, respectively, are involved in the polypseudorotaxane formation with one PEG chain in the pegylated insulin, suggesting that the PEG moiety of the pegylated insulin is embedded in the stacked α-CyD host channel, as shown in Figure 40.13, while in the case of the γ-CyD polypseudorotaxane, two PEG chains of a single or separate pegylated insulin molecules are in the host channel. On the other hand, β-CyD cavity is too large to include the PEG chain and too small to include two PEG chains giving no precipitates of

FIGURE 40.13 Photographs of precipitates of pegylated insulin/α- and γ-CyD polypseudorotaxanes and their interaction modes.

FIGURE 40.14 Plasma glucose levels after subcutaneous administration of pegylated insulin; (○), and pegylated insulin/γ-CyD polypseudorotaxane in the presence of 116 mg/mL; (△), or 232 mg/mL; (▲), γ-CyD in rats.

the polypseudorotaxanes. The pegylated insulin/α- and γ-CyD polypseudorotaxanes were less soluble in water, and the release rate of the drug decreased in the order of drug alone > the γ-CyD polypseudorotaxane > the α-CyD polypseudorotaxane. Figure 40.14 shows the plasma glucose level–time profiles after subcutaneous administration of the pegylated insulin in phosphate buffer (pH 7.4) or its γ-CyD polypseudorotaxane in the presence of γ-CyD.[130] When the pegylated insulin alone was injected, the minimal levels of glucose occurred about 2 h after injection and then the plasma glucose levels recovered within 6 h in basal levels. On the other hand, the γ-CyD polypseudorotaxane significantly sustained the hypoglycemic effect and its sustained effect became greater with increase in γ-CyD concentration added in the injection medium, e.g. the glucose level recovered after more than 12 h in the presence of 232 mg/mL γ-CyD, because the dethreading rate, i.e. release rate, of the pegylated insulin is suppressed by the addition of excess amounts of γ-CyD. Further, the area under the plasma glucose level–time curve was significantly increased by the administration of the γ-CyD polypseudorotaxane. The time to nadir blood glucose concentration (T_{nadir}) and the cumulative percentage of changes in the plasma glucose levels up to 12 h post-administration (AUC_G) values were as follows: 2 h and 265 (% h) for the pegylated insulin alone, 2 h and 305 (% h) for the γ-CyD polypseudorotaxane in the presence of 116 mg/mL γ-CyD and 4 h and 443 (% h) for that in the presence of 232 mg/mL γ-CyD. These results indicate that the pegylated insulin/CyD polypseudorotaxanes can work as a sustained drug release system, and the polypseudorotaxane formation may be useful as a sustained drug delivery technique for pegylated proteins and peptides.

H. Combined use of CyDs with additives

In the previous sections, the inclusion ability of CyDs has mainly been focused on modification of the pharmaceutical properties of drug molecules. However, practical formulations usually contain considerable amounts of pharmaceutical excipients and/or additives to maintain the efficacy and safety of the drug molecules. To extend their functions, attention should be given to the improvement of the properties of pharmaceutical additives by means of complexation with CyDs. Since CyDs are capable of modifying the physicochemical properties of hydrophobic guest molecules, a suitable combination of CyDs with additives will be particularly useful in the potentiation of drug efficacy.[75] In this section, therefore, a new strategy of CyDs applications will be described for the design of advanced dosage forms.

1. *Modified release of nifedipine*

Conventional formulations of nifedipine must be dosed either twice or three times daily because of the short elimination half-life, due to considerable first pass metabolism, producing significant fluctuations in plasma drug concentrations. To attain a prolonged therapeutic effect and a reduced incidence of side-effects, many attempts have been made to maintain a suitable plasma level of nifedipine for long periods with much reduced frequency of administration. In the design of a modified-release formulation of nifedipine the following desirable attributes were sought: (a) the release should be at least 90% within 6–8 h, which is the average passage time of a tablet in the human GI tract after oral administration (b) adequate release of nifedipine in the early stage is necessary to reduce the significant first pass metabolism in the intestine and liver, which may offer a more balanced bioavailability (c) nifedipine should be released according to zero-order kinetics over an extended period, to take advantage of such therapeutic advantages as duration of pharmacological effect and reduction of side-effects. With

FIGURE 40.15 Release profiles of nifedipine from double-layer tablets (25 mg as 5 mg nifedipine) consisting of nifedipine-HP-β-CyD with 3% HCO-60 as fast-release portion (FRP) and nifedipine-HPC-M as slow-release portion (SRP) in various weight ratios (FRP:SRP) in water at 37°C. ○, 1:3; ●, 1:2; △, 1:1; ▲, 2:1; □, 3:1.

FIGURE 40.16 Effects of topically applied PGE_1 or its β-CyD complexes in FAPG ointments (500 mg, 0.01% PGE_1) supplemented with HPE-101 (3 w/w%) on laurate-induced peripheral occlusive sequelae in ears of rabbits. ■, FAPG base; ○, PGE_1 alone; ●, PGE_1 with HPE-101; △, β-CyD complex with HPE-101; ▲, CME-β-CyD complex with HPE-101. Each point represents the mean ± SE of 3–6 rabbits. $^{*}p < 0.05$ versus PGE_1 with HPE-101.

these goals, a series of formulations of double-layer tablets containing a fast-release portion (FRP) and a slow-release portion (SRP) were designed.[131,132] The *in vitro* release rate of nifedipine from tablets containing nifedipine-HP-β-CyD (molar ratio of 1:1) and small amount of HCO-60®, non-ionic surfactant was much greater than those of drug or the complex alone, and this fast-release property was maintained for long periods of storage. In contrast, the release of nifedipine from hydroxypropylcellulose (HPC) solid dispersions was retarded with increasing viscosity of the HPC (L, M, H) or amount of each HPC. An optimal formulation of the double-layer tablet was tested by changing the ratios of the components (Figure 40.15).[133] A double-layer tablet consisting of HP-β-CyD with 3% HCO-60®/(HPC-L:HPC-M) in a weight ratio of 1/(1.5:1.5) was found to be an appropriate modified-release formulation because it showed zero-order release from the SRP after the burst release from the FRP. This preparation exhibited prolonged plasma nifedipine levels without decrease of AUC in dogs, and the retarding behavior was superior to that of a commercially available slow-release nifedipine product (Adalat L-20®). It is noteworthy that the release rate of nifedipine from this formulation was hardly affected by the pH of the medium and the rotation speed of the paddle in the dissolution test during accelerated storage (60°C, 75% RH). These findings suggest that a combination of HP-β-CyD, HCO-60® and HPCs can serve as a modified-release carrier of nifedipine and can be applied to other poorly water-soluble drugs with short elimination half-lives.

2. *Transdermal delivery of prostaglandin*

Transdermal delivery of PGE_1 as an alternative to parenteral injection has engendered much recent interest in the treatment of peripheral vascular disorders. Since PGE_1 is chemically unstable and permeates poorly into the skin, a topical preparation capable of overcoming these problems needs to be devised to realize the full therapeutic potential of PGE_1. CME-β-CyD markedly improved the chemical stability of PGE_1 in aqueous solution and ointments.[134] Topical application of an ointment containing PGE_1-CME-β-CyD complex supplemented with HPE-101, an oily penetration enhancer, on to the skin of hairless mice resulted in a significant increase in the cutaneous blood flow owing to the vasodilating action of PGE_1.[135] The combination of CME-β-CyD and HPE-101 enhanced the percutaneous penetration of PGE_1 in a synergistic manner; CME-β-CyD assisted the release of HPE-101 from the ointment base and its entry into the skin which may facilitate the percutaneous penetration of PGE_1.[136] Furthermore, this formulation suppressed the bioconversion of PGE_1 to give less pharmacologically active metabolites during the passage through the skin, which delivers intact PGE_1 more effectively to the site of action. Figure 40.16 shows the effects of topically applied PGE_1 or its CyD complexes in FAPG ointments supplemented with HPE-101 on the peripheral vascular occlusive sequelae induced by sodium laurate in the ears of rabbits.[137] After the intra-arterial injection of sodium laurate, vascular occlusive lesions became progressively aggravated finally resulting in degeneration or necrosis over the entire area of the ear. PGE_1 and its CyD complexes, when supplemented with HPE-101, significantly protected against the vascular occlusive sequelae. In particular, the ointment containing PGE_1–CME-β-CyD complex supplemented with HPE-101 showed the most marked inhibitory effect on the progress of the lesions; only regional discoloration was observed on day 7 after the injection of laurate. The inhibitory effect of

PGE_1 ointments was consistent with the sequence of magnitude of their vasodilating actions. Skin compatibility tests of PGE_1 ointments suggested no serious obstacles to their safe use. Accordingly, topical application of PGE_1 may be a rational choice for minimizing systemic side-effects and patient compliance for long-term therapy.

3. *Transmucosal delivery of morphine*

The potent opioid morphine has been used for pre-operative sedation and anxiolysis in pediatric patients and for the management of post-operative pain. In advanced cancer patients, there is a great clinical need for the development of long-acting rectal preparations of morphine in the treatment of intractable chronic pain. However, the use of prolonged release of drugs from suppositories is limited showing a wide inter-individual variation of rectal bioavailability. We have demonstrated that some hydrophilic CyDs enhance the rectal absorption of morphine in rabbits and xanthan gum, a polysaccharide-type polymer with high swelling capacity, sustains the *in vitro* release of morphine from the suppository.[138] An attempt was made to design a more effective rectal delivery system for morphine using α-CyD as an absorption enhancer and xanthan gum as a swelling hydrogel in Witepsol H-15 hollow-type suppositories, and this was tested in rabbits.[139] With the combination of α-CyD and xanthan gum, the plasma morphine level was significantly sustained with an improved bioavailability. Observation of the distribution behavior of suppositories in rabbit rectum and colon after rectal administration showed that the xanthan gum prevented the upward spread of the drug in the rectum. This suggests that the absorption site for morphine should be limited to the lower part of the rectum so as to eliminate the first pass metabolism of morphine in the liver. In addition, gross and microscopic observations indicated that this preparation was less irritating to the rectal mucosa. From the viewpoints of safety and efficacy, therefore, this retentive opioid preparation may promise excellent therapeutic potential for the treatment of severe malignant cancer pain offering an improvement in quality of life. Non-invasive transmucosal routes of administration for morphine have received much current attention because these routes could bypass the GI degradation and hepatic first pass metabolism of morphine as well as providing ease of administration. We have found that the nasal bioavailability of morphine in solution was greater than that of the rectal suppository in rats. Of the CyD derivatives tested, DM-β-CyD increased the rate of nasal absorption of morphine where its plasma drug level–time profile was similar to that obtained after intravenous administration. Furthermore, DM-β-CyD increased the cerebrospinal fluid (CSF) level of morphine after nasal administration, suggesting that DM-β-CyD may facilitate the transport of morphine not only to the systemic circulation but also to CSF. Interestingly, HP-γ-CyD sustained plasma levels of morphine (Figure 40.17) probably owing to the formation of less membrane permeable complex and/or the increase in the viscosity of the solution, since a viscosity-enhancing polymer, HPC-H sustained plasma levels of nasally administered morphine. These findings suggest that the combination of CyDs and viscous polymers may be useful for optimizing the transmucosal delivery of morphine and potentiating its analgesic action.

FIGURE 40.17 Plasma levels of morphine after nasal administration of morphine hydrochloride (MH; 1 mg/kg) with DM-β-CyD or HP-γ-CyD in rats. ○, MH alone; ●, with DM-β-CyD (50 mM); △, with HP-γ-CyD (200 mM). Each point represents the mean ± SE of 2–8 rats. $^{*}p < 0.05$ versus drug alone.

IV. CyD-BASED SITE-SPECIFIC DRUG DELIVERY

CyD complexes are in equilibrium with guest and host molecules in aqueous solution with the degree of the dissociation being dependent on the magnitude of the stability constant of the complex.[2,28] This property of CyD complexes is a desirable quality, because the complex dissociates to give free CyDs and drug at the absorption site, and only the drug in a free form enters into the systemic circulation.[6] However, the inclusion equilibrium is sometimes disadvantageous when drug targeting is to be attempted, because the complex dissociates before it reaches the organs or tissues to which it is to be delivered. One of the methods to prevent the dissociation is to bind a drug covalently to CyDs. A number of CyD derivatives have been synthesized for many purposes such as the construction of sophisticated enzyme models and reagents or catalysts for regio- and stereo-selective synthesis or chiral recognition.[140,141] CyD derivatives coupled to drugs such as enkephalin, doxorubicin, methicillin, ibuprofen and nucleobase were synthesized for modification of biological

FIGURE 40.18 Structures of BPAA/CyD amide and ester conjugates.

potencies.[142–146] In this section, recent results on site-specific delivery using CyDs are described.

A. Colon targeting

CyDs are known to be barely capable of being hydrolyzed and only slightly absorbed in passage through the stomach and small intestine. However, they are fermented by colonic microflora into small saccharides and thus absorbed as maltose or glucose in the large intestine.[147] Most bacteroid strains isolated from the human colon are capable of degrading CyDs, as shown by their ability to grow on CyDs using them as the sole carbon source and by the stimulation of cyclodextranase activity by exposure to CyDs.[148] This biodegradation property of CyDs is useful as a colon-targeting carrier, and thus CyD prodrugs may serve as a source of site-specific delivery of drugs to colon. Taking these factors into account, we have designed the CyD conjugates of a non-steroidal anti-inflammatory drug biphenylylacetic acid (BPAA),[149–151] and a steroidal drug prednisolone[152–154] and a short-chain fatty acid *n*-butyric acid,[155] anticipating new candidates for colon-specific delivery prodrugs.

The amide- and ester-type conjugates of BPAA with three CyDs (α-, β- and γ-CyDs) are shown in Figure 40.18. The release profiles of BPAA after incubation of the ester conjugates in rat GI tract contents, intestine and liver homogenates, and blood in isotonic buffer solutions were compared with those of ethyl ester biphenylylacetate, a simple ethyl ester of BPAA. Ethyl biphenylylacetate was easily hydrolyzed in liver and GI homogenates and also in blood. However, it was stable enough in the cecal and colonic contents. In sharp contrast, the α- and γ-CyD ester conjugates release BPAA significantly in cecal and colonic contents, while no appreciable drug release from the conjugates was observed on incubation with other contents of fluids. When the ester conjugates were incubated with rat cecal contents, the α- and γ-CyD conjugates produced BPAA quantitatively, while the β-CyD conjugate released the drug in only small amounts, probably because of its low solubility in water, compared with the α- and γ-CyD conjugates. Although the drug release patterns of the three CyD conjugates are different from each other, the *in vitro* data clearly suggest that the ester conjugates are firstly subject to the ring-opening of

FIGURE 40.19 Serum levels of BPAA after oral administration of the BPAA ester conjugates with: ○, α-CyD; △, with β-CyD; □, with γ-CyD; ●, BPAA alone or ◆, BPAA/β-CyD complex (equivalent to 10 mg/kg BPAA) to rats. Each point represents the mean ± SE of three experiments.

CyDs by bacterial enzymes to give the triose and maltose conjugates through longer linear oligosaccharide conjugates. Thus, the ester linkage of the small saccharide/BPAA conjugates could be highly susceptible to hydrolysis. On the other hand, the amide conjugates hardly released BPAA at all, despite fermentation to small saccharide/BPAA conjugates, because the amide linkage is resistant to the second hydrolysis. These results suggest the involvement of two type enzymes in the drug release of the CyD conjugates, i.e. firstly sugar-degrading enzymes and secondly ester-hydrolysing enzymes. This consecutive hydrolysis mechanism was demonstrated in the hydrolysis of the *n*-butyric acid/β-CyD ester conjugate catalyzed by α-amylase (*Aspergillus oryzae*) and carboxylic esterase (porcine liver).[145]Figure 40.19 shows the serum levels of BPAA after oral administration of BPAA/α-, β- and γ-CyD ester conjugates in rats compared with the drug alone and β-CyD complex (a non-covalent inclusion complex prepared by the kneading method in a molar ratio of 1:1). The fast dissolving form of the β-CyD complex shows a rapid increase and decrease in the serum drug levels compared with the drug alone. In the case of the β-CyD conjugate, little increase in serum level was observed probably due to the low aqueous solubility. However, the serum drug levels of the α- and γ-CyD conjugates increased after a lag time of about 3 h and reached maximum levels at about 9 and 8 h, respectively, accompanying a significant increase in the extent of bioavailability. The extents of bioavailability for the α- and γ-CyD conjugates were about four and five times larger than that of BPAA alone, respectively. The anti-inflammatory effect of the BPAA/γ-CyD ester conjugate after oral administration was evaluated using the model of carageenan-induced acute edema in rat paw. The β-CyD complex, a fast dissolving form of BPAA, showed the rapid, but short anti-inflammatory response (about 3 h), because the drug is absorbed mainly from the upper small intestine. On the other hand, the response of the CyD conjugates was

FIGURE 40.20 Structures of prednisolone amide (a) and ester (b) conjugates with CyDs.

highest at 12 h after the administration. These *in vivo* results indicate that the CyD conjugates need a lag time to exhibit the anti-inflammatory activity, because BPAA is produced after the conjugate had reached the cecum and colon.

Glucocorticoids are known to be effective in the treatment of various inflammatory and allergic disorders, and have been used for ulcerative colitis patients. However, oral and intravenous administrations of glucocorticoids to patients with severe ulcerative colitis are restricted, because of the undesirable systemic side-effects such as adrenosuppression, immunosuppression, hypertension, osteoporosis, etc. In order to reduce such side-effects, rectal application has been used for therapy of ulcerative colitis. However, the chronic use of glucocorticoids in high doses often causes systemic side-effects even under topical application. One of the methods to overcome this problem is to develop a prodrug which can be poorly absorbed from intestinal tracts and slowly releases the active drug at the site of action. The colon-specific drug delivery system may be particularly useful for the treatment of ulcerative colitis. Therefore, two kinds of prednisolone-appended CyD conjugates were prepared, i.e. prednisolone is covalently bound to CyD by ester and amide linkages through a spacer of succinic acid as shown in Figure 40.20.[152–154] The amide conjugate was prepared by binding prednisolone 21-hemisuccinate covalently to the amino group of mono(6-deoxy-6-amino)-β-CyD. This amide conjugate rapidly released prednisolone, due to the involvement of an effective intramolecular catalysis of the amide group, giving prednisolone and mono(6-deoxy-6-succimino)-β-CyD (half-life of the release at pH 7.0, 25°C = 6.5 min). Therefore, the amide conjugate can work as a fast-releasing prodrug of prednisolone, but is not suitable for the colon delivery system because the amide conjugate spontaneously releases the drug before it reaches the colon. On the other hand, the ester conjugate was prepared by coupling the carboxyl group of prednisolone 21-hemisuccinate to one of the secondary hydroxyl groups of CyDs. The drug release rate of the ester conjugate was much slower than the amide conjugate, thus working as a colon delivery prodrug. Interestingly, the ester conjugates are very soluble (>50% w/v) in water, the solubility being > 1,000 times than those of prednisolone and its succinate. This is in contrast to the decreased solubility of CyD derivatives conjugated at the primary hydroxyl groups. For example, the aqueous solubility of the BPAA and *n*-butyric acid conjugated at the primary hydroxyl group of β-CyD was about one-tenth of that of the parent drugs. The decreased solubility is attributable to strong intermolecular association between the drug moieties and the neighboring CyD cavity forming a stable columnar packing structure in the solid state. On the other hand, the increased solubility of the prednisolone conjugated at the secondary hydroxyl group of β-CyD is due to the formation of the self-inclusion complex, which inhibits the intermolecular stacking association of the conjugate. The prednisolone-appended α-CyD ester conjugate was intracolonically administered to 2,4,6-trinitrobenzenesulfonic acid (TNBS)-induced colitis rats, and its anti-inflammatory and systemic adverse effects were compared with those of prednisolone alone and prednisolone/HP-β-CyD complex (a non-covalent inclusion complex in a molar ratio of 1:1). The colonic damage score (CDS), the ratio of distal colon wet weight to body weight (C/B) and the myeloperoxidase (MPO) activity were evaluated as measures of the therapeutic effect of prednisolone, whereas the ratio of thymus wet weight to body weight (T/B) was evaluated as a measure of the side-effect of the drug. Healthy rats gave the CDS of 0, the C/B ratio of 0.0018, the MPO activity of 0.00016 units per milligram tissues and the T/B ratio of 0.0016. In the case of the control experiment where the saline solution without drugs was administered after the TNBS treatment, the CDS value, the C/B ratio and the MPO activity were 7.9, 0.0069 and 0.0145 units per milligram tissue, respectively. The HP-β-CyD (3.1% w/v) and α-CyD (2.2% w/v) solutions without drugs gave the therapeutic indices the same as those of the control experiments indicating no anti-inflammatory effects of the CyDs on the TNBS-induced colitis under experimental conditions. On the other hand, the intracolonic administration of prednisolone alone, the prednisolone/HP-β-CyD complex and the prednisolone/α-CyD conjugate (equivalent doses of 5 and 10 mg/kg of prednisolone) significantly decreased the CDS, the C/B ratio and the MPO activity.

The CDS value and the MPO activity were in the order of the complex < the conjugate < prednisolone alone indicating the higher anti-inflammatory activity of the HP-β-CyD complex followed by the conjugate and the drug alone. The thymus atrophy is known to be a typical systemic adverse effect in steroid therapy. The prednisolone/HP-β-CyD complex gave the smallest T/B ratio ($5.2–5.6 \times 10^{-4}$), indicating the significantly higher adverse effect. The decrease of the T/B ratio was also observed in the administration of prednisolone alone. On the other hand, the prednisolone/α-CyD conjugate gave the larger T/B ratio which was the same as that (1.16×10^{-3}) of the control experiment indicating the small systemic adverse effect. These results suggest that the conjugate can alleviate the adverse effect of prednisolone without reducing its therapeutic effect. The *in vitro* and *in vivo* hydrolysis studies indicated that the conjugate maintains the local concentration in the colon at a low but constant level for a long period due to the slow ester hydrolysis. The high anti-inflammatory effect with a low adverse effect was also observed in the oral administration of the ester conjugate to the TNBS-induced colitis in rats.[144b] Therefore, CyDs can serve promoieties for colon-specific targeting prodrugs and the CyD prodrug approach can provide a versatile means for construction of colon-specific delivery systems of certain drugs.

B. Cell targeting

β-Lactam antibiotics show their lethal effect by inhibiting the synthesis of bacterial peptidoglycan, thereby disrupting the cell morphology and eventually resulting in cell lysis and death. Because the antibacterial effect of β-lactam is affected by the permeability of the outer membrane of bacteria, it is important in the evaluation of drug efficiency to measure the diffusion rate of drugs across the membrane. However, the β-lactamase enzyme expressed at the cell surface interferes with the measurement. In order to solve this problem, a β-lactamase inhibitor that is water-soluble and could reach the cell membrane, but not penetrate it, must be prepared. Nedra Kurunarate et al.[144] chose β-CyD as a bulky moiety for the prevention of drug entry across channels in the bacterial outer membrane; thus, they prepared β-CyD conjugates linked to an antibiotic, methicilline, at the molecular terminus through spacer of different chain lengths. The extent of inhibition of the surface β-lactamase of *P. aeruginosa* in the presence of 10 mM of each of the conjugates was as follows: 6% using the conjugate with the spacer length of 12.0 Å, 20% using the 14.0 Å conjugate, 43% using the 16.8 Å conjugate, 77% using the 20.0 Å conjugate, 91% using the 22.6 Å conjugate and 100% using 23.7 Å conjugate. They concluded that the length of the spacer should be greater than 16 Å for optimum inhibition of β-lactamase in the outer membrane. Kim et al.[156] reported that the α-CyD conjugate with an anti-tumor sulfonylurea selectively blocks NADH oxidase activity at the external plasma membrane surface of HeLa cells.

Intercellular recognition events are fundamental to many biological processes in which oligosaccharides on cell surface, glycoproteins or lectins have been responsible for cell–cell recognition and adhesion. Therefore, CyD derivatives bearing small saccharides may be useful as a carrier for transporting active drugs to sugar receptors, such as lectins located on the cell surface. In accordance with this concept, several CyD derivatives with mono- and di-saccharides, such as glucose, galactose, mannose, fucose, etc., have been prepared and investigated for binding characteristics to sugar-specific receptors.[157–159] The β-CyD conjugates with galactose exhibited higher recognition by the galactose-specific *K. bulgaricus* cell wall lectin.[160] It is reported that some galactose and fucose conjugates have a significant cytotoxic effect on the human rectal adenocarcinoma cell line, with P-glycoprotein-positive multidrug resistance.[161] The sugar-substituted CyD derivatives may offer a new way of delivering drugs selectively to specific cell surfaces of organs such as liver and colon, although the uptake of drugs into cells may decrease due to the presence of a bulky, hydrophilic CyD moiety.

Péan et al.[162,163] reported that β- and γ-CyD derivatives coupled with the neuropeptide substance P bind to the neurokinin receptor of rat brain and cultured Chinese hamster ovary (CHO) cells transfected with the human neurokinin receptor gene, *in vitro*. The *in vivo* intracerebral injection of substance P/γ-CyD conjugate in the rat striatum induced a massive internalization of the receptor, and the conjugate recognized the neurokinin receptor-bearing neurons. Schaschke et al.[164] prepared the β-CyD conjugates linked covalently to tetra- and hepta-peptide analogs of gastrin through succinyl amide bond and investigated their binding characteristics of G-protein-coupled CCK-B/gastrin receptor expressed in CHO cells. The receptor affinity of the heptapeptide/β-CyD conjugate was identical to that of the unconjugated tetrapeptide, although it was weaker than that of corresponding unconjugated heptapeptide. However, the functional assay monitored by inositol phosphate production was potent as the corresponding parent peptides. They concluded by using docking experiments of the conjugates into the receptor model that the high hormonal potency of the heptapeptide conjugate is ascribed to a synergetic effect of the specific intermolecular interaction of the peptide with the receptor and the unspecific interaction of the β-CyD moiety with the receptor surface.

Gene therapy requires carriers that can efficiently and safely transfer the gene into the nucleus of the desired cells. Non-viral vectors have many advantages over viral vectors, such as ease of manufacture, safety, low immunogenicity and molecular attachment of targeting ligand. Among these non-viral vectors, starburst polyamidoamine dendrimers have received attention as cationic vector for gene transfection because of their unique molecular shape

FIGURE 40.21 Structures of CDE conjugates.

and the condensed positive charges. To improve the transfection efficiency, we prepared the α-, β- and γ-CyD conjugates with the starburst dendrimer (CDE conjugates, Figure 40.21) and investigated their transfection ability in NIH3T3 and RAW264.7 cells.[165] The CDE conjugates interacted electrostatically with plasmid DNA (pDNA) and protected the degradation of pDNA by DNAase I. The CDE conjugates showed a potent luciferase gene expression, especially in the α-CDE conjugate which provided the greatest transfection activity (approximately 100 times higher than those of the dendrimer alone and of the physical mixture of the dendrimer and α-CyD) among the three conjugates. These findings suggest that the α-CDE conjugate could be a new and preferable non-viral vector of pDNA.

C. Brain targeting

When a drug is covalently coupled to 1-methyl-1,4-dihydronicotinic acid through an enzymatically labile linkage, its lipophilicity increases and allows selective delivery of drugs into the brain across the blood–brain barrier, which is characterized by the endothelial cells of cerebral capillaries that have tight continuous circumferential junctions, thus restricting the passage of polar drugs to the brain.[166] After entry into the brain, the dihydropyridine moiety is oxidized by oxidoreductase to 1-methylpyridinium cation. This polar drug/1-methylpyridinium cation is trapped in the brain and releases the drug by action of other enzymes. This is an essential concept of Bodor's chemical delivery system and is applied to brain targeting of drugs such as steroids, anti-tumor agents and calcium-channel antagonists.[167] The problem, however, is that the prodrugs of the chemical delivery system are poorly water soluble due to the presence of the lipophilic moiety. HP-β-CyD solved this solubility problem by means of soluble complex formation, together with enhancing the chemical stability of dihydronicotinic acid in water. For example, the i.v. administration of estradiol chemical delivery systems (5 mg/kg) solubilized in 20% HP-β-CyD produced a higher concentration of the prodrug in rat brain, which was almost the same as that produced by the administration of the prodrug (15 mg/kg) solution in dimethylsulfoxide.[45] Unfortunately, application of CyD conjugates to brain targeting are few. One of the examples is the β-CyD conjugate with δ-opioid receptor peptide, *N*-leucine-enkephalin and its cyclic analog [*p*-I-Phe^4, D-Pen^2, D-Pen^5]enkephalin, where the carbonyl group of the C terminal leucine was coupled with 6-amino-6-deoxy-β-CyD and in the latter conjugate all hydroxyl groups of CyDs were further methylated to increase the lipophilicity.[168,169] Although the potency of the latter conjugate decreased in receptor binding assay and in the *in vitro* guinea pig intestine and mouse spermatic duct bioassay systems, it showed potent anti-nociceptive properties when given i.c.v. and i.v. in the mouse tail flick test and exhibited no toxicity. Furthermore, it showed prolonged action in the bioassay. Although the detailed transport mechanism of these conjugates to the brain has not been fully elucidated, this methodology can be applied to other neuropharmaceuticals, such as morphine. The β-CyD conjugate with *N*-leucine-enkephalin is of interest, because it has a vacant cavity and can include a neurotropic drug, dothiepine. This inclusion property of conjugates may be useful from the viewpoint of drug formulation, because two different drugs can be incorporated into the CyD molecule.

V. CONCLUSION

The natural CyDs and their synthetic derivatives have successfully been applied to improve undesirable drug properties such as solubility, stability or bioavailability. In addition, the enhancement of drug activity and selective transfer or the reduction of side-effects has been achieved by means of inclusion complex formation. The bioadaptable CyDs can serve as potent absorption enhancers or co-enhancers, and the release rate of water-soluble drugs, including peptides, can be appropriately modified by combination of hydrophobic, hydrophilic and ionizable CyDs. Since CyDs are also useful in extending the functions of pharmaceutical additives, the combination of molecular encapsulation with other carrier materials will be an effective and a valuable tool in the improvement of drug properties. Although the various CyDs described here have many advantages as novel drug carriers for development of advanced dosage forms, the toxicological issues together with their biological fates need to be investigated in detail.

REFERENCES

1. Szejtli, J. *Cyclodextrins and Their Inclusion Complexes*. Akademiai Kiadó: Budapest, 1982.
2. Saenger, W. Cyclodextrin inclusion compounds in research and industry. *Angew. Chem. Int. Ed. Engl.* 1980, *19*, 344–362.
3. Szejtli, J. *Cyclodextrin Technology*. Kluwer Academic Publishers: Dordrecht, The Netherlands, 1988.

4. Uekama, K. Pharmaceutical applications of cyclodextrin complexations. *Yakugaku. Zasshi* 1981, *101*, 857–873.
5. Duchêne, D. *Cyclodextrins and Their Industrial Uses*. Editions de Sante: Paris, 1987.
6. Uekama, K., Otagiri, M. Cyclodextrins in drug carrier systems. *CRC Crit. Rev. Ther. Drug Carr. Syst.* 1987, *3*, 1–40.
7. Duchêne, D. *New Trends in Cyclodextrins and Their Derivatives*. Editions de Sante: Paris, 1991.
8. Uekama, K., Hirayama, F., Irie, T. Cyclodextrin drug carrier systems. *Chem. Rev.* 1998, *98*, 2045–2076.
9. Szejtli, J. Dimethyl-β-cyclodextrin as parent drug carrier. *J. Inclusion Phenom.* 1983, *1*, 135–150.
10. Uekama, K. Pharmaceutical applications of methylated cyclodextrins. *Pharm. Int.* 1985, *6*, 61–65.
11. Pitha, J., Pitha, J. Amorphous water-soluble derivatives of cyclodextrins, non-toxic dissolution enhancing excipients. *J. Pharm. Sci.* 1985, *74*, 987–990.
12. Müller, B. W., Brauns, U. Solubilization of drugs by modified β-cyclodextrins. *Int. J. Pharm.* 1985, *26*, 77–88.
13. Yoshida, A., Arima, H., Uekama, K., Pitha, J. Pharmaceutical evaluation of hydroxyalkyl ethers of β-cyclodextrins. *Int. J. Pharm.* 1988, *46*, 217–222.
14. Yoshida, A., Yamamoto, M., Irie, T. *et al.* Some pharmaceutical properties of 3-hydroxypropyl- and 2,3-dihydroxypropyl-β-cyclodextrins and their solubilizing and stabilizing abilities. *Chem. Pharm. Bull.* 1989, *37*, 1059–1063.
15. Brewster, M. E., Simpkins, J. W., Hora, M. S., Stern, W. C., Bodor, N. The potential use of cyclodextrins in parenteral formulations. *J. Parent. Sci. Techn.* 1989, *43*, 231–240.
16. Koizumi, K., Utamura, T., Sato, M., Yagi, Y. Isolation and characterization of branched cyclodextrins. *Carbohyd. Res.* 1986, *153*, 55–67.
17. Yamamoto, M., Yoshida, A., Hirayama, F., Uekama, K. Some physicochemical properties of branched β-cyclodextrins and their inclusion characteristics. *Int. J. Pharm.* 1989, *49*, 163–171.
18. Hirayama, F. Development and pharmaceutical evaluation of hydrophobic derivatives as modified-release drug carrier. *Yakugaku. Zasshi* 1993, *113*, 425–437.
19. Hirayama, F., Kurihara, M., Horichi, Y. *et al.* Preparation of heptakis(2,6-di-O-ethyl)-[3-cyclodextrin and its nuclear magnetic resonance spectroscopic characterization. *Pharm. Res.* 1993, *10*, 208–213.
20. Uekama, K., Arima, H., Irie, T. *et al.* Sustained release of buserelin acetate, a luteinizing hormone-releasing hormone agonist, from the injectable oily preparation utilizing ethylated β-cyclodextrin. *J. Pharm. Pharmacol.* 1989, *41*, 874–876.
21. Uekama, K., Horiuchi, Y., Irie, T., Hirayama, F. *O*-Carboxymethyl-*O*-ethyl-cyclomaltoheptaose as a delayed-release type drug carrier. Improvement of the oral bioavailability of diltiazem. *Carbohyd. Res.* 1989, *192*, 323–330.
22. Folkman, J., Weisz, P. B., Joullie, M. M. *et al.* Control of angiogenesis with synthetic heparin substitutes. *Science* 1989, *243*, 1490–1493.
23. Uekama, K. Cyclodextrin inclusion compounds: effects on stability and bio-pharmaceutical properties. In *Topics in Pharmaceutical Sciences* (Breimer, D. D., Speiser, P., Eds). Elsevier: Amsterdam, 1987, pp. 181–194.
24. Coleman, A. W., Nicolis, I., Keller, N., Dalbiez, J. P. Aggregation of cyclodextrins: an explanation of the abnormal solubility of β-cyclodextrin. *J. Inclusion Phenom.* 1992, *13*, 139–143.
25. Okada, Y., Kubota, Y., Koizumi, K. *et al.* Some properties and inclusion behaviour of branched cyclodextrins. *Chem. Pharm. Bull.* 1988, *36*, 2176–2185.
26. Arima, H., Kondo, T., Irie, T. Use of water-soluble β-cyclodextrin derivatives as carriers of anti-inflammatory drug biphenylylacetic acid in rectal delivery. *Yakugaku. Zasshi* 1992, *112*, 65–72.
27. Horiuchi, Y., Hirayama, F., Uekama, K. Improvement of stability and bioavailability of 1-hexylcarbamoyl-5-fluorouracil (HCFU) by *O*-carboxymethyl-*O*-ethyl-β-cyclodextrin. *Yakugaku. Zasshi* 1991, *111*, 592–599.
28. Bender, M. L., Komiyama, M. *Cyclodextrin Chemistry*. Springer-Verlag: Berlin, 1978.
29. Hirayama, F., Kurihara, M., Utsuki, T., Uekama, K. Inhibitory effect of guest molecules on acid-catalyzed ring-opening of β-cyclodextrin. *J. Chem. Soc. Chem. Commun.* 1993, 1578–1580.
30. Marshall, J. J., Miwa, I. Kinetic difference between hydrolyses of γ-cyclodextrin by human salivary and pancreatic α-amylases. *Biochim. Biophys. Acta.* 1981, *661*, 142–147.
31. Yamamoto, M., Aritomi, H., Irie, T. *et al.* Biopharmaceutical evaluation of maltosyl-β-cyclodextrin as a parenteral drug carrier. *STP Pharm. Sci.* 1991, *1*, 397–402.
32. Brewster, M. E., Estes, K. S., Bodor, N. An intravenous toxicity of study of 2-hydroxypropyl-β-cyclodextrin, a useful drug solubilizer, in rats and monkeys. *Int. J. Pharm.* 1990, *59*, 231–243.
33. Frijlink, H. W., Visser, J., Hefting, N. R. *et al.* The pharmacokinetics of β-cyclodextrin and hydroxypropyl-β-cyclodextrin in the rat. *Pharm. Res.* 1990, *7*, 1248–1252.
34. Frank, D. W., Gray, J. E., Weaver, R. N. Cyclodextrin nephrosis in the rat. *Am. J. Pathol.* 1976, *83*, 367–382.
35. Irie, T., Uekama, K. Pharmaceutical application of cyclodextrins III. Toxicological issues and safety. *J. Pharm. Sci.* 1997, *86*, 147–162.
36. Ohtani, Y., Irie, T., Uekama, K. *et al.* Differential effects of α-, β-, and γ-cyclodextrins on human erythrocytes. *Eur. J. Biochem.* 1989, *186*, 17–22.
37. Gerloczy, A., Fonagy, A., Keresztes, P. *et al.* Absorption, distribution, excretion and metabolism of orally administered ^{14}C-β-cyclodextrin in rats. *Arznei.-Forschung.* 1985, *35*, 1042–1047.
38. Poelma, F. G. J., Tukker, J. J., Hilbers, H. W., Jansen, A. C. A. Intestinal absorption of drugs II. The effects of inclusion in cyclodextrins on the absorption of dantrolene. *J. Inclusion Phenom.* 1989, *7*, 423–430.
39. Arima, H., Kondo, T., Irie, T., Uekama, K. Enhanced rectal absorption and reduced local irritation of anti-inflammatory drug ethyl 4-biphenylyl acetate in rats by complexation with water-soluble β-cyclodextrin derivatives and formulations as oleaginous suppository. *J. Pharm. Sci.* 1992, *81*, 1119–1125.
40. Uekama, K., Imai, T., Maeda, T. *et al.* Improvement of dissolution and suppository release characteristics of flurbiprofen by inclusion complexation with heptakis(2,6-di-O-methyl)-β-cyclodextrin. *J. Pharm. Sci.* 1985, *74*, 841–845.
41. Arima, H., Adachi, H., Irie, T. *et al.* Enhancement of the anti-inflammatory effect of ethyl 4-biphenylylacetate in ointment by β-cyclodextrin derivatives: increased absorption and localized activation of the prodrug in rats. *Pharm. Res.* 1990, *7*, 1152–1156.
42. Zhang, M.-Q., Rees, D. C. A review of recent application of cyclodextrins for drug discovery. *Expert. Opin. Ther. Pat.* 1999, *9*, 1697–1717.
43. Dietzel, K., Estes, K. S., Brewster, M. E. *et al.* The use of 2-hydroxypropyl-β-cyclodextrin as a vehicle for intravenous administration of dexamethasone in dogs. *Int. J. Pharm.* 1990, *59*, 225–230.
44. Hoshino, T., Uekama, K., Pitha, J. Increase in temperature enhances solubility of drugs in aqueous solutions of hydroxypropylcyclodextrins. *Int. J. Pharm.* 1993, *98*, 239–242.
45. Brewster, M. E., Estes, K. S., Loftsson, T. *et al.* Improved delivery through biological membranes XXXI: Solubilization and stabilization of an estradiol chemical delivery system by modified β-cyclodextrins. *J. Pharm. Sci.* 1988, *77*, 981–985.
46. Millard, W. J., Romano, T. M., Bodor, N., Simpkins, J. W. Growth hormone (GH) secretory dynamics in animals administered estradiol utilizing a chemical delivery system. *Pharm. Res.* 1990, *7*, 1011–1018.
47. Merkus, F. W. H. M., Verhoef, J. C., Romeijn, S. G., Schipper, N. G. M. Absorption enhancing effect of cyclodextrins on intranasally administered insulin in rats. *Pharm. Res.* 1991, *8*, 588–592.

48. Kurihara, M., Hirayama, F., Uekama, K., Yamasaki, M. Improvement of some pharmaceutical properties of nocloprost by β and γ-cyclodextrin complexations. *J. Inclusion Phenom.* 1990, *8*, 363–373.
49. Utsuki, T., Imamura, K., Hirayama, F., Uekama, K. Stoichiometry-dependent changes of solubility and photo reactivity of an antiulcer agent, 2′-carboxymethoxy-4,4′-bis(3-methyl-2-butenyloxy) chalcone, in cyclodextrin inclusion complexes. *Eur. J. Pharm. Sci.* 1993, *1*, 81–87.
50. Hirayama, F., Kurihara, M., Uekama, K. Improving the aqueous stability of prostaglandin E_2 and prostaglandin A_2 by inclusion complexation with methylated β-cyclodextrins. *Chem. Pharm. Bull.* 1984, *32*, 4237–4240.
51. Yamamoto, M., Hirayama, F., Uekama, K. Improvement of stability and dissolution of prostaglandin E_1 by maltosyl-β-cyclodextrin in lyophilized formulation. *Chem. Pharm. Bull.* 1992, *40*, 747–751.
52. Hirayama, F., Kurihara, M., Uekama, K. Improvement of chemical instability of prostacyclin in aqueous solution by complexation with methylated cyclodextrins. *Int. J. Pharm.* 1987, *35*, 193–199.
53. Uekama, K., Fujinaga, T., Hirayama, F. *et al.* Improvement of the oral bioavailability of digitalis glycosides by cyclodextrin complexation. *J. Pharm. Sci.* 1983, *72*, 1338–1341.
54. Yoshida, A., Yamamoto, M., Hirayama, F., Uekama, K. Improvement of chemical instability of digitoxin in aqueous solution by complexation with β-cyclodextrin derivatives. *Chem. Pharm. Bull.* 1988, *36*, 4075–4080.
55. Uekama, K., Fujinaga, T., Hirayama, F. *et al.* Effects of cyclodextrins on the acid hydrolysis of digoxin. *J. Pharm. Pharmacol.* 1982, *34*, 627–630.
56. Seo, H., Uekama, K. Enhanced bioavailability of digoxin by γ-cyclodextrin complexation: evaluation for sublingual and oral administrations in human. *Yakugaku. Zasshi* 1989, *109*, 778–782.
57. Brittain, H. G. *Polymorphism in Pharmaceutical Solids*. Marcel Dekker: New York, 1999.
58. Bernstein, J. *Polymorphism in Molecular Crystals*. Oxford University Press: Oxford, 2002.
59. Davey, R. J., Cardew, P. T., McEwan, D. *et al.* Rate controlling processes in solvent-mediated phase transformation. *J. Cryst. Growth* 1986, *79*, 648–653.
60. Ostwald, W. Studien uber die Bildung und Umwandlung Fester Korper. *Z. Physik. Chem.* 1987, *22*, 289–302.
61. Sonoda, Y., Hirayama, F., Arima, H. *et al.* Cyclodextrin-based isolation of Ostwald's metastable polymorphs occurring during crystallization. *Chem. Commun.* 2007, 517–519.
62. Sonoda, Y., Hirayama, F., Arima, H. *et al.* Selective crystallization of the metastable form IV polymorph of tolbutamide in the presence of 2,6-di-O-methyl-β-cyclodextrin in aqueous solution. *Cryst. Growth Des.* 2007, *6*, 1181–1185.
63. Sugimoto, I., Kuchiki, K., Nakagawa, H. Stability of nifedipine-polyvinylpyrrolidone coprecipitate. *Chem. Pharm. Bull.* 1981, *29*, 1715–1723.
64. Uekama, K., Ikegami, K., Wang, Z. *et al.* Inhibitory effect of 2-hydroxypropyl-β-cyclodextrin on crystal-growth of nifedipine during storage: superior dissolution and oral bioavailability compared with polyvinylpyrrtolidone K-30. *J. Pharm. Pharmacol.* 1992, *44*, 73–78.
65. Kikuchi, M., Hirayama, F., Uekama, K. Improvement of chemical instability of carmofur in β-cyclodextrin solid complex by utilizing some organic acids. *Chem. Pharm. Bull.* 1987, *35*, 315–319.
66. Kikuchi, M., Uemura, Y., Hirayama, F. *et al.* Improvement of some pharmaceutical properties of carmofur by cyclodextrin complexation. *J. Inclusion Phenom.* 1984, *2*, 623–630.
67. Kikuchi, M., Hirayama, F., Uekama, K. Improvement of oral and rectal bioavailability of carmofur by methylated β-cyclodextrin complexations. *Int. J. Pharm.* 1987, *38*, 191–198.
68. Hirayama, F., Hirashima, N., Abe, K. *et al.* Utilization of diethyl-β-cyclodextrin as a sustained-release carrier for isosorbide dinitrate. *J. Pharm. Sci.* 1988, *77*, 233–236.
69. Uekama, K., Hirashima, N., Horiuchi, Y. *et al.* Ethylated β-cyclodextrins as hydrophobic drug carriers: sustained release of diltiazem in the rat. *J. Pharm. Sci.* 1987, *76*, 660–661.
70. Uekama, K., Horikawa, T., Yamanaka, M., Hirayama, F. Peracylated β-cyclodextrin as novel sustained-release carrier for water-soluble drug, molsiodomine. *J. Pharm. Pharmacol.* 1994, *46*, 714–717.
71. Uekama, K., Horikawa, T., Horiuchi, Y., Hirayama, F. In vitro and in vivo evaluation of delayed-release behaviour of diltiazem from its *O*-carboxymethyl-*O*-ethyl-β-cyclodextrin complex. *J. Control Release* 1993, *25*, 99–106.
72. Horiuchi, Y., Abe, K., Hirayama, F., Uekama, K. Control of theophylline by β-cyclodextrin derivatives: hybridizing of hydrophilic and ionizable β-cyclodextrin complexes. *J. Control Release* 1991, *15*, 177–182.
73. Matsubara, K., Kuriki, T., Arima, H. *et al.* Possible use of diethyl-β-cyclodextrin in preparation of sustained release oily injection of busenelin acetate (LHRH agonist). *Drug Deliver. Syst.* 1990, *5*, 95–99.
74. Matsubara, K., Irie, T., Uekama, K. Controlled-release of the LHRH agonist buserelin acetate from injectable suspensions containing triacetylated cyclodextrins in the oil vehicle. *J. Control Release* 1994, *31*, 173–180.
75. Uekama, K., Hirayama, F., Irie, T. Application of cyclodexterin. In *Absorption Enhancement: Concept, Possibilities and Limitations* (Boer, A. G., Ed.), Vol. 3. Harwood Publishers: Amsterdam, 1994, pp. 411–456.
76. Tokumura, T., Nanba, M., Tsushima, Y. *et al.* Enhancement of bioavailability of cinnarizine from its β-cyclodextrin complex on oral administration with DL-phenylalanine as a competing agent. *J. Pharm. Sci.* 1986, *75*, 391–394.
77. Seo, H., Tsuruoka, M., Hashimoto, T. *et al.* Enhancement of oral bioavailability of spironolactone by β- and γ-cyclodextrin complexations. *Chem. Pharm. Bull.* 1983, *31*, 286–291.
78. Uekama, K., Fujinaga, T., Otagiri, M. *et al.* Improvement of dissolution and chemical stability of proscillaridin by cyclodextrin complexations. *Acta Pharm. Suec.* 1983, *20*, 287–294.
79. Uekama, K., Otagiri, M., Uemura, Y. *et al.* Improvement of oral bioavailability of prednisolone by β-cyclodextrin complexation in humans. *J. Pharmacobio-Dynam.* 1983, *6*, 124–127.
80. Uekama, K., Narisawa, S., Hirayama, F., Otagiri, M. Improvement of dissolution and absorption characteristics of benzodiazepines by cyclodextrin complexation. *Int. J. Pharm.* 1983, *16*, 327–338.
81. Uekama, K., Matsuo, N., Hirayama, F. *et al.* Enhanced bioavailability of acetohexamide by β-cyclodextrin complexation. *Yakugaku. Zasshi* 1980, *100*, 903–909.
82. Miyaji, T., Inoue, Y., Acarrurk, F. *et al.* Improvement of oral bioavailability of fenbufen by cyclodextrin complexations. *Acta Pharm. Nord.* 1992, *4*, 17–22.
83. Irie, T., Tsunenari, Y., Uekama, K., Pitha, J. Effect of bile on the intestinal absorption of α-cyclodextrin in rats. *Int. J. Pharm.* 1988, *43*, 41–44.
84. Nakanishi, K., Nadai, T., Masada, M., Masada, M. Effect of cyclodextrins on biological membrane II. Mechanism of enhancement of the intestinal absorption of non-absorbable drug by cyclodextrins. *Chem. Pharm. Bull.* 1992, *40*, 1252–1256.
85. Horiuchi, Y., Kikuchi, M., Hirayama, F. *et al.* Improvement of bioavailability of menaquinone-4 by dimethyl-β-cyclodextrin complexation following oral administration. *Yakugaku. Zasshi* 1988, *108*, 1093–1100.
86. Ueno, M., Ijitsu, T., Horiuchi, Y. *et al.* Improvement of dissolution and absorption characteristics of ubidecarenone by dimethyl-β-cyclodextrin complexation. *Acta Pharm. Nord.* 1989, *2*, 99–104.
87. Otagiri, M., Imai, T., Uekama, K. Enhanced oral bioavailability of anti-inflammatory drug flurbiprofen in rabbits by tri-O-methyl-β-cyclodextrin complexation. *J. Pharmacobio-Dynam.* 1982, *5*, 1027–1029.

88. Nakai, Y., Yamamoto, K., Terada, K. *et al.* Interaction of tri-O-methyl-β-cyclodextrin with drugs, enhanced bioavailability of ketoprofen in rats when administered with tri-O-methyl-β-cyclodextrin. *Chem. Pharm. Bull.* 1983, *31*, 3745–3747.
89. Kikuchi, M., Uekama, K. Enhancement of antitumor activity of carmofur (HCFU) by dimethyl-β-cyclodextrin complexation in P-388 leukemia-bearing mice. *Xenobiotic Metab. Dispos.* 1988, *3*, 267–273.
90. Uekama, K., Horiuchi, Y., Kikuchi, M. *et al.* Enhanced dissolution and oral bioavailability of α-tocopheryl esters by dimethyl-β-cyclodextrin complexations. *J. Inclusion Phenom.* 1988, *6*, 167–174.
91. Duchêne, D., Wouessidjewe, D., Poelman, M.-C. Dermal uses of cyclodextrin and derivatives. In *New Trends in Cyclodextrins and Their Derivatives* (Duchene, D., Ed.). Editions de Sante: Paris, 1991, pp. 447–481.
92. Otagiri, M., Fujinaga, T., Sakai, A., Uekama, K. Effects of β- and γ-cyclodextrins on release of betamethasone from ointment bases. *Chem. Pharm. Bull.* 1984, *32*, 2401–2405.
93. Orienti, I., Zecchi, V., Bettasi, V., Fini, A. Release of ketoprofen from dermal bases in presence of cyclodextrins: effect of the affinity constant determined in semisolid vehicles. *Arch. Pharm. (Weinheim)* 1991, *324*, 943–947.
94. Uekama, K., Otagiri, M., Sakai, A. *et al.* Improvement in the percutaneous absorption of beclomethasone dipropionate by γ-cyclodextrin complexation. *J. Pharm. Pharmacol.* 1985, *37*, 532–535.
95. Arima, H., Adachi, H., Irie, T., Uekama, K. Improved drug delivery through the skin by hydrophilic β-cyclodextrins: enhancement of anti-inflammatory effect of 4-biphenylylacetic acid in rats. *Drug Invest.* 1990, *2*, 155–161.
96. Uekama, K., Maeda, T., Arima, H. *et al.* Possible utility of β-cyclodextrin complexation in the preparation of biphenylylacetic acid suppository. *Yakugaku. Zasshi* 1986, *106*, 1126–1130.
97. Arima, H., Irie, T., Uekama, K. Differences in the enhancing effects of water-soluble β-cyclodextrins on the release of ethyl 4-biphenylyl acetate, an anti-inflammatory agent from an oleaginous suppository base. *Int. J. Pharm.* 1989, *57*, 107–115.
98. Hermens, W. A. J. J., Deurloo, M. J. M., Romeyn, S. G. *et al.* Nasal absorption enhancement of 17β-estradiol by dimethyl-β-cyclodextrin in rabbits and rats. *Pharm. Res.* 1990, *7*, 500–503.
99. Schipper, N. G. M., Hermens, W. A. J. J., Romeyn, S. G. *et al.* Nasal absorption of 17β-estradiol and progesterone from a dimethylcyclodextrin inclusion formulation in rats. *Int. J. Pharm.* 1990, *64*, 61–66.
100. Stuenkel, C. A., Dudley, R. E., Yen, S. S. C. Sublingual administration of testosterone-hydroxypropyl-β-cyclodextrin inclusion complex simulates episodic androgen release in hypogonadal men. *J. Clin. Endocrinol. Metab.* 1991, *72*, 1054–1059.
101. Schipper, N. G. M., Verhoef, J. C., Merkus, F. W. H. M. The nasal mucociliary clearance: relevance to nasal drug delivery. *Pharm. Res.* 1991, *8*, 807–814.
102. Hermens, W. A. J. J., Belder, C. W. J., Merkus, J. M. W. M. *et al.* Intranasal estradiol administration to oophorectomized women. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 1991, *40*, 35–41.
103. Yoshida, A., Yamamoto, M., Itoh, T. *et al.* Utility of 2-hydroxypropyl-β-cyclodextrin in an intramuscular injectable preparation of nimodipine. *Chem. Pharm. Bull.* 1990, *38*, 176–179.
104. Irie, T., Otagiri, M., Uekama, K. *et al.* Alleviation of the chlorpromazine-induced muscular tissue damage by β-cyclodextrin complexation. *J. Inclusion Phenom.* 1984, *2*, 637–644.
105. Imai, T., Maeda, T., Otagiri, M., Uekama, K. Improvement of absorption characteristics and reduction of irritation on stomach of flurbiprofen by complexation with various cyclodextrins. *Xenobiotic Metab. Dispos.* 1987, *2*, 657–664.
106. Ishida, K., Hoshino, T., Irie, T., Uekama, K. Alleviation of chlorpromazine-photosensitized contact dermatitis by β-cyclodextrin derivatives and their possible mechanisms. *Xenobiotic Metab. Dispos.* 1988, *3*, 377–386.
107. Hoshino, T., Ishida, K., Irie, T. *et al.* An attempt to reduce the photosensitizing potential of chlorpromazine with the simultaneous use of β- and dimethyl-β-cyclodextrins in guinea pigs. *Arch. Dermatol. Res.* 1989, *281*, 60–65.
108. Uekama, K., Irie, T., Hirayama, F. Participation of cyclodextrin inclusion catalysis in photolysis of chlorpromazine to give promazine in aqueous solution. *Chem. Lett.* 1978, 1109–1112.
109. Hoshino, T., Ishida, K., Irie, T. *et al.* Reduction of photohemolytic activity of benoxaprofen by β-cyclodextrin complexation. *J. Inclusion Phenom.* 1988, *6*, 415–423.
110. Hoshino, T., Ishida, K., Irie, T. *et al.* Alleviation in protriptyline-photosensitized skin irritation by di-O-methyl-β-cyclodextrin complexation. *Int. J. Pharm.* 1987, *38*, 265–267.
111. Perrin, J. H., Field, F. P., Hansen, D. A. *et al.* β-Cyclodextrin as an aid to peritoneal dialysis. Renal toxicity of β-cyclodextrin in the rat. *Res. Commun. Chem. Pathol. Pharmacol.* 1978, *19*, 373–376.
112. Carpenter, T. O., Pettifor, J. M., Russell, R. M. *et al.* Severe hypervitaminosis A in siblings: evidence of variable tolerance to retinol intake. *J. Pediatr.* 1987, *111*, 507–512.
113. Irie, T., Fukunaga, K., Garwood, M. K. *et al.* Hydroxypropylcyclodextrins in parenteral use. II. Effects on transport and disposition of lipids in rabbits and humans. *J. Pharm. Sci.* 1992, *81*, 524–528.
114. Frijlink, H. W., Eissen, A. C., Hefting, N. R. *et al.* The effect of parenterally administered cyclodextrins on cholesterol levels in the rat. *Pharm. Res.* 1991, *8*, 9–16.
115. Irie, T., Fukunaga, K., Pitha, J. Hydroxypropylcyclodextrins in parenteral use. I. Lipid dissolution and effects on lipid transfers in vitro. *J. Pharm. Sci.* 1992, *81*, 521–523.
116. Uekama, K., Shiotani, K., Irie, T. *et al.* Protective effects of cyclodextrin sulphates against gentamicin-induced nephrotoxicity in the rat. *J. Pharm. Pharmacol.* 1993, *45*, 745–747.
117. Shimamoto, T. Pharmaceutical aspects: nasal and depot formulations of leuprolide. *J. Androl.* 1987, *8*, S14–S16.
118. Irie, T., Wakamatsu, K., Arima, H. *et al.* Enhancing effects of cyclodextrins on nasal absorption of insulin. *Int. J. Pharm.* 1992, *84*, 129–139.
119. Schipper, N. G. M., Verhoef, J., Romeijn, S. G., Merkus, F. W. H. M. Absorption enhancers in nasal insulin delivery and their influence on nasal ciliary functioning. *J. Control Release* 1992, *21*, 173–186.
120. Irie, T., Arima, H., Abe, K. *et al.* Possible mechanisms of cyclodextrin-enhanced nasal absorption of peptide and protein drugs. In *Minutes of the 6th International Symposium on Cyclodextrins* (Hedges, A. R., Ed.). Editions de Sante: Paris, 1992, pp. 503–508.
121. Maitani, Y., Igawa, T., Machida, Y., Nagai, T. Intranasal administration of β-interferon in rabbits. *Drug Des. Deliver.* 1986, *1*, 65–70.
122. Irie, T., Abe, K., Adachi, H. *et al.* Potential use of 2-hydroxypropyl-β-cyclodextrin in designing nasal preparations of insulin involving lipophilic absorption enhancer HPE-101. *Drug Deliver. Syst.* 1992, *7*, 91–95.
123. Brewster, M. E., Hora, M. S., Simpkins, J. W., Bodor, N. Use of 2-hydroxypropyl-β-cyclodextrin as a solubilizing and stabilizing excipient for protein drugs. *Pharm. Res.* 1991, *8*, 792–795.
124. Ressing, M. E., Jiskoot, W., Talsma, H. *et al.* The influence of sucrose, dextran, and hydroxypropyl-β-cyclodextrin as lyoprotectants for a freeze-dried mouse IgG_{2a} monoclonal antibody (MN12). *Pharm. Res.* 1992, *9*, 266–270.
125. Uekama, K., Yamamoto, M., Irie, T., Hirayama, F. Pharmaceutical evaluation of maltosyl-β-cyclodextrin as a drug carrier in parenteral formulation. In *Minutes of the 6th International Symposium on Cyclodextrins* (Hedges, A. R., Ed.). Editions de Sante: Paris, 1992, pp. 491–496.
126. Fukunaga, K., Hijikata, S., Ishimura, K. *et al.* Aluminium β-cyclodextrin sulphate as a stabilizer and sustained-release carrier

for basic fibroblast growth factor. *J. Pharm. Pharmacol.* 1994, *46*, 168–171.
127. Harris, J. M., Chess, R. B. Effects of pegylation on pharmaceuticals. *Nat. Rev. Drug Discov.* 2003, *2*, 214–221.
128. Veronese, F. M., Harris, J. M. Preface: introduction and overview of peptide and protein peglylation. *Adv. Drug Deliver. Rev.* 2002, *54*, 453–456.
129. Harada, A. Cyclodextrin-based molecular machines. *Acc. Chem. Res.* 2001, *34*, 456–464.
130. Higashi, T., Hirayama, F., Arima, H., Uekama, K. Polypseudorotaxanes of pegylated insulin with cyclodextrins: application to sustained release system. *Bioorg. Med. Chem. Lett.* 2007, *17*, 1871–1874.
131. Wang, Z., Ikegami, K., Hirayama, F., Uekama, K. Release characteristics of nifedipine from 2-hydroxypropyl-β-cyclodextrin complex during storage and its modification of hybridizing polyvinylpyrrolidone K-30. *Chem. Pharm. Bull.* 1993, *41*, 1822–1826.
132. Wang, Z., Hirayama, F., Uekama, K. Design and in vitro evaluation of a modified-release oral dosage form of nifedipine by hybridization of hydroxypropyl-β-cyclodextrin and hydroxypropyl-cellulose. *J. Pharm. Pharmacol.* 1993, *45*, 942–946.
133. Wang, Z., Hirayama, F., Uekama, K. In vivo and in vitro evaluation of modified-release oral dosage form of nifedipine by hybridization of hydroxypropyl-β-cyclodextrin and hydroxypropyl-celluloses in dog. *J. Pharm. Pharmacol.* 1994, *46*, 505–507.
134. Adachi, H., Irie, T., Hirayama, F., Uekama, K. Stabilization of prostaglandin E_1 in fatty alcohol propylene glycol ointment by acidic cyclodextrin derivative, *O*-carboxymethyl-*O*-ethyl-β-cyclodextrin. *Chem. Pharm. Bull.* 1992, *40*, 1586–1591.
135. Uekama, K., Adachi, H., Irie, T. *et al.* Improved transdermal delivery of prostaglandin E_1 through hairless mouse skin: combined use of carboxymethyl-ethyl-β-cyclodextrin and penetration enhancers. *J. Pharm. Pharmacol.* 1992, *44*, 119–121.
136. Adachi, H., Irie, T., Uekama, K. *et al.* Combination effects of *O*-carboxymethyl-*O*-ethyl-β-cyclodextrin and penetration enhancer HPE-101 on transdermal delivery of prostaglandin E_1 in hairless mice. *Eur. J. Pharm. Sci.* 1993, *1*, 117–123.
137. Adachi, H., Irie, T., Uekama, K. *et al.* Inhibitory effect of prostaglandin E_1 on laurate-induced peripheral vascular occlusive sequelae in rabbits: optimized topical formulation with β-cyclodextrin derivative and penetration enhancer HPE-101. *J. Pharm. Pharmacol.* 1992, *44*, 1033–1035.
138. Tobino, Y., Torii, H., Ikeda, K. *et al.* Release control of morphine hydrochloride in suppository. *Kyushu Yakugaku-Kai Kaiho* 1991, *45*, 15–20.
139. Uekama, K., Kondo, T., Nakamura, K. *et al.* Modification of rectal absorption of morphine from hollow-type suppositories with a combination of α-cyclodextrins and viscosity-enhancing polysaccharide. *J. Pharm. Sci.* 1995, *84*, 15–20.
140. König, W. A., Lutz, S., Wenz, G. Modified cyclodextrins-novel, highly enantioselective stationary phases for gas chromatography. *Angew. Chem. Int. Ed. Engl.* 1988, *27*, 979–980.
141. Breslow, R. Biomimetic chemistry artificial enzymes: catalysis by design. *Acc. Chem. Res.* 1995, *28*, 146–153.
142. Parrot-Lopez, H., Djedaini, F., Perly, B. *et al.* An approach to vectorization of pharmacologically active molecules: the covalent binding of Leuenkephalin to a modified β-cyclodextrin. *Tetrahedron Lett.* 1990, *31*, 1999–2002.
143. Tanaka, H., Kominato, K., Yamamoto, R. *et al.* Synthesis of doxorubicin-cyclodextrin conjugates. *J. Antibiot.* 1994, *47*, 1025–1029.
144. Karunaratne, D. N., Farmer, S., Hancock, R. E. W. Synthesis of bulky β-lactams for inhibition of cell surface β-lactamase activity. *Bioconjugate Chem.* 1993, *4*, 434–439.
145. Coates, J. H., Easton, C. J., van Eyk, S. J. *et al.* Chiral differentiation in the deacylation of 6^A-O-{2-[4-(2-methylpropyl)-phenyl] propanoyl}-β-cyclodextrin. *J. Chem. Soc. Chem. Commun.* 1991, 759–760.
146. Djedaini-Pilard, F., Perly, B., Dupas, S. *et al.* Specific interaction and stabilization between host and guest: complexation of ellipticine in a nucleobase functionalized cyclodextrin. *Tetrahedron Lett.* 1993, *34*, 1145–1148.
147. Flourie, B., Molis, C., Achour, L. *et al.* Fate of β-cyclodextrin in the human intestine. *J. Nutr.* 1993, *123*, 676–680.
148. Antenucci, R., Palmer, J. K. Enzymatic degradation of α-and β-cyclodextrins by bacteroides of the human colon. *J. Agric. Food Chem.* 1984, *32*, 1316–1321.
149. Hirayama, F., Minami, K. In vitro evaluation of biphenylylacetic acid–β-cyclodextrin conjugate as colon-targeting prodrug: drug release behavior in rat biological fluids. *J. Pharm. Pharmacol.* 1996, *48*, 27–31.
150. Uekama, K., Minami, K., Hirayama, F. 6^A-O-[(4-Biphenylyl) acetyl]-α-β-, and -γ-cyclodextrins and 6^A-deoxy-6^A [[(4-biphenylyl) acetyl]amino]-α-, -β-, and -γ-cyclodextrins: potential prodrugs for colon-specific delivery. *J. Med. Chem.* 1997, *40*, 2755–2761.
151. Minami, K., Hirayama, F., Uekama, K. Colon-specific drug delivery based on a cyclodextrin prodrug: release behavior of biphenylylacetic acid from its cyclodextrin conjugates in rat intestinal tracts after oral administration. *J. Pharm. Sci.* 1998, *87*, 715–720.
152. Yano, H., Hirayama, F., Arima, H., Uekama, K. Hydrolysis behavior of prednisolone 21-hemisuccinate/β-cyclodextrin amide conjugate: involvement of intramolecular catalysis of amide group in drug release. *Chem. Pharm. Bull.* 2000, *48*, 1125–1128.
153. Yano, H., Hirayama, F., Arima, H., Uekama, K. Preparation of prednisolone-appended α-, β- and γ-cyclodextrins: substitution at secondary hydroxyl groups and in vitro hydrolysis behavior. *J. Pharm. Sci.* 2001, *90*, 493–503.
154. (a) Yano, H., Hirayama, F., Arima, H., Uekama, K., Prednisolone-appended α-cyclodextrin: alleviation of systemic adverse effect of prednisolone after intracolonic administration in 2,4,6-trinitrobenzenesulfonic acid-induced colitis rats. *J. Pharm. Sci. 90*, 2103–2112; (b) Yano, H., Hirayama, F., Kamada, M. *et al.* (2002). Colon-specific delivery of prednisolone-appended α-cyclodextrin conjugate: alleviation of systemic side effect after oral administration. *J. Conrol Release 79*, 103–112.
155. Hirayama, F., Ogata, T., Yano, H. *et al.* Release characteristics of a short-chain fatty acid, *n*-butyric acid, from its β-cyclodextrin ester conjugate in rat biological media. *J. Pharm. Sci.* 2000, *89*, 1486–1495.
156. Kim, C., MacKellar, W. C., Cho, N. *et al.* Impermeant antitumor sulfonylurea conjugates that inhibit plasma membrane NADH oxidase and growth of HeLa cells in culture: identification of binding proteins from sera of cancer patients. *Biochim. Biophys. Acta* 1997, *1324*, 171–181.
157. Parrot-Lopez, H., Leray, E., Coleman, A. W. New β-cyclodextrin derivatives possessing biologically active saccharide antennae. *Supramol. Chem.* 1993, *3*, 37–42.
158. Robertis, L., Lancelon-Pin, C., Driguez, H. *et al.* Synthesis of new oligosaccharidyl-thio-β-cyclodextrins (CDS): a novel family of potent drug-targeting vectors. *Bioorg. Med. Chem. Lett.* 1994, *4*, 1127–1130.
159. Leray, E., Parrot-Lopez, H., Auge, C. *et al.* Chemical-enzymatic synthesis and bioactivity of mono-6-[gal-β-1,4-glcNAc-β-(1,6^0)-hexyl]amido-6-deoxy-cycloheptaamylose. *J. Chem. Soc. Chem. Commun.* 1995, 1019–1020.
160. Imata, H., Kubota, K., Hattori, K. *et al.* The specificity of association between concanavalin A and oligosaccharide-branched cyclodextrins with an optical biosensor. *Bioorg. Med. Chem. Lett.* 1997, *7*, 109–112.
161. Attioui, F., Al-Omar, A., Leray, E. *et al.* Recognition ability and cytotoxicity of some oligosaccharidyl substituted β-cyclodextrins. *Biol. Cell* 1994, *82*, 161–167.
162. Péan, C., Wijkhuisen, A., Djedaini-Pilard, F. *et al.* Pharmacological investigations of new peptide-cyclodextrins. In *Proceedings of the 9th*

International Symposium on Cyclodextrins (Torres Labandeira, J. J., Vila-Jato, J. L., Eds). Kluwer Academic Press: Dordrecht, 1998, pp. 387–390.

163. Péan, C., Fischer, J., Doly, A. *et al.* In vitro and in vivo investigations of the specific binding of substance P-γ-cyclodextrin adducts (SB-γ-CD) to rat brain NK1 receptors. In *Proceedings of the 9th International Symposium on Cyclodextrins* (Torres Labandeira, J. J., Vila-Jato, J. L., Eds). Kluwer Academic Press: Dordrecht, 1998, pp. 403–406.
164. Schaschke, N., Fiori, S., Weyher, E. *et al.* Cyclodextrin as carrier of peptide hormones: conformational and biological properties of β-cyclodextrin/gastrin constructs. *J. Am. Chem. Soc.* 1998, *120*, 7030–7038.
165. Arima, H., Kihara, F., Hirayama, F., Uekama, K. Enhancement of gene expression by polyamidoamine dendrimer conjugates with α-, β-, and γ-cyclodextrin. Bioconjugate. Chem. 12: 476–484; Kihara, F., Arima, H., Tsutsumi, T. *et al.*. (2002) Effects of structure of polyamidoamine dendrimer of gene transfer efficiency of the dendrimer conjugate with α-cyclodextrin. *ibid*. 13(2001) 1211–1219.
166. Begley, D. J. The blood–brain barrier: principles for targeting peptides and drugs to the central nervous system. *J. Pharm. Pharmacol.* 1996, *48*, 136–146.
167. Bodor, N. Prodrugs versus soft drugs. In *Design of Prodrugs* (Bundgaard, H., Ed.). Elsevier: Amsterdam, 1985, pp. 333–354.
168. Djedaini-Pilard, F., Désalos, J., Perly, B. Synthesis of a new carrier: N (Leuenkephalin)yl-6-amino-6-deoxycyclomaltoheptaose. *Tetrahedron Lett.* 1993, *34*, 2457–2460.
169. Hristova-Kazmierski, M. K., Horan, P., Davis, P. *et al.* A new approach to enhance bioavailability of biologically active peptides: conjugation of a δ-opioid agonist to β-cyclodextrin. *Bioorg. Med. Chem. Lett.* 1993, *3*, 831–834.

Chapter 41

Chemical and Physicochemical Approaches to Solve Formulation Problems

Camille G. Wermuth

I. INTRODUCTION
II. INCREASING CHEMICAL STABILITY
III. IMPROVED FORMULATION OF PEPTIDES AND PROTEINS
IV. DEALING WITH MESOMORPHIC CRYSTALLINE FORMS
V. INCREASING THE MELTING POINT
A. Salt or complex formation
B. Covalent derivatives
C. Introduction of symmetry
VI. GASTROINTESTINAL IRRITATION AND PAINFUL INJECTIONS
A. Gastrointestinal irritability
B. Avoidance of painful injections
VII. SUPPRESSING UNDESIRABLE ORGANOLEPTIC PROPERTIES
A. Odor
B. Taste
REFERENCES

"You may readily infer that such substances as agreeably titillate the sense (of smell) are composed of smooth round atoms. Those that seem bitter and harsh are more tightly compacted of hooked particles and accordingly tear their way into our sense and rend our bodies by their inroads."

Lucretius, 47 BC[1]

I. INTRODUCTION

Medicinal chemists aim to discover new lead compounds and to optimize them to highly potent, selective, and non-toxic new drug candidates. However, the selected drug candidates can still present some drawbacks of physicochemical nature (insufficient stability, inappropriate aggregation state) or which might affect the patient's compliance (causticity, irritation, painful injections, undesirable organoleptic properties). To render the compound marketable, the pharmacists must then invent an adequate pharmaceutical formulation, stable on conservation and perfectly adapted to its clinical use. In leaving this task to the formulation pharmacists, the chemists are not always aware of their own role in the formulation phase and the final presentation of a new drug. They also probably overestimate the intervention possibilities of the formulation pharmacists. Of course, new pharmaceutical technologies such as micro-encapsulation or cyclodextrin complexation can overcome various drawbacks, but they are costly and often replaceable by simple chemical derivations. Medicinal chemists should therefore also feel concerned with these problems and tackle them in an early research stage, before the investment in a moderately satisfying molecule has become too large for an alternative to be sought. Besides the pharmaceutical approach to formulation problems there is place for chemical solutions.

II. INCREASING CHEMICAL STABILITY

The shelf-life of an organic chemical is the time taken for its pharmacological activity to fall by an acceptable amount. Whilst this cannot be universally defined, about 10% decomposition may be considered acceptable, unless the decomposition products are toxic. Among the possible

Copyright © 2008, Elsevier Ltd
All rights reserved.

FIGURE 41.1 Stabilized derivatives of prostaglandin E2: ethylene ketal and crystalline ester with para-hydroxyacetophenone.[6,7]

decomposition paths, the following ones are the most frequent: hydrolysis, oxidation, photochemical degradations, racemization, thermal decomposition, chemical interactions, and microbial degradations.[2] Usually, selected pharmaceutical prevention procedures and manufacturing practices can overcome decomposition problems. For example, hydrolyzes are avoided in dissolving the active principle in an anhydrous solvent, oxidations are prevented in replacing atmospheric oxygen by an inert gas or by adding antioxidants and so on. However, it can happen that chemical derivatizations are needed to achieve satisfying drug formulations. Some of them are illustrated by the examples described below. More detailed reviews on the basic physical and chemical principles that determine the stability do drugs are available in specialized monographs.[3–5]

Pressing problems with prostaglandin E_2 (dinoprostone) and many other prostaglandins are chemical instability and difficulties in handling a liquid compound. Effectively, prostaglandin E_2 is a crystalline solid (mp 63°C), stable at room temperature for short periods, but it liquefies and decomposes rapidly after a few months. The instability of dinoprostone is due to the 9,11β-ketol system in which the activated C_{11} hydroxyl undergoes facile elimination to give prostaglandin A_2. The corresponding ethylene ketal derivative (Figure 41.1) possesses much improved solid-state stability. It may be an orally useful prodrug form, since it readily undergoes an acid-catalyzed hydrolysis back to the parent prostaglandin under conditions similar to that prevailing in the stomach.[6]

The crystalline ester with *para*-hydroxy-acetophenone (Figure 39.1), or with some other phenolic compounds, represent another means to overcome these difficulties.[7] Storage of the crystalline esters at room temperature for 22–30 months resulted in no detectable degradation as shown by silica TLC; the esters remained as white solids. The free acid, on the other hand, underwent 44–59% degradation in 12 months at room temperature.

Cycloserine forms a bioreversible condensation product with acetylacetone (Figure 41.2). This greatly depresses the

FIGURE 41.2 Derivatization of the primary amino function of cycloserine. A bioreversible enamine, probably stabilized by an intramolecular hydrogen bond, is formed by reaction with acetylacetone.[8]

dimerizing tendency of the parent drug, which occurs particularly fast in concentrated aqueous media and can even take place in the solid state.[8]

Pilocarpine is widely used as a topical miotic for controlling the elevated intraocular pressure associated with glaucoma. Beside its low lipophilicity, which stimulated the search for prodrugs,[9] pilocarpine has a short duration of action, its lactonic ring being rapidly opened to yield pilocarpic acid.

The synthesis of its isosteric carbamate (Figure 41.3), which is as effective as pilocarpine, has greatly improved the stability of the former lactonic ring.[10]

Although the phthalimido-bearing pyridinone **1** (Figure 41.4) is a potent reverse transcriptase inhibitor (IC_{50} = 30 nM), hydrolytic instability under physiological conditions precludes the demonstration of its antiviral activity in cell culture. This instability was recognized to be due to the aminal structure of the exocyclic NH.

Subsequent efforts led to the substitution of the NH with a methylene unit and the replacement of the phtalimido

FIGURE 41.3 Stabilization of the lactonic ring of pilocarpine. In preparing the carbamic analog of pilocarpine, the sensitivity of the carbonyl group toward nucleophilic attacks is greatly decreased.[10]

FIGURE 41.4 Improved chemical stability through suppression of an aminal function.[11]

moiety by a benzoxazole. One of the prepared compounds, 3-[(benzoxazol-2-yl)ethyl]-5-ethyl-6-methylpyridin-2(1*H*)-one **2** (Figure 41.4), is as potent (IC_{50} = 23 nM) as the phthalimido derivative but is 10 times more stable towards acid hydrolysis and shows good oral bioavailability.[11]

Cefoxitin is a broad spectrum, semi-synthetic cephamycin antibiotic. The free acid form is a white, crystalline, practically insoluble solid and therefore the decision was made to use the sodium salt to provide a sterile solution for intravenous administration. It was soon evident that the sodium salt of cefoxitin would have limited stability in solution and would not lend itself to the formulation of a marketable solution product. Stability studies highlighted the advantages of sterile crystalline solid over amorphous freeze-dried product, and also the need for rubber stopper screening studies to eliminate interactions between sodium cefoxitin powder and rubber, and finally the profound effect of oxygen on the coloration rate of the product.[12]

The *N*-nitrosoureas such as lomustine, carmustine, and tauromustine are important alkylating antineoplastic agents that have demonstrated activity against a wide spectrum of tumors. Although some *N*-nitrosoureas (e.g. lomustine) are sufficiently stable to be administered orally, they are generally very unstable in aqueous solutions. Addition of an equimolar amount of Tris (tris-hydroxymethyl aminomethane) forms a complex that is stable in aqueous solution and which increases shelf-life of the nitrosoureas. The rate of degradation of the drugs in the complex is 1.5 to 2.5 times slower than outside the complex.[13]

FIGURE 41.5 Formation of a magnesium salt stabilizes omeprazole.

In a similar way the antiulcerative drug omeprazole is claimed to be more stable if prepared as a magnesium salt.[14] Probably the compound forms a 2:1 complex (Figure 41.5).

Cyclodextrin complexation (See Chapter 40) can also represent a way to improve the stability and the solubility of sensitive drugs such as thalidomide. Thalidomide is currently in clinical use for the treatment and prevention of graft-versus-host disease in leukemia patients after bone marrow transplant. However, this drug is sparingly soluble in aqueous solutions (50 κg/mL) and is readily hydrolyzed. Complexation with hydroxypropyl β-cyclodextrin increases the solubility to 1700 κg/mL and extends the half-life of a dilute solution from 2.1 to 4.1 h.[15] Other vulnerable and sparingly soluble drugs stabilized by means of cyclodextrin complexation are the non-steroidal antiinflammatory drugs diclofenac, piroxicam, and indomethacin[16] and the anthracycline antibiotic daunorubicin.[17]

The presence of an alkyl substituent at N3 in the oxazaphosphorine ring stabilizes *N*-substituted 4-(alkylthio) cyclophosphamides from "spontaneous" decomposition. On the basis of this finding, several *N*-methyl-4-thio-cyclophosphamide derivatives were synthesized and examined

FIGURE 41.6 Stable precursor of 4-hydroxycyclophosphamide.[18]

FIGURE 41.7 Thioester derivatives of leinamycin are prodrugs with improved stability.[19]

as prodrugs of 4-hydroxycyclophosphamide, the activated species of cyclophosphamide.[18]

All prodrugs are stable in aqueous buffer but undergo *N*-demethylation when incubated with rat hepatic microsomes, forming alkylating species. *N*-Methyl-4-(diethyldithiocarbamyl) cyclophosphamide particularly (Figure 41.6) shows notable *in vitro* toxicity against mouse 3T3 cells and human tumor cells.

The half-life of the antitumor antibiotic leinamycin in aqueous solution at pH 7 at 37°C is about 6 h and the hydrolysis of the reactive dithiolanone moiety could be one of the reasons of instability. However, this dithiolanone moiety of leinamycin has been shown to be essential for its DNA-cleaving activity and antiproliferative activity. Accordingly, Kanda *et al.*[19] devised more stable thioester derivatives, resulting from various alkylations in presence of potassium carbonate (Figure 41.7).

These derivatives have a unique 3-isothiazolidinone-1-oxide moiety and can act as prodrugs that provide dithiolanone compounds in biological media. Particularly compound "3c" was found to show potent antitumor activity against human tumor xenograft, such as lung, liver, ovary, prostate, and colon carcinomas.[19]

III. IMPROVED FORMULATION OF PEPTIDES AND PROTEINS

Enhancement of proteolytic stability of peptides can be achieved by well-established procedures such as synthesis of retropeptides, isosteric replacements of the peptidic bond, *N*-methylation and use of unnatural D-amino acids.[20,21] Prodrugs of peptides can also be helpful, see Chapter 36.

FIGURE 41.8 *N*-acetylation of Z-Gly-ProNH2 protects against the hydrolytic activity of prolyl endopeptidase.[21]

FIGURE 41.9 Condensation of acetone with Met-enkephalin yields a 4-imidazolidinone with increased stability toward aminopeptidases.[22]

However, sometimes very simple derivatives can afford protection against enzymatic degradation. Thus the *N*-acetyl derivative of Z-glycyl-prolylamide (Figure 41.8) is six times more resistant to gut prolyl endopeptidase (which normally hydrolyzes the terminal primary amide) than the non-acetylated parent molecule.[21]

In a analogous way, Leu-enkephalin and Met-enkephalin were rendered resistant to aminopeptidases[22] by condensation with various aldehydes and ketones to form 4-imidazolidine derivatives (Figure 41.9).

Interesting case histories about formulation, characterization and stability of protein drugs are discussed in the book edited by Pearlman and Wang.[23]

IV. DEALING WITH MESOMORPHIC CRYSTALLINE FORMS

Polymorphism is a keyword of considerable importance in the life sciences and especially in the pharmaceutical industry. It abridges the fact that a solid compound can exist in different crystalline forms that can have different physical chemical properties. To ensure no variations in the product to be due to different solid-state properties, care must be taken in selecting the most appropriate solid-state form for the substance and in ensuring a reproducible production of this form.

As concerns the pharmaceutical industry, it has been shown that more than half of the drug substances described in monographs crystallize in more than one solid-state form, being it either polymorphs, solvates, or both.[24] A list of about 600 polymorphic drug substances was published and discussed by Giron.[25] The solid-state form of a drug substance can influence a variety of properties, namely the solubility and rate of dissolution or the chemical stability or stability against excipients. Thus, the regulatory bodies require an exhaustive search for polymorphic forms of a drug substance.[26,27] The manufacturer is required to make a substantiated choice for one of the forms, or a defined mixture of forms. Changes in the polymorphic form of the batches produced are seen as indicative of changes in the production process, also requiring the reproducible crystallization of a certain solid-state form. The choice of the solid-state form of a new drug substance is up to the applicant and should be made by considering all aspects, such as chemical stability and stability against excipients, dissolution behavior and bioavailability, and last but not least thermodynamic stability of the solid-state form and ease and reproducibility of production.

Some of the possible crystalline states are metastable and can be converted into more stable forms with different physicochemical properties. Two types of transformation are possible: the reversible enantiotropic transformation in which the polymorphic forms can change from one into the other and vice versa, and the irreversible monotropic transformation, which is now most often met and which involves the change from a thermodynamically unstable form to a more stable one. The different crystalline forms of a given compound can be distinguished by their melting point and solubility, by scanning calorimetry, by thermogravimetric analysis, by infrared spectrometry, by X-ray powder diffraction, and by scanning electron microscopy. A complete list of the physical properties that differ among polymorphs is given in Table 41.1.[28]

As a rule metastable polymorphs will tend to have an increased solubility and a faster dissolution rate than a stable polymorph. Although fourfold increases in solubility are sometimes observed between the two forms,[29,30] a 50–100% increase in dissolution rate represents the most usual situation.[31] Riboflavin is an exception. It has three polymorphs with solubilities of 60, 810, and 1200 mg/mL.[32] When a metastable form is placed in contact with solvents, it can change progressively to the most stable but least soluble form. An example is provided by the acidic and amorphous form of novobiocin that can hardly be administered as a suspension because of its tendency to turn into a much less soluble crystalline form.[33] Spray drying of pure drugs often results in amorphous products with increased rates of dissolution, eventually an auxiliary ingredient can be added to yield solid spray dispersions of drugs.[34]

The transformation from one crystalline form to the other can also occur during the manufacturing processes. Chloroquine diphosphate crystals for example can be obtained anhydrous by storing the hydrate at elevated temperatures. Dehydration is favored by grinding the raw

TABLE 41.1 Physical Properties that Differ Among Polymorphs[28]

1.	Packaging properties Molar volume and density Refractive index Conductivity, electrical and thermal Hygroscopicity
2.	Thermodynamic properties Melting and sublimation temperatures Internal energy (i.e. heat content) and entropy Heat capacity Free energy and chemical potential Thermodynamic activity Vapor pressure Solubility
3.	Spectroscopic properties Electronic, vibrational, rotational, and nuclear spin transitions
4.	Kinetic properties Dissolution rate Rates of solid state reactions Stability
5.	Surface properties Surface free energy Interfacial tensions Habit (i.e. shape)
6.	Mechanical properties Hardness, compatibility, and tabletting properties Tensile strength, flow, blending, and handling

BOX 41.1 Importance of Pharmaceutical Polymorphism: the Case of Ritonavir

Ritonavir (Norvir®; Abbott) is a drug for treating patients infected with human immunodeficiency virus-1 (HIV-1). In solid state the compound consists exclusively of one kind of monoclinic crystals. As this form, now called form I, was not sufficiently bioavailable by the oral route, a liquid formulation containing dissolved ritonavir in a hydroalcoholic solution was developed. It gave satisfactory results during the development and early manufacturing phases. However, in 1998, 2 years after the launch, a second form (form II) precipitated during shelf life. The new form is thermodynamically more stable and about 50% less water-soluble than form I. Within weeks–days, this new form was produced in all the production lines and Abbott had to stop the production of ritonavir. Finally, after considerable effort and expense, a new formulation of ritonavir was developed.

material. The transition of anhydrous chloroquine diphosphate into a second hydrated form is possible by storing the drug at a high relative humidity. Compression of the raw material resulted in the formation of another crystal form.[35] In a similar way, grinding the polymorphic form A of chloramphenicol stearate in the presence of colloidal silica converts it into form B.[36] Taken together, these findings emphasize the necessity for standardization of the manufacturing processes and closer examination of the solid drug as part of the quality control.

V. INCREASING THE MELTING POINT

A less frequent but nonetheless interesting problem arises in the chemical modification of liquid, or low-melting, active principles in solid prodrugs, suitable for tablet or capsule preparation. Indeed solid dosage forms are still the most widely used for the administration of medicines, as well for patient acceptability and convenience for product stability and ease of manufacture. Their preparation implies that the active principle can itself be handled as a stable solid, an objective that is usually attained by one of the following strategies: formation of a salt or a molecular complex, formation of a crystalline covalent derivative, introduction of symmetry.

A. Salt or complex formation

Many basic active compounds, belonging to various therapeutic classes (antihistaminics, neuroleptics, local anesthetics, etc.) are oily liquids. Salification with an appropriate acid (hydrochloric, phosphoric, tartaric, etc.) represents a good mean to convert them into the desired aggregation state (see Chapter 37). Neutral liquids can sometimes be converted into solids through formation of molecular complexes. The central and respiratory stimulant nikethamide is a slightly viscous oil or a low-melting (mp 24–26°C) crystalline solid. With calcium thiocyanate it forms a solid combination made of two molecules of *N,N*-diethylnicotinamide and one molecule of Ca(SCN)2 that can be used for tabletting.[30] Chloral hydrate, another low-melting solid, yields a crystalline and odorless 2:1 complex with phenazone;[37,38] see below. Chloral formamide (Merck Index) and Chloral *N*-acetylglycinamide[39] are other solid complexes.

B. Covalent derivatives

The methodology consists in linking the liquid active principle, by means of a covalent bond, to a non-toxic moiety having a strong tendency to crystallize. Liquid or greasy fatty acids can be easily converted into phenacyl esters[40] and substances related to prostaglandins give, as shown above, crystalline esters with *para*-hydroxy-acetophenone (Figures 41.1 and 41.10) or with some other phenolic compounds[7]. Chloral yields a crystalline hemiacetal (Figure 41.17) with the phenolic hydroxyl of paracetamol.[41]

Trichloroethanol can be converted to its phosphate ester, that gives a crystalline monosodium salt,[42] the same

FIGURE 41.10 Crystalline derivatives of fatty acids.

FIGURE 41.11 Covalent solid derivatives of liquid active principles.

possibility exists for the low-melting (mp 28°C) guaiacol (Guaiacol phosphate: Merck Index) (Figure 41.11).

Phenols can be acylated (guaiacol benzoate, phenylacetate, or valerate) or alkylated (guaifenesin, guaiapate) (Figure 41.11).

C. Introduction of symmetry

It is well known that symmetric molecules have higher melting points than non-symmetric ones, for example, para-nitrotoluene is a solid, ortho- and meta-nitrotoluenes are liquids. For chemical formulation of pharmaceuticals, symmetry can be obtained in preparing twin drugs (see Chapter 18). Low molecular weight duplication reagents are formaldehyde or glyoxylic acid (methylene-bis type derivatives), phosgene or phosphorus oxychloride (carbonate or phosphate diesters), and small bi-functionalized molecules (ethylene glycol, ethylene-diamine esters, or amides).

Symmetric carbonates of guaiacol or ethyl salicylate are crystalline solids (Figure 41.12), pentaerythritol tetranitrate is the solid counterpart of the liquid trinitrine and petrichloral is again a stabilized chloral derivative. Methenamine, formulated as an enteric-coated tablet, is used as a urinary tract antibacterial. Following absorption, the compound is eliminated in the urine where formaldehyde is generated in the acidic environment.

VI. GASTROINTESTINAL IRRITABILITY AND PAINFUL INJECTIONS

A. Gastrointestinal irritability

Gastrointestinal disturbancies can occur through several causes: (1) Direct contact of the drug with the gastric mucosa, producing a localized irritating and necrotizing effect, (2) Irritation of the gastric mucosa through an indirect mechanism, for example, stimulation of the gastric secretion, (3) Irritation of the intestinal mucosa, (4) Inhibition of the mucopolysaccharide biosynthesis, and (5) Destruction of the intestinal flora.

Chemical formulations can mainly deal with the direct irritating properties of drugs containing phenolic or acidic groups. For example, to eliminate gastric irritation produced by salicylic acid, aspirin was developed. Another possibility is given by carbonate esters (*n*-hexyl carbonate) that are rapidly hydrolyzed *in vivo* and might have hydrophobic properties adequate to allow absorption and distribution over a greater area of the gastrointestinal (GI) tract, thus reducing local irritation.[43]

For non-steroidal antiinflammatory carboxylic acids such as mefenamic acid[44] or *N*-(7-chloro-4-quinolyl)-anthranilic acid,[45] glyceryl esters are claimed as less irritating. Alternatively ulcerogenicity of indomethacin derivatives

was reduced by means of the ester with glycolic acid[46,47] or the peptide with serine.[48,41] Another indomethacin-related antiinflammatory drug, sulindac, is an inactive sulfoxide and becomes only activated after absorption and reduction into the corresponding sulfide. Thus, the initial exposure of gastric and intestinal mucosa to the active drug is circumvented (see Chapter 36: Prodrugs).

B. Avoidance of painful injections

When an initially painful intravenous or intramuscular injection must be administered repetitively, patient reluctance develops. Injection pains are usually accompanied by hemorrhage, edema, inflammation, and tissue necrosis.[49] Among the factors responsible for painful injections, the most important are the drug solubility in aqueous medium, the viscosity, the pH and the hypo- or hyperosmotic character of the injected drug solution, the amount of the injected volume, the site of injection, the pain tolerance of the patient, and the technique of administration. Other factors include precipitation of the drug at the injection site, and localized cell lysis.[49]

Too high concentrations of the active compound at the injection site (initial peak concentrations) are avoided in masking the irritating agent through complexation. The free concentrations of iron or of calcium are reduced in using calcium gluconate or levulinate for intravenous injections and iron-sorbitol citrate for intramuscular injections.[50]

The low aqueous solubility of the antibiotic clindamycin hydrochloride is responsible for the pain experienced on intramuscular injection. Phosphorylation improves the aqueous solubility from 3 to >150 mg/mL and avoids the pain resulting from injection of the parent clindamycin.[51,52] The phosphorylated drug (Figure 41.13) possesses little or no intrinsic antibacterial activity,[53] but, due to the enzymatic action of phosphatases, it is rapidly converted to the parent drug. The half-life for hydrolysis *in vivo* is approximately 10 min and only 1–2% of an intravenous dose is eliminated in the urine as unchanged prodrug.[54]

The anticonvulsant drug phenytoin (5,5-diphenylhydantoin) is sparingly soluble in water (0.02 mg/mL) and is therefore formulated as sodium enolate for intravenous or intramuscular injections. The solvent is a mixture of 40% propylene glycol, 10% ethanol and 50% water and the pH of the solution is alkaline (pH ≈ 12), due to the weak acidity

FIGURE 41.12 Compounds for which the solid state results from molecular symmetry.

FIGURE 41.13 Phosphate esters of clindamycin and of 3-hydroxymethyl-phenytoin (as sodium salts) are freely soluble in water and allow painless injections.

(pK_a = 8.3) of the drug.[55] Such a formulation allows concentrations as high as 50 mg/mL, but there is a serious risk of phenytoin precipitation at the physiological pH 7.4 of the injection site. On the other hand propylene glycol is a cardiac depressant. The disodium salt of the phosphate ester of 3-hydroxymethyl-phenytoin was found to be extremely soluble (4,500 times phenytoin), pharmacologically inert and able to regenerate free phenytoin rapidly and quantitatively with no apparent irritation.[56]

VII. SUPPRESSING UNDESIRABLE ORGANOLEPTIC PROPERTIES

The use of flavors and flavor modifiers represents the first attempt to improve pharmaceuticals in masking undesirable organoleptic properties such as taste, odor and feeling factors (For a review see Ref. [57]). When this approach reveals to be ineffective chemical modifications have to be considered.

A. Odor

Despite its interesting antiseptic properties, a compound such as iodoform was rejected by the medical community because of its disagreeable odor. A first historical attempt towards an odorless substitute was bismuth iodosubgallate (Airoform®) which is practically odorless (Figure 41.14).

Bismuth iodosubgallate was followed by the introduction of iodochlorhydroxyquin (Vioform®). Unfortunately this latter compound has been linked with the occurrence of subacute myelo-optic neuropathy (SMON syndrome) in Japan. The decisive progress came with the iodine-free chlorquinaldol (Sterosan®).

Thiamine chloride has a slight, but disagreeable and penetrating odor and causes often vomiting when incorporated to pediatric polyvitaminic preparations. The corresponding monophosphoric ester (Figure 41.15) is totally odorless and shows increased stability.[58]

The development of diethyl dithio*iso*phtalate (Etisul®) as an antituberculosis and antileprosy drug originated from the observation that ethyl mercaptan, an evil smelling, low boiling, inflammable liquid was active as an antitubercular agent.

Among a number of derivatives prepared and which would be expected to be metabolized to ethyl mercaptan *in vivo*, the most successful was diethyl dithio*iso*phtalate (Figure 41.16), an almost bland odorless oil, which is effective in experimental tuberculosis parenterally.[59,60]

Besides its bad taste and to its gastric and intestinal irritating properties chloral hydrate has an unpleasant odor. All these disadvantages have been overcome by complexing two molecules of chloral hydrate with one molecule of phenazone to give dichloralphenazone.[37,38]

FIGURE 41.15 Thiamine chloride and its odorless monophosphate ester.[58]

FIGURE 41.16 Diethyl dithio*iso*phtalate an odorless administration form of ethyl mercaptan[59,60].

FIGURE 41.14 Change from iodoform to odorless antiseptics.

Dichloralphenazone (Figure 41.17) possesses all the hypnotic and sedative properties of chloral hydrate, and is less toxic and free from the unpleasant odor and taste of chloral. Another masked form of chloral, cloracetadol, results from a hemiacetalic bonding between chloral hydrate and paracetamol (Figure 41.17). As such, this combination represents three achievements: attenuation of the bad odor, obtaining of a solid state, and associative synthesis.[41] Trichloroethanol, the active metabolite of chloral, is deodorized as a symmetric ester with carbonic acid. The ester is prepared from trichloroethanol and phosgene. It is a tasteless, crystalline solid with limited solubility in water. Pharmacological comparisons show its equivalence with trichloroethanol.[61]

B. Taste

Taste is a complex combination of sensations including gustation, olfaction, tactility, and responses to heat and cold. In humans the sensation of taste is confined primarily to the dorsal face of the tongue, the soft palate, the epiglottis, and parts of the gullet. In children, however, the taste receptors are distributed over larger areas of the mouth.[49] Basically there are four so-called primary tastes: sour, bitter, salty, and sweet. Others include astringent, metallic, and alkaline (soapy). For pharmaceuticals, the main problem resides in masking the bitterness of some orally administrated pediatric formulations, such as antibiotic containing suspensions and sirups. Two strategies are then possible: create a decrease in drug solubility, enough to lie under the threshold value for taste receptor activation, or modify the shape and the electronic potential of the molecule. The first strategy is well illustrated by the formation of very sparingly soluble salts or esters (Table 41.2). In the GI tract or following absorption the derivatives are cleaved with formation of the parent drugs.

FIGURE 41.17 Chloral hydrate yields an odorless 2:1 complex with phenazone[37,38] and a hemiacetal with paracetamol[41]. Trichloroethanol is deodorized as carbonate ester.[61]

However, insolubility in water is not always a prerequisite and in some cases water-solubilizing groups achieve as well the suppression of bitterness. Examples are *N*-arylanthranilic glyceryl esters[44] or lincomycin phosphate.[68,69]

An illustration of the second strategy is provided by the para-cyano analogs of the non-nutritive sweetener suosan (Figure 41.18). The replacement of the planar carboxylic group by the bioisosteric tetrazolyl group yields less potent, but still sweet compounds.[70]

However, replacement of the carboxylic group by the tetrahedral sulfonic group yields an *antagonist* of the sweet taste response that inhibits the sweet taste perception of a variety of sweeteners without having any effect on the sour or salty taste response. Surprisingly the sulfonic antagonist antagonized also the bitter taste response to reference compounds such as caffeine, quinine, and naringine.[71]

Tinti and Nofre[72,73] observed that the combination of the cyano analog of suosan with the elements of aspartame yields "superaspartame" which has a potency 14,000 times that of sucrose (aspartame has a potency of about 180 times that of sucrose).

Subsequently, these authors discovered sweet guanidines such as the compound SC-45647 which has a potency 28,000 times that of sucrose.[73,74] More recently, they reported *N*-alkyl-substituted aspartame derivatives such as neotame which has a potency 10,000 times that of sucrose.[75] For these

FIGURE 41.18 Suosan analogs.[70,71]

FIGURE 41.19 Aspartame and aspartame-suosan hybrids.[72–75]

TABLE 41.2 Sparingly Soluble Salts or Esters of Bitter Drugs

Parent molecule	Derivative	Reference
Dextropropoxyphene	Napsylate salt	Gruber *et al.*[62]
Tetracycline	3,4,5-Trimethoxybenzoate salt	Rocca and Rusconi[63]
Chloramphenicol	Palmitic ester	Glazko *et al.*[64]
Sulfafurazole (sulfisoxazole)	N′-Acetyl	McEvoy[65]
Erythromycin	Ethylsuccinate	Murphy[66]
Clindamycin	Dialkylcarbonate esters	Sinkula *et al.*[67]

three high-potency sweeteners (Figure 41.19), Walters *et al.* were able to identify a five-point pharmacophore model in which the carboxylate, the two hydrophobic groups and two NH groups match well sterically and with respect to the orientation of hydrogen bond donors and acceptors. More recently Bassoli *et al.* developed a pseudoreceptor model for sweet compounds from a training set of 24 sweet compounds belonging to different families.[76] Their model provided a correlation coefficient (r^2) of 0.985 between the calculated and the experimental free energies of binding related to the molar relative sweetness.

A series of qualitative and empirical rules was collected by Sinkula.[77] They may provide some clues to the medicinal chemist for improving the taste quality of objectionably tasting drugs (Table 41.3).

TABLE 41.3 Empirical and Qualitative Structure–Taste Relationships[77]

1.	Compounds containing several hydroxyl groups are usually *sweet* (glycols, sugars, and other carbohydrates)
2.	An increase in molecular weight, that is, extending a homologous series, often changes from *sweet* to *bitter*. Further increases in homologation reduce drug solubility below taste threshold values
3.	Nitro groups present in a molecule are usually indicative of *bitter* taste (chloramphenicol, picric acid)
4.	Alkylation of an amine or amide usually produces a *sweet* tasting substance
5.	Alkylation of a hydroxyl soup (etherification) *destroys sweet* taste
6.	Alkylation of an imide *eliminates sweet taste*; N-alkyl saccharins are tasteless
7.	Addition of a phenyl group to a drug molecule causes or increases bitter taste and may be due to enhanced lipophilicity and/or preferential binding with bitter taste receptors
8.	An increase in chain branching decreases sweet taste and increases *bitter* taste
9.	Esterification *enhances sweetness*, for example, ethyl butyrate, aspartyl dipeptide esters
10.	Introduction of unsaturation into a molecule *increases bitterness and pungency*
11.	Primary amines *enhance sweet taste* especially if in near proximity to an electronegative group (-COOH), that is, in certain D- and L-amino acids
12.	Secondary, tertiary amines and quaternary ammonium salts are *bitter*, especially alkaloids, antibiotics and quaternary ammonium drugs

(Continued)

TABLE 41.3 (Continued)

13. Monosubstituted ureas are usually *sweet* although some are tasteless. Urea itself is *bitter*. Symmetrical molecules are often *bitter*
14. Aliphatic polyhalogenated compounds are sweet (chloroform, 1,2-dichloroethane, dichloromethane)
15. Aromatic halogen substitution increases bitterness and is a function of the atomic weight of the halogen (see e.g., halogenated saccharins)
16. The presence of sulfur in an aliphatic molecule is usually associated with *bitterness* (—SH, —S—, —S—S—, C═S)
17. Sulfonic acids are usually *bitter* or *tart*
18. Aldehydes are *sweet*, Aldehydo-semicarbazones are not as *sweet* as the parent aldehydes, Aldehydo-phenylhydrazones *are not sweet*
19. Oximes are usually *sweet* or *tasteless*

REFERENCES

1. Lucretius, T. C. *The Nature of the Universe (Translated by Latham)*. Penguin Books: London, 1951. 47 bc.
2. Shotton, E., Ridgway, K. *Physical Pharmaceutics*. Clarendon Press: Oxford, 1974. pp. 313–331.
3. Connors, K. A., Amidon, G. L., Stella, V. J. *Chemical Stability of Pharmaceutics: A Handbook for Pharmacists*, 2nd Edition. John Wiley & Sons: New York, 1986.
4. Essig, D., Hofer, J., Schmidt, P. C., Stumpf, H. *Stabilisierungst echnologie- Wege zur haltbaren Arzneiform*. Wissenschaftliche Verlagsgesellschaft mbH: Stuttgart, 1986.
5. Carstensen, J. T., Rhodes, C. T. *Drug Stability – Principles and Practices*, 3rd Edition. Culinary and Hospitality Industry Publication Services, 2000. p. 792.
6. Cho, M. J., Bundy, G. L., Biermacher, J. J. Prostaglandin prodrugs. 5. Prostaglandin E2 ethylene ketal. *J. Med. Chem.* 1977, *20*, 1525–1527.
7. Morozowich, W., Oesterling, T. O., Miller, W. L., Lawson, C. F., Weeks, J. R., Stehle, R. G., Douglas, S. L. Prostaglandine prodrugs I: stabilization of dinoprostone (prostaglandin E2) in solid state through formation of crystalline C1-phenyl esters. *J. Pharm. Sci.* 1979, *68*, 833–836.
8. Jensen, N. P., Friedman, J. J., Kropp, H., Kahan, F. M. Use of acetylacetone to prepare a prodrug of cycloserine. *J. Med. Chem.* 1980, *23*, 6–8.
9. Bundgaard, H. *Design of Prodrugs*. Elsevier: Amsterdam, 1985. pp. 55–61.
10. Sauerberg, P., Chen, J., WoldeMussie, E., Rapoport, H. Cyclic carbamate analogues of pilocarpine. *J. Med. Chem.* 1989, *32*, 1322–1326.
11. Hoffman, J. M., Smith, A. M., Rooney, C. S., Fisher, T. E., Wai, J. S., Thomas, C. M., Bamberger, D. L., Barnes, J. L., Williams, T. M., Jones, J. H., Olson, B. D., O'Brien, J. A., Goldman, M. E., Nunberg, J. H., Quintero, J. C., Schleif, W. A., Emini, E. A., Anderson, P. S. Synthesis and evaluation of 2-pyridone derivatives as HIV-1-specific reverse transcriptase inhibitors. 4. 3-[(Benzoxazol-2-yl) ethyl]-5-ethyl-6-methylpyridin-2(1 *H*)-one and analogues. *J. Med. Chem.* 1993, *36*, 953–966.
12. Portnoff, J. B., Henley, M. W., Restaino, F. A. The development of sodium cefoxitin as a dosage form. *J. Parenteral Sci. Technol.* 1983, *37*(5), 180–185.
13. Loftsson, T., Fridriksdottir, H. Stabilizing effect of tris(hydroxymethyl) aminomethane on *N*-nitrosoureas in aqueous solutions. *J. Pharm. Sci.* 1992, *81*, 197–198.
14. Brandstrom, A.E. Base addition salts of omeprazole. United States patent 4, 738, 974 Aktiebolaget Hassle, Sweden.
15. Krenn, M., Gamcsik, M. P., Vogelsang, G. B., Colvin, O. M., Leong, K. W. Improvements in solubility and stability of thalidomide upon complexation with hydroxypropyl-β-cyclodextrin. *J. Pharm. Sci.* 1992, *81*, 685–689.
16. Backensfeld, T., Müller, B. W., Kolter, K. Interaction of NSA with cyclodextrins and hydroxypropyl-cyclodextrin derivatives. *Int. J. Pharm.* 1991, *74*, 65–93.
17. Suenaga, A., Bekers, O., Beijnen, J. H., Underberg, W. J. M., Tanimoto, T., Koizumi, K., Otagiri, M. Stabilization of daunorubicin and 4-demethoxydaunorubicin on complexation with octakis(2,6-di-*O*-methyl)-γ-cyclodextrin in acidic aqueous solution. *Int. J. Pharm.* 1992, *82*, 29–37.
18. Moon, K.Y., Kwon, C.H., Shirota, F.N., Baturay, N.Z. In *Design, Synthesis, and Evaluation of N-Methyl-4-Thiocyclophosphamide Derivatives as Chemically Stable, Alternative Prodrugs of 4-Hydroxycyclophosphamide*, 84 Meeting, 1993, p. 267.
19. Kanda, Y., Ashizawa, T., Kakita, S., Takahashi, Y., Kono, M., Yoshida, M., Saitoh, Y., Okabe, M. Synthesis and antitumor activity of novel thioester derivatives of leinamycin. *J. Med. Chem.* 1999, *42*, 1330–1332.
20. Plattner, J. J., Norbeck, D. W. Obstacles to drug development from peptide leads. In *Drug Discovery Technologies* (Clark, C. R., Moos, W. H., Eds). Ellis Horwood Limited: Chichester, 1990, pp. 92–126.
21. Møss, J., Bundgaard, H. Prodrugs of peptides. 17. Bioreversible derivatization of the *C*-terminal prolineamide residue in peptides to afford protection against prolyl endopeptidase. *Int. J. Pharm.* 1992, *82*, 91–97.
22. Rasmussen, G. J., Bundgaard, H. Prodrugs of peptides. 15. 4-Imidazolidinone prodrug derivatives of enkephalins to prevent aminopeptidase-catalyzed metabolism in plasma and absorptive mucosae. *Int. J. Pharm.* 1991, *76*, 113–122.
23. Pearlman, R., Wang, Y. J. *Formulation, characterization, and stability of protein drugs. Case histories*, Vol. 9. Plenum Press: New York, 1996.
24. Henck, J. O., Griesser, U. J., Burger, A. 50% of all drugs are polymorphs. *Pharm. Ind.* 1997, *59*, 65–169.
25. Giron, D. Thermal analysis and calorimetric methods in the characterization of polymorphs and solvates. *Thermochim. Acta* 1995, *248*, 1–59.
26. Byrn, S., Pfeiffer, R., Ganey, M., Hoiberg, C., Poochikian, G. Polymorphs and regulatory bodies. *Pharm. Res.* 1995, *2*, 945.
27. Beckmann, W. Seeding the desired polymorph: background, possibilities, limitations, and case studies. *Org. Process Res. Dev.* 2000, *4*, 372–383.
28. Chawla, G., Bansal, A. K. Challenges in polymorphism of pharmaceuticals. *CRIPS* 2004, *5*, 9–12.
29. Lin, S. L. Preformulation investigation II: Dissolution kinetics and thermodynamic parameters of polymorphs of an experimental antihypertensive. *J. Pharm. Sci.* 1972, *61*, 1423–1430.
30. Aguiar, A. J., Zelmer, J. E. Dissolution behavior of polymorphs of chloramphenicol palmitate and mefenamic acid. *J. Pharm. Sci.* 1969, *58*, 983–987.
31. Shefter, E. Solubilization by solid-state manipulation. In *Techniques of Solubilization of Drugs* (Yalkowsky, S. H., Ed.). Marcel Dekker, Inc.: New York, 1981, pp. 159–182.
32. Shotton, E., Ridgway, K. *Physical Pharmaceutics*. Clarendon Press: Oxford, 1974. p. 337.
33. Mullins, J. D., Macek, T. J. Some pharmaceutical properties of novobiocin. *J. Amer. Pharm. Assoc., Sci. Ed.* 1960, *49*, 245–248.
34. Nürnberg, E. Darstellung und Eigenschaften pharmazeutisch relevanter Sprühtrocknungsprodukte, eine Übersicht. *Acta Pharm. Technol.* 1980, *26*, 39–67.
35. Bjerga Bjaen, A. K., Nord, K., Furuseth, S., Agren, T., Tonnesen, H. H., Karlsen, J. Polymorphism of chloroquine diphosphate. *Int. J. Pharm.* 1993, *92*, 183–189.

36. Forni, R., Coppi, G., Iannucelli, V., Vandelli, M. A., Cameroni, R. The grinding of the polymorphic forms of chloramphenicol stearic ester in the presence of colloidal silica. *Acta Pharm. Suec.* 1988, *25*, 173–180.
37. Willis, G. C., Arendt, E. C. Relative efficacy of sodium amytal, chloral hydrate and dichloralphenazone. *Canad. Med. Assoc. J.* 1954, *71*, 126–128.
38. Rice, W. B., McColl, J. W. A comparison of dichloralphenazone and chloral hydrate. *J. Amer. Pharm. Assoc., Sci. Ed.* 1956, *45*, 137–141.
39. Bruce, W.F., Chloral derivatives and methods for their preparation. *US Pat. 2,784,237 (5 March 1957, to American Home Products)* 1957.
40. Vogel, A. I. *Vogel's Textbook of Practical Organic Chemistry*, 5th Edition. Longman Scientific & Technical: London, 1989. pp. 1261–1265
41. Anonymous, Médicaments à base de paraacétylaminophénoxy-1 trichloro-2 éthanol. Fr. Pat. M 2031 (May 24, 1962; Brevets Pharmaceutiques et Cosmétologiques S.A., Switzerland); C.A. 60, 9 201a. 1963.
42. Hems, B. A., Atkinson, R. M., Early, M., Tomich, E. G. Trichloroethyl phosphate. *Brit. Med. J.* 1962, *1*, 1834–1835.
43. Dittert, L. W., Caldwell, H. C., Ellison, T., Irwin, G. M., Rivard, D. E., Swintosky, J. V. Carbonate ester prodrugs of salicylic acid. *J. Pharm. Sci.* 1968, *57*, 828–831.
44. Anonymous, Glyceryl-N-(Substituted Phenyl) Anthranilates in the Treatment of Inflammation. *US Pat. 3,767,811 (Oct. 23, 1973; Schering Corporation)* 1973.
45. Allais, A., Rousseau, G., Girault, P., Mathieu, J., Peterfalvi, M., Branceni, D., Azadian-Boulanger, G., Chifflot, L., Jequier, R. Sur l'activité analgésique et antiinflammatoire des 4-(2'-alcoxycarbonyl phénylamino) quinoléïnes. *Eur. J. Med. Chem.* 1966, 65–70.
46. Boltze, K.-H., Brendler, O., Dell, H.-D., Jacobi, H. Novel Substituted Indole Compound, Process for the Preparation and Therapeutic Compositions Containing it. *Ger. Pat. 3,910,952 (Oct. 7, 1975, Troponwerke Dinklage & Co., Cologne)* 1975.
47. Boltze, K.-H., Brendler, O., Jacobi, H., Opitz, W., Raddatz, S., Seidel, P.-R., Vollbredht, D. Chemische Struktur und antiphlogistische Wirkung in der Reihe der substituierten Indol-3-essigsäuren. *Arzneim.-Forsch./Drug Res.* 1980, *30*(8a), 1314–1325.
48. Anonymous, Dérivés d'indolylacétylamino-acide, leur procédé de préparation et les compositions pharmaceutiques comprenant ces composés. *Belg. Pat.826,711 (March 14, 1975; Schering Aktiengesellschaft, Berlin und Bergkamen)* 1975.
49. Sinkula, A. A., Yalkowsky, S. H. Rationale for design of biologically reversible drug derivatives: prodrugs. *J. Pharm. Sci.* 1975, *64*, 181–210.
50. Berger, F. M. The anticonvulsant activity of carbamate esters of certain 2,2-disubstituted-1,3-propanediols. *J. Pharmacol. Exp. Ther.* 1952, *104*, 229–233.
51. Edmondson, H. T. Parenteral and oral clindamycin therapy in surgical infections: a preliminary report. *Ann. Surg.* 1973, *168*, 637–642.
52. Gray, J. E., Weaver, R. N., Moran, J., Feenstra, E. S. The parenteral toxicity of clindamycin 2-phosphate in laboratory animals. *Toxicol. Appl. Pharmacol.* 1974, *27*, 308–321.
53. Brodasky, T. F., Lewis, C. An *In Vivo–In Vitro* Comparison of the 2- and 3-phosphate esters of clindamycin. *J. Antibiot.* 1972, *25*, 230–238.
54. DeHaan, R. M., Metzler, C. M., Schellenberg, D., Van Denbosch, W. D. Pharmacokinetic studies of clindamycin phosphate. *J. Clin. Pharmacol.* 1973, *13*, 190–209.
55. Stella, V. J., Charman, W. N. A., Naringrekar, V. H. Prodrugs do they have advantages in clinical practice? *Drugs* 1985, *29*, 455–473.
56. Varia, A. A., Stella, V. J. Phenytoin prodrugs. VI: *In Vivo* evaluation of a phosphate ester prodrug of phenytoin after parenteral administration to rats. *J. Pharm. Sci.* 1984, *73*, 1087–1090.
57. Adjei, A. L., Doyle, R., Reiland, T. Flavors and flavor modifiers. In *Encyclopedia of Pharmaceutical Technology* (Swarbrick, J., Boylan, J. C., Eds), Vol. 6. Marcel Dekker, Inc.: New York, 1988, pp. 101–139.
58. Wenz, A., Göttmann, G., Koop, H. Verfahren zur Herstellung der freien Base und der Salze des Aneurin-orthophosphor-säureesters. *Ger. Pat. 1 110 646 (July 13, 1961; E. Merck A.G. Darmstadt)* 1961.
59. Davies, G. E., Driver, G. W. The antituberculous activity of ethyl thiolesters with particular reference to diethyl dithiol*iso*phtalate. *Br. J. Pharmacol.* 1957, *12*, 434–437.
60. Davies, G. E., Driver, G. W. Inhibitory action of ethyl mercaptan on intracellular tubercle bacilli. *Nature (London)* 1958, *182*, 664–665.
61. Caldwell, H. C., Adams, H. J., Rivard, D. E., Swintosky, J. V. Trichloroethyl carbonate I. Synthesis, physical properties, and pharmacology. *J. Pharm. Sci.* 1967, *56*, 920–921.
62. Gruber, J. C. M., Stephens, V. C., Terrill, P. M. Propoxyphene napsylate: chemistry and experimental design. *Toxicol. Appl. Pharmacol.* 1971, *19*, 423–426.
63. Rocca, G., Rusconi, L. Procédé pour la préparation de triméthoxybenzoate de tétracycline et les médicaments qui en dérivent. *French Pat. 2,099,449 (March 17, 1972, to Officina Therapeutica Italiana, SRL)* 1972.
64. Glazko, A. J., Edgerton, W. H., Dill, W. A., Lenz, W. R. Chloromycetin palmitate: A synthetic ester of chloromycetin. *Antiobiot. Chemotherap.* 1952, *2*, 234–242.
65. McEvoy, J. P. *American Hospital Formulary Service*, Vol. 8. American Society of Hospital Pharmacists: Washington, DC, 1974. p. 24
66. Murphy, H. W. Esters of erythromycin. *Antibiot. Annu.* 1954, 500–521.
67. Sinkula, A. A., Morozowich, W., Rowe, E. L. Chemical modifications of lincomycin: synthesis and bioactivity of selected 2,7-dialkylcarbonate esters. *J. Pharm. Sci.* 1973, *62*, 1106–1111.
68. Morozowich, W., Sinkula, A. A., Karnes, H. A., MacKellar, F. A., Lewis, C., Stern, K. F., Rowe, E. L. Synthesis and bioactivity of lincomycin-2-phosphate. *J. Pharm. Sci.* 1969, *58*, 1485–1489.
69. Morozowich, W., Sinkula, A. A., MacKellar, F. A., Lewis, C. J. Synthesis and bioactivity of lincomycin-2-monoesters. *J. Pharm. Sci.* 1973, *62*, 1102–1105.
70. Owens, W. H. Tetrazoles as carboxylic acid surrogates in the suosan sweetener series. *J. Pharm. Sci.* 1990, *79*, 826–828.
71. Muller, G. W., Culberson, J. C., Roy, G., Ziegler, J., Walters, D. E., Kellogg, M. S., Schiffman, S. S., Warwick, Z. S. Carboxylic acid replacement structure–activity relationships in suosan type sweeteners. A sweet taste antagonist. 1. *J. Med. Chem.* 1992, *35*, 1747–1751.
72. Tinti, J.-M., Nofre, C. Synthetic sweeteners. *French Demande FR 2,533,210, Chem. Abstr.* 1984, *101*. 152354k
73. Tinti, J.-M., Nofre, C. Design of sweeteners. A rational approach. In *Sweeteners: Discovery, Molecular Design, and Chemo-reception* (Walters, D. E., Orthoefer, F. T., DuBois, G. E., Eds). American Chemical Society: Washington, DC, 1991, pp. 88–99.
74. Nofre, C., Tinti, J.-M., Chatzopoulos-Ouar, F. Preparation of (phenylguanidino)- and [[(1-phenylamino)ethyl]amino]acetic acids as sweeteners. *Eur. Pat. Appl. EP 241,395 Chem. Abtr.* 1988, *109*. 190047k 1987.
75. Nofre, C., Tinti, J.-M. *N*-substituted derivatives of aspartame useful as sweetening agents. *U.S. Pat. 5,480,668; Chem. Abstr*. 1996, *112*, 235846. 1996.
76. Bassoli, A., Drew, G. B. D., Merlini, L., Morini, G. General pseudoreceptor model for sweet compounds: a semiquantitative prediction of binding affinity for sweet-tasting molecules. *J. Med. Chem.* 2002, *45*, 4402–4409.
77. Sinkula, A. A. Design of improved taste properties through structural modification. In *Design of Biopharmaceutical Properties through Prodrugs and Analogs* (Roche, E. B., Ed.). American Pharmaceutical Association – Academy of Pharmaceutical Sciences: Washington, DC, 1977, pp. 422–445.

Part VIII

Development of New Drugs: Legal and Economic Aspects

Bryan G. Reuben

Section Editor

Chapter 42

Discover a Drug Substance, Formulate and Develop It to a Product

Bruno Galli and Bernard Faller

I. INTRODUCTION
II. DISCOVER THE DRUG SUBSTANCE
A. Exploratory research (target finding)
B. Early discovery program (lead finding)
C. Mature discovery program (lead optimization)
D. Research–development interface
E. Learning experiences
III. DEFINING EXPERIMENTAL FORMULATIONS, THE CREATIVE PHASE
A. Basic thoughts on oral formulation
B. What is the purpose of a formulation?
C. Suggested sequence of activities prior to start formulation
D. Biopharmaceutical classification of compounds
E. How do we proceed at a practical level?
F. Which formulation principles are used?
IV. PHARMACEUTICAL DEVELOPMENT IN INDUSTRY
V. FIXING THE QUALITY AND DEVELOP THE PRODUCT IN A REGULATED ENVIRONMENT
REFERENCES

I. INTRODUCTION

Drug-development companies need to improve efficiency to sustain growth and profitability. The rate of technical development must be high and consistent. Formulating and adapting a future drug product as fresh information is provided by researchers in different disciplines is a continuous challenge. There is always something to improve. Because of this continuous iterative adaptation, formulation development can easily become the critical stage in the overall development process.

This chapter presents information on the discovery and development processes. It draws attention to some strategic technical points that are important if a project is to progress rapidly and without repetition. It indicates hidden pitfalls. It aims to provide background information regarding formulation development to enable researchers to ask the right questions. Links are provided to the essential regulatory documents of the European Regulatory Agency, the FDA and the ICH organization. The development of a standard oral dosage form is taken as an example.

II. DISCOVER THE DRUG SUBSTANCE

Knowledge of the properties of a drug substance increases during the development process. Usually lead compounds are obtained from high-throughput screening, the literature or competitors' compounds. Initial work focuses on potency and selectivity. Later, properties influencing drug disposition become important and need to be optimized in parallel to potency. Finally, some substance characteristics critical for developability are determined in early development or at an interface between research and development. It has become common practice in early development phase to investigate a package of compounds as opposed to a single substance. The relative merits of the different compounds are then studied.

Drug discovery can be divided into three steps:

1. exploratory research,
2. early discovery program,
3. mature discovery program.

Copyright © 2008, Elsevier Ltd
All rights reserved.

A. Exploratory research (target finding)

The aim of the exploratory phase is to identify new therapeutic targets or mechanisms of action. Historically, this phase was mainly driven by comparison of phenotypes and genotypes, elucidation of biochemical pathways and luck …. In the past, therapeutic breakthroughs often emerged from unexpected experimental results rather than rational drug discovery. Nowadays, molecular genetics and pathway screening are new technologies which are helping to identify new targets of interest. Having identified such a target, one needs to estimate how difficult would it be to interact with it to produce a commercial drug.

Drugability is not a word.

Many factors need to be considered:

- What is the amount of structural information available on the target?
- Does one need to seek an agonist or an antagonist?
- Where in the body is the target located (extracellular versus intracellular, free versus membrane bound)?
- Are predictive animal models available?
- Is the disease of interest acute or chronic?
- How clinically relevant might the target be?

Given a candidate drug, one must then estimate the market value of the new compound. Is it an improvement over an existing medicine? What are the potential advantages over the existing treatment? Is it likely to be a life-saving drug (i.e. cancer) or is it designed to improve quality of life (i.e. obesity)? What is the size of the patient population? What is the targeted drug delivery pathway?

B. Early discovery program (lead finding)

Once a target has been identified, one can start an early discovery program. In this period, one tries to validate the concept (or the target) in order to decide whether it is worth investing further resources.

This includes assay(s) development, synthesis of reference or competitor compounds and setting up the basic animal models needed in pharmacology. Validation of the target can often not be done with full predictive power, first because animals can behave differently from humans and second, the models are always a simplification of the real pathophysiological condition encountered in patients. Once the assay(s) are in place, one tries to identify lead compounds in the compound collection of the company. The size, quality and chemical diversity of the accessible compound collection obviously plays a major role. Common problems in high-throughput screen are the stability and purity of the solution archived and the compound solubility. Poor solubility can significantly alter the results as much lower concentrations than expected are truly in solution. Once hits are obtained from high-throughput screen, they need to be validated or confirmed. The assay is repeated with fresh stock solutions, which also undergo structure confirmation. In some occasions re-synthesis is necessary.

C. Mature discovery program (lead optimization)

Once the concept has been validated using reference compounds and suitable animal models, other activities can be initiated. The main goals of this phase of discovery are to optimize the lead compound(s) with respect to potency, selectivity and pharmacokinetic properties, to avoid unwanted side effects and to minimize toxicity. The difficulty is to optimize everything simultaneously because the objectives often clash. For example, for protease inhibitors, potency and permeability are often in conflict because of the chemical constraints imposed by the pharmacophore. Binding is often increased through addition of hydrogen bond donor or acceptor atoms; the same features usually hinder permeability and thus oral absorption. Similarly, there is a balance between binding and solubility. Binding is usually gained during lead optimization by increasing lipophilicity (grease is a cheap way to gain binding) which in turn decreases aqueous solubility; and hence drug disposition . In the past, one first optimized potency and selectivity. The pharmacokinetic and toxicity aspects – the ADME properties (administration, distribution, metabolism and elimination) – were addressed at a later stage. The drawback of this approach is that, when one eventually focuses on the ADME properties, it is usually too late to introduce changes in the molecule. Today one tries to optimize ADME properties and potency in parallel. This new strategy calls for the development of tools, which allow high-throughput profiling and have substance requirements compatible with drug discovery. It is usually helpful to set up a flow chart to indicate what assay is done when, and which operations are done sequentially and which in parallel. Because many drug candidates are sparingly soluble, it is often necessary to devise a dosage form even in the early phase of drug discovery. This preformulation is important to get good quality data on dose–response and exposure in animal models, particularly when the therapeutic dose is relatively high ($>\sim 5$ mg/kg). These so called PK/PD experiments (pharmacokinetics/pharmacodynamics) initiate the transition from the Research to Development stages. A reasonable, well-tolerated formulation principle in place at this stage allowing continuity in the subsequent stages significantly adds confidence to the results obtained. Preformulation ingredients should ideally not affect the animal model itself and should be tolerated with repeated and even chronic administration.

The correct time to file the patent is a subtle balance between too early (the patent protection clock already starts) and too late (competition). It is at this stage that one usually applies.

D. Research–development interface

Initially one synthesizes only a few milligrams of drug substance for *in vitro* assays. The initial batch is rarely crystalline as, nowadays, most early discovery compounds are purified by HPLC and not recrystallized. Initial solubility determinations therefore tend to be over-optimistic. Interesting compounds are then re-synthesized on a slightly larger scale (50–500 mg) for testing in animal models. Different salts can be evaluated and early polymorphism screening initiated, even with relatively small amounts of material (10 mg range). A few compounds will be selected for further evaluation and larger (1–100 g) quantities made. At that point, one starts to characterize the physicochemical properties of the drug substance in greater detail, a procedure that can lead to disappointments. For example, solubility may change even by an order of magnitude compared with earlier determinations done with non-crystalline, less pure batches. A drug substance can have many polymorphs and pseudo-polymorphs, which need to be characterized with respect to their solubility, stability, hygroscopicity, dissolution rate and other physicochemical properties. Some of the more obscure properties start to become important, like "adhesiveness," which can lead to substantial loss of material through physical adsorption on the walls of vials, tubes or pipettes. Particularly highly potent lipophilic compounds being subject to such phenomena may lead to, for example, erratic dosing. What does this mean?

The current bottleneck is that many of the drug substance properties important for developability are only obtained in early development, at a time where it is late to introduce chemical changes in the structure of the molecule.

All trends in industry go into the direction to frontload the investigations performed with fractions of milligram and to identify such characteristics as soon as possible and to take action as long as taking influence on the chemistry is still possible. Frontload is not a word.

E. Learning experiences

One first needs to stress that risk-taking is inherent to pharmaceutical discovery and therefore one often need to take conscious risks when one selects a development candidate. Although challenging compounds have been successfully developed (with massive investments in technical R&D) in the past (cyclosporine, tamoxifen, amiodarone …), the question is how many of these difficult compounds one can afford in a drug development pipeline? For one successful compound, how many did fail despite significant (or sometimes desperate) efforts? From our experience in the past few years a few points emerge:

1. Solubility alone is not necessarily meaningful: it must be considered with the intended dose, and the solubility/lipophilicity balance needs to be taken into account (low soluble, low lipophilic compounds are the hardest). In general one should try to investigate early what limits the solubility of the compound (lipophilicity, wettability, crystal packing, other factors) as it translates into different types of additional investigations and risks.
2. Extreme lipophilicity ($\log D$ (7.4) > 5) continues to be a cause of compound termination and disappointment for research teams because pharmacological activity observed in early animal models cannot be translated to larger animal species.
3. Unrealistic expectations for salt formation, in particular with weak bases ($pK_a < 5$).

III. DEFINING EXPERIMENTAL FORMULATIONS, THE CREATIVE PHASE

A. Basic thoughts on oral formulation

Various factors determine the bioavailability of a compound to what extent and for how long a drug substance is present in the blood circulation and in the actual target tissue.

"Pfizer's rule of five" (see Lipinsky 1997[1] and previous chapter) primarily allows positioning the compound with respect to potential problems of oral absorption.

On the other hand it might be of conceptual interest at this point also to know what substance properties the person encharged with the formulation would ideally like to obtain:

- a Log P of <3 and > 0;
- a permeability in the CaCo-2 model better than, for example, mannitol as a reference;
- a compound withstanding over years at room temperature to moisture, oxygen, and light;
- an anticipated dose for humans of <500 mg;
- a dose number of <20 (according to Amidon *et al.*[2]).

B. What is the purpose of a formulation?

A formulation is mainly intended for accurate dosing, preservation of the physical and chemical integrity of the compound. Additionally it aims to preserve the degree of dispersion of the molecule or of the crystals embedded in excipients herewith assuring the dissolution properties defined.

Figure 42.1 shows the main interrelated factors, which lead individual substances to the formulation and subsequently to a dosage form. At early stages the efforts will primarily consist in trying to understand and combine "Basic Factors" or the physicochemical properties such as to readily set up a useful, versatile and reliable formulation

FIGURE 42.1 Main interrelated factors surrounding an oral formulation.

principle. For an oral solid product, this translates into a maximization of solubility and with it possibly also of its dissolution rate.

The mentioned "Derived Factors" will predominantly have to be investigated, understood gradually and validated later, possibly in parallel or latest once fixing the quality of the drug substance (see also Figure 42.4).

Only to a minor extent the task of a formulation is to improve the absorption process which is normally and mainly driven by the passive diffusion of the compound. This is somehow in contrast with the cases where special heroic formulation efforts allowed repeatedly to rescue compounds and to compensate for their poor physicochemical properties. A good and integrated formulation approach should therefore always consider physicochemistry of the compound first relating it to *THE* key information of any formulation activity namely to the envisaged *DOSE*.

However, there are also biological factors such as metabolism in, for example, liver or the gastrointestinal tissues as well as active molecule transporters like P-gP or MRP in, e.g., the gastrointestinal tract which influence the bioavailability and can mislead the prediction of the dose. The influence of such active transports is particularly noted if the main driving forces of the absorption process, physicochemical parameters like solubility, dissolution rate, and permeability are low.

The anticipated dose drives also the selection of the route of administration just because of limitations with respect to mass or volume. Table 42.1 provides a rough idea about the achievable doses related to the administration route.

TABLE 42.1 Approximate Dosages in Humans Depending on the Administration Route Chosen[6]

Administration route	"Realistic" dosages related to drug substance (humans)
Oral	0.25–1,000 mg
Passive transdermal	0.05–5 mg
Inhalation	0.01–5 mg
Nasal, buccal, and sublingual	0.05–2 mg
Injection	Mainly depending on volume that can be administered and on physicochemical properties

C. Suggested sequence of activities prior to start formulation

Prior to start multiple formulation efforts and work-intensive *in vivo* formulation optimization experiments, an assessment should be made on whether drug metabolism and transport are relevant pharmacokinetic driving factors for the substance to be formulated.

This is done stepwise using following procedures:

1. "*In silico*" estimation of the physicochemical properties like calculated pK_a/Log P and polar surface area (see Palm K. *et al.*[3]).

2. Experimental determination of physicochemical properties like pK_a, Log P, solubilities, chemical stability or even special amphiphilic properties (for the latter see Seelig *et al.*[4]).
3. Assessment of the ability of the compound to pass through membranes using *in vitro* models like PAMPA/CaCo-2/MDCK cell layers always together with a well-documented reference compound like, for example, mannitol (see Arthusson *et al.*[5]).
4. Assessment *in vitro* of the metabolic pattern, biological stability, protein binding and blood/plasma partitioning *in vitro* and on a need basis followed by a verification *in vivo*.

D. Biopharmaceutical classification of compounds

Provided that metabolism and active transport systems are not the dominating factors for the drug absorption process, there are – in general terms – two factors that fairly well characterize the oral absorption: the solubility in combination with the permeability through membranes. Consequently, four classes of substances have been defined by Amidon G. *et al.*:[2]

1. *High solubility – high permeability (e.g. Metoprolol, L-Dopa)*

Formulation principles chosen for Class I compounds are normally straightforward. Aqueous solutions and *in situ* suspensions cover the needs for an early formulation.

2. *Low solubility – high permeability (e.g. Carbamazepine, Nifedipine)*

Generations of pharmacists have struggled to overcome the difficulties related to this class of compounds. Scientific progress in the field of formulation delivered for the Class II compounds, many interesting formulation principles in the last two decades. Thus, handling of this class is now relatively well understood. The success is mainly due to improvements in controlling the solubility and dissolution rate by measures such as the selection of the appropriate salt, the use of the appropriate crystal form, micronization, controlling of the wettability, the addition of surfactants and the use of microemulsions or solid dispersions just to provide a couple of examples. These principles have become the basis for adequate formulations. Highly dosed compounds (low potency) of this class of substances may face formulation limits regarding feasibility and attractiveness due to the massive use of special or new excipients needed to counterbalance the unfavorable properties.

3. *High solubility – low permeability (e.g. Atenolol, Terbutaline)*

The bioavailability of Class III compounds can practically not be influenced by a formulation. Rarely, the bioavailability can be improved by increasing the dose or by modifying the permeability of the gastrointestinal membrane. Modification of the gastrointestinal membrane structure by penetration or absorption enhancers always bears a safety risk because of also improving the uptake of all kind of toxins from the intestinal lumen and potentially causing severe irritations. What may be acceptable for an acute therapy may not be suitable for chronic treatment.

4. *Low solubility – low permeability*

Real Class IV compounds are not likely to be formulated. For highly potent drugs there might be the opportunity to raise the dose, but in general one starts seeking for suitable backups and to switch to a different chemical class. Changing the route of administration might be viable (e.g. parenteral), but this has consequently a strong impact on the attractiveness of the product or limitations in dosing, but it might be helpful for life saving drugs.

E. How do we proceed at a practical level?

Solubility always has to be considered together with the envisaged dose (anticipated needed dose and reference volume of a stomach where it has to dissolve). This leads to define a

$$\text{Dose number} = \frac{M_o}{V_o} C_s$$

M_o = dose to be administered
V_o = 250 mL (by definition the reference stomach volume)
C_s = solubility

Dose numbers of >20 are by definition related to "low soluble and dissolution rate limited" substances (see Amidon *et al.*[2]). In such cases special formulation approaches are to be considered from the scratch in the formulation strategy.

A permeability assessment at an early stage can be done using, for example, the CaCo-2 cell model and a reference substance. Low permeable substances can be defined as to be less permeable than, for example, mannitol.

Low permeable compounds always create serious formulation problems, whether it is due to the physicochemistry or due to active transport systems. For both the CaCo-2 model provides useful information. However, the crucial question at this stage is to sort out in view of the evaluation of alternative routes whether the observed low bioavailability is due to metabolism or low permeability or due to both.

Thus, permeability is always complex and not straightforward to assess.

In summary this leads to (see also Lipper [6]):

- Poor biopharmaceutical properties may sometimes be corrected by formulation, but at the expense of time and resources.
- Poor solubility and stability may be amenable to being fixed by the formulation.
- Poor permeability is very difficult to correct by formulation.
- First-pass metabolism problems cannot be fixed by oral formulations.

F. Which formulation principles are used?

At early stages, stable solutions are preferred because of their defined physical state assumed as to be monomolecularly dispersed systems.

Simple micellar dispersed solutions are also directly usable by the experimenter.

Complex micellar systems like microemulsions or special solid systems (suspensions or solid dispersions) usually require a basic pharmaceutical development strategy and sufficient understanding of the technical and biological concept to ensure appropriate quality, safety, tolerability, and reproducibility of the formulation. Such sophisticated systems can usually be properly elaborated only at later development stages also because of lacking of sufficient drug substance at early stages!

In early preclinical drug testing, it is preferable to use right from the beginning excipients that can and will be used in later development phases. Whenever possible, the formulation principle and the excipients should be used consistently throughout the preclinical testing.

This applies to all relevant studies with experimental animals namely pharmacological, toxicological and pharmacokinetic studies.

Toxicology studies are special as the chosen high doses of poorly soluble compounds open a permanent conflict between the maximal volume which can be administered and the safety/tolerability of the excipients needed. The chances in such exercises consist of having to sort out early on serious biopharmaceutical problems and to get them resolved iteratively in various pharmacokinetic experiments with many "lessons learned" for the human formulation.

To switch without a rationale to and to use different formulation principles may significantly affect the consistency of *in vivo* results such as bioavailability or in the worst case even pharmacodynamics. Not to adhere to this simple rule has resulted many times in false interpretations of *in vivo* results and delays in drug development at later, more expensive stages (Figure 42.2).

With the complexity and the multitude of open biological and physicochemical questions at early stages, it becomes imperative to closely work and interact in a multidisciplinary structure with strong emphasis on communication and information.

Iteratively, information is on one hand gathered from stability tests based on exposing the formulation to different physical stresses (temperature, humidity, light over time) and on the other hand information is obtained sequentially out of *in vivo* tests. Of particular interest is a consistent response/exposure to the wide range of doses administered from the pharmacological up to the toxicological experiments.

In sum, the main difficulties consist in properly assessing and to distill as early as possible useful information out of these experiments and to integrate them into a formulation, dosage form or drug product concept. Boundless information flow and contribution of the formulator to the different running protocols is of particular importance as by tradition the units in different Departments may not share congruent objectives, budget and time schedules. In essence these efforts should lead to a good guess of the first human dose (even if only extrapolated from animals at this point), with

FIGURE 42.2 Iterative loops supporting the optimization of a formulation.

a sufficient safety margin, the right route of administration and a consistent formulation approach for the subsequent expensive clinical trials (Bullet point 1 in Figure 42.4).

IV. PHARMACEUTICAL DEVELOPMENT IN INDUSTRY

In the last decade free flow of information, pressure from Health Authorities and considerable efforts in science have driven Drug Development toward "evidence-based medicine", cost savings and speed. An excellent review on the state of the science in the field Drug Development and its different approaches together with critical assessments are provided by Lesko *et al*. An attempt of how the future of Pharma could look like has recently been published by PriceWaterhouseCoopers on one hand, but also from Health Authorities like the FDA (see Refs. [7,8,9]).

Figure 42.3 provides a general overview over the main drug development stages, an approximate overall average duration and the likelihood of success for a drug as a function of its development stage.

With "creativity at early stages" it is meant that beside setting up within this frame an innovative therapeutic concept there is also a lot of information generated allowing not only to feed the main product development strategy. The concentrated available information should at this stage also be used to proactively trigger ideas and concepts for the future life cycle management and a related Drug Product strategy. However, it is very important to proceed with a balanced risk/resource management always considering the chance of success at the relative stage. In this situation, it is mandatory to constantly monitor the competitive situation as well.

At a later stage, from clinical Phase II onwards it will be much more difficult to perform changes on the drug product concept and its quality for regulatory and logistic reasons. Narrowing specifications, the process parameters and fixing quality in view of the preparation for the submission and forthcoming inspections by the Health Authorities will be the main drivers of the process from here on. Normally at the beginning of the Phase IIb clinical studies, the market form of an oral should be defined to consistently support the big Phase III confirmatory clinical trials and to avoid time consuming, risky bridging studies.

The huge costs for the large clinical Phase III trials led many companies in the recent past to implement rigorous selection processes allowing to identify and select the most promising projects. An interesting review on the different current philosophies is provided by Shillingford *et al*.[10] For decision making and prioritization of the selected concepts and criteria, a target product profile needs to be figured out early on to align the commercial and clinical strategy. This triggers beneficial early interactions between marketing/commercial and R&D functions leading to implement criteria allowing a constant challenge of the project regarding its viability. From the cited author a table of go/no-go criteria is taken and slightly modified. It reflects the thought resulting from the collaboration at early stages of Research, Development and Marketing/Business functions. (Table 42.2)

The activities in the projects have become increasingly creative and a specialized discipline of Early Phase Development was created here and there. With the challenging goal to become faster and to better predict the chance of success Proof of Concept studies meant to measure surrogate markers, biomarkers, results from functional imaging, creative early clinical readouts and endpoints have been

FIGURE 42.3 Main stages of drug development.

TABLE 42.2 Criteria and Data Supporting Go or No-go Decision Making

Criterion	Go!	No-Go!
PD activity at tolerable doses	Reproducible and related to clinically relevant dose	Absent, variable, poorly dose related
PD duration	Allows dosing regime acceptable to patients	Requires inconvenient complex dosing regime
PK characteristics	Linear with dose and time, low inter/intra subject variability	Non-linear with dose and time, high inter/intra subject variability
PK/PD relationship	Well-defined, predictable	Inadequate, unpredictable
Safety, tolerability profile	Predictable, wide therapeutic range	Unpredictable, narrow therapeutic range, potential drug–drug interactions
Bioavailability	Acceptable, allowing therapeutic levels, predictable, low inter/intra subject variability, low food interaction	Unpredictable, therapeutic levels not reached, high inter/intra subject variability
Physicochemical properties	Allows adequate exposure in human studies in an appropriate formulation. Reasonable stability.	Adequate exposure in human studies in an appropriate formulation not reached. Instability
Commercial viability	Adequate market size and/or price profile to achieve return of investment	Market too small and/or low priced to ensure return of investment

Abbreviations: PK: *Pharmacokinetic*; PD: *Pharmacodynamic*

implemented. Such studies involve generally small number of volunteers or patients to support a process toward fast selection and decision making. However, there are risks as well associated with these fast investigations kept within a tight frame of resources. Prior to start investigations in man at preclinical stage these risks might be:

- a validation of clinical relevance of the chosen target is often not possible in the animal models;
- the metabolism in man may differ from the one observed in animals;
- the indication chosen may be quickly accessible, but commercially not attractive affecting the risk-benefit assessment and interests;
- the profile of the compound could suffer from a lack of an optimal formulation and thus show a too low bioavailability;
- the forecast regarding the market conditions, competitive situation, and projected return of investment assessment may be wrong. In the case of an innovative new approach, it may be poorly assessable and thus leading to a low interest.

Once first human data on efficacy are becoming available at clinical Phase I or II stage threats may come from e.g.:

- a lack of understanding of the relationship dose–pharmacokinetics–pharmacodynamics in the patient;
- the relevance of any biological markers used to clinical outcomes;
- the uncertainty of the therapeutic range;
- the uncertainty on clinical relevance of the effect;
- the relevance of unexpected findings;
- the relative return of investment forecasted based on other competing approaches, or from the profile of a more attractive backup, followup compound.

Bullet point 2 in Figure 42.4 shows the impact of such findings and related decisions: new data or a new interpretation thereof may tell that, for example, an expensive synthesis campaign 2 manufactured at risk has been futile or, it may immediately trigger the synthesis of campaign 3 of the drug substance and in parallel the development of the Market Form, respectively, of long-term TOX studies. It illustrates how starting activities in parallel increases the risk of a premature or wrong investment at a stage with still a relatively low chance of success (see also Figure 42.3). It goes without saying that the alternative to minimize the risk by starting all activities in sequence may on the other hand save resources, but at the price of becoming extremely slow. To find the right and dynamic balance between starting the investments at risk herewith speeding up activities and waiting in other cases has become very complex and is to be handled with a high degree of experience, mutual understanding and transparent information of the involved disciplines.

FIGURE 42.4 Scheme for a fast drug product development showing the interdependence of selected key activities.

V. FIXING THE QUALITY AND DEVELOP THE PRODUCT IN A REGULATED ENVIRONMENT

As already alluded commenting Figure 42.3, it is vital that after the positive outcome of the clinical Phase IIa or Proof of Concept studies and once the whole organization really wants to develop the product, the full development process starts as soon as possible. From now on the whole process is run to gradually fix the variables by narrowing the specs, fixing the analytical methods, defining the launch sites, the equipment, the production batch sizes, the final market product and to try to start generating the long-term stability data as soon as possible to avoid becoming time-critical (bullet point 6). In parallel, optimization of the profitability runs concomitantly with setting up and assuring the supply chain. It is very important to quickly come to a controlled, robust process in order to timely validate and to consistently meet the expectations relevant to the Health Authorities during the approval and inspection process regarding: Efficacy, Safety, Quality.

Bullet point 3 relates the start of fixing the quality of the drug substance, the main part of the synthesis, the definition of the launch site to the outcome of a mid-size drug substance campaign 4. The experience gained during the campaign 3 and maybe first hints of the human dose (Phase IIa or IIb results) allow already to anticipate the target production volume and based on these information decisions on the production site must be taken. The latter is particularly important as it is part of the filing information and it is inspected by Health Authorities prior to the approval.

The quality of the drug substance at this stage is also very important for the long-term and carcinogenicity TOX studies to be initiated. It would be detrimental to have to repeat at a later time-point extensive TOX studies because of e.g. new by-products, which were not present in the drug substance previously. This is particularly important for the carcinogenicity studies, which are usually time-critical for filing (Bullet point 5).

As a general rule one tries to become gradually "cleaner and more consistent/robust" with forthcoming batches. However, to become "cleaner" may have a strong impact on the properties relevant to the formulation principle chosen for drug product. It is therefore very important to involve both disciplines (Chemical and Pharmaceutical development) in this evolutionary decision-making process

of fixing the quality in both disciplines and mutually agree on the properties and constraints.

It is important that after campaign 4 when Chemistry and Pharmaceutical Development have upscaled and validated their processes that the parameters of the drug substance and the drug product remain fixed (Bullet point 4). Validation according to the definition means to provide documented evidence that the processes consistently work. In the logics of validation of processes and products, the analytical methods need to be validated and fixed first.

Last but not least, bullet point 6 shows the impact of the long-term stability tests of the drug product on the submission and filing date. Together with the carcinogenicity studies and the readout of the Phase III clinical studies their availability triggers the submission process and therefore, they are critical for the project. The outcome of these stability data are going to define the packaging, shelflife and storage conditions, and contribute significantly to the product attractiveness and costs.

Today regulatory information is readily and freely accessible. As the information is updated and consistent, it is highly recommended to periodically check the latest information provided and also to periodically dedicate a few minutes in reading published documents still being in the review process. They are accessible as well and should be a mandatory lecture for all those involved in drug development process.

The repositories of the essential regulatory documents of the European Regulatory Agency (EMEA), the US Food and Drug Administration (FDA) and of the tripartite International Committee of Harmonization (Europe, USA and Japan) (ICH) organization can be found with the following links:

EMEA: http://www.emea.eu.int/index/indexh1.htm
FDA: http://www.fda.gov/cder/guidance/index.htm
ICH: http://www.ich.org/cache/compo/276-254-1.html

Health Authorities like the FDA nowadays take innovation very seriously and intend to contribute in providing guidance to successful drug development. To illustrate this, the following quote from a recent Whitebook on Innovation/Stagnation is provided: "FDA intends to lead in the development of a national Critical Path Opportunities List … to bring concrete focus to the tasks that lie ahead."

REFERENCES

1. Lipinsky, C.A., Lombardo, F., Dominy, B. W., Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development setting. *Adv. Drug Deliv. Rev.* 1997, *23*, 3–25.
2. Amidon, G., Lennernas, H., Shah, V. P., Crison, J. R. Theoretical basis for a biopharmaceutical drug classification: correction of in vitro drug product dissolution and in vivo bioavailability. *Pharm. Res.* 1995, *12*(3), 413–420.
3. Palm, K. et al. Polar molecular surface properties predict intestinal absorption in humans. *Pharm. Res.* 1987, *14*, 568–571.
4. Seelig, A., Gottschlich, R., Devant, R. M. A method to determine the ability of drugs to diffuse through the blood–brain barrier. *Procl. Natl. Acad. Sci. USA* 1994, *91*, 68–72.
5. Arthusson, P. et al. Correlation between oral absorption in humans and apparent drug permeability coefficient in human intestinal epithelial (CaCo-2) cells. *Biophys. Res. Commun.* 1991, *157*, 880–885.
6. Lipper, R. *Modern Drug Discovery* 1999, *2*(1), 55–60.
7. Lesko, L. J., Rowland, M., Peck, C. C., Blaschke, T. F. Optimizing the science of drug development: opportunities for better candidate selection and accelerated evaluation in humans. *Pharm. Res.* 2000, *17*(11), 1335–1344.
8. PriceWaterhouseCoopers. Pharma 2020 the vision – which path will you take? 2007.
9. FDA Whitepaper. Innovation/Stagnation-Challenge and Opportunity on the Critical Path to New Medical Products. 2004.
10. Shillingford, C. A., Vose, C. W. Effective decision-making: progressing compounds through clinical development. *Drug Discov. Today* 2001, *6*(8).
11. Kola, I., Landis, J. Can the pharmaceutical industry reduce attrition rates? *Drug Discov.* 2004, *3*, 711–715.

Chapter 43

Drug Nomenclature

R. G. Balocco Mattavelli, J. C. Dong, S. Lasseur and S. Kopp

I. INTRODUCTION
II. TRADE NAMES AND NONPROPRIETARY NAMES
III. DRUG NOMENCLATURE
A. International nonproprietary names for pharmaceutical substances (INN)
B. Common names selected by the International Standards Organization (ISO)
IV. USE AND PROTECTION OF NONPROPRIETARY NAMES
A. Use of nonproprietary names
B. Protection of nonproprietary names
V. SUMMARY
REFERENCES
ANNEX

I. INTRODUCTION

The World Health Organization (WHO) has a constitutional mandate to "develop, establish, and promote international standards with respect to biological, pharmaceutical, and similar products." One way in which this mandate is discharged is through the program for International Nonproprietary Names (INNs) for pharmaceutical substances. This program was established in 1950 by resolution WHA3.11 of the World Health Assembly and became operational in 1953. At its heart is the concept that there should be a single name globally for a unique pharmaceutical active substance. A substance may have several designations: an invented name, a systematic chemical name, a laboratory code, a trivial name, a Chemical Abstracts number. Assignment of an INN allows one name to be used in literature and in different circumstances so that there is a common basis for identification.

II. TRADE NAMES AND NONPROPRIETARY NAMES

Usually, pharmaceutical preparations are also marketed by industry under proprietary names or trademarks ("TMs"). In many countries trade names, also called trademarks or brand names, are used when prescribing, dispensing, selling, promoting, or buying a medicine. Trade names are usually selected by the owner of the product and registered in national trademark or patent offices. They are private property and can be used only with the consent of the owner of the trademark.[1,2]

In most cases, brand names are chosen for a finished pharmaceutical product, that is, for one or various active drug substances in a defined dosage form and formulation. Therefore, pharmaceutical preparations containing the same active drug substance are frequently sold under different brand/trade names, not only in different countries, but even within the same country (see Figure 43.1).

In practice, this means that the number of trade names in one country is usually much higher than the number of active drug substances marketed and used. Nonproprietary names, also called generic or common names, are intended to be used as public property without restraint, that is, nobody should own any rights on their usage. These names are usually designated by national or international nomenclature commissions. Both trade names and nonproprietary names are normally published first in the form of proposals. Comments may be made and objections raised for a certain time period before final publication. Although, both nonproprietary names and trade names may appear similar in form to an outsider, there is, in fact, a big difference. Firstly, nonproprietary names are designations to identify

Copyright © 2008, Elsevier Ltd
All rights reserved.

FIGURE 43.1 Various trade names for one substance, example *paracetamol*.

FIGURE 43.2 Various common names for one substance, example *paracetamol*.

FIGURE 43.3 One international name for one substance, example *paracetamol.*

the active pharmaceutical drug substance rather than the final product. Secondly, the selection of a nonproprietary name follows established rules so that the name itself communicates to the medical and pharmaceutical health professional to which therapeutic or pharmacological group the active drug substance belongs.

III. DRUG NOMENCLATURE

A. INNs for pharmaceutical substances

1. *History*

During the 20th century, the rapid development of pharmaceutical chemistry has brought with it the need to identify large numbers of active drug substances by unique, universally available and accepted names. The systematic chemical name, codified by international bodies, including the International Union for Pure and Applied Chemistry (IUPAC) and International Union of Biochemistry (IUB) has the advantage of unambiguously defining a specific chemical substance, but it is often very long, difficult to memorize and practically incomprehensible for the non-chemist. Moreover, it gives no indication as to the therapeutic action of the substance.

In order to avoid citation of difficult chemical names, generic names came into being. However, in the beginning different names were independently assigned to the same substance in different countries. For example, not everybody would know that acetaminophen, *N*-(4-hydroxyphenyl) acetamide, 40-hydroxyacetanilide, *p*-acetamidophenol, *N*-acetyl-p-aminophenol, acetaminophen and paracetamol are the same substance (see Figure 43.2).

When WHO started the INN program, the experts had to coordinate the activities of existing national nomenclature programs, which were especially active in France, the Nordic countries, the United Kingdom, and the United States. As a result of these national activities, many substances already had different well-established national names. Members of the newly established international nonproprietary nomenclature program were faced with the difficulty of choosing a single name in these instances – *paracetamol* in the example given below (see Figure 43.3).

Since then, the activities of national nomenclature commissions have been coordinated in order to achieve international standardization in nomenclature under the auspices of WHO according to article 2a and 2u of its constitution:[3]

> In order to achieve its objective, the functions of the World Health Organization shall be: (a) to act as the directing and coordinating authority on international health work;… (u) to develop, establish and promote international standards with respect to food, biological, pharmaceutical and similar products … procedure (application, detailed steps).

Requests for recommended INNs are submitted on a form to the WHO INN program,[4] These requests are then submitted by the WHO Secretariat on behalf of the Director-General to the members of the WHO INN Expert Group. The following information has to be provided on the form:

- Name and address of manufacturer and/or originator, name of responsible person.
- Suggested nonproprietary name(s) (various proposals possible).
- Chemical name (following IUPAC rules) or description, *ad hoc* information on biological products.
- Molecular formula.
- Graphic formula.
- Stereochemical information.
- Therapeutic use and pharmacological mode of action.
- Code, trademark (known or contemplated).
- Date of commencement of clinical trials.
- Letter from the Chemical Abstract Service (CAS) with CAS Registry Number and CA Index Name.

Each name suggested by the originator is then examined to determine whether it complies with the rules and guiding principles for the selection of an INN. When all members of the WHO INN Expert Group designated to select nonproprietary names agree to a name, it is first published as a proposed INN. During a 4-month period, any person can forward comments or a formal objection, for example, on the grounds of similarity to existing trademarks. When there is no objection, the name is published a second time as a recommended INN and the Director-General of WHO gives notice to member states that the name has been selected by the WHO as a recommended international nonproprietary name.

2. *Selection process and selection criteria*

General rules were established at the beginning of the INN Program in order to allow health professionals to understand the rationale for a number of new names for pharmaceutical substances. At first some countries used shortened chemical names as generic names, but this system was found to be very limited, since many molecules contain similar elements and groups, such as phenol, chlorine, methyl, or benzyl-rings, in their chemical structures. In addition, in most cases, a name that indicates relationship to a group of pharmacologically similarly acting substances is more meaningful to users.

The following principles should in general be applied when selecting an INN. The name should: (1) be distinctive in sound and spelling, (2) not be too long, and (3) show relationship to substances with the same pharmacological action. In addition, the new name should not conflict with any existing common names or trademarks, and patients should not be confronted with nonproprietary names that are likely to have anatomical, physiological, or pathological connotations: for example, a name starting *cancer-* would not be acceptable (see Figure 43.4)

In principle, INNs are given only to the active base or the active acid. Names for different salts or esters of the same active substance should differ only in respect of the name of the inactive moiety of the molecule. For example, *oxacillin* and *ibufenac* are INNs and their salts are named *oxacillin sodium* and *ibufenac sodium*. The latter are also called modified INNs (INNM). (Note that before the existence of this rule, some INNs were published for salts; the term modified INN may, therefore, sometimes be used for a base or acid; for example, *levothyroxine sodium* was published as INN and *levothyroxine* may thus be referred to as INNM.)

To facilitate the transliteration and pronunciation of INNs for pharmaceutical substances certain letters, such as "h" and "k," should be avoided. Preference is given to "f" instead of "ph," "t" instead of "th," "e" instead of "ae" or "oe" and "i" instead of "y." The INN for *amphetamine* is, therefore spelt *amfetamine*.

When devising an INN it is important to be aware of possible language problems. Since the name is used worldwide, not only should certain letters be avoided, but experts need to be aware of unsuitable connotations in the major languages spoken in the world. A name may appear excellent for an English speaker, but unacceptable in another language. For example, the name *inglicretin* could remind a French speaker of the term *crétin anglais* (stupid Englishman) and might therefore not be the best choice for naming a pharmaceutical substance.

As INNs should show relationship to other substances of similar pharmacological action, common stems have been created. A large number of such common stems are in use, and new stems are created when necessary.[5] Some examples are given in Table 43.1.

Examples of INNs ending *-entan*, stem selected for designating endothelin receptor antagonists are: a*mbrisentan*, *atrasentan*, *avosentan*, *bosentan*, *clazosentan*, *darusentan*, *edonentan*, *enrasentan*, *fandosentan*, *feloprentan*, *nebentan*, *sitaxsentan*, *tezosentan*, and *zibotentan*.

When requesting selection of an INN, the manufacturer has often not yet finalized the precise indications for the therapeutic use of the compound. A name is usually requested during the development phase of a new compound, which

General principles for guidance in devising international nonproprietary names for Pharmaceutical substances

1. International nonproprietary names (INN) should be distinctive in sound and spelling. They should not be inconveniently long and should not be liable to confusion with names in common use.
2. The INN for a substance belonging to a group of pharmacologically related substances should, where appropriate, show this relationship. Names that are likely to convey to a patient an anatomical, physiological, pathological, or therapeutic suggestion should be avoided.

These primary principles are to be implemented by using the following secondary principles:

3. In devising the INN of the first substance in a new pharmacological group, consideration should be given to the possibility of devising suitable INN for related substances, belonging to the new group.
4. In devising INN for acids, one-word names are preferred; their salts should be named without modifying the acid name, e.g. "oxacillin" and "oxacillin sodium," "ibufenac" and "ibufenac sodium."
5. INN for substances which are used as salts should in general apply to the active base or the active acid. Names for different salts or esters of the same active substance should differ only in respect of the name of the inactive acid or the inactive base.
   For quatemary ammonium substances, the cation and anion should be named appropriately as separate components of a quatemary substance and not in the amine-salt style.
6. The use of an isolated letter or number should be avoided; hyphenated construction is also undesirable.
7. To facilitate the translation and pronunciation of INN, "f" should be used instead of "ph", "t" instead of "th", "e" instead of "ae" or "oe," and "i" instead of "y;" the use of the letters "h" and "k" should be avoided.
8. Provided that the names suggested are in accordance with these principles, names proposed by the person discovering or first developing and marketing a pharmaceutical preparation, or names already officially in use in any country, should receive preferential consideration.
9. Group relationship in INN (see Guiding Principle 2) should if possible be shown by using a common stem. The following list contains examples of stems for groups of substances, particularly for new groups [list see text]. There are many other stems in active use. Where a stem is shown without any hyphens it may be used anywhere in the name.

FIGURE 43.4 General principles for guidance in devising INNs for pharmaceutical substances.

means that the request is submitted to WHO during the clinical trials phase. A name is, however, needed as soon as an application for registration of a product is forwarded to the national authorities. This means that the naming process is close to all new scientific developments in the pharmaceutical field. External expertise is often needed for specific questions concerning new therapeutic groups and new types of products.

During the last few years the selection process has become more complex. New receptors and pharmacological actions are discovered more and more frequently. This means in many cases that new stems have to be created. However, there is sometimes a structural relationship to existing molecules and experts have to decide whether an existing stem may be used or whether a new one must be established. Fibrinogen receptor antagonists are an example. These substances act as platelet aggregation inhibitors for which the stem *-grel* existed for several years. The nomenclature experts have to decide whether the same stem should be used for the fibrinogen receptor antagonists or whether the group of new molecules is so important that a new stem needs to be established.

On the other hand, a new mode of action is sometimes discovered for an existing substance. If further substances are developed with a similar mode of action, the question arises whether a new stem is needed, which would mean modifying the "old" name for the first compound in the series. For example, *albifylline* and *pentoxifylline* are *N*-methylxanthine derivatives and the stem *-fylline* was therefore chosen for their names. These substances have been found to also suppress tumor necrosis factor-α.[6] The experts decided to retain the stem *-fylline* in this case, since the "new" action was nevertheless based on the typical xanthine-mediated inhibition of phosphodiesterase.

New approaches to naming pharmaceutical substances may be needed in the near future because of the increasing research using molecular design. "Simple" derivatives of known compounds are becoming more and more rare. Chemistry based on receptor structure and molecular design focuses more on synthesizing compounds to fit receptor binding sites. This will mean that nomenclature will have to move in the same direction. Chemical relationship will need to be looked at from a different standpoint, and the pharmacological activity might have to be considered in

TABLE 43.1 Common stems in INN formation

Stem	Pharmacotherapeutic group
-ac	Anti-inflammatory agents, ibufenac derivatives
-adol/-adol-	Analgesics
-ast	Antiasthmatic, antiallergic substances not acting primarily as antihistaminics
-astine	Antihistaminics
-azepam	Diazepam derivatives
bol	Steroids, anabolic
-cain-	Class I antiarrhythmics, procainamide and lidocaine derivatives
-caine	Local anaesthetics
cef-	Antibiotics, cefalosporanic acid derivatives
-cillin	Antibiotics, derivatives of 6-aminopenicillanic acid
-conazole	Systemic antifungal agents, miconazole derivatives
cort	Corticosteroids, except prednisolone derivatives
-coxib	Selective cyclo-oxygenase inhibitors
-entan	Endothelin receptor antagonists
Gab	Gabamimetic agents
gado-	Diagnostic agents, gadolinium derivatives
-gatran	Thrombin inhibitors, antithrombotic agents
gest	Steroids, progestogens
gli-	Antihyperglycaemics
io	Iodine-containing contrast media
-metacin	Anti-inflammatory substances, indometacin derivatives
-mycin	Antibiotics, produced by Streptomyces strains
-nidazole	Antiprotozoal substances, metronidazole derivatives
-olol	Beta-adrenoreceptor antagonists
-oxacin	Antibacterial agents, nalidixic acid derivatives
-platin	Antineoplastic agents, platinum derivatives
-poetin	Erythropoietin type blood factors
pril(at)	Angiotensin-converting-enzyme inhibitors
-profen	Anti-inflammatory substances, ibuprofen derivatives
prost	Prostaglandins
-relin	Pituitary hormone release-stimulating peptides
-sartan	Angiotensin II receptor antagonists, antihypertensive (non-peptidic)
-vaptan	Vasopressin receptor antagonists
vin-/-vin-	Vinca-type alkaloids

almost all cases as a basis for assigning a given substance to a group.

3. *Procedure*

The Annex reproduces the *Procedure for the Selection of Recommended INNs for Pharmaceutical Substances* as adopted by the WHO Executive Board in its resolution EB15.R7 and as amended in 2005 by resolution EB115.R4.

4. *Publication*

After acceptance of the selected name by the request originator, it is included in a list of proposed INN, which is published in the *WHO Drug Information*, a journal published by WHO.[7] For example:

alogliptinum	
alogliptin	2–6–3R)-3–aminopiperidin-1-yl]-3-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl}methyl)benzonitrile *antidiabetic*
alogliptina	2-({6-[(3R)-3-aminopipéridin-1-yl]-3-méthyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl}méthyl)benzonitrile *antidiabétique*
alogliptina	2-({6-[(3R)-3-aminopiperidin-1-il]-3-metil-2,4-dioxo-3,4-dihidropirimidin-1(2H)-il}metil)benzonitrilo *hipoglucemiante*
	$C_{18}H_{21}N_5O_2$ 850649-61-5

Two lists of proposed INNs are published per year. If no objection has been raised during a 4-month period following publication, the proposed name is published a second time as a recommended INN as shown in the following example:

alogliptinum	
alogliptin	2-({6-[(3R)-3-aminopiperidin-1-yl]-3-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl}methyl)benzonitrile
alogliptine	2-({6-[(3R)-3-aminopipéridin-1-yl]-3-méthyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl}méthyl)benzonitrile
alogliptina	2-({6-[(3R)-3-aminopiperidin-1-il]-3-metil-2,4-dioxo-3,4-dihidropirimidin-1(2H)-il}metil)benzonitrilo

$C_{18}H_{21}N_5O_2$

The statements indicating action and use are largely based on the information supplied by the manufacturer. This information is merely meant to provide an indication on the potential use of new substances at the time they are accorded proposed INN. Because of their provisional nature, these descriptors will neither be revised nor included in the recommended and cumulative lists of INN.

Lists of both proposed and recommended INNs are sent together with a circular letter to WHO Member States (at present 193), to national pharmacopoeia commissions and to other bodies designated by member states. In this letter, the WHO requests that member states should take such steps as are necessary to prevent the acquisition of proprietary rights in the published names, including prohibiting registration of these names as tradenames.

Up to now nearly 8,000 INNs have been selected. All names ever selected are published in a cumulative list of INNs, which is updated periodically.[8] The generic names are presented in alphabetical order by Latin name. Each entry includes:

- INNs in Latin, Arabic, Chinese, English, French, Russian, and Spanish.
- A reference to the INN list in which the name was originally proposed or recommended.
- Reference to substances that have been abandoned or never been marketed.
- Reference to national nonproprietary names.
- Reference to pharmacopoeial monographs or similar official references.
- Names issued by the International Organization for Standardization (ISO).
- The molecular formula.
- Its Chemical Abstracts Service (CAS) number.
- Its chemical structure (when available).

This cumulative list is published now only as a CD-ROM which contains the list in a PDF format and a searchable database (by INN name (all or part), list number, CAS, alternate names, Anatomical Therapeutic Chemical (ATC) Classification, etc …).

All INNs and related information are also now available on the internet as shown in Figure 43.5.

INN web site: http://www.who.int/medicines/services/inn/en/
GUIDANCE
  Guidelines on the Use of INNs for Pharmaceutical Substances
INFORMATION
  INN consultations and meetings
    Dates, deadlines and Executive Summaries
  Requesting an INN
    INN application form
  Selection process of INNs
    The procedure for the selection of recommended INNs for pharmaceutical substances in the six official languages
  Published INN Lists
    Proposed and Recommended Lists
  Related Publications
  Biosimilars
ONLINE ACCESS
  To access the INN database on Mednet: http://mednet.who.int/

FIGURE 43.5 Web-links to relevant information and documents on INN.

5. *Recent developments*

Biotechnology-derived medicine is one of the fastest growing sectors of the pharmaceutical market and has begun to play an increasingly important role in health care. With the scientific and technical developments currently taking place, many new biological products are being introduced for the prevention, diagnosis, or treatment of human disease. How to name these substances is therefore raising a novel challenge for the nomenclature committee.[9] New schemes and concepts need to be developed on a worldwide basis. One example is the scheme for naming monoclonal antibodies [5] as given below:

1. **General stem**
   -mab
2. **Substems for source of product**

-amab	rat origin
-emab	hamster origin
-imab	primate origin
-omab	mouse origin
-umab	human origin
-ximab	chimeric origin
-zumab	humanized origin

3. **Substems for disease or target group**

-ba(c)-	bacterial
-ci(r)-	cardiovascular
-fung-	fungal
-le(s)-	inflammatory lesions
-li(m)-	immunomodulator
-os	bone
-vi(r)	viral

Substems for tumors

-co(l)-	colon
-go(t)-	testis
-go(v)-	ovary

-ma(r)- mammary
-me(l)- melanoma
-pr(o)- prostate
-tu(m)- miscellaneous

Whenever there is a problem in pronunciation, the final letter of the substems for diseases or targets may be deleted, e.g. *-co(l)-*, *-vi(r)-*, *-li(m)-*, etc.

4. **Prefix**
   The prefix should be random, e.g. the only requirement is to contribute to a euphonious and distinctive name.
5. **Second word**
   If the product is radiolabeled or conjugated to another chemical, identification of this conjugate is accomplished by use of a separate, second word or acceptable chemical designation. For monoclonals conjugated to a toxin, the infix *-toxa-* can be inserted either into the first (main) name or included in the second word.
6. If the monoclonal antibody is used as a carrier for a radioisotope, the latter will be listed first in the INN.
   Examples of INNs are: *abagovomab, abciximab, altumomab, efungumab, nebacumab, sevirumab, technetium (^{99m}Tc) pintumomab, telimomab aritox, urtoxazumab*, etc.

In 2005, a nomenclature system for gene therapy products[5,10,11] was developed by WHO within the INN framework. This system is based on a two-word scheme. The first word describes the expression gene, and the second word the vector component (see below):

1. Word one is for gene component.
   (a) The prefix contributes to the distinctive name.
   (b) The infix identifies the gene using, when available, existing infixes for biological products or using similar infix as for the protein for which the gene codes.
   -ermin- growth factors
   -kin- interleukins
   -lim- immunomodulators
   -mul- multiple genes
   -tusu- tumor suppression
   (c) The suffix is -(a vowel)gene, e.g. -(*o*)*gene*.
2. Word two is for vector component.
   (a) The prefix contributes to the distinctive name.
   (b) For viral vector, the infix will be:
   -adeno- adenovirus
   -cana- canarypox virus
   -herpa- herpes virus
   -lenti- lentivirus
   -retro- other retrovirus
   -vari- vaccinia virus
   (c) The suffix will be *-repvec* for replicating viral vector, *-vec* for nonreplicating viral vector or *-plasmid* for plasmid vector.
3. In case of naked DNA, there is no need for a second word in the name.

Until now, five gene therapy substances have been published: *alferminogene tadenovec, beperminogene perplasmid, contusugene ladenovec, sitimagene ceradenovec,and velimogene aliplasmid*.

Since 2002, five meetings dealing with aspects of INNs for biologicals have been convened at WHO headquarters. The relevant documents for these meetings can be downloaded from the following web pages:

"International nomenclature on Gene Therapy Products": http://www.who.int/druginformation/(*WHODrugInformation* 2003, *19*(2), 103)

"WHO Informal Consultation on INN Policy for Biosimilar products": http://www.who.int/medicines/services/inn/inn_bio/en/index.html

The last meeting held in April 2007 was dedicated to an in-depth review of the current INN policy for naming and defining biologicals to ensure consistency of approach and to identify and address anomalies. Specific areas for review included INN policies on post-translational modifications of proteins, monoclonal antibodies, vaccines, gene, cell and tissue therapy products, blood products, biologicals derived through transgenic processes, and enzymes. Recommendations made in this meeting were finally adopted by the INN Expert Group[12]. One of these recommendation is that more information should be requested at the time of application for an INN. This has been addressed by adding a special section to the INN application form as shown below:

Additional instructions for applications related to biological and biotechnological substances

Information also needs to be provided:

1. *For a protein substance, the complete mature amino acid sequence in a format that can be copied for analysis (Word or in the text of an e-mail) and with spaces between groups of 10 characters, and the sequence length; the general information on disulfides and all post-translational modification.*

   For a glycoprotein/glycopeptides, the glycosylation pattern (site, the type of sugar, etc.).
   For a monoclonal antibody, the precursor nucleotide sequence with spaces between codons and translation and with numbers per line; CDR-IMGT; the origin of each chain; sites of disulfide-bridges; Ig class and subclass; name/structure of the antigen against which the monoclonal antibody is directed.
   For a recombinant DNA protein, expression system; comparison with the native sequence.

2. *For a substance comprising nucleic acid (e.g. a DNA vaccine, oligonucleotide, gene therapy product), the full nucleotide sequence of the nucleic acid with the pertinent regions (e.g. reading frame) delineated.*

As substances produced by biotechnology are more and more complex and challenging, the INN Expert Group had

requested the WHO INN Secretariat to prepare a working document intended to summarize and review the past and present INN situation in this field. This document presents an inventory of the policy decisions taken by the INN Expert Group during all these years of change, and of the names assigned to biological and biotechnological substances. Considering the potential for further developments in the field of biologicals, this review is intended to be a living document which will be regularly updated to include new policies, and future INNs assigned. This review[11] can be downloaded from the following web page: http://www.who.int/medicines/services/inn/INN_Biorev11-06.pdf.

B. Common names selected by the International Standards Organization

1. *Selection principles*

The International Organization for Standardization (ISO) has laid down principles for selecting common names for pesticides and other agrochemicals.[13] These principles are comparable to the guiding principles for selecting INNs and have a similar purpose: to provide short, distinctive, and easily pronounced names for substances whose full chemical names are too complex for convenient use. The names chosen should not be permitted to become privately owned trademarks. ISO names are also given for salts and complex esters, as well as mixtures of isomers. The work of the INN and ISO committees sometimes overlaps, especially in the field of veterinary medicine. The two committees collaborate to avoid different names used for the same compound.

2. *National nomenclature*

Since the INN program came into existence, WHO has coordinated the activities of national nomenclature commissions. Several INN experts are secretaries to national nomenclature commissions and in most cases the WHO secretariat also acts as corresponding member of these commissions. Differences between national and international nomenclature have become rare.

In most countries national nomenclature commissions are part of or closely linked to the national pharmacopoeia. Some countries no longer have a commission and publish the INNs directly as national names in their legal publications, for example, in Germany and the Nordic countries; the latter published previously Nordic Pharmacopoeia Names (NFN). In others, the national nomenclature commission adopts INNs in the language of the country as national names.

All European directives include INNs as "usual terminology" (point 3 of Article 4 (2) of Directive 65/65/EEC). The European Pharmacopoeia uses INNs in the main titles of monographs.

National nomenclature commissions select and publish the following national nonproprietary names (address of responsible authority given in brackets):

- BAN (British Approved Names) (The Secretary, British Pharmacopoeia Commission, Market Towers, 1 Nine Elms Lane, London SW8 5NQ, UK).
- DCF (Dénominations Communes Françaises) (Secretariat of the French Pharmacopoeia Commission at the Drug Agency, Direction des Laboratoires et des Contrôles, Unité Pharmacopée, 145–147, Boulevard Anatole France, 93200 Saint-Denis, France).
- JAN (Japanese Accepted Names) (Japanese Ministry of Health and Welfare, New Drugs Division, Pharmaceuticals Affairs Bureau, 1-2-2, Kasumigaseki, Chiyoda-ku, Tokyo 100, Japan).
- USAN (United States Adopted Names) (United States Adopted Names Council, American Medical Association, P.O. Box 10970, Chicago, Illinois 60610, USA).

IV. USE AND PROTECTION OF NONPROPRIETARY NAMES

The WHO INN Program has been actively providing nonproprietary names since 1953. During this period nearly 8,000 names have been published. New pharmaceutical substances are continually being developed and some 100–150 new names are published every year.

In order to avoid confusion, it is strongly recommended by the WHO that new drug substances are identified by codes rather than arbitrary names until nonproprietary name have been designated.

A. Use of nonproprietary names

Nonproprietary names are intended to be used in pharmacopoeias, labeling, advertising, drug regulation, and scientific literature, and as product names, for example, for generics.

Some countries have defined the minimum size of characters in which the generic name must be printed under the trademark labeling and advertising. In Canada, the USA, and Uruguay, the generic name must appear prominently in type at least half as large as that used for the proprietary or brand name. Certain countries, for example Mexico, have even gone as far as abolishing trademarks for the public sector.

B. Protection of nonproprietary names

The consequence of introducing INN common stems into trademarks, which seems to be increasingly popular, hampers the selection of new nonproprietary names within the established system. Given that all new INNs should be distinctive from existing INNs, without similarity to trademarks, this practice causes confusion to the health professional, may

be the source of serious errors in prescribing and dispensing and hinders the selection of future names for compounds in the same group of substances. Based on recommendations made by the WHO Expert Committee on the Use of Essential Drugs, a resolution[14,15] was adopted during the 46th World Health Assembly requesting Member States to:

> "*enact rules or regulations, as necessary, to ensure that international nonproprietary names … are always displayed prominently; to encourage manufacturers to rely on their corporate name and the international nonproprietary names, rather than on trademarks, to promote and market multisource products introduced after patent expiration; to develop policy guidelines on the use and protection of international nonproprietary names, and to discourage the use of names derived from INNs, and particularly names including INN stems in trademarks*"

V. SUMMARY

The existence of an international nomenclature for pharmaceutical substances, in the form of INNs, has proved since 1953 to be important for the safe prescription and dispensing of medicines to patients, and for communication and exchange of information among health professionals worldwide. INNs identify pharmaceutical substances by unique names that are globally recognized and are public property. They are also known as generic names.

Common stems are developed for the selection of INNs to communicate to the health professionals the type of pharmaceutical product in question. National and international nomenclature commissions collaborate closely to select a single name of worldwide acceptability for each active substance that is to be marketed as a pharmaceutical.

To avoid confusion, which could jeopardize the safety of patients, nonproprietary names and their common stems should not be used in trademarks. The selection of further names within a series should not be hindered by the use of a common stem in a brand name.

The authors alone are responsible for the views expressed in this article.

REFERENCES

1. Wehrli, A. Pharmaceuticals: trademarks versus generic names. *Trademark World* 1986, *16*, 31–35.
2. WHO. Trademarks versus generic names for pharmaceuticals. A conflict that requires resolution. *WHO Drug Information* 1987, *1*, 39–40.
3. WHO. *Basic Documents*, 39th Edition. World Health Organization: Geneva, 1992.
4. WHO. INN Application Form to be sent to: World Health Organization, INN Programme, 20 avenue Appia, 1211 Geneva 27, Switzerland 2007.
5. WHO. *The use of stems in the selection of INN for pharmaceutical substances*. World Health Organization: Geneva, 2006.
6. Semmler, J., Gebert, U., Eisenhut, T. *et al*. Xanthine derivatives: comparison between suppression of tumor necrosis factor-α production and inhibition of cAMP phosphodiesterase activity. *Immunology* 1993, *78*, 520–525.
7. Examples of published INN lists: List 96 of proposed INNs: *WHO Drug Information* 2006, *20*(4); List 58 of recommended INNs: *WHO Drug Information* 2007, *21*(3). World Health Organization, Geneva.
8. WHO. *Cumulative List No. 12 of International Nonproprietary Names (INN) for Pharmaceutical Substances*. World Health Organization: Geneva, Switzerland, 2007.
9. Wehrli, A. Generic names for biotechnology-derived products. *Drug News Perspect.* 1992, *5*, 55–58.
10. International nomenclature and gene therapy products.*WHO Drug Information* 2005,*19*(2) 103.
11. International Nonproprietary Names (INN) for Biological and Biotechnological Substances (a Review), INN working document 05.179, 2006.
12. 44th INN Consultation-Executive Summary on International Nonproprietary Names (INNs) for Pharmaceutical Substances, INN Working Document 07.214.
13. ISO Standard Pesticides and Other Agrochemicals. *Principles for the Selection of Common Names (ISO/DIS 257, 1988)*. International Organization for Standardization: Geneva, 1988.
14. Forty-sixth World Health Assembly Resolution on Nonproprietary Names for Pharmaceutical Substances (WHA46.19) (1993) World Health Organization, Geneva. (website)
15. International Nonproprietary Names, Information Leaflet for Trademark Departments (2004) (web site).

ANNEX

Procedure for the Selection of Recommended INNs for Pharmaceutical Substances

The following procedure shall be followed by the World Health Organization (hereinafter also referred to as "WHO") in the selection of recommended INNs for pharmaceutical substances, in accordance with resolution WHA3.11 of the World Health Assembly, and in the substitution of such names.

1. Proposals for recommended INNs and proposals for substitution of such names shall be submitted to WHO on the form provided therefore. The consideration of such proposals shall be subject to the payment of an administrative fee designed only to cover the corresponding costs of the Secretariat of WHO ("the Secretariat"). The amount of this fee shall be determined by the Secretariat and may, from time to time, be adjusted.
2. Such proposals shall be submitted by the Secretariat to the members of the Expert Advisory Panel on the International Pharmacopoeia and Pharmaceutical Preparations designated for this purpose, such designated

members hereinafter referred to as "the INN Expert Group," for consideration in accordance with the "General principles for guidance in devising International Nonproprietary Names for Pharmaceutical Substances," annexed to this procedure. The name used by the person discovering or first developing and marketing a pharmaceutical substance shall be accepted, unless there are compelling reasons to the contrary.

3. Subsequent to the examination provided for in article 2, the Secretariat shall give notice that a proposed international nonproprietary name is being considered.

   (a). Such notice shall be given by publication in *WHO Drug Information*[3] and by letter to Member States and to national and regional pharmacopoeia commissions or other bodies designated by Member States.
      (i) Notice shall also be sent to the person who submitted the proposal ("the original applicant") and other persons known to be concerned with a name under consideration.

   (b). Such notice shall:
      (i) set forth the name under consideration;
      (ii) identify the person who submitted the proposal for naming the substance, if so requested by such person;
      (iii) identify the substance for which a name is being considered;
      (iv) set forth the time within which comments and objections will be received and the person and place to whom they should be directed;
      (v) state the authority under which WHO is acting and refer to these rules of procedure.

   (c). In forwarding the notice, the Secretariat shall request that Member States take such steps as are necessary to prevent the acquisition of proprietary rights in the proposed name during the period it is under consideration by WHO.

4. Comments on the proposed name may be forwarded by any person to WHO within 4 months of the date of publication, under article 3, of the name in *WHO Drug Information*.

5. A formal objection to a proposed name may be filed by any interested person within 4 months of the date of publication, under article 3, of the name in *WHO Drug Information.*

   Such objection shall:
      (i) identify the person objecting;
      (ii) state his or her interest in the name;
      (iii) set forth the reasons for his or her objection to the name proposed.

6. Where there is a formal objection under article 5, WHO may either reconsider the proposed name or use its good offices to attempt to obtain withdrawal of the objection. Without prejudice to the consideration by WHO of a substitute name or names, a name shall not be selected by WHO as a recommended international nonproprietary name while there exists a formal objection thereto filed under article 5 which has not been withdrawn.

7. Where no objection has been filed under article 5, or all objections previously filed have been withdrawn, the Secretariat shall give notice in accordance with subsection (a) of article 3 that the name has been selected by WHO as a recommended international nonproprietary name.

8. In forwarding a recommended international nonproprietary name to Member States under article 7, the Secretariat shall:
   (a) request that it be recognized as the nonproprietary name for the substance;
   (b) request that Member States take such steps as are necessary to prevent the acquisition of proprietary rights in the name and to prohibit registration of the name as a trademark or trade name.

9. A. In the extraordinary circumstance that a previously recommended international nonproprietary name gives rise to errors in medication, prescription, or distribution, or a demonstrable risk thereof, because of similarity with another name in pharmaceutical and/or prescription practices, and it appears that such errors or potential errors cannot readily be resolved through other interventions than a possible substitution of a previously recommended international nonproprietary name, or in the event that a previously recommended international nonproprietary name differs substantially from the nonproprietary name approved in a significant number of Member States, or in other such extraordinary circumstances that justify a substitution of a recommended international nonproprietary name, proposals to that effect may be filed by any interested person. Such proposals shall be submitted on the form provided therefore and shall:
      (i) identify the person making the proposal;
      (ii) state his or her interest in the proposed substitution;
      (iii) set forth the reasons for the proposal;
      (iv) describe, and provide documentary evidence regarding the other interventions undertaken in an effort to resolve the situation, and the reasons why these other interventions were inadequate.

   Such proposals may include a proposal for a new substitute international nonproprietary name, devised in accordance with the general principles, which takes into account the pharmaceutical substance for which the new substitute international nonproprietary name is being proposed.

   The Secretariat shall forward a copy of the proposal, for consideration in accordance with the procedure described in subsection (b) below, to the INN Expert Group and the original applicant or its successor (if different from the person bringing the proposal for substitution and provided

that the original applicant or its successor is known or can be found through diligent effort, including contacts with industry associations).

In addition, the Secretariat shall request comments on the proposal from:

(i) Member States and national and regional pharmacopoeia commissions or other bodies designated by Member States (by including a notice to that effect in the letter referred to in article 3(a);

(ii) Any other persons known to be concerned by the proposed substitution.

The request for comments shall:

(i) state the recommended international nonproprietary name that is being proposed for substitution (and the proposed substitute name, if provided);

(ii) identify the person who submitted the proposal for substitution (if so requested by such person);

(iii) identify the substance to which the proposed substitution relates and reasons put forward for substitution;

(iv) set forth the time within which comments will be received and the person and place to whom they should be directed;

(v) state the authority under which WHO is acting and refer to these rules of procedure.

Comments on the proposed substitution may be forwarded by any person to WHO within 4 months of the date of the request for comments.

B. After the time period for comments referred to above has elapsed, the Secretariat shall forward any comments received to the INN Expert Group, the original applicant or its successor and the person bringing the proposal for substitution. If, after consideration of the proposal for substitution and the comments received, the INN Expert Group, the person bringing the proposal for substitution and the original applicant or its successor all agree that there is a need to substitute the previously recommended international nonproprietary name, the Secretariat shall submit the proposal for substitution to the INN Expert Group for further processing.

Notwithstanding the foregoing, the original applicant or its successor shall not be entitled to withhold agreement to a proposal for substitution in the event the original applicant or its successor has no demonstrable continuing interest in the recommended international nonproprietary name proposed for substitution.

In the event that a proposal for substitution shall be submitted to the INN Expert Group for further processing, the INN Expert Group will select a new international nonproprietary name in accordance with the General principles referred to in article 2 and the procedure set forth in articles 3–8 inclusive. The notices to be given by the Secretariat under article 3 and article 7, respectively, including to the original applicant or its successor (if not the same as the person proposing the substitution, and provided that the original applicant or its successor is known or can be found through diligent effort, including contacts with industry associations), shall in such event indicate that the new name is a substitute for a previously recommended international nonproprietary name and that Member States may wish to make transitional arrangements in order to accommodate existing products that use the previously recommended international nonproprietary name on their label in accordance with national legislation.

If, after consideration of the proposal for substitution and the comments received in accordance with the procedure described above, the INN Expert Group, the original applicant or its successor, and the person bringing the proposal for substitution do not agree that there are compelling reasons for substitution of a previously recommended international nonproprietary name, this name shall be retained (provided always that the original applicant or its successor shall not be entitled to withhold agreement to a proposal for substitution in the event that the original applicant or its successor has no demonstrable continuing interest in the recommended international nonproprietary name proposed to be substituted). In such an event, the Secretariat shall advise the person having proposed the substitution, as well as the original applicant or its successor (if not the same as the person proposing the substitution, and provided that the original applicant or its successor is known or can be found through diligent effort, including contacts with industry associations), Member States, national, and regional pharmacopoeia commissions, other bodies designated by Member States, and any other persons known to be concerned by the proposed substitution that, despite a proposal for substitution, it has been decided to retain the previously recommended international nonproprietary name (with a description of the reason(s) why the proposal for substitution was not considered sufficiently compelling).

Chapter 44

Legal Aspects of Product Protection

What a Medicinal Chemist Should Know About Patent Protection

Maria Souleau

I. INTRODUCTION
A. History of the patent system prior to 1883
B. Main conventions and treaties
II. DEFINITION OF A PATENT – PATENT RIGHTS
III. KIND OF INVENTIONS
IV. SUBJECTS OF PATENTS: BASIC AND FORMAL REQUIREMENTS FOR FILING A PATENT
A. Basic requirements
B. Formal requirements
V. LIFETIME OF PATENTS
VI. OWNERSHIP OF PATENTS
VII. INFRINGEMENT OF A PATENT
VIII. PATENTS AS A SOURCE OF INFORMATION
IX. PATENTING IN THE PHARMACEUTICAL INDUSTRIES
X. CONCLUSION
REFERENCES

I. INTRODUCTION

A patent is a grant to an inventor of the exclusive right to use and profit from his invention for a limited term. In exchange, the inventor discloses the invention in such a way that an expert could follow it without doing research. Patents play an important role in the progress of national economies and in progress in general.

During the last two centuries, the USA and some western European countries have had a phenomenal industrial and economic development. In this process, research and innovation played a predominant role and patents were instrumental, inasmuch as inventions were no longer kept secret, but were made known to the public, thus inducing new research and innovation. The post-war development of Japan's industrial potential is an example which demonstrates the role and the importance of the patent system.

A. History of the patent system prior to 1883

Originally patents were intended to encourage the introduction of new technologies into a country so as to develop native industry, rather than to provide a just reward to the inventor. In order to encourage and reward the publication of new knowledge, privileges were granted giving the owner of the technology the exclusive right to exploit that knowledge in the relevant territory for a limited period. Until the 15th century, the various states were scarcely concerned with furthering the development of industry, and generally speaking manufacturing secrecy remained the rule. The statute for inventors decreed by the Council of the Venetian Republic on 19 March 1474 can be considered as the first law on patents.[1] This statute for inventors stipulated that:

> "*whoever will make in this City [Republic of Venice] any new and ingenious artifice, not made previously in our state, will be obliged to register it at the office of our proveditors of the Commune, as soon as it will be reduced to perfection, so that it will be possible to use and apply it. It shall be forbidden to anyone else in any our land and place to make any other artifice to the image and similarity of that one without consent and license of the author during the term of ten years.*"

Studies of the detailed records of early grants in other countries suggest that the idea of a patent system spread over Europe from Italy with emigrating Venetian glassworkers.

In England, as from 1315, the grant of monopolies by the Crown had been widely used as a means of promoting

Copyright © 2008, Elsevier Ltd
All rights reserved.

and regulating trade and as a source of revenue. As a means of attracting new industries from overseas, letters of protection were also granted to groups of foreign workers. From the beginning, an essential condition of the validity of English patents for invention rights was that the invention should actually be worked in the country. The practice of granting patent monopolies became firmly established in England in the reign of Elizabeth-I. Between 1561 and 1590 more than 50 monopolies were granted to encourage English industry and to make the country self-sufficient. Some of these patents were patents of invention for new industries, but the Queen's need for money also led her to grant other monopolies giving the right to control established industries. Such patents were clearly against the public interest and provoked vigorous protest in the House of Commons. So, in 1601, the Queen issued a proclamation revoking the objectionable patents and enforced the monopoly rights grants for invention. In 1623, under the reign of James I, the "Statute of Monopolies" prescribed the cases for which there could be a monopoly. According to this statute an invention had to be novel and non-detrimental to the general interest. This law established also a time limit for the "duration of the monopoly". After the passage of the Statute of Monopolies, there was no further legislation in England on patents of invention for over 200 years.

The Statute of Monopolies was followed by similar legislation in the American colonies, where settlers began to petition the legislature or governor for monopoly patents similar to those they had been familiar with in England. In 1641 the General Court of Massachusetts adopted a "Body of Liberties" prohibiting the granting of monopolies except for "such new inventions as are profitable to the country and that for a short time". After the American War of Independence, the old concept of a patent granted as an act of grace and favor by the Crown began to be replaced by the idea that an inventor had a natural right to his property. In South Carolina, an act of 1784 provided for 14 years monopoly rights to inventors and most of the other new states continued and extended the colonial practice of rewarding inventors with the grant of patent rights. When the delegates of the states met to draft the Constitution of the United States of America, it was recognized that it would be more effective to replace the separate state patents by a single form of protection for the whole country. In 1790 Thomas Jefferson drafted the first American law of patents. This law stipulated a monopoly to inventors of 17 years from the grant for novel inventions useful to the country.

In France, the early practice of the royal grant of patents (from 1543) for inventions was confirmed by an edict of Louis XV in 1762 which prohibited permanent privileges, but provided for inventors patents to be limited to 15 years. In 1789, with the revolution, the privileges of inventors were abolished. Nevertheless on 7th January 1791, the Constitutional Assembly passed the first law on patents. In the words of the preamble of the new French law:

> *"every novel idea whose realisation or development can become useful to Society belongs primarily to the man who conceived it, and it would be a violation of the rights of man in the very essence if an industrial invention were not regarded as the property of its creator."*

According to this French law for a disclosure of the invention, the state granted to the author a "manufacturing monopoly" for a limited period. Under the law of 1791, French patents were granted for 5, 10 or 15 years. Patents became null and void if not worked within 2 years, or if the patentee sought to take out a patent abroad. Examination of applications was still required, but in 1844 this was abolished. According to the 1844 law, the inventor was not obliged to delimit his monopoly by claims, but the inventor was obliged to describe his invention to the best of his ability.[2]

In the 17th and 18th centuries, with the development of commerce and the beginnings of industry, a certain number of states thus adopted patent laws which gave their inventors protection against the infringement of their inventions. In Europe subsequent developments in the 19th century were greatly influenced by competing economic interest. Industrialists who recognized the value of patents in protecting their businesses increasingly pressed for the patent system to be strengthened and expanded. A uniform patent law for the whole of Germany was adopted in 1877 with strict examination for "novelty" and "level of invention."

Before enumerating the main conventions and treaties still in force, it must be noted that almost every country has now adopted a patent law.

B. Main conventions and treaties

In the middle of the 19th century, it was already apparent that industrial property rights should extend beyond national boundaries.

In 1883, in order to facilitate the transfer of technology, a large number of states created the *Paris Convention*, an international convention. The Paris Convention, established major basic rules concerning industrial property and was a remarkable piece of legislation for its time. The Convention was revised in subsequent international conferences in Brussels in 1900, Washington in 1911, The Hague in 1925, London in 1934, Lisbon in 1958, and Stockholm in 1967. The main provisions which apply between the member countries are now:[3]

1. Foreigners receive in each country the same treatment as nationals of that country.
2. An applicant for a patent in one country receives the benefit of his original filing date for any subsequent application for that invention, filed in another country within 1 year (this is the priority clause – *Priority Right*).

3. Patents in different countries for the same invention are independent, so that they are not affected by refusal, revocation, or expiry in any other country.
4. Importation by the patentee of goods produced in one country is not to entail forfeiture of patent protection for those goods.
5. Each country may take steps to prevent abuses resulting from the exclusive rights arising from patents, for example, failure to use the invention, but the patent may be revoked only if compulsory licensing is not a sufficient remedy.

Concerning the Priority Right, no extension of the 12 month period can be allowed. After 12 months, the Priority Right is lost.

1. *European Patent Convention and the Patent Cooperation Treaty*

In 1978, two treaties came into force which had a great impact on the world patent system unequalled since the Paris Convention of 1883.

These two treaties were the Convention of the Grant of European Patents, named the European Patent Convention (EPC), signed in Munich in 1973, set up by the Council of Europe, and the Patent Cooperation Treaty (PCT), signed in Washington on 1970, established by the World Intellectual Property Organization. Furthermore, the Council of the European Economic Community has, for some time, been endeavoring to bring into force a convention for the European Patent for the Common Market (the Community Patent Convention) for which the basic text was signed in Luxembourg in December 1975 by the EEC countries at that time. This has been supplemented by a Community Patent Agreement, initiated by the 12 EEC countries in December 1985.

The EPC is in addition to the national patent laws which exist in each EPC-contracting state, even though the contracting states had to modify their laws in order to harmonize them with the EPC.

The EPC and the PCT each establishes a system for filing a single patent application to obtain, ultimately, national patents in countries designated in the application.

As concerns the EPC, the possibility of filing a European patent application is open to countries both inside and outside the EPC member states. According to the EPC, it is possible to obtain national patents in the contracting states by means of a direct single filing at the European Patent Office (EPO) or an indirect single filing through certain EPC-member states national patent offices, if the law of the contracting state permits, it is possible to obtain national patents in the contracting states.

In January 2008, the contracting states were: Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Monaco, Netherlands, Norway, Poland, Portugal, Romania, Spain, Slovakia, Slovenia, Sweden, Switzerland and Liechtenstein, Turkey, and the United Kingdom.

The European Patent Organization has also concluded agreements on cooperation in the field of patents and on extending the protection conferred by European patents (Extension Agreements) with a number of states which are not party to the EPC. These agreements form the basis of an extension system providing European patent applicants with a simple and cost-effective way of obtaining patent protection in these countries. At the applicant's request patents can be extended to these countries, where they will have the same effects as national applications, and patents will enjoy substantially the same protection as patents granted by the EPO for the member states of the European Patent Organization. At present, extension to the following states may be requested: Albania, Bosnia, and Herzegovina, the former Yugoslav Republic of Macedonia and Serbia.

According to the EPC, a European patent application is submitted to a single examination system. An applicant files a patent application in one of the official languages of the EPC, namely German, English, and French. This application is examined at the EPO for formalities and then for prior art. The invention must not have been disclosed by prior publication or in prior patents. It must not have been in public use or on sale. The patent application and the search report are then published.

After publication of the application and the search report, the applicant has to request substantive examination for patentability to be performed at the EPO. The examination of the application is performed by the Examining Division of the EPO. On completion of the examination, which includes an examination for inventive step as well as for novelty (novelty and inventive step are basic requirements for patentability; see Section IV. A.), either the application is rejected or the European patent is granted. The specification as granted is published and the grant of the European patent leads to national patents being registered in the designated EPC countries.

Opposition can be filed at the EPO after the European patent has been granted. The results of the opposition, i.e. amendment or revocation of the European patent, apply equally to all the national patents resulting from the European patent. Furthermore, each national patent can be amended or revoked by national courts under the relevant national law.

The EPO has set out to establish a practice which is a compromise between various traditions and practices (German, British, etc.). The officials in charge of the development procedure have attempted to avoid the worst features of the existing national systems. The EPO regards itself as having the function of granting patents rather that refusing them, and with the proviso that the claimed invention is

patentable, an application will be allowed without prolonged arguments and numerous official actions.[4,5]

The PCT is open to all countries of the world which have acceded to the Convention of Paris. A majority of the world's countries are signatories to the PCT, including all of the major industrialized countries (with a few exceptions, including Argentina and Taiwan). As of February 2008, there were 138 Contracting States to the PCT. The PCT entered into force on 27 December 2007 for Angola, the 138th Contracting State.

The procedures of the PCT are under the control of World Intellectual Property Organization (WIPO, Geneva).

Under the provisions of the PCT an applicant files a single patent application named an "international application" at the national patent office in any country which has acceded to the PCT and in a language acceptable by that country. The national office at which the international application is filed is known as the "receiving office." The application must designate the countries for which national patents are required and must request a search to be performed, named as an "international search" by a competent Patent Office. It is also possible to designate a European application, the so-called "regional designation."

The international application is examined with the required formalities. A prior art search is performed by an International Searching Authority (ISA), and a search report is sent to the applicant. The application and the international search report are published. This is known as the "international phase" of the procedure (Chapter I of the PCT).

Articles 31–42 of PCT provide for an "international preliminary examination" to be performed during the international phase (Chapter II of the PCT). Countries having acceded to the PCT can make a declaration that they will not be bound by Chapter II of the PCT. An applicant resident in a country which is bound by Chapter II of the PCT can make a request for an "international preliminary examination" to be performed for selected patent offices, known as "elected offices."

After the issue of the search report (and the international preliminary examination, if applicable and requested) the application passes to the national patent office of each of the designated countries, and this is the start of the "national phase" of the procedure.

The PCT does not lead to the grant of an "international patent". This does not exist.

Finally, at 30 months, from the filing of date of the international application or from the earliest priority date of the application if a priority is claimed, the international phase ends and the international application enters in national and regional phase. (The 30-month time limit to enter the national phase is not applicable to all countries. A small number of countries still request the entering of the national phase under Chapter I at 20 months.).

However, any national law may fix time limits which expire later than 30 months. For instance, it is possible to enter the European regional phase at 31 months from the earliest priority date. National and regional phases can also be started at earlier the express request of the applicant.

If the entry into national or regional phase is not performed within the prescribed time limit, the international application generally ceases to have the effect of a national or regional application.

At the request of the applicant a substantive examination takes place in each of designated countries, under the laws of that country or regional organisation such as the EPO, under the law of the regional conventions.

After examination in each designated country or regional organisme, the application can be rejected or a patent can be granted. Oppositions against the issued national patents can be filed individually, if such procedure is provided for under national laws or conventions and the issued national patents can be amended or revoked, if these are provided for under national laws.

Compared with direct filing at the national offices or at the EPO, the PCT route affords the advantage of postponing for at least 18 months the payment of the national or regional filing fees, and search, designation and claims fees if any, although the total costs are eventually greater[4,6,7]

Currently there are a number of choices available for filing a patent.

An applicant can file national applications. If an applicant wishes to obtain a national patent in one or more countries which are EPC contracting states or PCT contracting states, it is possible to file a PCT application designating all the selected countries which are contracting states (including a European application) or an application at the EPO designating all the selected countries which are EPC contracting states.

2. *The* TRIPS *agreement*

The agreement on Trade Related Aspects of Intellectual Property Rights (TRIPS) is the Annex. 1C of the Marrakesh Agreement establishing the World Trade Organization (WHO), signed in Marrakesh, Morocco on 15 April 1994.

The TRIPS Agreement is an international agreement administered by the WTO that sets down minimum standards for many forms of intellectual property regulation. It was negotiated at the end of the Uruguay Round of the General Agreement on Tariffs and Trade (GATT) in 1994.

Specifically, TRIPS contains requirements that nations' laws must meet for copyright the granting of rights. Copyright includes the rights of performers, producers of sound recordings and broadcasting organizations, geographical indications, including appellations of origin, industrial designs, integrated circuit layout-designs patents, monopolies for the developers of new plant varieties, trademarks, trade dress, and undisclosed or confidential information. TRIPS also specifies enforcement procedures, remedies, and dispute resolution procedures.

The TRIPS agreement introduced intellectual property law into the international trading system for the first time, and remains the most comprehensive international agreement on intellectual property to date.

TRIPS has been criticized by the anti-globalization movement regarding, for example, its consequences with regards to the AIDS pandemic in Africa.

Despite the role which patents have played in maintaining higher drug costs for public health programs across Africa, this controversy has not led to a revision of TRIPS. In fact, developing countries concerned that developed countries were insisting on an overly narrow reading of TRIPS, initiated a round of talks that resulted in the Doha Declaration in November 2001, a WTO statement that clarifies the scope of TRIPS. The Doha Declaration states that TRIPS can and should be interpreted in light of the goal "to promote access to medicines for all" and should not prevent states from dealing with public health crises.

II. DEFINITION OF A PATENT – PATENT RIGHTS

A patent is "piece of property" issued by a state or a state organization which grants to the owner (author) the exclusive rights of the claimed invention[8] in exchange for the disclosure of the details of the invention.

In the majority of countries, a patent bestows a right of preventing any third party, from manufacturing, offering, selling, using, importing or stocking the product which is the subject of the invention without the consent of the owner of the patent.

The exclusive rights given by a patent do not extend, in general, to the acts carried out in private and for non-commercial purposes, nor usually experimentation with a product when the final aim is not the marketing of the said product. Experimentation in order to make an improvement had not been considered as an act of infringement.[9] But use of patented research tools, as for example receptors, with the aim of screening new drugs, is regarded in some countries, in particular in the US and in European countries, as an act of infringement.

The patent can be considered as a right to "exclude." On the other hand, a patent taken out in any country has several important limitations. The most significant are:

- *The territorial limitation*. The patent provides a legal protection only within the frontiers of the state in which it is granted.
- *The time limitation*. A patent has a limited life, which in a great number of countries is further restricted by its being subjected to the payment of maintenance fees or the like, so that failure to pay such fees automatically invalidates the patent.
- *The exploitation limitation*. The claimed invention according to the patent specification may only be made, used, exercised, and sold.

It must be noted that once a patent has been granted, this does not mean that the patentee is free to go ahead and make use of or sell the subject of the invention.

If the invention is an improvement on an earlier product, process or piece of apparatus which is the subject of an earlier non-expired patent having a claim which would be infringed by the improvement, it would be necessary to get a license from the earlier patentee.

III. KIND OF INVENTIONS

Inventions can be divided into four main categories:

1. inventions concerning products or machines,
2. inventions concerning processes,
3. inventions concerning the use of products already known for other uses (in pharmaceuticals, this means primary therapeutic use and secondary therapeutic uses),
4. inventions concerning compositions or combinations.

IV. SUBJECTS OF PATENTS: BASIC AND FORMAL REQUIREMENTS FOR FILING A PATENT

Since a patent is a creation of a statute in every country, it is essential to examine the statutes for which a patent may be granted.

To be able to file a patent application, a certain number of basic and formal requirements have to be met. Basic requirements are quite similar in the main industrial countries. Formal requirements vary a good deal from one country to another. Inventors, for example chemists, are mainly concerned with basic requirements and some formal requirements like disclosure of the invention, that is, description (see Section IV. B. 2.).

A. Basic requirements

Depending on the country, there are two or three such requirements. Concerning the EPC, the criteria for patentability of an invention to be considered are laid down in Articles 52–57 EPC.[4,5]

Article 52(1) EPC defines three basic criteria for an invention which must be satisfied if a patent is to be granted for the invention:

- that the invention must be susceptible of industrial application,
- must be new,
- must involve an inventive step.

- Article 52(2) EPC defines certain items which are not regarded as inventions. Discoveries, scientific theories and mathematical methods, aesthetic creations, schemes, rules and methods for performing mental acts, playing games or doing business and programs for computers, and presentation of information are not considered as inventions. It must be noted that according to the wording of Article 52(2) EPC, this list is not exhaustive.
- Article 52(3) is important when considering what is excluded from patentability, since it explicitly states that the exclusion relates only *"to the extent to which a European patent application ... relates to such subject matter or activities as such."* Thus under certain circumstances other principles will apply if the subject-matter or activities listed in Article 52 (2) are associated with a technical use. For example, the discovery of X-rays as such is not patentable, but the use of X-rays in a process could be protected by means of a patent. A number of EPO decisions follow the general principles of the EPC in that inventions are excluded from patentability only to the extent that applications relates to the excluded subject matter of Article 52(2) as such. There is no problem in patenting an invention which includes technical and non-technical features.
- Article 52(4) EPC [Article 53 (c) of EPC 2000 in force since 13 December 2007] defines certain items which shall not be regarded as inventions susceptible of industrial application. Article 52(4) of the EPC (or Article 53 (c) of EPC 2000), which defines patentable inventions, uses the conditions of industrial application to reject a category of inventions. This article states that methods for treatment of the human or animal body by surgery or therapy and diagnostic methods practised on the human or animal body are not regarded as inventions susceptible of industrial application. A very important decision of the Enlarged Board of Appeal of EPO is the Decision GR 01/83. With this decision the Enlarged Board of Appeal has confirmed that a claim to "the use of a substance or composition of the treatment of the human or animal body by therapy" is not allowable, but a claim to "the use of a substance or composition for the manufacture of a medicament for a specified new and inventive use" is allowable. Article 52(4) EPC or Article 53 (c) of EPC 2000 does not exclude substances or compositions for use in methods of treatment of the human or animal body by surgery or therapy from being as inventions susceptible of industrial applications. Apparatus for such treatment are also patentable.
- Article 53 of EPC defines certain inventions for which patents will not be granted. Two types of invention for which patents cannot be granted can be identified:
  1. Inventions the publication or exploitation of which would be the contrary to the "ordre public" or morality. Exploitation is not deemed to be contrary to the "ordre public" or morality, merely because, it is prohibited by law or regulation in some or all of the EPC contracting states. This prohibition is only likely to be invoked if it is probable that the invention is regarded as so abhorrent that the grant of patent rights would be socially unacceptable.
  2. Inventions of plants or animal varieties or essentially biological processes for the production of plants or animals. However microbiological processes and the products thereof are patentable.

  NB: The new Article 53 (c) of CBE 2000 corresponds to the Article 52(4) of the former CBE and concerns method of treatment of the human or animal body by surgery or therapy and diagnostic methods.

  It must be mentioned that the patentability of animals has given rise to much discussion. The United States Patent and Trademarks Office (USPTO) following the decision in "Ex parte Allen" 2 USPQ 2nd 1425 (1987), regards unnaturally occuring living organisms including animals, as patentable.

  Concerning EPO, the Examining Division of the EPO rejected the application for the "Harvard mouse" *inter alia* as excluded from patentability by the Article 53(b) of EPC. This decision was appealed and the Board of Appeal disagreed with the reasoning of the Examining Division. The board agreed that the process claims for the production of transgenic non-human mammals through chromosomal incorporation of an activated oncogene sequence into the genome of the non-human mammal did not involve an "essentially biological process" within the meaning of Article 53(b):

- Article 54 EPC indicates that an invention shall be considered to be new if it does not form part of the state of the art.
- Article 55 EPC defines certain disclosures of an invention which shall not be taken into consideration when determining its novelty under Article 54.
- Article 56 EPC defines inventions involving an inventive step.
- Article 57 EPC specifies the criteria for an invention to be considered as susceptible of industrial application.

Concerning the United States patent law (Patent-Act US Code Title 35-Patents), Section 101 of the Patent Act states:

> *"whoever invents or discovers any new and useful process, machine, manufacture or composition of matter, or any new and useful improvement thereof, may obtain a patent therefore, subject to the conditions and requirements of this title."*

- Section 101 of the Patent Act expresses the novelty requirement in general terms.
- Section 102 of the Patent Act attempts to spell out just what is not to be considered novel in the patent law sense.

– With regard to the utility, this is only considered in the Section 101 and in the first paragraph of Section 112 which states:

"The specification shall contain a written description of the invention, and the manner and process of making and using, it..."

Section 101 and the above proviso are the only ones in the Patent Act which make mention of utility. Nevertheless, courts have almost uniformly regarded, demonstration of utility –along with novelty and unobviousness – to be the affirmative requisites of every valid patent.[10]

– In adition to novelty and utility, Section 103 of the Patent Act makes patentability depend on the non-obvious nature of the subject matter sought to be patented.

Japanese law defines also novelty, inventive step, and industrial application as the basic criteria required for an invention.

1. Novelty

Novelty is the *sine qua non* of every invention. The requirement for novelty is very severe, since it has an absolute character and since any prior public knowledge is destructive of novelty.[11] An invention may be protected if it is new and thus enriches technology. However, within this broad definition, the term may be subjected to different interpretations. It appears simple enough but has developed quite differently under the various patent systems.

Article 54(1) and (2) of the EPC stipulates:[5]

"An invention shall be considered to be new if it does not form part of the art.

The state of the art shall be held to comprise everything made available to the public by means of a written or oral description, by use, or in any other way, before the date of filing of the European patent application."

The definition of *the state of the art* according to Article 54(2) encompasses all conceivable forms of disclosure, and expressly includes written as well as oral description as well as use. In order to cover all other possible forms of disclosure, this definition also includes making information available.

Under European law and under the law of most countries in the world, except the United States, a disclosure of an invention by the inventor before an application is filed (first application allowing to the Priority Right) will destroy novelty, i.e. patentability.

For a document "to be made available to the public" it does not matter whether any member of the public actually reads it. The point at issue is whether, on the balance of probabilities, any member of the public had access to it. A document available to the public in a library,[12,4] or in a patent office file[13,4] would represent a disclosure even if it could be proved that no member of the public had actually read the document. A document is not made available to the public upon posting, but on delivery.[4]

Concerning chemical composition of a product, the chemical composition of a product is the state of the art when the product as such is available to the public and can be analyzed and reproduced by a skilled person, irrespective of whether or not particular reasons can be identified for analyzing the composition.[14]

In order to ascertain if there is prior knowledge, three criteria are taken into account: public disclosure, sufficient disclosure, and date of disclosure.

It will thus be investigated whether, at the filing date of the patent application, anyone could know about the invention. It will be also taken into consideration whether the disclosure is sufficient for an expert to be able to perform the invention without an inventive act. Finally the date of the disclosure will be considered, which must be certain and not in doubt. In order to destroy novelty, Prior knowledge must of course be prior to the filing date of the patent application.

As noted above, according to the Article 54 EPC, the definition of the state of the art is very wide. There is neither restriction as to where and when a disclosure was made available to the public, nor are there any restrictions as to the language and manner of the disclosure. There is however, certain exceptions relating to the disclosure in breach of confidence and displays at certain international exhibitions defined in Article 55 EPC.[4] Since the obligation to secrecy is particularly important during licensing negotiations and demonstrations prior to patent application filing, various principles have been adopted and are taken into consideration by the EPO. This is also the case in the law of many countries). According to the EPO's practice the subject matter of a patent is not made available to the public if there is an express or tacit agreement to secrecy. If a person under the obligation to secrecy does make the information available to the public in abuse of this obligation, then from that moment this information forms part of the state of the art and thus destroys novelty. However, such information is not prejudicial for the invention.

A prior disclosure in a document or by use destroys the novelty of any claimed invention that is directly and unambiguously derivable there from.[15] The disclosure not only destroys the novelty of specifically mentioned features, but also destroys the novelty of features that are implicit to those skilled in the art (an expert). A specific disclosure of a chemical compound destroys the novelty of a generic claim embracing it.[16] For example, a disclosure of the use of rubber in circumstances where its elastic properties are used, would destroy novelty of the use of an elastic material in general.[17]

On the other hand, a generic disclosure does not usually destroy the novelty of a specific example falling within it or even of a subgeneric claim. Thus the disclosure of a genus does not disclose its species and the disclosure of a range does not disclose elements within that range. In this case,

the claimed species or subgenus must not include, and must show an unexpected advantage over, prior disclosed species of the genus. This is similar to the situation in many countries, and within EPO, it was confirmed by the very first decision of Technical Board Appeal.[18]

Consideration must be given not only to specifically described features, but also to whether by following the prior disclosure there would be an inevitable result.[4]

Concerning enantiomers, the EPO's Technical Board of Appeal confirmed that the disclosure of a racemate did not destroy the novelty of the enantiomers, even though they would be understood as being present in the racemate (unseparated form). According the EPO's decision, the enantiomers had not been "individualised" in the prior art.[4,19]

To attack novelty, it is not permitted to combine two or more prior disclosures.

It may in some cases be difficult to prove the date relating to oral disclosure. A written confirmation even updated of the oral disclosure enforces it.

The concept of absolute novelty at the filing date of the patent application is now to be found in practically all patent laws, with the exception however, as said before, of the law of the United States.

Article 102 of the American law on patents,[20] which establishes the criteria of novelty in the United States in effect is as follows:

> *"A person shall be entitled to a patent unless*:
>
> (a) *the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for patent,*
> (b) *the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than 1 year prior to the date of the application for patent in the United States,*
> (c) *he has abandoned the invention,*
> (d) *the invention was first patented or caused to be patented, or was the subject of an inventor's certificate, by the applicant or his legal representatives or assigns in a foreign country prior to the date of the application for patent in this country on an application for patent or inventor's certificate field more than 12 months before the filing of the application in the United Sates,*
> (e) *the invention was described in a patent granted on an application for patent by another filed in the United States before the invention thereof by the applicant for patent, or on an international application by another who has fulfilled the requirements of paragraphs (1), (2), and (4) of Section 371(c) of this title before the invention thereof by the applicant for patent,*
> (f) *he did not himself invent the subject matter sought to be patented,*
> (g) *before the applicant's invention thereof the invention was made in this country by another who had not abandoned, suppressed, or concealed it. In determining priority of invention there shall be considered not only the respective dates of conception and reduction to practice of the invention, but also the reasonable diligence of one who was first to conceive and last to reduce to practice, from a time prior to conception by the other."*

Paragraph (b) and (d) of this article provide a grace period of one year which enables an inventor in the USA to test an invention without such testing being regarded as destroying novelty. So the invention can therefore be freely developed without the possibility of personal disclosures being counted against the inventor. During this grace period of one year counting from the first disclosure, the inventor can then file a patent application.

It must be emphasized here that all other countries in the world have opted for the criterion of absolute novelty at the filing date of the patent application and disclosures of the inventors before filing applications destroy the novelty of their own invention. Thus, an inventor in the United States taking advantage of this possibility offered by Article 102 of the American law will no longer have the chance to file in foreign countries, since there is no longer novelty at the filing date of the patent application.

Article 102 of the American law refers moreover to the invention and not to the filing date of the patent application. This article thus created the so-called *"first inventor system"* as contrasted with the so-called *"first depositor system"* used in almost all other countries.

Paragraph (g) of Article 102 of the US law, lays down a ruling on the question of "double patenting" different from that of other countries and the corresponding Article of EPC [Article 54(3)]. According to paragraph (g) of Article 102 of the US law, in the case of "double patenting," the patent will be granted to the inventor who not merely has the earliest date for conception and for putting the invention into practice, but who also has displayed diligence in putting the invention to use. Under 54(3) EPC, the content of the art includes the content of other European patent application field earlier than, but published on or after the filing date of a European application. Concerning co-pending application with the same priority date, there is no express provision in the EPC. The EPO will allow both to proceed to the grant. But if applicants are the same, the EPO will require amendment in order to avoid double patenting.

Effect of "priority right." According to the Paris Convention, a person who has filed an earlier application in any state which is party to the Paris Convention may claim for a later filed application for the same invention priority from the date of filing of the earlier application. As a result, the date of filing of the earlier application becomes the effective date of filing of the later application, if this later application is filed in a state party to the Paris Convention. This is also stipulated in Article 87(1) EPC.

Thus, a document, or oral disclosure or co-pending European patent application will be cited by the EPO as part of the state of the art, if such a document or oral disclosure is believed to have been publicly available or such an application is believed to disclose matter having a priority date before the filing date of the application.[15] The test for determining whether subject matter is entitled to priority earlier than that of the date on which the application containing

that matter was filled is the same as the test novelty, that is, the disclosure in a single earlier application from which priority is claimed must be sufficient to destroy the novelty of a claim to that subject matter in the later application, for which priority is claimed.

2. *Inventive step*

A patentable invention requires the exercise of the inventive faculty or the exercise of more than the expected skill of the art. In other words, the invention must have been non-obvious to a person skilled in the art when it was made.

Even in the first patent laws, the criterion of novelty was regarded as insufficient. German Patent Law therefore demanded that for a patent to be granted, it had to have a certain "inventive height" (or "inventive level") or a certain "technical advance." The United States Patent Laws speaks of "unobviousness."

Article 52(1) EPC requires an invention to involve an inventive step if it is to be patentable, and Article 56 EPC states that:

> *"An invention will be considered as involving an inventive step if, having regard to the state of the art, it is non-obvious to a person skilled in the art..."*

The state of the art for this purpose is defined as not including documents under Article 54(3) EPC, that is, patent applications which were not published at the filing date of the patent application under consideration. It cannot really be expected that an invention "B" should display an inventive step as compared with another invention "A" which was totally unknown at the filing date of the application "B" being considered. (As to novelty, the contents of European patent applications still unpublished at the filling date are retained, to avoid granting two patents for the same invention, i.e. double patenting).

Since the criterion for an inventive step is expressed in a negative manner, no positive virtues are required in order to satisfy Article 56. In other words, there is no requirement for a technical advance. However a technical advance may be useful in arguing for the existence of an inventive step.[4] As noted by Dr. R. Singer[21] (Chairman of the Legal Board of Appeal of the European Patent Office):

> "in order to understand the term inventive step it is useful to divide technical knowledge into three levels. The first level includes all that is known, the mentioned state of the art, i.e. nothing new. The second level contains everything that is new but does not involve an inventive step, i.e. every-day developments of the state of the art. This is the so-called patent-free area. Finally, the third level includes everything that is new and which is also inventive and can be protected by patents."

As with any assessment based on a qualitative judgement, often there is a difference of opinion between the different people involved in the assessment of the inventive step (the applicant, the representative if any, the examiner and the opponent if any).

During the examination by the EPO an inventive step may arise from devising a solution to a known problem. It may also arise from the identification of a problem to be solved and its solution, even though the solution is obvious once the problem has been identified. Each case should be carefully analyzed on its own merits.

To reach a conclusion as to whether an invention includes an inventive step, it is necessary to determine the difference between it and the whole of the state of the art. As cited above, the state of the art comprises everything made available to the public before the filing of the application or its priority date.

When considering whether there is inventive step, the question must always be asked: Who is the person skilled in the art? Who is presumed to know all the relevant prior art? *A person skilled in the art* is presumed to have access to the entire state of the art and, in particular, it is presumed to possess common general knowledge applicable to his field of interest. Furthermore, it is part of the normal activities of the skilled person to select from the material known to him the most appropriate item. The skilled person is not regarded as having any inventive ingenuity. So, the teaching of a document may have narrower implications for a skilled person and broader implications for a potential inventor who first perceives the problem which the future invention is intended to solve.

The skilled person is a "practical person." According to an EPO Decision,[22] if the problem prompts the skilled person to seek a solution in another technical field, the person of ordinary skill in that other field is the appropriate person to solve the problem. The assessment of inventive step must therefore be based on the skill of that latter person. EPO's directives[15] point out that the skilled person need not be an individual. In certain circumstances depending on the nature of the invention, it may be more appropriate to think of a group of specialist pooling their knowledge.

The assessment of an inventive step is achieved by starting from the state of the art viewed objectively.[23] In an evaluation of the state of the art to define an inventive step, it is permissible to combine the disclosures of two or more documents, but only where to do so would have been obvious to the skilled person. Thus, documents should not be combined if it is unlikely that a skilled person would combine them to solve his problem, especially where there is an inherent incompatibility.

If an invention differs from the state of the art by the substitution of an alleged equivalent for one of its elements, it is not necessary to combine a secondary reference of the prior art in related subject matter to establish lack of inventive step, with the proviso, that the type of substitution is generally known and a skilled person would expect the general effects given by the invention. But, if the substitution gives an unexpectedly greater effect or a non-obvious additional effect, an inventive step exists. Additionally if the inventor was confronted with many possible potential

solutions to his problem, and he had no particular reason to select the one leading to the invention, an inventive step should be recognized.

A situation encountered in chemical cases is where two distinct substances, each comprising two or more structural elements, have a common property, and a claim is presented for a new substance combining structural elements from each of the known substances and having the same property. If the common property of the two known substances is indicated as being associated with each substance as a whole and not with the structural elements selected for the new substance, the new substance may be not obvious from the two known substances.[4,24]

In a combination of known features, an inventive step should be recognized if the combined features mutually support each other in their effect and so a new technical result is achieved. According the EPO's guidelines[15] for examination of patent applications, there is an inventive step "if in a mixture of medicines consisting of a analgesic and a sedative (tranquillizer), it was found that through the addition of the sedative – which intrinsically appeared to have no analgesic effect – the analgesic effect of the analgesic was intensified in a way which could not have been predicted from the known properties of the active substances."

An unexpected effect is one of the arguments for the existence of an inventive step. The unexpected effect must be established by evidence. Very often, it is necessary to present appropriate comparative tests with materials from the closest prior art. If a particular effect or property would be expected qualitatively, applicants are often asked to supply evidence to support an inventive step by demonstrating a significant quantitative improvement.

In many case a technical advantage may be foreseeable and therefore, not unexpected. In these cases there is a lack of inventive step.

The EPO's Technical Board of Appeal has held that if the prior art very strongly and clearly directs the skilled person in a particular direction, there is no inventive step involved in following that direction, irrespective of the results. This is known as the "one-way street" situation and has led to considerable discussion amongst commentators. "Obvious to try" is not a standard for denying the existence of an inventive step. In accordance with recent practice in national Courts in Europe and in the United States, the EPO is reluctant to maintain an "obvious to try" objection in the face of unexpected results.

An unexpected effect in one species or subgenus selected from a known group and it is sufficient to render the species or subgenus patentable, because an inventive step can be acknowledged. Such inventions are called "selection inventions." The principle of selection invention has been upheld by the EPO in many decisions.

Regarding chemical intermediates, a chemical intermediate which has no inherent useful properties other than as a starting material for a useful end-product claimed in the same application may depend on the useful end-product for its inventive step because it is part of the same invention.[4] On the other hand, according a decision of the EPO's Technical Board of Appeal, a new chemical intermediate may be inventive if it is prepared in the course of a multistage process, providing that the intermediate itself makes a contribution to the quantitative effect of the process.[25]

According the American law of patents, the criterion of inventive step is replaced by the criterion of unobviousness. Non-obvious subject matter and conditions for patentability are set forth in Section 103 of the US Patent Law which states that:

> *"A patent may not be obtained though the invention is not identically disclosed or described as set forth in Section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made."*

In general, in determining whether a patentable invention has been made or not, the USPTO and a court ascertain, among other points, whether the invention[1] produced new, improved, or unexpected results; satisfied a long-felt want; solved an outstanding problem; was contra-indicated in the prior art; succeeded over unsuccessful efforts of others; went into extensive use and enjoyed public acquiescence; had commercial success. On the other hand, inventions are ordinarily held to be unpatentable when they involve nothing more than a change of form, size or location; a substitution of materials or equivalents; a change of proportions or ingredients of a composition without a surprising result; superior or excellent workmanship.

3. *The invention must be capable of industrial application*

If novelty is the *sine qua non* of invention, then industrial application or utility is its *"raison d'être"*. The laws of almost every country require that the invention should be capable of industrial application.

Industrial application is defined by Article 57 of the EPC as follows:[5]

> *"An invention shall be considered as susceptible of industrial application if it can be made or used in any kind of industry including agriculture."*

It would seem that the term "industrial" will be interpreted broadly so as to include most commercial activities. For example, according to EPO decisions,[26,27] professional use of a cosmetic process in a beauty parlor is found to fall within the meaning of "industrial application."

The criterion of "utility" found in the American law is quite close to the criterion of "industrial application" which is found in the laws of a good number of countries and in particular in the EPC. A chemical product for which neither

industrial application nor any pharmacological property had been found would not be patentable under the United States patent law, because it would not have any utility.

Of interest is the decision of the US Supreme Court in the Manson case (383 US 519), which overruled the utility doctrine enunciated by the Court of Customs and Patent Appeals in the "Nelson case" (47 C.C.P.A. 1031), and which adjudicated the question of what constitutes utility in a chemical process. The invention involved a process for making certain steroids.[1] The Court held that the utility of the product must be established and that merely providing a product to be available for "use testing" by the trade was insufficient. In the decision, it was pointed out that: *"It was never intended that a patent be granted upon a product, or a process producing a product, unless such product be useful."*

The Court said that it did not mean to disparage: *"the importance of contributions to the fount of scientific information short of the invention of something "useful", or that we are blind to the prospect that what now seems without "use" may tomorrow command the grateful attention of the public. But a patent is not a hunting license. It is not a reward for the search but compensation for its successful conclusion. A patent system must be related to the world of commerce rather than to the realm of philosophy."*

In view of the Supreme Court decision, it is clear that as many affirmative statements and positive results should be included in a patent specification, as possible.

B. Formal requirements

The patent application must have a single subject. This principle is known as *unity of invention*. According to EPC's Article 82:

> *"The European patent application shall relate to one invention only or to a group of inventions so linked as to form a single general inventive concept."*

Unity of invention is required for documentary research reasons. In the United States, the requirement of unity of invention (single subject per patent application) is investigated with the greatest strictness.

If the patent application relates to more than one invention and the applicant wishes to protect two or more inventions, the "invention or inventions" can be the subject of one or more *divisional applications.* The divisional application may be filed at any time before the grant of the "parent application." A divisional application cannot be filed if the parent application has been withdrawn or is deemed to have been withdrawn. Lack of unity is not a cause of invalidity of a granted patent.

The patent application must contain a *description* explaining the invention.

The description shall:

(a) specify the technical field to which the invention relates
(b) indicate the background art which as far as it is known to the applicant, can be regarded as useful for understanding the invention and it will preferably cite the documents reflecting such art,
(c) disclose the invention as claimed and state any advantageous effect of the invention with reference to the background art,
(d) describe in detail at least one way of carrying out of the invention claimed using examples where appropriate, and referring to the drawings if any,
(e) indicate explicitly, when it is not obvious from the nature of the invention, the way in which the invention is capable of exploitation in industry,
(f) briefly describe the figures in the drawings if any.

In fact, the patent application must explain the invention in a sufficiently clear and complete way for a person skilled in the art to be able to reproduce the invention without having to carry out research or lengthy and delicate operations.

Applications to be filed in the United States must in addition give the best method of realising the invention.

In general, it is not necessary to give detailed results of trials of use of the invention, for example, biological pharmacological or clinical results for pharmaceuticals. But sometimes, it can be useful to present comparative results between the products of the application and the products of the prior art, in order to demonstrate the advantage of the invention.

The patent application must contain *claims*. According to Article 84 of EPC:

> *"The claims shall define the matter for which protection is sought. They shall be clear and concise and be supported by the description."*

In fact, the claims shall define the matter for which protection is sought in terms of the technical features of the invention.

The EPC enumerates four kinds of claims corresponding to the four main categories of invention. But in fact, there are two basic types of claims:

1. claims for products or machines,
2. claims for activities (processes, utilization).[15]

Any claim stating the essential features of an invention may be followed by one or more claims concerning particular embodiments of that invention.

It is possible to have several independent claims in a European patent application. This is in particular permitted in order to claim certain permutations of what are termed as different categories.

A patent application to be filed at the EPO can have different types of claim. The different categories permissible are:

- a product, a process adapted for the manufacture of the product and a use of the product,
- a process and an apparatus or means specifically designed for the carrying out the process,
- a product, a process specially adapted for making the product and an apparatus or means specially designed for carrying out the process.

In American practice, in an application for invention concerning compounds, only products and methods of use claims are generally admitted by the examiners of the Patent Office. For a process for the manufacture of products, a divisional application is often required. In all cases, the claims cannot extend more widely than the description, completed if appropriate with the drawings.

Products claims bestow a protection that can be regarded as absolute. Until the beginnings of 1990's, in most of Eastern Europe patents systems, it was not possible to claim new compounds.

This was also the case especially for pharmaceuticals in Canada, Austria, Spain, Greece, in most Scandinavian countries except Sweden, and in South American countries. Pharmaceuticals were protected only by process claims. Now the patent laws have been changed, and pharmaceutical compounds can be claimed in most countries.

A product or a family of products can be defined as well by a general chemical formula (Markush formula), as by a specific chemical name.

In chemistry, *a process claim* describes a series of reactions leading to the synthesis obtention of a given chemical product. Process claims have the effect of preventing a third party from using the same process. With regard to the product, a process claim is ineffective against a third party who might use a different process to prepare the same product.

Sometimes, a patentable product cannot be defined by reference to its structure, composition or other reproducible property. Such products are protected (claimed) by "product by process" claims. A product by process claim covers a product independently of the process used to obtain it.

Concerning *the first therapeutic use*, for a known compound which is shown to have a specific therapeutic use, it is possible according to the EPO's practice, to have a broad claim in the form "substance or composition X for use in medicine." Formulation of claims such as "the use of a substance X for the therapeutic treatment of man or animals" are not permitted, because they are regarded as methods of medicinal treatment.

With regard to *the second therapeutic use*, as noted above, the EPO's Enlarged Board of Appeal approved a wording such as "The use of a substance X for the manufacture of a medicine for therapeutic application."[15]

In the USA, the patentability of "second therapeutic uses" has been recognized for a long time and many methods of patents use may be granted for the same product. It is now possible to protect second therapeutic uses in many countries. Claims for second non-medical uses are also admitted by the Patent Laws of many countries.

As regards the effects, a claim covering the novel use of a known product will only cover the claimed use. Other uses of the compound and the compound as such are not protected.

Compositions or combinations, if they yield an unexpected result of their own or a technical effect of their own are patentable, but the mere juxtaposition of two compounds is not by itself patentable. *Composition claims* are admitted in the patents laws of most countries. A patent covering a particular composition or a given composition will cover the composition or combination effectively claimed, and not each constituent of the composition or the combination.

Another formal requirement is that the patent application must *designate the true inventors*. The true inventors are the persons who have actually participated in achieving the invention. People who have only carried out routine tests (e.g. biological tests relating to screening) should not be regarded as true inventors.

A failure to designate any inventor in the request of a grant of a European patent creates a formal deficiency. According to the EPC, there is no penalty if an incorrect designation of inventors exists apart from procedures for correction of the designation. So, the resulting patent cannot be opposed or revoked merely on the grounds of wrong inventor designation.

Identifying the true inventor is very important in the United States. According to the US patent law, any failure to designate the true inventor is a cause for invalidation of a patent.

Patent applications must be filed in the *official language* or one of the official languages of the country in which the application is filed. As mentioned above, European patent applications must be filed in one of the three official languages of the EPO, that is, German, English, or French.

Filing a patent application also gives rise to *payment of fees*, that is, filing fees, examination fees (if any), and to payment of annuities. Payment of annuities or periodical fees must be paid throughout the whole life of the patent. Fees must be paid in due time. Failure to pay fees automatically invalidates the patent or the patent application.

For filing a patent application in foreign countries, a *legal representative* is required.

Concerning the European patent application, parties to proceedings may act in relation to EPO in the following alternative ways:

(a) a real person may act for himself,
(b) a real person may act through an authorized representative, who may be a professional representative, a legal practitioner or any employee of the real person,
(c) a legal person (i.e. a body corporate) may act for itself, that is to say by the actions of any responsible officer of the legal person,
(d) a legal person may act through an authorized representative who may be a professional representative, a legal practitioner or any employee of the legal person.

Natural or legal persons having neither a residence nor a principal place of business in any EPC contracting state must act through a representative. This representative must be a professional representative or a legal practitioner.

Patenting is a complicated task. Applicants without experience of patenting should prepare patent applications through professionals (patent attorney, patent lawyers).

V. LIFETIME OF PATENTS

At the beginning of the 1900's, the lifetime of patents varied enormously from one country to another. With the introduction of the European Patent System, the EPC member countries ratifying the EPC opted for a lifetime of 20 years from the filing date. Since then, other countries not participating in the EPC have also opted for the lifetime of 20 years. Before 8 June 1995, a US patent had a term of 17 years from the issue date of the patent. Under the new law, enacted 8 December 1994, a patent issuing from an application filed on or after 8 June 1995 has a term that begins when the patent is issued, but ends 20 years after the earliest effective US filing date. As for patents that are in force on, or issue from the US applications filed before, 8 June 1995, such patents have a term that is the greater of 17 years from the patent's issue date or 20 years from the application's earliest effective US filing date, whichever is longer. Canada used to have the same system until 1989. From January 1989, the lifetime of a Canadian patent is 20 years from filing. In Japan, the lifetime of a patent was 15 years from the publication of the application, but could not exceed 20 years from filing. Under the new law, which will be enforced as from 1 July 1995, the lifetime of a patent is 20 years from the filing date of the patent application.

Today, the lifetime of a patent is 20 years. According the TRIPS Agreement, patents must be granted in all "fields of technology," although exceptions for certain public interests are allowed (Articles 27.2 and 27.3 [1]) and must be enforceable for at least 20 years (Article 33).

Information concerning patent law and lifetime of patents for all countries throughout the world are given in *"Manual for the Handling of Applications for Patents, Designs and Trade Marks Throughout the World"* (Octroibureau Losen Stiger, Holland).

In 1984, the USA adopted a law which enables the lifetime of certain American patents to be extended. According to this law, it is possible to obtain a maximum extension of 5 years to holders of pharmaceutical patents when the marketing of a pharmaceutical product has been delayed in its registration by the FDA.

In 1988, Japan has adopted a law which enables the life of pharmaceutical patents to be extended for a period of 2–5 years.

In 1990, France and Italy adopted a Certificate of Complementary Protection of Pharmaceutical Patents.

The laws concerning the French and Italian Certificate of Complementary Protection were abrogated in January 1993. Since this date the *Community Supplementary Protection Certificate* has been in force. According to the Community Certificate, it is possible to obtain a maximum extension of 5 years to holders of pharmaceutical patents. This certificate concerns all the EEC and EPC contracting States.

VI. OWNERSHIP OF PATENTS

A patent is a grant from the State and constitutes a property. As such, it is subject to all of the laws and regulations of property. The grant of a patent gives to the patentee a title to an intangible and incorporate right in the nature of a franchise. The title of a patent constitutes a right to the patentee until it is divested from him or by voluntary grant or by other legal means.

A patent can be conveyed like any other piece of property. It can be assigned and transferred to another party.

A patent may be licensed under the terms which give to the licensee the right to make, use, or sell the patented invention. It is in this respect similar to a piece of property which may be leased and may be also sold.

A license may be exclusive or non-exclusive. It may be restricted to a certain geographical territory, or to certain industrial fields, or may even be restricted to a certain specific factory or other place of commerce and industry.

In the USA, the titular of a patent can be only the inventor who may assignee his patent application to a company or to an organization, whereas in the other countries the titular may be the inventor, a company or an organization.

VII. INFRINGEMENT OF A PATENT

An act of infringement is an attack of the rights of the patentee in the territory where a patent is in force and effect. A person may be guilty of "direct infringement" when he either makes use of or sells the entire invention as defined by the claims of a patent. A person may be guilty of "contributory infringement" if he co-operates with another, in infringing a patent. For example, a chemical company which sells substances to another for infringing a patent can be sued just like a direct infringer.

An action for infringement can only be launched from the day the patent application is published. The reason for this rule is to protect third parties who might have infringed a patent or patent application of which they were unaware.

According to the laws of certain countries, it is possible to notify a suspected infringer of the existence of a patent application. Concerning infringements, the Judges' rule is according to the claims of the patent as granted.

Suits for the infringement of patents must be instituted in national tribunals. Infringement of a national patent issued from a European patent is dealt with under the national law of the contracting states and may be heard by the national

Courts. Generally, proof of infringement must be provided by the owner of the patent. Such proof can be witnesses, documents, purchase of products, and so on.

In every case where a patent has been held to be valid and infringed, the patent owner is entitled to have an injunction issue against the infringer.

The injunction will restrict the infringer from making, using, or selling the invention. Additionally the patent owner may obtain the payment of compensation to make up for the loss of profit or for the loss suffered. According to the law of some countries patented articles may be withdrawn and delivered up to the proprietor of the patent. In some countries relief for the infringing action may be provided by criminal penalties including a prison sentence.

In order to determine infringement, the subject matter protected by a patent is compared with the act alleged to constitute an infringement. Slavish reproduction of the subject matter protected by a patent constitutes an act of infringement. The reproduction of a patented invention by varying only details or by using equivalents (variants of execution) also constitutes infringement.

VIII. PATENTS AS A SOURCE OF INFORMATION

Patent literature is a useful source of technical, legal, and commercial information. It is extremely rich in technical information. In a patent application, applicants must disclose the invention in a sufficiently clear and complete way for a person skilled in the art to be able to reproduce the invention. New and inventive concepts give rise to a considerable number of refinements, improvements and modifications. Often, the importance of an invention is related to the number of patents and patent applications which protects it.

Complete information about the patents of the competitors is useful before the beginning of a new research program. Liberty of exploitation (freedom to use) has to be evaluated before the commercialization and preferably, before the development of a new product, a process, or a machine. In this case, patent applications, patents and the corresponding patents statutes of competitors must to be studied in detail.

The negotiating and conclusion of license agreement have become a specialised profession. Information obtained by patents can help companies to find new products or processes for expansion or diversification of their manufacturing activities.

IX. PATENTING IN THE PHARMACEUTICAL INDUSTRIES

In certain trades and industries and in particular in the pharmaceutical industries, it is generally not possible to inhibit competition effectively by keeping key formulae, products, and techniques secret. Pharmaceuticals are a classical case of an industry dependent on patents. In fact, patent protection is of great importance to the pharmaceutical research and development process. Practically all new drugs which were developed after World War II were developed by private pharmaceutical firms of the United States, the Federal Republic of Germany, Switzerland, France, Great Britain, the Nordic European countries, and Japan. These countries provide the strongest known patent protection for pharmaceuticals. Countries with similar industrial standards and possibilities, not having provided patent protection for pharmaceuticals, cannot show any drug innovations of importance.

In any case, having obtained a patent, the patentee should not think he is heir to a fortune. Many worthwhile inventions have failed commercially for lack of interest or funds to promote them. Additionally, the grant of a patent is no guarantee that the invention is significant or that the patent is valid. It should therefore be of no surprise that one of the factors considered by pharmaceutical management in deciding whether to develop a new drug, in addition to its medical and commercial potential, and cost of manufacturing, is the strength and duration of patent protection.

The best kind of protection, and really the only truly effective kind, is that which protects the product itself. A patent covering a product, as opposed to one covering only a process for making the product or a use for the product can prevent others from manufacturing or selling the product for any purpose, no matter how it is made. Since chemists can often find new processes for making a compound, a patent on a process can be circumvented by use of new processes, and the owner of a patent no longer has an exclusive position with respect to the drug. The same is true concerning a patent covering only a particular use. A new use may permit a copier to sell the product free of the original use- limited patent, even though the product may be prescribed and used for the original purpose. Patents of new use are only strong patents if the new use of a compound requires a special and new formulation.

Another aspect of protection, in addition to the type of protection available, is the term of the patent.

For competitive reasons, a patent application must be filed as soon as possible. In fact, a patent application for a new compound or series of compounds must be filed as soon as the new compound or series of compounds have been made and found to have significant and interesting activity. This is necessary because others may have made the same invention, and failure to file promptly may result in a total loss of patent rights (in most of countries, except the United States patents are granted to the first party to file a patent application on the subject matter). There are many cases in which the same invention has been independently made by more than one party within as short a period as 2 or 3 weeks.

However, in filing an application as soon as possible, a lot of problems can be generated. The term of a patent starts to run even before the product is marketed. As a period of at least 7–10 years is required for the development of a new drug after a compound and its analogs have been prepared and biologically tested, the patent terms can expire before the product begins to produce any real profits.

Under the Paris Convention, initial filing can be followed up in other countries where protection is desired. The benefit of the initial early filing date is obtained if applications are filed within a year of this date. So, the choice of countries must be decided before knowing the full importance of an invention.

Commercial interest and effective protection against infringers or competitors offered by the "local" patent law must be taken into consideration in the choice of countries where patent protection is sought.

Often, at the time a patent application is filed for a new compound, research and development staffs do not know which compound in a series will turn out to be the product to be developed. In addition, the chosen compound or compounds may well be chemically related to compounds patented earlier, and this fact complicates the question of "patentability." Even before the full importance of an invention is known, a number of needs must be met at this early stage when a patent application is filed. In order to adequately protect an invention, the applicant must describe how to realize and how to use it. Normally, a chemist will invent a series of compounds, all having the same type of activity to a greater or lesser degree. Inventors cannot afford to simply patent for the best compound, even if they know which compound is the best. Competitors would read the patent and make the next closest relative of the compound, thereby depriving the inventor of part of his invention. Thus examples must be described of the group of the invented compounds. Moreover a company may change its interest to another compound in the series as it learns more about the compounds. Potency, stability, side-effect profile, bioavailability, and difficulty or cost of manufacturing may indicate that later prepared compounds in a series are preferable for development. For this reason, it is also necessary for a variety of compounds in the series to be described.

The other significant problem in obtaining valid and effective patent protection relates to requirements of the patent law variously referred to in different countries as the requirement for non-obviousness, or attainment of a sufficient inventive level. Normally, the requirement that an invention be useful or capable of industrial application is not a problem. The novelty requirement is also not usually a problem. However, even in the usual case where compounds are novel, few compounds are of such distinctiveness that no compound of similar structure has ever been previously published. Thus, it must be established, either by arguments of chemical difference or of biological properties, that the new compounds merit patent protection. Inventors' previous patents and publications often complicate the situation and it is most distressing when inventors have difficulty patenting their chosen compounds because of their earlier public disclosure of compounds long since discarded.

The opportunity for filing a patent application has to be well evaluated. Sometimes, it is preferable to keep an invention secret, if the infringement of the future patent related to this invention will be difficult to demonstrate. This concerns especially process patents for preparing commercially successful compounds not protected by a "compound patent." Infringement of such patents can generally be demonstrated by identifying impurities due to the process (if any) or by witnesses. If the demonstration of infringement is not possible, competitors may use the process they are interested in, after the publication of the patent application, without any jeopardy or sanction.

There is a correlation between Research and Development and industrialization. Patents provide a stimulus for investment in Research and Development. Investment in the type of development which has brought in the world great progress in medicinal chemistry requires a healthy patent strategy.

X. CONCLUSION

Patent systems throughout the world are today based on the principle of territorial rights. As an invention will be protected only in a state in which a patent has been applied for and granted, inventors and applicants are forced to apply for patents for an invention in several states. This involves the use of different laws and rules and a considerable financial outlay. Inventors and applicants are often confronted with the legalistic problems and phraseology of the world of patent protection.

Even if the phraseology can be easily understood and adopted, legalistic problems require a good knowledge of the law of the countries and the corresponding practice and jurisprudence.

Patents provide a stimulus and assure possibility for investment in Research and Development. Without investment, it is not possible to have an efficient Research and Development Programme. A healthy patent strategy is the key to scientific progress.

REFERENCES

1. C. H. Greenstrat (Liebesny, F., Ed.)*Mainly on Patents: The Use of Industrial Property and Its Literature*. Butterworth and Co: London, 1972.
2. Mathély, P. Le Droit français des brevets d'invention. Journal des Notaires et des Avocats, 6 rue de Mézières. 1992.
3. Bodeyhausey, G. H. C. Guide d'application de la Convention de Paris pour la Protection de la Propriété Industrielle. BIRPI. Genève, Suisse, 1969.
4. Chartered Institute of Patent Agents. European Patents Handbook Release 10. Longman Group Ltd: Essex, UK, 1991.

5. Druck R. European Patent Convention. European Patent Office: Kirchheim, Germany, 2002.
6. Traité de Coopération en matière de brevets (PCT) et règlement d'exécution du PCT OMPI Genève, Suisse, 1998.
7. PCT-Guide du déposant. n° 6 Organisation Mondiale de la Propriété Industrielle (OMPI) Genève, Suisse, 2000.
8. Chavanne, A., Burst, J. J. Droit de la Propriété Industrielle 3ème edt. Dalloz, 11 rue Soufflot, 75240, Paris cedex 05 France, 1990.
9. PIBD. III235. Tribunal de Grande Instance de Paris, 8 Juillet, 1982.
10. Rosenberg, P. *Patent Law Fundamentals*. Clark Boardman Company Ltd: New York, 1975.
11. Decision of the Technical Boar of Appeal 3.3.1. T31/84. *Official Journal of the European Patent Office*. 4 May 1985, *11*(86), 369–372.
12. Decision of the Technical Board of Appeal 3.3.1. T381/87. *Official Journal of the European Patent Office*. 10 November 1988, *5*(90), 213–223.
13. Decision of the Technical Board of Appeal T444/88 12 IPD Decision IPD 13222. (1990)
14. Opinion of the Enlarged Board of Appeal G1/92. *Official Journal of the European Patent Office*. 18 December 1992, *5*(93), 277.
15. *Directives relatives à l'examen pratique à l'Office Européen des brevets*. Publication Office Européen des brevets Pindar Infotek, 11 Melrose Yard Walmgate York 2001, UK.
16. Decision of the Technical Board of Appeal 3.3.1. T188/83. *Official Journal of the European Patent Office*. 30 July 1984, *11*(84), 555–562.
17. Decision of the Technical Board of Appeal 3.3.1. T181/82. *Official Journal of the European Patent Office*. 28 February 1984, *9*(84), 401–414.
18. Decision of the Technical Board of Appeal 3.3.1. T01/80. *Official Journal of the European Patent Office*. 6 April 1981, *9*(81), 349.
19. Decision of the Technical Board of Appeal 3.4.1.T296/87. *Official Journal of the European Patent Office*. 30 August 1988, *5*(90), 195–212.
20. *Patents and Trademark-Laws* The Bureau of National Affairs: Washington, DC, 1989.
21. Singer R. (Devons, D. J., Ed.) *The New European Patent System*. Seminar Services International: Swiss, 1981.
22. Decision of the Technical Board of Appeal 3.2.1. T195/84. *Official Journal of the European Patent Office*. 10 October 1985, *5*(86), 121–128.
23. Decision of the Technical Board of Appeal 3.3.1. T24/81. *Official Journal of the European Patent Office*. 13 October 1982, *4*(83), 133–142.
24. Decision of the Technical Board of Appeal 3.3.1. T20/83. *Official Journal of the European Patent Office*. 17 March 1983, *9*(86), 295–301.
25. Decision of the Technical Board of Appeal 3.3.2. T163/84. *Official Journal of the European Patent Office*. 21 August 1986, *7*(87), 301–308.
26. Decision of the Technical Board of Appeal 3.3.1. T36/83. *Official Journal of the European Patent Office*. 14 May 1985, *9*(86), 295–301.
27. Decision of the Technical Board of Appeal 3.3.1. T144/83. *Official Journal of the European Patent Office*. 27 March 1986, 9(86), 301–305.

Chapter 45

The Consumption and Production of Pharmaceuticals

Bryan G. Reuben

I. "IMPORTANT" DRUGS
A. The top-earning drugs
B. The most widely prescribed drugs
C. National differences in prescribing
II. SOURCES OF DRUGS
A. Vegetable sources
B. Animal sources
C. Biological sources
D. Fermentation
E. Chemical synthesis
III. MANUFACTURE OF DRUGS
A. Good manufacturing practice
B. Plant design
C. Formulation and packaging–sterile products
D. Choice of reagents
E. "Green" chemistry
F. Downstream processing
G. Outsourcing
IV. SOCIAL AND ECONOMIC FACTORS
A. Pattern and cost of innovation
B. Patents
C. Orphan drugs
D. Generic pharmaceuticals
E. Parallel trade
F. Cost containment measures
G. Pharmacoeconomics
V. THE FUTURE OF THE PHARMACEUTICAL INDUSTRY
A. Trends in pharmaceuticals
B. Conclusion
REFERENCES

The end of the medical art is health and that of economics is wealth.

Aristotle (Ethics)

Medicinal chemistry is of great academic and intellectual interest. The elucidation of the arachidonic acid cascade, for example, is one of the most fascinating bits of chemistry of our generation. The study of the chemistry of the brain, so that we are able to understand how it is that we can understand, is one of the great frontiers of science. To the author, at least, it is of far greater relevance than the problems of the origins of the universe, the descent of man or the existence of the Higgs boson, the last of which increasingly resembles the search for the Holy Grail.

Unlike astrophysics and evolutionary theory, however, medicinal chemistry is an applied science. It is lavishly funded not because of its philosophical centrality, but because it provides the hope that human disease can be cured or alleviated. Its paymasters intend that mankind, or at least those sections of it with access to advanced medical care, live longer and more comfortably. Its practitioners are judged by this criterion. There are few Nobel prizes for those who discover, say, the biochemical origins of a rodent-specific dermatitis.

As the test of success is pragmatic, serendipity plays an important role in medicinal chemistry. The discoverers of sulfonamides thought that dyestuffs might prove efficacious because they bonded specifically to certain tissues, as in Ehrlich's classic experiment. In the end, Prontosil worked not because it was a dye but because it cleaved in the gut to *p*-aminobenzenesulfonic acid. We still honor Domagk, just as we honor Fleming, whose chance discovery of penicillin would have been meaningless had not Florey and Chain solved the problem of its purification and Coghill and his co-workers (who did not get Nobel prizes) solved the problem of its large-scale production.

Medicines are thus judged by their results. A successful drug can be manufactured reasonably easily, has negligible

Copyright © 2008, Elsevier Ltd
All rights reserved.

side effects, is widely prescribed, makes a lot of money and is perceived as making a major contribution to health care. In this chapter, we shall discuss which drugs are important on which criteria and say something about the technical, social and economic problems of the pharmaceutical industry.

I. "IMPORTANT" DRUGS

To the person whose life is saved by it, any drug is important. None the less, some are more significant than others, either because they provide better returns for their manufacturers, or because they are more frequently prescribed.

A. The top-earning drugs

From the point of view of the manufacturer, the money side is the most important. A research-oriented drug company needs adequate cash flow to maintain its organization, fund its research and keep its shareholders happy. The world market for drugs in 2006 was $607.9 billion. Table 45.1 shows the ten best-sellers in the United States and the world. Most of them are still under patent, or their patents have only recently expired; thus their price and profitability remains high. Five of the top ten find a place in both lists, indicating that prescribing patterns throughout the developed world are similar, and that the cheaper drugs in poorer countries do not have sufficient sales to find a place.

Four of the fifteen drugs in the two lists are for cardiovascular disease (Lipitor, Simvastatin, Plavix and Norvasc), reflecting that this is the major cause of death in the developed world, but their prominence is less than a few years ago because there have been few dramatic advances since the lipid regulators. Three are for schizophrenia, reflecting the continuing prevalence of mental illness and the lack of any definitive way of curing it. Three (Aranesp, Epogen and Enbrel) are genetically engineered biotechnology products, reflecting the maturing of the biotechnology sector. Three more – fluticasone, salmeterol and montelukast – are for asthma, the only disease for which developed world mortality is rising. Two (Nexium and Lansoprazole) are for ulcers. Until the mid-1970s, ulcers could be treated only by surgery and were generally uncomfortable rather than fatal. The importance of these drugs is that they improve the quality of life rather than saving it.

The above compounds are the glamorous pharmaceuticals, the results of fairly recent research and mostly based on a clear idea of the enzyme systems they are supposed to influence. The most widely consumed pharmaceuticals, however, are paracetamol (called acetaminophen (APAP) in the United States), aspirin and vitamin C. These can be bought over-the-counter. Annual world consumption is of the order of 50,000 tonnes each, an order of magnitude higher than consumption of, say, penicillin V. At the other extreme come materials such as the anti-cancer drug, methotrexate, with annual production of about 150 kg. The prices of the large tonnage drugs are so low, however, that US prescription sales of APAP, even when compounded with hydrocodone, propoxyphene and codeine etcetera are only about $2.3 billion. This figure, however, omits over-the-counter (OTC) products, where these compounds are sold directly to the consumer as such, or dressed up as cold cures or analgesics and promoted by extensive advertising. The number of people who take these drugs is astronomical. In the United Kingdom, some years ago, it appeared that consumption of aspirin-containing tablets averaged 200 per year per man, woman and child. This figure has decreased, reflecting the substitution of aspirin by paracetamol, but modern figures are not available.

B. The most widely prescribed drugs

More important in terms of public health than the top-earning drugs are the drugs that are most widely prescribed. Table 45.2 shows the 50 most frequently prescribed drugs in the United States. Seven (Effexor, Lipitor, Nexium, Norvasc, Singulair, salmeterol and simvastatin) occur in the top-earning list. The others are mainly out-of-patent and often made by a number of manufacturers. Hydrocodone (formulated with APAP) comes high on the list. It is a mild analgesic but, because hydrocodone is an opioid compound, it is available only on prescription in the United States, although many European countries permit its sale over-the-counter. Hydrochlorothiazide is a mild diuretic formulated alone or with a range of other heart drugs. Amoxicillin is a vitally important antibiotic and, together with its analogs, has played a central role in the conquest of infectious diseases. Amoxicillin is marketed by numerous companies, and only the aggregation of their data brings it near the top of the prescriptions list. Metoprolol, a β-blocker for relief of angina, is a drug where US practice differs; atenolol or propranolol are likely to be preferred in Europe.

Figure 45.1 shows world sales of drugs by therapeutic class.

These figures are similar to those for the United States alone, the latter allocating 4.5% more to central nervous system (CNS) drugs and, surprisingly, 2% less to antineoplastic drugs. The most important drug groups and the 2006 US expenditure in the above categories are lipid regulators, $35.2 billion; oncologics, 34.6; respiratory agents, 24.6; proton pump inhibitors, 24.1; antidiabetics, 21.2; anti-depressants, 20.6; antipsychotics, 18.2; angiotensin-II antagonists, 16.5; erythropoietin products, 13.9 and antiepileptics, 13.1.

Several features of these figures compared with 20 years ago are significant. Top of the list are lipid regulators, which are hydroxymethyl-glutaryl-CoA (HMG-CoA)

TABLE 45.1 Top 10 Products by the United States and Global Sales (2006)

Rank	Product	Generic name	Therapeutic class	US sales ($ billion)	Rank	Product	Generic name	Therapeutic class	Global sales ($ billion)	% Growth 2005–2006
1	Lipitor	Atorvastatin	Lipid regulator	8.6	1	Lipitor	Atorvastatin	Lipid regulator	13.6	4.2
2	Nexium	Esomeprazole	Proton pump inhibitor	5.1	2	Nexium	Esomeprazole	Proton pump inhibitor	6.7	16.9
3	Advair	Fluticasone + salmeterol	Anti-asthma	3.9	3	Advair	Fluticasone + salmeterol	Anti-asthma	6.3	10.3
4	Aranesp	Darbepoetin alpha	Anaemia from kidney disease or chemotherapy	3.9	4	Plavix	Clopidogrel	Anti-coagulant	5.8	−3.4
5	Prevacin	Lansoprazole	Proton pump inhibitor	3.5	5	Norvasc	Amlodipine	Calcium antagonist	5	−0.5
6	Epogen	Epoetin alpha	Anemia from kidney disease	3.2	6	Aranesp	Darbepoetin alpha	Anemia from kidney disease or chemotherapy	5	35.6
7	Zocor	Simvastatin	Lipid regulator	3.1	7	Zyprexa	Olanzapine	Schizophrenia	4.7	−0.4
8	Enbrel	Etanercept	Tumor necrosis factor for rheumatoid arthritis	3.0	8	Risperdal	Risperidone	Schizophrenia	4.6	12.3
9	Seroquel	Quetiapine fumarate	Schizophrenia	3.0	9	Enbrel	Etanercept	Tumor necrosis factor for rheumatoid arthritis	4.5	18.4
10	Singulair	Montelukast	Asthma prophylaxis	3.0	10	Effexor	Venlafaxine	Anti-depressive	4	2.7
	Total			40.3		Total			60.2	8.0

Source: IMS Health, IMS National Sales Perspectives™, March 2007.

TABLE 45.2 Top Fifty Pharmaceuticals by Prescription Numbers, United States (2006)

Generic name	Typical brand name	Therapeutic class	Prescriptions ('000)
1. Acetaminophen (paracetamol)	Tylenol	Analgesic	169703
2. Hydrocodone	Vicodin (with APAP)	Analgesic	112041
3. Hydrochlorothiazide	HydroDIURIL	Diuretic	99583
4. Amoxicillin	Augmentin (with KClav), Amoxil	Antibiotic	74136
5. Atorvastatin	Lipitor	Hypolipaemic	62311
6. Metoprolol	Toprol XL, Lopressor	β-Blocker	61888
7. Levothyroxine	Levoxyl	Thyroid hormone	58577
8. Lisinopril	Zestril	ACE inhibitor	55039
9. Amlodipine	Norvasc, Lotrel (with benazepril)	Calcium antagonist	47413
10. Atenolol	Tenormin	β-Blocker	46186
11. Simvastatin	Vytorin (with ezetimibe), Zocor	Hypolipaemic	43605
12. Azithromycin	Zithromax	Macrolide antibiotic	40752
13. Albuterol (salbutamol)	Ventolin	β-2 agonist for asthma	38313
14. Alprazolam	Xanax	Anxiolytic	37327
15. Furosemide	Lasix	Diuretic	36830
16. Metformin	Glucophage	Anti-diabetic	34815
17. Oxycodone	OxyContin	Analgesic	30958
18. Fluticasone	Advair (with salmeterol)	Corticosteroid	29268
19. Potassium chloride	Klor-Con	Hypokalaemia	28364
20. Sertraline	Zoloft	SSRI	28060
21. Ibuprofen	Brufen	NSAID	27662
22. Levothyroxine	Synthroid	Thyroid hormone	27209
23. Esomeprazole	Nexium	Proton pump inhibitor	26917
24. Escitalopram	Lexapro	SSRI	26098
25. Ezetimibe	Zetia, Vytorin (with simvastatin)	Anti-cholesterol	25434
26. Montelukast	Singulair	Asthma prophylactic	24611
27. Warfarin	Coumadin	Anti-clotting	24113
28. Cephalexin	Keflex	Cephalosporin antibiotic	22789
29. Prednisone oral	Deltacortril	Corticosteroid	22657
30. Triamterene	Dyazide (with HCTZ)	Diuretic	22436
31. Propoxyphene	Darvocet (with APAP)	Analgesic	21940
32. Fluoxetine	Prozac	SSRI	21733
33. Lansoprazole	Prevacid	Proton pump inhibitor	21180
34. Zolpidem tartrate	Ambien	Non-benzodiazepine hypnotic	20344

(Continued)

TABLE 45.2 (Continued)

Generic name	Typical brand name	Therapeutic class	Prescriptions ('000)
35. Lorazepam	Ativan	Benzodiazepine	19789
36. Alendronic acid	Fosamax	Osteoporosis	18305
37. Bupropion	Wellbutrin XL	Anti-depressant	18274
38. Salmeterol	Serevent	Anti-asthma	18190
39. Clonazepam	Rivotril	Benzodiazepine	18152
40. Cetirizine	Zyrtec	Non-sedating antihistamine	17218
41. Cyclobenzaprine	Flexeril	Muscle relaxant	17183
42. Ventlafaxine	Effexor XR	SNRI	17101
43. Clopidogrel	Plavix	Anti-platelet	16248
44. Paroxetine	Paxil	SSRI	16242
45. Fexofenadine	Allegra	Anti-histamine	16213
46. Gabapentin	Neurontin	Antiepileptic	16150
47. Pantoprazole	Protonix	Proton pump inhibitor	16089
48. Tramadol	Ultram	Analgesic	15906
49. Ciprofloxacin	Cipro	Quinolone antibiotic	15815
50. Conjugated oestrogenic hormone	Premarin	Hormone replacement therapy	14804

Prescriptions for two substances are counted toward the totals for each substance.
APAP = acetaminophen; SSRI = selective serotonin reuptake inhibitor; SNRI = serotonin and noradrenaline re-uptake inhibtor; HCTZ = hydrochlorothiazide; NSAID = non-steroid anti-inflammatory drug.
Source: Author's estimates based on Verispan VONA, www.drugtopics.com/drugtopics/data/articlestandard/drugtopics/092007/407652/article.pdf.

FIGURE 45.1 Global drug sales by therapeutic class. Data are for the year ending March 2006 and apply to the United States, Europe and Japan. *Sources*: http://www.samedanltd.com/homepage/ict/Summer2006/SusanMG.pdf, IMS Health; Prescriptions dispensed in the community, UK Office of National Statistics.

reductase enzyme inhibitors. They control cholesterol and blood lipids, and the first one (lovastatin) was introduced only in 1987. The first angiotensin II antagonist (losartan) was introduced in 1995, at a time when thiazides, β-blockers, calcium antagonists and angiotensin-converting enzyme (ACE) inhibitors were already available to treat high blood pressure. Proton pump inhibitors such as omeprazole date from about 1990, having overtaken the H_2-receptor antagonists, cimetidine and ranitidine.

The new expensive anti-cancer drugs (oncologics) are high on the list, reflecting advances in technology. For the first time in over 70 years, cancer deaths in the United States began to fall in 2003, and in 2004 that trend continued. Three thousand and fourteen fewer people died of cancer in 2004 compared with 369 fewer in 2003. Cancer death *rates* have been declining and stabilizing for decades, but because of increasing incidence this is the first time *total deaths* have decreased.

Antidiabetics and respiratory agents retain high places, diabetes and asthma being among the few diseases whose incidence is increasing. The presence on the list of four categories of CNS drugs – anti-depressants, antipsychotics, anti-epileptics and anxiolytics/hypnotics – is unexpected. Prescriptions for anxiolytics and appetite suppressants have been dropping since the mid-1980s, as physicians have recognized the problems of tolerance and addiction. The

so-called specific serotonin re-uptake inhibitor (SSRI) and serotonin and norepinephrine re-uptake inhibitor (SNRI) anti-depressants have to some extent replaced the anxiolytics but there is concern, especially in the United Kingdom, at the rate at which prescriptions are increasing. These drugs seem to be prescribed as first line therapy rather than as an adjunct to counseling, cognitive therapy and outdoor exercise. They occasionally trigger suicidal thoughts and self-harm and are not recommended for the under-18s.

The disappearance of antibiotics from the US top subgroups above comes as a surprise. On a world basis, antibiotics are the most important drugs, but they are cheap and are prescribed for a single bout of infection, whereas many drugs are taken for the remainder of a sufferer's life. The global market was valued at $22.6 billion in 2006, 2% down on 2004.[1]

While only steroids against asthma come high on the above lists, the chemical class as a whole plays an unusual role, being the active component of oral contraceptives; some anti-asthma drugs; some anti-inflammatories; and conjugated estrogenic hormone, taken as hormone replacement therapy by menopausal women. Some other chemical classes resemble steroids in being therapeutically active in a number of areas, and the most important of these are shown in Table 45.3.

C. National differences in prescribing

In 2007, world sales of pharmaceuticals topped $600 billion for the first time. Figure 45.2 shows the sales by country.

TABLE 45.3 Chemical Groups with a Range of Pharmaceutical Applications

Chemical class	Example	Therapeutic use
Phenylpiperidines	Meperidine (Pethidine)	Opioid analgesic
	Loperamide	Anti-diarrheal
	Haloperidol	Neuroleptic
	Terfenadine	Anti-histamine
Sulfonamides/ sulphones	Sulphamethoxazole	Antibacterial
	Acetazolamide	Carbonic anhydrase inhibitor used in glaucoma and (infrequently) as a diuretic and in epilepsy
	Glyburide	Hypoglycaemic
	Dapsone	Antileprotic/antimalarial
Steroids	Ethynyloestradiol, Norethindrone, Medroxyprogesterone	Sex hormones used in oral contraceptives and for menstrual and menopausal disorders
	Prednisone	Anti-inflammatories in arthritis; immunosuppressives in transplant operations
	Beclomethasone dipropionate,	Anti-asthma
	Fluticasone propionate	
	Estramustine phosphate	Prostate cancer
	Testosterone esters	Anabolic hormones used for body-building after surgery
Prostaglandins	Prostaglandin F2 and E2	For abortions and to induce labor F2 reduces intraocular pressure and may be of use in glaucoma
	Misoprostol	For gastric ulcers
	Prostaglandin E1, iloprost, cisaprost	Vasodilator used to treat vascular disease of the leg. Reynaud's disease. Also for ductal-dependent congenital heart disease in new-born babies
	Alprostadil	Erectile disfunction following surgery
	Iloprost (prostaglandin PGI2)	For pulmonary arterial hypertension
	Prostacyclin (epoprostenol)	Inhibitor of platelet aggregation when blood contacts nonbiological systems, as in renal dialysis, as a pulmonary antihypertensive, and as a vasodilator

North America predominates. The Canadian contribution is small, so the United States accounts for about 45% of the global market. Europe accounts for another 30%. Even with Australia added in, the contribution of the developing world is tiny.

National differences in prescribing can be identified and are related to national wealth, national culture and attitudes toward the role of drugs in medical treatment. Typical countries, whether rich or poor, spend rather over 1% of GDP on pharmaceuticals, but there are wide variations. The United States tops the list with 2.02% in 2004. France comes second with 1.39%, and Switzerland and the Netherlands are the least extravagant in the OECD with 0.91and 0.74% respectively.[2]

FIGURE 45.2 World pharmaceutical sales 2006 (total $607.9 billion). *Source*: Based on IMS health; Figures are in $ billion.

Drugs represented, on average, around 18% of total health spending in OECD countries in 2003. The share ranged from highs of almost 30% or more in the Slovak Republic and Hungary, to lows of around 10% in Denmark and Norway. Drugs are cheap in France and Italy, expensive in the United States and Japan. Prescription numbers per head per year are 32 in France,[3] 14 in the United Kingdom, 12 in the United States and 8 in Germany.

Expenditure per head per year in different countries varies dramatically as shown in Figure 45.3. The figure also illustrates the proportion of healthcare costs paid privately, indicating that the Americans, the Canadians and the Mexicans pay a high proportion of their own costs while the Irish, Slovaks and Luxembourgeois (and the British although they are not shown on the figure) are treated from the public purse.

Inevitably, spending in third world countries is low. In 2004, the ten Asian countries China, Hong Kong, India, Indonesia, Malaysia, Philippines, Singapore, South Korea, Taiwan and Thailand, with a total population of about 3 billion, spent about $10 per head on drugs. Bottom of the league comes Bangladesh with a spending of $4/head.

It is perhaps harsh to attribute low expenditure in poor countries to national culture, but it certainly needs to be invoked to explain differences between developed countries. A belief in the prophylactic benefits of taking medicines is widespread in Belgium, France, Greece, Italy and

FIGURE 45.3 Drug expenditure per capita, OECD countries 2004. *Note*: (2) 2002; (3) 2003. Data are expressed in US dollars as purchasing power parities (PPPs), which allows countries to be compared on a basis that compares their values with the cost of a given basket of commodities in each country rather than standard exchange rates. *Source*: OECD Health Data 2006; http://www.oecd.org/dataoecd/5/24/36984917.pdf.

Spain but less prevalent in Britain, Ireland, the Netherlands and the Scandinavian countries. Germany, Austria and the United States have a culture with elements in common with both groups.[4]

Precisely what constitutes the cultural differences is unclear. It has been suggested that it is related to the proportion of Roman Catholics in the population, but that scarcely explains the gap between Italy and Ireland. There is a loose correlation with climate, people from more temperate climates requiring fewer drugs, but that does not explain the difference between Belgium and the Netherlands. The perception of physician and patient of their roles, however, is at the root of the discrepancies. A patient visiting a French physician will emerge with a prescription in 80% of cases.[5] In Germany and the United Kingdom, this drops to below 70%, and in the Netherlands it falls to 43%.

The giving of a prescription is seen to a greater or lesser extent in high and low-prescribing countries as a way of signaling the conclusion of a consultation. To hand out a prescription reinforces the physician's image of him or herself as a physician (that's what physicians are supposed to do) and the patient's image of him or herself as a patient (patients are given medicine; by giving me a prescription, the physician is confirming that I am ill, have not troubled him or her unnecessarily and am justified in taking the day off work to visit the surgery).

The differences between countries are less clear-cut, none the less, than the above figures imply. The drugs discussed so far are those that must be prescribed by a physician. Some drugs are classified as "generally regarded as safe," however, which means that they can be sold direct to the public, that is OTC, rather than on prescription. Diet supplements are apparently much more popular in Germany than in the United States but, in fact, the Americans buy them OTC instead of having them prescribed. This reflects both culture and systems of health insurance reimbursement.

Apart from differences in quantity, there are disparities between countries in the types of drugs prescribed, although patterns of prescribing in the developed world are drawing closer together. Meanwhile, in the United States and United Kingdom, almost all the leading drugs are therapeutically active; in Germany, France and Italy, many prescriptions are for "comfort" drugs that are generally agreed to have little therapeutic value. The Germans show enthusiasm for ointments based on non-steroid anti-inflammatories and salicylic acid derivatives. They also take peripheral vasodilators and anti-hypotensives, drugs that are rarely used by the British and Americans. Germans are concerned, as are the French, with problems of the digestive system and take drugs that supposedly reinvigorate the gall bladder. The French also take peripheral vasodilators and numerous drugs for the *crise de foie.*

In comparison, the British and Americans take large numbers of antibiotics. These are indeed therapeutically active, but the number of prescriptions compared with Germany suggests that in many cases they too are prescribed as "comfort" drugs for conditions that are either self-limiting or viral infections insensitive to antibiotics or both. Patients with a sore throat demand antibiotics, despite the fact that studies have shown antibiotics to be of little or no benefit for minor throat infections, and general practitioners prescribe them anyway. The problem has not been ignored by the authorities and there was a campaign at the end of the 1990s. In the United Kingdom, antibiotics prescriptions dropped by 20% between 1999 and 2001.[6] In Germany prescriptions for products "of disputed efficacy" dropped by 13.4% between 1998 and 1999. Over that period the Germans saved DM4.8 billion (about 2% of the drug budget) by bringing pressure on physicians.[7] Between 1992 and 2004, the number of prescriptions plunged from 1,063 million to 570 million (but back up to 648 million in 2005).[8]

That useless drugs (notably those that were sanctioned before the regulations were changed in 1976) are still prescribed is scandalous. The fact remains that herbal and homeopathic cures enjoy considerable political support, not only in Germany but throughout the developed world. The situation in the third world is much worse. A 2003 survey indicated that 39% of the drugs supplied by German companies to the third world were irrational, admittedly a reduction from 67% in 1984/1985. Some drugs were actually dangerous. In the developed world, aspirin is contraindicated for infants because of the danger of Reyes disease. Bayer none the less sold it in Brazil in 2003 as a pediatric formulation.[9]

A problem where it is more difficult to assign blame is that many drugs do not work for many patients. Allen Roses, Head of Genetics at GlaxoSmithKline, wants to test the DNA of patients in order to identify those for whom a particular drug will work – the "responders." Such a pharmacogenomic approach would enable physicians to avoid giving useless drugs to "non responders."[10]

As noted, the above tables exclude OTC drugs. OTC sales vary between countries, Switzerland having the highest level recorded in the literature at 28.8% of drug sales and Spain coming lowest at 5.1%. The OTC market in 2006 amounted to $93 billion, about 12% of the total market.

In most developed countries, the therapeutically active constituents of OTC drugs are largely confined to non-narcotic analgesics, anti-histamines, antacids and anti-inflammatory steroid ointments. In some countries, antibiotics are available without prescription, and opioid analgesics such as codeine are sometimes permitted mixed with aspirin or paracetamol. As governments try to increase the proportion of drug costs paid by patients, the range of OTC drugs increases. Compounds that have moved from prescription only to OTC in the United Kingdom in recent years include the antiviral, acyclovir; the H2-antagonists, ranitidine and famotidine; the steroid ointment, clobetasone; the hair-growing compound, minoxidil; amorolfine

against fungal nail infections; and various nicotine replacement therapies. A weight loss drug, orlistat, has been approved for OTC use in the United States, and there are moves to extend the basis for OTC classification from drugs for mild conditions to those for chronic illnesses – high blood pressure, high cholesterol and osteoporosis.[11]

The other major group of drugs omitted from the prescriptions list is those given in hospitals. Typically for developed countries they amount to about 10% of drug sales, although in the United States it rises as high as 22%. Drugs that are largely confined to hospitals include numerous injectable antibiotics (e.g. third generation cephalosporins, lincomycin and vancomycin) used to counter potentially resistant bacteria found in hospitals, together with blood products (e.g. albumin), intravenous products (e.g. total parenteral nutrition, dextrose) and the anticoagulant heparin.[12]

II. SOURCES OF DRUGS

Drugs are obtained from five sources: animal and vegetable extracts, biological sources, fermentation and chemical synthesis. Our ancestors, lacking synthetic skills, drew primarily on animal and vegetable sources. Macbeth's witches brewed up "eye of newt, toe of frog and liver of blaspheming Jew," which must have had a largely psychological effect, not to mention encouraging anti-Semitism, but they remembered to add pharmacological activity with "root of hemlock picked in the dark."

A. Vegetable sources

Plants continue to this day to provide a range of medicinal alkaloids (e.g. papaverine, atropine, codeine, quinidine), glycosides (e.g. digoxin), steroids (e.g. diosgenin, stigmasterol, sitosterol, the raw materials for medicinal steroids) and vitamins (e.g. vitamin E from soya bean oil footstocks). This is discussed in Chapter 8. Traditional Chinese medicine is providing leads for new drugs,[13] and Artemisinin from the herb qinghaosu or sweet wormwood is already used against malaria. The anti-cancer drug paclitaxel (Taxol), a diterpenoid ester, was at one time produced in 0.02% yield from the stem bark of the Pacific yew tree *Taxis brevifolia*. This could not be an adequate source, and Bristol-Myers-Squibb developed a plant cell culture method to grow cell lines from yew trees in large fermentation tanks, using only water, sugars, vitamins and trace elements.[14]

B. Animal sources

Animal sources are less convenient than vegetable sources for production of commercial quantities of drugs, but some pharmaceuticals are obtained from slaughterhouse wastes. Heparin and bile acids are examples. Cow and pig insulin have largely been displaced by human insulin, generated using recombinant DNA. Conjugated estrogenic hormone is obtained from the urine of pregnant mares, with the bonus that the animal does not have to be slaughtered. Gelatine is used to encapsulate active materials.

Traditional Chinese medicine draws heavily on animal products to the point at which it threatens the survival of tigers, rhinoceros and other large animals. It none the less has provided a number a leads for new drugs. Black bear bile is a particular source of active compounds. Tauroursodeoxycholic acid is already in use against human cholestatic disease, and 4-phenylbutyric acid has been approved for treatment of urea-cycle disorders. Other bile compounds could be used to treat type 2 diabetes, but all these materials, if they are to be widely used, require synthetic methods rather than the mass slaughter of black bears.[13]

Human blood, given by blood donors, is a source not only of cellular material (red cells, white cells, platelets) but also of blood plasma, which is subjected to protein fractionation to give albumin, anti-hemophilia factors and immunoglobulins. The possibility of transmission of Creutzfeld-Jakob disease (CJD) – prions are not really understood – and acquired immune deficiency syndrome (AIDS), if sterilization procedures fail, combined with the feasibility of manufacturing proteins by recombinant DNA technology, mean that there is uncertainty about the long-term future of the blood products industry. Indeed, the problems with bovine spongiform encephalitis (BSE) and the related fatal human brain disorder, CJD, have led to unease about cattle-derived substances.

C. Biological sources

Biological sources are primarily used for vaccines. Vaccines are suspensions of living or killed micro-organisms, components or products thereof. They are produced in living systems. Eggs are the most widely used medium at present, notably for influenza vaccines, but there is increasing use of animal cell cultures. Bacterial vaccines, for example, the diphtheria vaccine, can be cultured. There is an overlap here with fermentation processes (Section II.D.).

Recombinant DNA processes can be used to make subunit vaccines. These can be classified either as biological or fermentation processes. Hepatitis B vaccines are made by genetic manipulation of yeast or animal cells. The cells then express the hepatitis B virus outer coat protein, which is harmless to the host because there is no DNA present. Similarly, in 2007, two companies launched a vaccine for cervival cancer caused by the human papilloma virus (HPV).

There are two types of immunization, active and passive. In active immunization, the vaccine is administered to the patient, whose immune system responds by producing

antibodies and cells sensitized to the vaccine and hence conferring immunity to the disease. In passive immunization, a donor who has had a specific disease donates blood, and an immunoglobulin is isolated from the blood plasma. This is then injected into the patient.

Some immunoglobulins cannot be obtained from normal, healthy blood donors and are obtained instead by inoculation of volunteers, so that they develop the required antibody. Such procedures are used for immunoglobulins against vaccinia, tetanus, measles, hepatitis A, Rhesus incompatibility and rabies.[15] The biological "factory" for the immune globulin, in such cases, is the body of the volunteer.

Monoclonal antibodies are coming onto the market at an increasing rate. For example, abciximab is an anti-platelet drug, trastuzumab is active against breast cancer, and adalimumab and inflixamab treat rheumatoid arthritis. The suffix "mab" is a convenient guide to products in this category.

D. Fermentation

Fermentation in its simplest forms, such as the production of bread or wine, is very old technology. In its application with recombinant DNA to produce pharmacologically active proteins (erythropoetin, human growth hormone, human insulin), it is central to the modern biotechnological revolution. In between, it is the method for production of antibiotics and vitamin B_{12} and provides a step in the production of vitamin C and the synthesis of cortisone from diosgenin.

Plant cell culture might be classified as fermentation or production from a biological or vegetable source. It is used for Taxol production (see above) and, by the Japanese, to produce shikonin, a red pigment with anti-inflammatory properties.

As noted above, fermentation processes overlap with biological processes in the production of biopharmaceuticals. These rely on the effective part of a living organism, typically a fragment of DNA, being removed from a plant or animal and then, after being purified, being inserted into host cells that can be induced to promote its growth. The most important hosts are bacteria such as the satisfactorily large *E. coli* cell, mammalian cells, and viruses. These become manufacturing factories producing the desired biopharmaceutical. Once produced, however, the product has to be separated; in particular, all traces of the host cells must be removed. The biopharmaceuticals are produced at high dilution and the whole process is complex. Biopharmaceuticals typically cost 10–20 times the amount of conventional pharmaceuticals.

E. Chemical synthesis

In spite of the significance of the above methods, chemical synthesis is the most important method of drug manufacture. It is responsible for almost all heart drugs, drugs for the CNS, anti-ulcer drugs, analgesics and anti-histamines. Many pharmaceuticals are made simply from readily available bulk chemicals. For example, the non-steroid anti-inflammatory agent, ibuprofen, is made from toluene and propylene, two of the seven basic building blocks of the petrochemical industry.[16] Chemical synthesis is also used to modify materials made from other sources, and chemical methods of extraction are used in the downstream processing of products from other sources.

The combination of methods illustrates the eclecticism of the pharmaceutical industry. The use of fermentation in the production of vitamin C and cortisone was mentioned above. In the production of the semi-synthetic penicillins, penicillin G or V is first made by fermentation. It is then cleaved to 6-aminopenicillanic acid by an immobilized enzyme, and a new side chain is added by chemical means.

III. MANUFACTURE OF DRUGS

The largest tonnage organic chemical is ethylene, made in the United States and in Western Europe on a scale of about 26 and 22 million tonnes/year, respectively. It sells for about $0.5/kg. The largest medicinal chemicals are aspirin, paracetamol (APAP) and vitamin C, made worldwide on a scale of 30,000–50,000 tonnes/year. A total of about 149,000 tonnes of medicinal chemicals was manufactured in the United States in 1992 (more recent figures are not available). Ninety-two thousand tonnes of these were sold for $2.4 billion at an average price of $25.90/kg.[17] These medicinal chemicals were formulated into products, however, which sold for about $35 billion, amounting to almost $200/kg of raw material. A Merck executive remarked "We buy chemicals at the price of silver and sell them at that of gold."

The major drug producers are the United States, Japan, France, the United Kingdom, Germany, Italy, Switzerland, Ireland and Spain, with Sweden, the Netherlands, Denmark and Belgium being smaller but still significant.

The manufacture of pharmaceutical chemicals in some ways resembles a scaled-up version of the organic chemical syntheses carried out in the laboratory and, in others, a scaled down version of the processes used in the heavy chemicals industry. The heavy chemicals industry manufactures commodity chemicals that are largely undifferentiated and have to be sold at the ruling market price. The individual company, having little control over prices, is therefore preoccupied with reducing costs. A major source of cost reduction is economies of scale, and those in turn are related to the use of continuous rather than batch processes. The former are well-nigh universal in the heavy chemicals sector.

The pharmaceutical industry, in contrast, manufactures a large number of chemicals in small quantities to a high level of purity, which is of greater importance than high yield.

Economies of scale are relatively inaccessible. The product, especially if it is under patent, is highly differentiated and can be sold, like other speciality chemicals, at a price that reflects its value-in-use rather than its manufacturing cost. The industry is preoccupied with quality control, and this is more easily achieved in batch equipment. There is, none the less, increasing interest in continuous processes where, for example, the control of inventory levels enables highly reactive species to be handled safely.

The priority given to quality leads to certain design features not normally found in chemical plants. An analytical department intervenes at every stage to monitor the progress of a drug through the system. There is a code of good manufacturing practice (GMP) for the pharmaceutical industry, which lays down guidelines within the European Community,[18] and this is matched by regulations of the Food and Drug Administration (FDA) in the United States. These dominate the organization of pharmaceutical manufacturing sites. Uniform standards are desirable, and a common technical document for submissions for new drug approval and an agreement on common GMP guidelines were agreed at San Diego in November 2000. The implementation of these, so that different countries would mutually recognize each others plant inspection procedures, is proving more difficult.

A. Good manufacturing practice

Medicinal products must be fit for their intended use, comply with the requirements of Market Authorization and not place patients at risk because of inadequate safety, quality or efficacy. These objectives are easy to state, but to achieve them requires a comprehensively designed and correctly implemented system of quality assurance. It is important that this is documented and its effectiveness monitored. Records must be kept that are open to inspection by validating bodies, and there must also be procedures for self-inspection and quality audit that allow appraisal of the quality assurance system. Management must be adequately trained and their responsibilities minutely defined. Two key posts are the Heads of Production and of Quality Control. These posts are required by the guidelines to be independent of one another.

Quality control is part of GMP. It is concerned with sampling, specification and testing, and with the organization, documentation and release procedures that ensure that tests have indeed been carried out. No materials must be released for use or sale until their quality has been judged satisfactory. For example, the purity of most drugs is greater than 99.9%, and the content of the remaining 0.1% is usually specified in the Pharmacopoeia and in the quality control specifications. Absorbent cotton fibers may present a problem, and ointments are specified to contain less than 1 fiber/cm^3, the fiber being less than 1 mm long.

B. Plant design

A large pharmaceuticals manufacturer will operate several different plants. The plant for the manufacture of active ingredients by chemical synthesis will be separate from the plant that formulates the active ingredients into finished medicines. Fermentation plant will usually be on a totally different site geographically from a chemical synthesis plant. The reason for this is to minimize contamination by the ubiquitous *Penicillium* micro-organism, among others. Its concentration in finished products must be kept below 1 ppm to avoid allergic reaction in penicillin-sensitive patients. Bacteria insensitive to an antibiotic may grow in a medium containing it, and most countries demand the absence of harmful varieties (e.g. salmonella) and specify a limit for total bacterial count. Thus the US Pharmacopoeia quotes 5,000 cm^{-3} for a gelatine base. There is also likely to be a geographically separate for plant making biological medicinal products, and these have separate regulations (Ref. 18, p. 87).

As far as equipment is concerned, the pharmaceutical industry rarely follows the chemical industry pattern of purpose built plant, carefully optimized and with the larger sections fabricated on site. In general, equipment is bought "off the shelf" and assembled like a child's erector set.

Optimization is sacrificed to availability of plant items. A secondary benefit to emerge from this is that lead times (i.e. the time taken to build plant) are lower than in the chemical industry. Because kilogram quantities of a drug are required for testing before it is known whether a drug will be a success or not, the initial manufacturing process usually involves a batch preparation on general purpose equipment. There may be a need for changes to the initial method subsequently on economic or safety grounds (e.g. one of the reagents may be pyrophoric), but the need to register a validated process means that there will be a conservative attitude to process change. Any new process or important modification of an existing process must be validated, and critical phases of an existing manufacturing process must be regularly revalidated. If the manufacturer wants to add a recycle stream, for example, clearance must first be obtained.

The equipment supplier will therefore aim to offer versatile, multi-purpose equipment that can be used in a range of processes. Most of it will be stainless steel or glass lined, to avoid problems with contamination or corrosion. Wherever possible, a single reaction vessel is dedicated to one particular process, and usually each preparation is performed in a self-contained area. This situation has eased somewhat in recent years, and more multi-use equipment is being installed. There is still not the flexibility in equipment use that one might expect in comparison, say, with the dyestuffs industry, but there is a trend in that direction.

Effluent disposal is a central problem. Yields of pharmaceuticals are low compared with the heavy chemicals

industry, and waste streams, perhaps containing pharmacologically active materials, must be incinerated or scrupulously purified before discharge. This is known as "end-of-pipe" technology. Much preferred is the design of synthetic processes to yield fewer or less toxic effluents, and an example is given in Section III.E.

An exception to the rule of small, general purpose equipment arises with fermentation processes. A reactor producing penicillin G will certainly be dedicated and may have a capacity of 250 tonnes and a 30,000 hp stirrer. Because yields are so low, the weight of penicillin produced per run will be two orders of magnitude smaller. One of the problems of downstream processing is to reduce the volume of the penicillin-containing stream to more manageable proportions and to dispose of the effluent fermentation liquid without imposing an unacceptable biochemical oxygen demand on the aqueous environment.

The availability of online microcomputers and microprocessor controlled equipment is extending many of the benefits of continuous processing to the batch processor. The pharmaceutical industry has been a leader in the employment of these new techniques.[19,20]

C. Formulation and packaging–sterile products

Plant for the production of active ingredients resembles plant for the production of other small tonnage chemicals. Formulation and packaging plant, in which the active ingredient is converted into a saleable form, is less familiar to the industrial chemist. The reason is the need for a controlled environment. Some straightforward chemical products to be taken by mouth do not need to be tableted in a sterile environment. For others, such as injectable antibiotics, it is essential. Although the regulations do not demand sterility for all products, most companies would carry out all their formulating in a controlled environment, the severity of which would depend on the nature of the product. Walls and floors are rounded off to allow efficient cleaning. Sinks and drains are avoided as far as possible. Packaging is carried out in individual cubicles to avoid cross-contamination if there is spillage. Production of sterile medicines is carried out in clean areas whose entry is through airlocks both for people and goods. The areas are supplied with suitably filtered air and are classified A, B, C or D according to the required characteristic of the air. These are shown in Table 45.4. The actual regulations are more detailed and the table merely summarizes a few of the points. In particular, it illustrates the need for personnel to wear suitable clothing that will prevent contamination by viruses, bacteria or particulate matter.

Manufacturing operations for sterile products are divided into two groups, one where the product is sterilized terminally, that is at the end of processing after it has been sealed in its container, and the other where some of the processing must be conducted aseptically. For terminally sterilized products, a grade C environment is usually

TABLE 45.4 Air Classification System for the Manufacture of Sterile Products

Grade	Maximum permitted numbers of particles per m^3		Maximum permitted number of viable micro-organisms	Approximate equivalent US Federal Standard	Clothing for each grade
	0.5–5 μm	>5 μm			
A[a]	3,500	None	<1[b]	100	
B	3,500	None	<5[b]	100	Clean sterilized protective garments provided at least once a day. Gloves disinfected during operations. Masks and gloves changed every working session. Disposable clothing a possibility.
C	350,000	2,000	100	10,000	Hair covered. Two piece non-fiber-shedding trouser suit gathered at wrists and with high neck required.
D	3,500,000	20,000	500	100,000	Hair covered. Protective clothing and shoes required.

The number of air changes should generally be higher than 20 per hour in a room with good air flow pattern and appropriate high efficiency particle air (HEPA) filters.
[a]Laminar air-flow work station. Systems should provide homogeneous air speed of 0.3 M/s vertical and 0.45 M/s horizontal flow.
[b]Only reliable when a large number of air samples is taken.
Source: Ref. 10.

adequate. Aseptic preparation can be performed in a grade C environment provided the product is sterile filtered later in the process. If not, a grade A zone with a grade B background is required.

The provision of sterile areas in a packaging and formulation plant means that such a plant may cost as much as a conventional manufacturing plant for the active ingredient. The tableting and packaging equipment are relatively inexpensive, and most of the money is spent on equipment to produce a controlled environment.

The formulation of active ingredients into a satisfactory dosage form is a science in itself. It is beyond the scope of this chapter, but a brief discussion is given in Ref. 16.

D. Choice of reagents

The special characteristics of the pharmaceutical industry also affect choice of reagents. In general, processing costs are higher than in the chemical industry as a whole, and capital costs are lower. Because of the high unit value of pharmaceuticals, expensive chemicals are practicable. For example, if a reducing agent is required, lithium aluminium hydride might be used while, in the petrochemical industry, only hydrogen would be economic.

Pharmaceutical companies are involved in long, multistep synthetic processes. A modest way in which they try to keep costs within bounds and to reduce the number of steps they have to carry out themselves is by the use of standard intermediates – synthons – that can be used for a number of products. For example, all the different chemically modified penicillins are based on 6-aminopenicillanic acid, which can consequently be manufactured on a much larger scale than any of the individual penicillins. Synthons and other high quality raw materials, reagents and key intermediates are frequently purchased from fine chemical manufacturers. These manufacturers are often prepared to develop reagents and intermediates specifically for one customer in the expectation that they will have further uses in the future. Reagents that increase reaction selectivities are often identified in academic laboratories but cannot be used on a tonnage scale because of lack of availability or cost. Fine chemical companies can find niche markets by developing methods of making these reagents cheaply available. Examples are Hünig's base, *N*,*N*-diisopropyl-*N*-ethylamine[21] (i-$C_3H_7)_2NCH_2CH_3$, and oxalyl chloride[22] $(COCl_2)_2$. An improvement in selectivity is economically attractive, not only because it gives higher yields, but also because it produces less waste; the reduction of waste is a key environmental issue.

Companies may contract out some intermediate stages in a drug synthesis (see Section III.G.) or conduct them in an isolated part of the plant, if they offer particular hazards. Nitrations would probably be carried out in a special nitrating plant, with equipment designed to prevent runaway reactions; phosgene would be handled only under carefully controlled conditions because of its toxicity. Experience is similarly needed with organofluorine compounds, reactions at cryogenic temperatures and fermentation routines for steroids and antibiotics. A firm that has been engaged in such operations for many years accumulates know-how, and this gives them a great advantage over new entrants to the industry.

Chirality is an issue. In May 1992, the FDA published a policy statement "strongly urging companies to evaluate racemates and enantiomers for new drugs" and declaring that drugs be sold only in their enantiomerically effective form. This hard line has softened, and the marketing of single enantiomers is not yet mandatory. Since Monsanto's pioneering asymmetric synthesis in the early 1980s of L-dopa, a drug for Parkinson's disease, chiral syntheses and separations have been of intense interest. This was accentuated by experiments with mice showing that it was the D-isomer of thalidomide that was responsible for birth defects, while the L-isomer was the hypnotic. With rabbits, however, both forms were active. Meanwhile, thalidomide has returned to the market with applications in leprosy, Hansen's disease, tuberculosis, lupus and cachexia (the weight loss and metabolic wasting associated with AIDS).[23] It has orphan drug status (Section IV.C.) in some of these applications. Furthermore, Astra-Zeneca were able to patent Nexium (esomeprazole), the active (*S*)-isomer enantiomer of their proton pump inhibitor Prilosec (omeprazole). With 2006 global sales of $6.7 billion, this was no mean achievement, although the cynic might ask whether the ingestion of the less active enantiomer ever harmed anyone.

Sometimes both enantiomers of a drug are pharmacologically active and, in such cases, there is no reason to favor one over the other. In other cases, a chiral drug, when administered, racemises in the body anyway. Nonetheless, there are numerous examples of drugs where the administration of a racemic mixture means that patients are needlessly exposed to pharmacologically inactive chemicals. Even if the racemic mixture is synthesized but separated in the production stage, half of the production may be wasted. It might in the end be cheaper to make the correct enantiomer to start with. Hence the interest in chiral intermediates and chiral catalysts. There are many specialist firms trying to find niche markets in this area.

Choice of reagents has been detailed encyclopedically in a remarkable 1,300-page single-author book by an Indian chemist C. Someswara Rao.[24] Readers who wish to move from broad generalizations to specific situations will find no better compendium than this.

E. "Green" chemistry

Concerns about the spread of physiologically active chemicals in the environment have prompted attempts to devise

"green" routes to desired chemicals that generate fewer and less toxic effluents. This is not always possible but is a desirable goal. An example is the synthesis of the antidepressant sertraline. An alternative synthetic pathway was devised in which the three steps in the original process were reduced to one (Figure 45.4).[25]

In the original synthesis,[26] 3,4-dichlorobenzoyl chloride and benzene were converted to a benzophenone by a Friedel Crafts catalyst. This was originally aluminum chloride, and it is not disclosed if it has been replaced by a zeolite. A side chain was built onto the ketone group and cyclized, again in a Friedel Crafts reaction, to give (±)-4-3,4-dichlorophenyl)-3,4-dihydro-1(2*H*)-naphthalenone. The three stages that followed were later converted to one. The carbonyl group of the naphthalenone was converted to an imine with methylamine in tetrahydrofuran or toluene. Titanium tetrachloride was used as a dehydrating agent to drive the equilibrium toward the imine. The imine was isolated and hydrogenated with hydrogen and a palladium/carbon catalyst in tetrahydrofuran to form isomers (*cis:trans*, 6:1). Fractional crystallization isolated the racemic *cis* hydrochloride. This was finally resolved with d-mandelic acid in ethanol to give the desired (S,S)-*cis* isomer. The sertraline mandelate was then converted to the hydrochloride in ethyl acetate.

In the new process, the three stages starting with the naphthalenone are carried out in ethanol without isolating the intermediates (red arrow in the figure). Only the environmentally benign ethanol is used as solvent, and the unreacted methylamine is recovered by distillation. Titanium tetrachloride is no longer required to drive the first step, since the imine is poorly soluble in ethanol and precipitates. Palladium on calcium carbonate provides better selectivity than palladium/carbon, giving a *cis:trans* isomer ratio of 18:1. Overall yield has been nearly doubled to 37%, and raw material usages for methylamine, (±)-4-3,4-dichlorophenyl)-3,4-dihydro-1(2*H*)-naphthalenone and D-mandelic acid have been cut by 60%, 45% and 20%, respectively. Solvent requirements have been reduced from 60,000 gallons to 6,000 gallons per tonne of sertraline. 440 tonnes/year of titanium dioxide-methylamine hydrochloride acid waste, 150 tonnes of 35% hydrochloric acid waste and 100 tonnes of 50% sodium hydroxide waste have been eliminated.[27]

A second example is the Bristol Myers-Squibb amoxicillin process which, during its first 5 years, was expected to eliminate 32 tonnes of hazardous wastes, six drying steps and ten solvents.[14]

F. Downstream processing

The reagents used by the pharmaceutical industry are more complex and sophisticated than those employed by the heavy organic chemical industry. The same is true of the purification and isolation of the products. The small simple molecules of the petrochemical industry can usually be separated from impurities by distillation. In the pharmaceutical industry, this remains true for solvents, less true for intermediates and rarely true for end products. In the case of medicinal products from biological sources or made by fermentation, it is never true.

Much downstream processing is performed by the well documented but less widely used unit processes of chemical engineering–solvent extraction, filtration, leaching, adsorption and crystallization. Freeze drying, centrifugation, ion exchange and preparative chromatography are valuable options. Cross-flow filtration has become important especially since the development of anisotropic membranes. These permit adequate flow rates but are still sufficiently robust to withstand high pressures and the wear and tear of industrial scale operations. Cross-flow techniques may be used for microfiltration, reverse osmosis and ultrafiltration.

Sterilization is a unit process not found in the heavy chemical industry. The aim of sterilization is to kill and/or remove bacteria and pyrogens. Pyrogens are fragments of bacterial cell walls, which may produce fever as a reaction to foreign proteins if they enter the bloodstream of a sensitive person.

For heat-stable materials, sterilization by steam for 15 min at 121°C kills all bacteria and viruses. Pasteurization for a longer period at a lower temperature is also effective. Dry heat at a temperature above 250°C destroys pyrogens.

Heat-sensitive materials may be sterilized by radiation, providing the product and its container are not radiation sensitive. It is particularly used with plastics, and more widely used with containers than with products. An alternative is ethylene oxide gas. This presents problems in that the gas must be brought into contact with the cell walls of contaminating bacteria. Ethylene oxide is highly reactive and can only be used when it can be proved not to react with the product.

Filtration through a 0.22 μm filter removes bacteria and moulds but not all viruses or mycoplasma. It is not considered adequate when terminal sterilization is practicable.

G. Outsourcing

Outsourcing, that is contracting out of all or part of one's business, has been a major feature of the pharmaceutical industry's policies since the 1990s. In 2004, it accounted for some $30 billion of business. Initially companies outsourced low technology activities such as cleaning and security. Then came stages of the manufacturing process that the outsourcing company preferred not to have to deal with, then parts of the testing. Testing drugs on humans had always been done in hospitals, away from company premises, but animal testing was also outsourced. More complex chemicals for use as raw materials

FIGURE 45.4 Synthesis of sertraline.

were purchased, so that the early stages of synthesis were accomplished by specialist firms. Companies making patent-protected drugs usually preferred to perform the final synthetic stages and the formulation in-house. In addition to outsourcing, there was also a trend toward collaborative projects. The drying up of the pipelines of many multinational drug companies has also led to the purchase of initial ideas from small "ideas" companies. These are then developed in-house.

The main reasons given for outsourcing were, first, that it enabled projects to go ahead where in-house resources were lacking. No money was saved, but the project was speeded up. Second, the company might gain access to specialized skills, for example, experience with particular synthetic reactions or with the formulation of such things as transdermal patches. Third, by outsourcing some activities, the company could focus on its core competencies. Fourth, a company could avoid setting up a specialist department to undertake a particular function, which might not be needed once a project was concluded. Meanwhile, cynics wonder about the possibility of a virtual pharmaceutical company, based in the Cayman Islands, that outsources all its functions.

IV. SOCIAL AND ECONOMIC FACTORS

The economics of the research-based pharmaceutical industry is as remote from classical economics as its technology is from that of the heavy chemical industry. Neither the manufacturers nor the consumers conform to the concept of a free market.

Producers of ethical pharmaceuticals are monopolists to the extent that they have patents on their products. Instead of conventional competition on price and quality, there is competition in innovation between rival producers, so that a number of manufacturers may offer, for example, different H_2-antagonists for the treatment of gastric ulcers. Insofar as the patient requires a particular product when it is under patent, however, he or she is faced by a monopoly producer.

At the consumer's end, the situation is even more confused. Drugs are taken by the patient, prescribed by the physician and paid for (to a varying extent in different countries) by the state or by insurance companies. Thus, the usual market constraints are lacking. On the other hand, the state, insurance companies and hospitals sometimes have much of the power of monopoly purchasers – so-called monopsonists. In many countries, the state licenses individual pharmaceuticals and sets their prices. In the negotiations with drug companies, the usual price determinants of supply and demand are replaced by political considerations.

The above factors mean that the market for pharmaceuticals is subject to severe distortions. Certain aspects will be discussed in this section, notably the pattern and cost of innovation, the role of patents, orphan drugs, the rise of generic drugs, parallel trade and the attempts by governments at cost containment.

A. Pattern and cost of innovation

The modern pharmaceutical industry dates back to the discovery of the sulfonamide drugs in 1935. The industry has grown since 1950 typically at 7–15% per year, and over 95% of drugs available originated after that date. An astonishing series of innovations has increased life expectancy and improved the quality of life in developed and, to a lesser extent, in developing countries. The industry has moved from a "molecular roulette" system, by which new drugs were developed on the basis of trial and error, to a sophisticated structure hinging on an understanding of the enzyme systems the drugs are intended to influence and, more recently, on a knowledge of the structure of the human genome.

The industry was surprisingly unregulated in its early days but, since the thalidomide disaster in 1960, regulation has become increasingly strict, and the cost of developing new drugs has escalated. Figure 45.5 shows the number of new products launched worldwide each year since 1994. The number averaged about 40 in the 1980s, reached a peak in 1996 and has since declined to not much over 20 per year. Biologicals, of which so much was hoped, are only slowly justifying their enormous research cost.[28]

The cost of introducing a new molecular entity (NME) to the drug market in the early 2000s was estimated at $800 million, measured in 2005 dollars.[29] An NME is a truly novel product, not a line extension of a drug already on the market. The figure includes the cost of failures, that is, drugs that were eliminated during testing. This high cost is accentuated by the increased time required for drug discovery and testing. The average new chemical entity takes about 12 years to reach launch. This has two consequences. First, the patent life remaining to the drug, during which its development costs can be recovered, is much shortened. Second, the return on the R&D outlay is delayed. To get a true cost for the development of a new drug, one must discount the expenditure to the moment of launch. The above figure assumes a discount rate of 9%. Various uncertainties and assumptions in the above type of calculation have been discussed.[30] A rough estimate is that the actual costs of producing the drug are about half the cost once it is discounted, and a lower discount rate would give a much lower cost.[31]

This huge figure explains why only large multinationals with generous cash flow from other projects can be serious players in the game of pharmaceutical innovation. Smaller companies may come up with bright ideas that they sell to the multinationals, but only the giants can undertake the

FIGURE 45.5 New molecular entities and biopharmaceuticals. *Source*: EFPIA.

testing and registration procedures. Indeed, not only is pharmaceutical innovation confined to large companies, but it is also confined to countries that can provide the resources and manpower for such high technology operations.

The countries where drug discovery has traditionally taken place are the United States, the United Kingdom, Germany, Japan, Switzerland and France. In 1987, six of the leading fifteen products were American, five British, two Japanese, one Swiss and one German. It was expected that, with technology transfer, the number of non-American contributors would increase, perhaps to involve Scandinavia, India or China. The reverse has happened. By 2006, seven of the top ten products were American, two British and one French. The European and Japanese industries have declined, and their research spend has dropped. Between 1990 and 2006, R&D investment in United States grew 5 times while in Europe it only grew 2.9 times. The United States has thus maintained its flow of new products, but at ever-increasing expense, so that it now accounts for over half of world pharmaceutical R&D. In December 2007, the European Commission and the European pharmaceutical industry launched an Innovative Medicines Initiative, in an attempt to reverse this trend, by increasing public and private research funding.[32–35]

Meanwhile, in terms of foreign sales of pharmaceutical products,[36] Switzerland, France, Germany, the United Kingdom, Sweden and the Netherlands showed a positive balance of trade in 2005, while the United States and Japan and everyone else were net importers. The reason for the large United States deficit is that United States companies are happy to manufacture abroad and, in particular, see Ireland as their gateway to Europe. The Japanese, although innovative, concentrate on their large home market and prefer to license abroad products that they might otherwise export.

The governments of the leading drug-producing countries are in a curious position. On the one hand, they are under pressure to limit the social security costs of drugs at home. On the other hand, they are anxious to promote foreign trade in pharmaceuticals and to ensure that their industry remains prosperous. The drug producers are collaborating on global measures which, they hope, will save money without wrecking the industry.[37] Countries with less drug-producing capability may have ambitions to build up their own industry, but otherwise are more determined to limit drug expenditure. The drug companies have been successful in fighting cost containment measures in terms of their overall profits, but these are made almost exclusively in the United States and Japanese markets.[38] The rest of the world contributes to overheads but little else. There were fears that the Medicare Drug Improvement and Modernization Act of 2003, implemented in 2006, would damage the crucial United States market. The act extended Medicare coverage to include prescription drugs. Ten million seniors gained drug coverage for the first time, and millions more obtained better coverage than they had before. After a year, the measure seems to have increased drug industry sales, but the downside is that the federal government is paying for nearly 60% of prescription drugs in the United States, and that worries the industry.

B. Patents

"I knew that a country without a patent office … was just a crab," said Mark Twain, "and couldn't travel any way but sideways or backwards." Patent protection is recognized as a cornerstone of the success of the industrialized market economies of Western Europe and North America in

the 19th and 20th centuries. Without it, individuals would have been discouraged from investing in innovation fearing that, even if their research was a success, they would never recover development costs because of cut-price competition from imitators.

As described in Chapter 44, a patent provides the discoverer of a "non-obvious" invention with a monopoly to exploit it for a specified period after its registration, in return for disclosure of details of the invention. The period in the United States is 17 years from the granting of the patent, while the European Community has adopted a European Patent Convention, which grants 20 years from the filing of the patent application. This applies also in some countries outside the Community, for example, Switzerland. In Japan, a patent lasts for 20 years from filing or 15 years from publication, whichever is shorter. As there is an interval between the filing and granting of a patent, all these patent lives are similar. Some countries, especially in the third world, do not have patent laws and their citizens are not bound by foreign patents. Such countries have difficulty gaining access to first world technology and are unable to export to countries that subscribe to patent conventions. Trade sanctions may be imposed on offenders. Meanwhile, the Trade-Related Intellectual Property (TRIPS) agreement, which came into force in 1995, declared that at World Trade Organisation countries had to recognize patents, 50 of them not having done so previously.[39] While the agreement was a landmark, its full implementation will take some time.[40]

The pharmaceutical industry invariably seeks patent protection for new drugs. The problem is that the time taken to develop and test a new drug has lengthened. In 1962, the process of invention and testing took on average only 2 years. By the late 1980s, this had lengthened to about 12 years and has remained more or less in this region.[41] At the end of the 1980s, Lis and Walker[42,43] found that effective patent life had shortened on average to between 6 and 8 years. The industry campaigned with some success for patent term restoration and achieved possible extensions of up to 5 years in most developed countries.

Problems now center on developing countries, and especially India and China, with their huge populations and ambitions to manufacture pharmaceuticals themselves. India joined the World Trade Organisation in 1995 and had to agree to start granting patents. The passage of the Indian Patents (Amendment) Bill of 2005 was seen as a step in the right direction, but a new provision excluded "incremental innovations." That was used as a reason for refusing a patent on Novartis's cancer drug Glivec, because a patent had previously been granted for its active ingredient imatinib. Novartis lost an appeal against the ruling in August 2007, but is now accusing India of not complying with the TRIPS agreement by the agreed deadline.[44] China is even more of a loose cannon.

An interesting scientific question raised by the rise of biotechnology is the point at which two drugs should be regarded as "identical." The definition affects not only whether a novel drug could be patented as a NME, but also whether an out-of-patent generic drug could be approved on the basis of an abbreviated new drug application (ANDA) (Section IV.D.).

With small molecules, there was no problem. If structures were different, apart from salts and esters, then the drug was a new chemical entity. With macromolecules, however, small and insignificant differences might occur which, it was felt, did not constitute genuine innovation. For example, two protein drugs would be considered the same if the only difference between them was due to post-transitional events; infidelity of transcription or translation; or minor differences in amino acid sequence. Different glycosylation patterns or tertiary structures would only be considered significant if the drug were shown to be clinically superior.

At the time of writing, there is no simple pathway by which biogenerics can be approved in the United States. This is of escalating importance. The biopharmaceutical market had reached $56 billion by 2006 and is expected to reach $100 billion by 2010.[45]

C. Orphan drugs

The spiraling cost of drug development also creates problems with drugs for rare diseases. About 4,000 rare diseases are described in the literature. Many are known by the names of those who first described them – Huntingdon's chorea, Paget's disease, Tourette's syndrome and so on. Many might or do respond to chemotherapy. It is in the public interest that such diseases should be treated, but the problem is who should pay. One might argue that social security funds should be willing to pay large sums for the relief of a tiny minority of the population, but there is a counter-argument that limited funds should be spread where they can do most good to most people. Whatever view is taken, research on rare diseases must appear uneconomic to pharmaceutical companies planning their research expenditure. Even if they find a drug to cure one, the disease may be so rare that they cannot hope to recoup their expenditure. Alternatively, a drug may be available, but they do not want to produce it, either because it is not patentable or because the cost of testing is too high. For a truly rare disease, there are problems assembling enough patients for statistical analysis of test results. Add to these factors that the companies fear litigation, damages and loss of their good name if things go wrong.

Compounds that no one wants to develop are known as orphan drugs.[46] Because of the difficulties surrounding their development, there was a vigorous campaign in the United states to alter the law. One of the principal spokesmen was David Abelow, who suffered from neurofibromatosis, which became widely known as a result of the film

The Elephant Man. John Merrick, the subject of the film, had it in a severe form.

As a result of the campaign, the "Orphan Drugs Act" was passed in 1983, which allows for grants, provides tax credits, eases regulatory pressures and permits 7 years' exclusive marketing rights for non-patentable drugs to companies making them. It is possible for the US Secretary of State to hold the product license and for the supplier not to be responsible for the safety of the product. An orphan drug was defined as one used for the treatment of a disease or condition affecting fewer than 200,000 people in the United States. If more than 200,000 people are affected, it could alternatively be shown that the cost of developing the drug and making it available would not be covered by sales in the United States.

It was 1999 before the European Community followed suit. They agreed full or partial exemption from registration fees, 10 years exclusive marketing, and help with chemical R&D for drugs for life-threatening or chronically debilitating illnesses with a patient population less than 5 per 10,000.[47]

In the United States, from January 1983 to June 2004, a total of 1,129 different orphan drug designations were granted by the Office of Orphan Products Development (OOPD), and 249 orphan drugs received marketing authorization in the United States. In contrast, in the decade prior to 1983, fewer than ten such products came to market.

In 2003, the leading orphan drug by worldwide sales revenue was Amgen's erythropoetin (Epogen), with US sales of $3.2 billion. It obtained orphan drug status for treatment of anemia associated with azidothymidine (AZT) therapy, prematurity and end stage renal disease in both dialyzed and non-dialyzed patients.

Clearly, erythropoietin is no longer a truly orphan drug, and there have been concerns that the act was not always working as intended. In some cases, compounds were registered as orphan drugs when a disease was rare, but then it became an epidemic, and the number of sufferers rose well above the 200,000 mark. For example, the AIDS virus was identified in 1981. By 1986 there were only 23,000 cases in the United States, and the drug AZT was genuinely orphan. Since then, many millions of cases have been diagnosed.

Even allowing for government grants and tax concessions, the cost of developing orphan drugs is high, and the price is bound to be substantial. The original price set for AZT drew widespread protests and was, in fact, reduced on the grounds that sales were likely to be much higher than had been anticipated.

There are other diseases whose incidence is likely to remain static. Cystic fibrosis is the commonest genetic disorder in Caucasian populations. It afflicts between 1 person in 1,600 and 1 in 8,000. In the United Kingdom, about 400 infants per year are born with the disorder and fewer than 25% would be expected to live until their thirties. The population of sufferers depends on survival rates but, with a United States incidence of 1 in 3,800,[48] a figure of the order of 10,000 is probable, well within the orphan drugs limit. Recombinant DNase, Pulmozyme (dornase alpha) is an orphan drug that breaks down the thick mucous secretions associated with the disease. It gained regulatory approval in 1994 but the cost per patient is about $10,000 per year.[49]

An orphan drug that caused a furore was alglucerase (imiglucerase, Cerezyme), a treatment for Gaucher's disease. Gaucher's disease is characterized by the presence of enlarged lipid-containing histiocytes (Gaucher cells) in the bone marrow. The cells cause bone pain and necrosis of some bones and make the patient liable to bone fractures. The disease (to which Ashkenazi Jews are genetically disposed) is due to a deficiency of the enzyme glucocerebrosidase. There are about 2,000–3,000 sufferers in Europe, and a similar number in the United States, who can benefit from enzyme replacement therapy. Ceredase is a modified preparation of glucocerebrosidase derived from placentas. A year's supply in the United Kingdom used to cost about $500,000. It was described as "the orphan drug that broke the camel's back,"[50] yet was none the less available on the UK National Health Service. It has since largely been replaced by Cerezyme, made by recombinant DNA technology, but the cost has not dropped substantially.

Thus, the Orphan Drug Act has been a success in that companies have been encouraged to develop drugs for rare diseases. On the other hand, the United States Government is currently paying the bill and, in certain cases, feels that it is being asked to pay excessively for drugs that could legitimately be asked to stand on their own. Whether it proves possible to "fine-tune" the legislation to make it more cost-effective remains to be seen. The issue has to some extent been overshadowed by the high cost of the new biologics (Section V.A.).

D. Generic pharmaceuticals

Innovative drugs, particularly orphan drugs, present economic problems for health administrators. The same is true at the other end of the spectrum, for drugs that have been on the market for many years and are out-of-patent. The original manufacturer is then open to imitation and competition. Out-of-patent pharmaceuticals fall into four categories[51]:

1. Those manufactured by their inventors under their brand names. For example, Hoechst still sells Lasix even though furosemide is out-of-patent.
2. Those marketed by non-originating companies under their own brand names. For example, amoxicillin is sold as Larotid, Polymox, Trimox, Utimox and Wymox, as well as under the originator's name, Amoxil. These are known as branded generics.
3. Out-of-patent pharmaceuticals marketed by non-originating companies under a generic name, plus a

company name or prefix or suffix. Furosemide is sold under its own name by Geneva, Lederle and Rugby, and the company name features on the label. The companies undertake limited promotional activity. In some statistics, these are included in branded generics.

4. Out-of-patent pharmaceuticals marketed by non-originating companies under a generic name, with minimum mention of the company's name. The bottle may be labeled furosemide BP or furosemide USP, and the producer appears only in small letters.

Categories 2 and 3 are the true generics. Group 4 (minimum name generics) are illegal in most European countries, since the physician must specify the source. Counterfeit generics pose a problem. In 2005, the FDA estimated that fake drugs comprised roughly 10% of the global medicine market. The situation has been exacerbated by the availability of drugs online over the internet. In some developing countries, over half the drugs on the market are fakes. An estimated 192,000 patients were killed by fake drugs in China in 2001.

Although some generics companies try to develop new dosage forms, calendar packs and so on, generic drugs in general sell entirely on price. This is resented by the research-based companies, who see their profit margins being eroded after what they regard as an inadequate period of patent protection. They present the public with a picture of hole-in-the-wall generics companies, flooding the market with inferior, inadequately tested drugs, made on the cheap, because the companies do not need to pay the costs of R & D. In fact, most generics are made by the same companies that manufacture ethical pharmaceuticals.

The claim that generics companies can manufacture more cheaply is doubtful. In general, the originating companies can manufacture out-of-patent products most cheaply. They have already amortized their plant, they have personnel already available, they have optimized the process, and the product can share analytical and distribution facilities with in-patent products. The reason that the research-based companies *appear* to have higher costs is an accounting convention that loads the costs of research and development of *new* products onto products already being sold. This is legitimate in that the cash flow for research and development must come out of current income, but it is an illusion to see it as a higher cost. Rather it is a question of cross-subsidy.

A generation ago, the issue of generics was not significant. Physicians prescribed drugs by the names they knew and loved, which were usually the brand names given by the originators.

In the United States, the situation was revolutionized by the Waxman-Hatch Act of 1984. Generics companies were granted the right to market out-of-patent drugs under their generic names, without having to repeat all the tests performed by the originator. Instead, they were able to submit an ANDA in which they had to prove that the chemical composition and bioavailability of the drug matched the standards previously agreed by the FDA and the originator. Products regarded as equivalent are listed in the "Orange Book."[52]

Generics accounted for 63% of medicines dispensed in the United States in 2006, at an average cost of $29.82, giving a total market of $54.1 billion. The market for branded drugs was worth $220.6 billion at an average of $101.71 per prescription. Thus, far more generics are being prescribed but at about 30% of the cost.[53] The proportion of prescriptions written generically is certainly much lower – perhaps only 15–20% – but laws in all 50 States permit the pharmacist to substitute a less expensive generic even when the physician prescribes a brand name. In some States, physicians must write "prescribe as written" if they want the patient to receive the branded product and, in other States, they must sign on one of two lines. Always they must specify that the branded product be supplied if that is what they wish.

In United Kingdom medical schools, physicians are trained to write generic names rather than brand names. By 2005, the proportion of prescriptions written generically had risen to 80.1% and the number dispensed generically to 59.3%, the 20.8% difference being where the prescription was written with a generic name, but the product was still under patent. Because it is the older, cheaper products that are available as generics, the 59.3% of prescriptions written generically were responsible for only 20% of expenditure.[54]

In France, the generics market has been largely confined to hospitals. About 3% of prescriptions by volume and value were filled generically in 2000, and this had risen to 13% by volume and 7% by value by 2005. Even in-patent drugs in France are cheap, and French physicians have hitherto been hostile to generics, although this is changing.

Generics established themselves from the early 1980s in Germany. Reference pricing, together with other extensive and complicated health service reforms, have boosted the generics sector so that it holds about 68% of the out-of-patent market by volume and 33% by value. The German generics market is the biggest in Europe, accounting for a little over half of its $11 billion value.

From a simplistic viewpoint, the distinction between originators' drugs and generics looks like the difference between, say, Kellogg's corn flakes and a supermarket's own brand. Two factors modify this picture. First, the consumer is able to tell if one brand of corn flakes tastes as good as another, and how much the difference is worth financially. Second, if the consumer does not pay directly for the product, then there is no reason to choose a cheaper one with an unpromoted brand name. As in all economics problems, it is the organization that pays that complains, and governments are anxious to promote cheaper drugs whether generic or otherwise.

In January 2008, the EU's Competition Commissioner raided major pharmaceutical companies to determine

whether they were acting illegally to prevent new drugs from coming onto the market. It seems unlikely that innovatory drugs are being delayed, in spite of the drop in their number, but the question was raised as to whether the companies were delaying the introduction of generic alternatives.[55]

E. Parallel trade

In the past, barriers to international trade were erected in almost every country. Imports of drugs were restricted by tariff or non-tariff barriers, to economise on imports, keep down the national health care bill and protect the local pharmaceutical industry. This is no longer the situation. The GATT agreement of 1994 established the World Trade Organisation and provided a framework for freer trade in drugs (and other goods). With the United States acting as the bulldozer, other traditional barriers may eventually be demolished. However, a different problem remains: the parallel trade in drugs. This situation occurs when the same product has significantly different prices in two countries. As noted in Section I.C., prices of a given pharmaceutical vary widely between countries. Companies tend to set higher prices in richer countries, although government intervention may distort the pattern. Traders can then buy the drug in the low-priced market and sell it in the high-priced one; this form of arbitrage forces prices to converge downward. The threat to the incomes of international pharmaceutical companies is obvious.

A long series of court cases in Europe has established, however, that parallel trade is legal within the EU, even when the price differences that make it worthwhile result from the actions of governments rather than companies. Priority was given to Article 30 of the Treaty of Rome, which forbids all measures that establish quantitative restrictions on imports, rather than to Article 36 of the treaty, which permits such restrictions on the grounds of public health or the infringement of intellectual property rights. Accordingly, parallel trade remains a major cause of anxiety for the research-based firms. Although it accounted for only 3.4% (about $5.125 billion) of the EU pharmaceutical market in 2005, the proportion was much higher in certain countries. In the United Kingdom it was 14.5% and in Germany 5.9%.[56] Furthermore, most of the products affected are popular and profitable patent-protected medicines, such as esomeprazole and atorvastatin.

The real worry for pharmaceutical companies has always been the possible extension of parallel trade. The most immediate threat is in Europe, where supply problems have in practice limited it. European wholesalers are mostly oriented to their national markets; they are therefore often unable to provide traders with the volume of products desired, and the pharmaceutical companies astutely try to supply each country only with sufficient supplies of a drug to meet its own needs. Early in 2007, Pfizer declared that it would distribute drugs in the United Kingdom only through a single wholesaler, Unichem. This was primarily an anti-counterfeiting measure, Pfizer having been damaged by fake Lipitor, but it also gives them greater control over the supply chain. As pan-European wholesalers emerge, the whole system may be reorganized; indeed, such wholesalers may practice parallel trade themselves.

Could parallel trade develop outside Europe? The TRIPS agreement specifically excludes this type of situation from the treaty's scope. Drug prices are high in the United States, lower in Canada, and very much lower in Mexico, so parallel importing from Mexico to the United States would be attractive. At present, the US Prescription Drug Marketing Act prevents the reimport of a patented product by a company other than the US patent holder. As neither the World Trade Agreement nor the North American Free Trade Agreement (NAFTA) deals with this issue, United States law prevails for the time being. There has none the less been dissatisfaction expressed that some United States drugs (e.g. Pfizer's Xalatan), whose discovery was partly funded by the Federal Government, cost many times more in the United States than in other developed countries, and various suggestions have been made to deal with the problem.[57]

There is also a question about the supply of drugs cheaply to developing countries, which might then be re-exported to more attractive markets. The South African Government decided in 2001 to import cheap "pirate" AIDS drugs. An Indian generics company offered to supply combination therapy at $350/year compared with the United States cost of $10,400/year.[58] The snag is that, in much of Africa, even the lower priced therapy costs several times the per capita income. The major pharmaceutical companies decided not to sue. Similarly, early in 2007, there was a dispute between Abbott Laboratories and the Thai government over the latter's declared intention to ignore the patent on Abbott's anti-AIDS mixture of lopinavir and ritonavir. Abbot dropped its price from $181 to $107 per patient per month to pre-empt such a move.

The long-term consequences are still unclear. On the one hand, if the decision allows Africa's and Asia's many AIDS sufferers to receive treatment, it cannot be a bad thing. On the other hand, particularly if parallel trade develops and the patent system is undermined, there is a much reduced incentive for companies to invest in treatments for tropical diseases.

F. Cost containment measures

Governments are increasingly worried about the rise in health care expenditure. The pharmaceutical industry would say that drugs provide the best possible value within that system and permit treatment at home of patients who would otherwise be costing much more in hospital. Drugs

in the United States and Europe average about 18% of health care spending (up from 10%, 20 years ago), so that savings on drugs have to be huge to make much impact on overall health care costs. Much smaller cuts in hospital costs in percentage terms would have a larger effect. On the other hand, medical and *para*-medical personnel enjoy high status and public support. The pharmaceutical industry does not: its reputation is equivocal and its very success in commercial terms make it more vulnerable to attack. Official controls on pharmaceutical expenditure are found everywhere in Europe and North America.[59]

Most European countries control the permitted prices of individual prescription drugs. A price is agreed at the time a new product is introduced, and it may not be increased even to compensate for inflation without official permission. The exceptions to this system are Denmark, which permits free pricing, and Ireland, which permits free pricing in principle but sets an upper limit based on average prices in other countries. Germany and the Netherlands permit free pricing but have reference prices (see below). The United Kingdom regulates prices indirectly via limits on profitability. The United States has free pricing but stimulates generic competition. In all countries, OTC drugs and prescription drugs that are not reimbursed by the national health care system are free from controls.[60]

In the Pharmaceutical Price regulation Scheme (PPRS), the UK Department of Health negotiates with companies a maximum rate of return on capital employed in their sales to the National Health Service. An overall rate is fixed for the industry as a whole and, within this figure, an individual firm is awarded a particular rate "having regard to the scale and nature of the company's relevant investments and activities and the associated long-term risks." Annual growth from 1995–2004 ranged from 5.5% to 12.5%.The system was modified in 2005, and the new scheme required that all companies that sold more than £1 million of branded medicines to the National Health Service to reduce prices by 7%. For the first time, year on year expenditure dropped, the saving being 1.8%.

The German and Dutch reference pricing systems set reference prices for three groups of pharmaceutical products:

1. Containing the same active ingredients (i.e. multi-source products).
2. Containing active ingredients that are comparable therapeutically and pharmacologically (e.g. benzodiazepine anxiolytics).
3. That have comparable effects, especially if combination products (e.g. anti-hypertensive-diuretic combinations).

The reference price is related to the average price of the various products on the market. A physician may prescribe as he or she pleases but, if the product costs more than the reference price, the patient has to pay the difference. The result of reference pricing was that virtually all producers dropped their prices to the reference price and, where they were in a position to do so, raised the price of their in-patent (therefore not reference-priced) products. This did not necessarily work to the advantage of the generics producers, because the cut in the price of the branded product undermined their own major selling point.

An alternative to price controls is to stimulate competition from generics producers, and this is the path that has been followed in the United States and Canada. This was discussed in Section IV.D. Germany, Denmark and the Netherlands are the European countries that permit generic substitution.

Pharmaceutical spending may also be controlled by regulation of the drugs that may be prescribed. Denmark, France, Germany and Japan have positive lists. The United Kingdom, the Netherlands and Germany have negative lists. Both reduce the proportion of prescriptions that are reimbursed. In principle, a positive list should indicate a more considered approach to prescribing than a negative list but in practice other factors seem much more important. It is difficult to deny reimbursement to popular drugs of doubtful efficacy.

A way for governments to save money is for part of the burden to be shifted to the patient, that is patient copayments. These are more or less universal. Two systems operate: a flat-rate copayment unconnected with the price of the drug as in Ireland, the Netherlands and the United Kingdom, or a proportion of the price of the drug as in Spain, Denmark Belgium and France. Exemptions are common. For example, in the United Kingdom 87.6% of prescriptions were exempt in 2005. Encouragement of OTC drugs is another way in which patients can be induced to contribute. Self-diagnosis and prescription worries physicians, but there has been an increase in the scope of drugs licensed for OTC sale and this may be expected to increase.

The final method for discouraging lavish prescribing is drug budgets for general practitioners. These have been tried in Germany and the United Kingdom but their effect has largely been swamped by other factors, and they have been discontinued.

The research-based pharmaceutical industry has fought against cost containment measures. To some extent, the patent term restoration legislation was intended as a sweetener to diminish industry anxiety about the other initiatives. Financial pressures have persuaded a number of research-based companies to take over generics companies, to provide them with some sort of insurance against a swing in that direction. The irony of this is that the added respectability given to generics by such policies has increased their rate of penetration of the market.[61] In addition to takeovers, there have been numerous mergers that will enable companies to spread the research risk, although there is evidence that mergers result in a market share lower than the joint market shares of the participants.[39] The number of medium-sized companies is shrinking, while the number of small start-up companies selling ideas to the giants is

increasing. There are elaborate tie-ups between universities, large and small firms and even regulatory authorities, which may or may not improve efficiency or involve conflicts of interest.[62] In the end, however, the industry will continue to flourish only if it continues to produce a stream of innovative drugs (Section V.A.).

G. Pharmacoeconomics

Nowadays a drug company has not only to show its paymasters – governments, insurers and so on – that its new product is safe and works, but also that it is cost-effective. In Australia, this has been spelled out in legislation. Since 1993, any drug submitted for approval must be accompanied not only by the results of clinical trials but also by an economic impact analysis. In 1999, the United Kingdom set up a National Institute for Clinical Excellence (NICE) to advise the National Health Service on the cost-effectiveness of health care technologies. Other countries ask formally or informally for pharmacoeconomic analysis.[63] Economic impacts can be measured in a variety of ways,[64] for example, cost-effectiveness, cost–utility or full cost–benefit studies.

Pharmacoeconomic analysis must be based on comparisons with existing treatments. It must cover not only the spending on the drug but possible savings on other medical services, for example, cost of a period in hospital or of providing social services for an elderly patient at home.

The exercise is relatively simple if a drug is only a marginal improvement on its competitors. Small improvements in hospital bed occupancy and so on can readily be calculated. On the other hand, drugs that offer only marginal improvements will rarely be able to command a high enough price to justify development costs. Thus, the exercise is mainly of importance to companies trying to justify very high prices for "breakthrough" products.

Cost-effectiveness studies measure the cost in relation to outcome.

The classic cost-effectiveness study was carried out on the pioneering anti-ulcer drug, cimetidine. The figures are shown in Table 45.5. The major saving in cimetidine-treated patients was in hospital costs. It was also shown that among non-hospitalized patients who were on older drugs, 2 days of work per week were missed on average, while among the cimetidine-treated group only 1 day was missed. A similar study showed the benefits of the anti-cancer drug carboplatin, in spite of its costing 10 times as much as the older drug cisplatin. The point is that the latter has to be given to inpatients while the former can be used as outpatient therapy. When hospital costs are taken into account, the more expensive drug offers higher cost-effectiveness.[65]

A third example compared the cost in Sweden of treating patients who had had myocardial infarction with the β-blocker metoprolol. The direct health service costs for treatment were reduced from 17,120 to 12, 310 Swedish Kröner, a saving of 28%.[66] A fourth example was used to justify hepatitis B vaccination in Japan. A net saving of 16 billion yen was estimated as a result of vaccination costing 1.3 billion yen.[67]

TABLE 45.5 Costs and Benefits of Cimetidine. Average Annual Michigan Medicaid Expenditures per Patient with Duodenal Ulcer

	Control group ($)	Cimetidine treated group ($)
Physician	109	57
Hospital	602	97
Drugs	10*	66
Total	721	220

*Does not include antacids that are excluded from Michigan Medicaid.
Source: Patterson, M. L. *Manage. Decis. Econ.*, 1983, 4, 50.

The problem for hospital administrators is that the largest savings come from a reduction in time spent in hospital. In countries with long hospital waiting lists, such as the United Kingdom, however, beds do not remain empty. They are reoccupied at once. During a patient's stay in hospital, the first few days are the most expensive, and subsequently only minimal care and nursing are required. Consequently, by hastening the turnover of patients, a hospital will have more patients in the early expensive stage of their stay. Although the average cost per patient treated will decrease, the overall hospital costs will increase, and so will the cost per patient per day. Governments might complain, and the efficient hospitals will be the ones that are criticized. With "inefficient" care, one patient is treated at a cost of $x. With efficient care, three patients are treated at a cost of $2x. More patients benefit but at higher total cost. (In a country with too many hospital beds, such as France, the unoccupied bed might remain empty, in which case, although the direct costs of hospitalization are saved, the overheads and depreciation will be spread among fewer patients, and the cost per patient per day will again appear to rise.)

The Health Service reforms in the United Kingdom in 1991 avoided this trap and used instead the cost–utility concept. Cost-effectiveness takes into account only the cost of a single standardized outcome; cost–utility looks at the costs of a range of different outcomes, such as extra years of survival and improved quality of life. A Dutch study of two cholesterol-lowering drugs, for example, showed that cholestyramine cost 131,000 Dutch Guilders per year of life saved compared with simvastatin which cost 31,500 Dutch Guilders.[68]

The above example is attractive in that it compares two forms of therapy. What it does not do is place a value on an extra year of life gained or, indeed, on the quality of life.

Economists have worked hard at producing measures to quantify these two concepts. A number of indexes have been produced, the meaning of which is uncertain and discussion of which is beyond the scope of this chapter. They include the Nottingham Health Profile, the American "Sickness Impact profile" and the Rosser Health Index.[69] The Rosser Index has been used to produce a "Quality Adjusted Life Year" or QALY. Each year of life is discounted by a factor reflecting the disability and distress suffered during the year. The factor is obtained by asking people to estimate what number of years of perfect health would be equivalent to 10 years of life in a particular state of disability. Related to the QALY is the ICER, the incremental cost-effectiveness ratio, which is the cost per QALY gained. Table 45.6 shows the estimated ICERs for a range of new drugs submitted to NICE.[70] NICE asks whether an intervention is cost-effective and should be funded free at the point of prescription.

The technique is clearly subjective, and there is room for much development and validation. Nonetheless, it makes possible some worthwhile comparisons. Simple preventive measures, such as changing diet or giving up smoking, produce the cheapest benefits. Reduction of blood pressure to avert strokes and selective testing for cholesterol are

TABLE 45.6 NICE Decisions[a] Ranked by Incremental Cost-effectiveness Ratio (ICER)

	ICER(£)		ICER(£)
GP advice to stop smoking	430	Irinotecan and oxaliplatin for advanced colorectal cancer 3	22,500
Topecatan inhibits topoisomerase I in cancer cases	1,000	Sibutramine for obesity combined with treatment plan	22,500
Cytology for detecting cervical cancer	**1,100**	Overall treatment of juvenile arthritis	22,500
Asthma inhalers	5,000	Recurrent laparascope hernia	25,000
Taxane for ovarian cancer	8,271	Stent insertion and treatment	25,000
Glycoprotein[a]	9,250	Implantable cardioverter defibrillators	28,500
Non-small cell lung cancer (first line) clinical trial	9,475	**Advanced colorectal cancer 1**	29,000
Methylphenidate (CNS stimulant for treatment of attention-deficit hyperactivity disorder and narcolepsy	12,500	Overall treatment of adult arthritis	31,000
Gemcitabine for non-small cell lung cancer, pancreatic, bladder and breast cancer	12,950	Temozolamide as second line treatment for malignant glioma (brain cancer)	35,000
Other non-small cell lung cancer treatments	14,000	Trastuzumub combination therapy for HER + breast cancer	37,500
Alzheimer's disease, e.g. donepezil	15,000	Zanamavir (Relenza) for all	38,000
Taxane for breast cancer (second line)	15,250	Riluzole for amyotrophic lateral sclerosis (Lou Gehrig's disease)	38,750
Taxane for breast cancer	15,500	Orlistat for obesity (lipase inhibitor)	46,000
Taxane for breast cancer (first line)	**19,000**	**Primary laparascope hernia**	**50,000**
Trastuzumub monotherapy (herceptin) for HER2 + breast cancer	19,000	**COX II anti-inflammatories (routine use)[b]**	**150,000**
Zanamavir (Relenza) for influenza for at risk patients	20,400	**Beta interferon for multiple sclerosis**	**187,000**
Ribavirin for hepatitis C	20,500		

Drugs not approved in 2003 are in bold.
[a]Note that some drugs are approved for one sub-group of patients but not another.
[b]Cox II anti-inflammatories for at risk patients were approved apparently without cost-effectiveness data.

also cost-effective. Asthma treatment is relatively cheap and effective; cancers as a whole are expensive and, while there have been huge improvements in survival rates in the past generation, they remain a major source of suffering and death. NICE distinguishes between various subsets of patients as indicated in the table. For a time, NICE refused to allow Aracept for early stage Alzheimer's disease, but changed its mind between 2001 and 2003. Herceptin was accepted for advanced breast cancer in 2002 and for the early stages in 2006. The cost to the National Health Service was huge, and it is difficult to believe that NICE's decisions hinged solely on economic factors and not on media pressure.

As Devlin and Parkin[67] point out, the decision as to what is an acceptable or unacceptable cost should, in principle, be a step function with a clear threshold but, because of the many extraneous factors (uncertainty and the burden of disease are cited), it is a smeared out sigmoid curve. NICE states that its "range of acceptable cost-effectiveness" lies between £20,000 and £30,000 per QALY but it is probably somewhat higher.

The final type of analysis is the full cost–benefit study. The term is often used to cover all methods described above but strictly it applies only to studies where every cost and every benefit are taken into account. This includes such things as the contribution to the national economy of people whose lives are saved or prolonged. Such a study would be immensely complicated, but a partial study of hypertension and stroke in Britain has been performed.[71] Between the mid-1950s and the mid-1980s, new cases of stroke dropped from 2.4 to 1.75 per 1,000 population. The cost of stroke to the Health Service in 1984 was £550 million. Had the incidence been at the 1950s level, it would have cost $550 \times 2.4/1.75 =$ £754 million. In addition, there was a contribution to the national economy of £322 million through a reduction in premature mortality. A total gain of £526 million can be set against the cost of anti-hypertensive drugs (at manufacturers' prices) of £185 million. The social benefit of control of hypertension is evident.

Economics is not an exact science, and the question of the cash value of a year of life or a quality adjusted year of life (whose? where? under what conditions?) is one that no one would find it easy to answer, if there is an answer, and many would find it offensive even to ask the question. Devlin and Parkin[67] discuss many of the shortcomings in the QALY concept. For example, is the life of someone who has a slight disability worth less than that of a similar person in robust health? Is a child (who has a longer life expectancy) more deserving of intervention than (say) a mature Professor of Medicinal Chemistry? Reinhardt[72] gives the example of the vertical economist (who is in good health) who considers that the resource-intensity of medical treatment should be increased to the point at which the marginal cost of further intervention becomes less than the marginal benefit derived from it. By contrast, the horizontal economist (who is seriously ill) will wish for any increase in resource-intensity, no matter what its marginal cost, if it will make him or her feel better. The problem is that, with limited resources, a community has to channel cash to where it will do the most good. And even getting a society to agree on what constitutes "most good" is a formidable if not a impossible task.

A second approach, related to pharmacoeconomics but leading to fewer moral problems, is outcomes analysis. Physicians, notably those collaborating with Health Management Organisations (HMOs) in the United States but also elsewhere, are required to report their diagnoses and prescriptions to the HMO, which analyses the data and is then in a position to recommend the most effective treatment for a particular diagnosis. The benefit of this is that it should show up marginal differences between treatments and provide solid statistical advice for participating physicians. The drawback, if it is a drawback, is that it reduces a physician's autonomy. The physician might say "I have known the patient for 30 years and he or she is suffering from W and should be prescribed X," while the database says that he or she is suffering from Y and should be prescribed Z. Skilled physicians would wish to back their own judgment but, given the litigation currently surrounding medical intervention, they might very well prefer to practice defensive medicine and take the computer's advice.

The question then arises as to why to bother with a physician in the first place. Diagnostic packages are readily available, and uninsured people (and over 50 million Americans lack full medical insurance) might well opt for a visit to their local internet café, where they could obtain diagnosis and prescription, also via the internet, for perhaps a tenth of what a physician would charge. Perhaps the pharmaceutical would come from a "pirate" source, but there is still a question as to whether this low grade treatment is better than no treatment at all. The above example is an extreme one, but outcomes analysis is here to stay, and medical autonomy is certain to decrease as physicians are squeezed between the empowerment of their clients and the demands of their paymasters.

V. THE FUTURE OF THE PHARMACEUTICAL INDUSTRY

The years since 1935 have seen the rise of the pharmaceutical industry to its present eminence. Despite the thalidomide disaster and a handful of other tragic episodes, the industry has flourished. A formulary of life-saving and life-improving drugs has been discovered. Two questions are now paramount. First, can society go on spending ever more of its national product on health care and, second, can the momentum be maintained? Have all the easy illnesses been dealt with so that only the difficult ones are left?

A. Trends in pharmaceuticals

The quick answer is a gloomy one. There has been a drop in the number of NMEs reaching the market; costs are escalating, and almost every large pharmaceutical company has a problem with its pipeline. Insufficient new drugs are being developed to maintain the industry, once the patents on the present blockbusters run out. Between 2010 and 2011, there is a "generics cliff" when "big pharma" will lose 28% of its current sales.

Looked at from the point of view of the scientist rather than the shareholder, however, the outlook is much brighter. There has been a third chemotherapeutic revolution in the past 20 years[73] that opens up whole new worlds of drug targets and possibilities. The technological breakthroughs have been numerous. There has been a many-orders-of-magnitude increase in the sensitivity of analytical methods. Techniques have been developed to produce proteins, receptors, enzymes etcetera with the aid of Taq polymerase. Combinatorial chemistry and high throughput screening permit evaluation of whole libraries of compounds. Computer graphics techniques enable models of test molecules to be created on a screen to see if they bond to receptors. The human genome has been deciphered. New sciences of genomics and proteomics have been developed. The determination of the structure of DNA and the use of target validation and biomarkers offer the possibility of determining if a person will be susceptible to a particular disease or respond to a particular drug. For example, Herceptin (trastuzumab) is a suitable treatment for breast cancer in the 15%–25% of patients whose cancer cells have a large amount of a protein called HER2. Such "personalized" drugs offer especial hope for the future.

Biologics have taken some time to get off the ground but are making slow and steady progress. Table 45.7 shows the global sales of the top ten. These add up to $33 billion, compared with $60 billion for the top ten conventional drugs. Thus, they are serious players in the pharmaceutical market.

The costs per patient, however, are formidable. In the United Kingdom in 2007, a course of alemtuzubab for chronic lymphocytic leukemia was said to cost about $19,000/month; pemetrexed for mesothelioma $16,000/month and even the relatively cheap bevacizumab for colon and rectal cancer $3,700/month.

A characteristic of the new biologics is that they have broken out of the constraints of the Lipinski rule of five (Chapter 39). There has also been a renewal of interest in vaccines, another drug group falling outside those rules. Cervarix (GlaxoSmithKline) and Gardasil (Merck) act against the HPV, responsible for cancer of the cervix (the second most common cancer among women) and genital warts. Vaccines are also sought against hepatitis C and the avian flu virus (H5N1). In addition, Sanofi-Aventis and Medimmune are trying to develop a method of producing vaccines on a plant scale by cell culture rather than by inoculation of eggs.

TABLE 45.7 Sales of Top 10 Biologics, United States (2006)

Brand name	Generic name	Indication	Sales 2006 ($ million)
Enbrel/Embrel	Etanercept	Rheumatoid arthritis	4,474
Aranesp	Darbepoetin alpha	Renal and cancer anaemia	4,121
Rituxan/ Mabthera	Rituximab	Rheumatoid arthritis and non-Hodgkin's lymphoma	3,912
Remicade	Infliximab	Rheumatoid arthritis	3,764
Procrit/Eprex	Epoetin alpha	Anaemia	3,180
Herceptin	Trastuzumab	Metastatic breast cancer	3,175
Epogen	Epoetin alpha	Renal and cancer anaemia	2,844
Neulasta	Pegfilgrastim	Neutropaenia	2,710
Actrapid/ Novolin	Human insulin	Diabetes	2,653
Avastin	Bevacizumab	Metastatic coloreactal cancer	2,395

Figures differ slightly from those in Table 45.1 because they come from different sources using different exchange rates.
Source: http://www.pipelinereview.com/Top_20_Biologics_by_La_Merie.pdf.

Vaccines are also sought against tropical diseases. Problems of compliance among indigenous African and Asian populations mean that a one-shot vaccine is much more likely to be effective than drug treatment.

A third group of (sometimes) non-Lipinski compounds are antibiotics and antibacterials. Until the 1990s, there had been no new group of antibiotics for several decades, and they did not appear in the top-20 drugs in research in 1993. The proliferation of antibiotic-resistant super-bugs, notably methicillin resistant *Staphyloccus aureus* (MRSA), vancomycin resistant *Enterococci* (VRE) triggered research. Synercid, a synergistic mixture of the streptogramins quinupristin and dalfopristin, appeared in 1999. They are active against MRSA and some strains of VRE. A totally new group of antibiotics, the oxazolidinones, is also active against resistant bacteria; linezolid was licensed in 2000 and had sales of $782 million in 2006. The lipopeptide, daptomycin, was launched in 2003 for intravenous use. Ketek (telithromycin), a macrolide antibiotic, was launched

in 2004, although there are recently recognized side effects. Retapamulin, a member of another new class of antibiotics called pleuromutilins, was approved in 2007. The economic problem with many new antibiotics is not the expiry of patent rights but the development of bacterial resistance.[1]

All this does not mean that traditional pharmaceuticals are on the way out. They still make up the majority of new medical entities licensed. They too benefit from the new platform technologies. Although the number of medicines produced by these technologies is disappointingly small at present, their quality is high and there are fewer "me-too" drugs. Refractory diseases such as malaria are being tackled, as are the degenerative diseases of old age such as Alzheimer's and Parkinson's.

B. Conclusion

There is a mismatch between what the drug companies are hoping to deliver and what the paymasters – governments, insurance companies and so on – are prepared to pay for. How this conflict will be resolved remains in doubt. Unquestionably there are many interesting and important problems that can potentially be solved by the pharmaceutical industry, but at a cost. All drugs are expensive to develop, and drugs for rare diseases are prohibitively expensive on a per capita basis unless there is a government subsidy. Biologics, to add to the worries, are also expensive to manufacture.

Since the 1990s, for the first time as an overt policy issue, society has questioned the amount of money spent on health care. There is general agreement that some sort of rationing is inevitable but none on who is to do the rationing or how. Will a balance be struck on the basis of ability to pay or on some sort of cost–benefit analysis? Consider the following situation:

It appears that the consumption of health care increases exponentially with age and is 5–10 times higher at 80–84 than at 60–64.[74] The number of old people is increasing rapidly throughout the developed world. Demand for additional health care comes, and will come increasingly in the future, from the elderly. On a cost–utility analysis, however, treatment of the elderly is not an attractive option. The elderly are less likely to recover from their illnesses. If they do, the fact that they are elderly to start with means that the years of life gained will be smaller. The elderly will have already retired so that their contribution to national income is zero or (because of their pensions) negative.

If health care is to be rationed by cost, then physicians will have the task of turning away patients who cannot pay (and if the physicians are soft-hearted, then the hospitals will turn them away instead). If care is to be allocated on a cost–utility basis, then the physician will have to find out the patient's age before deciding whether to offer a coronary by-pass, a packet of anti-angina pills or a loaded revolver.

There is thus an ethical challenge to governments, to economists and, worst of all because they are the ones to see the patients, the medical profession. Physicians do not want to be the people to do the rationing, particularly if it is to be done overtly against a series of guidelines, so that they can be criticized afterwards. They have taken the Hippocratic oath and are committed to "carry out regimen for the benefit of the sick and keep them from harm and wrong." But, if a physician does not take the decision, will it be taken by an economist or a politician? Will a patient abandoned by a physician be permitted to seek a second opinion, and at whose expense? Will there be a health care sector for people whose treatment under the social security system would not be cost-effective but who are willing and able to pay? A National Health Service patient in the United Kingdom in 2008 was controversially refused permission to pay for a drug authorized by NICE.

At the beginning of this chapter, I quoted Aristotle's distinction between health and wealth. My late father, Jacob Victor Reuben, to whom this chapter is dedicated, qualified as a physician in 1920. For his first 15 years in practice, there was scarcely a drug he could prescribe that would cure disease. The most he could do in many cases was to advise patients to stay in bed, take some aspirin and plenty of fluids, and allow nature to run its course. If the patients died, he could say in all honesty that he had done everything in his power to save them. Is his successor 90 years later going to have to say to patients, at any rate the elderly ones, that there are many things he or she *could* do to save them, but they are all too expensive, and why not stay in bed, take some aspirin and plenty of fluids, and allow nature to run its course?

REFERENCES

1. Wilkinson, M. Resisting the call for new antibacterials. *Drugresearcher.com* May 2007, *25*.
2. Association of British Pharmaceutical Industry, World Review.
3. Author's figures based on www.ameli.fr/fileadmin/user_upload/documents/ Tableau_medicaments_rembourses_2002__2005.xls
4. O'Brien, B. *Patterns of European Diagnosis and Prescribing*; Office of Health Economics: London, 1984. Lyn Payer, *Medicine and Culture* 1990, Gollancz; Teeling-Smith, G. *Patterns of Prescribing*. Office of Health Economics: London, 1991.
5. Based on www.sante.gouv.fr/drees/etude-resultat/er440/er440.pdf.
6. Wise, R. Antibiotic resistance: a hazard to public health. *Eurohealth* 2001, *7*, 29–31.
7. Arzneiverordnungs-Report 2000, Springer Verlag: Berlin and Heidelberg, 2001.
8. Arzneiverordnungs-Report 2000, Springer Verlag: Berlin and Heidelberg, 2006.
9. Schaaber, J. *et al. German Drugs in the Third World*. BUKO Pharma-Kampagne: Bielefeld, 2004.
10. Connor, S. *Independent/UK* December 2003, *8*.
11. FDA re-examines basis for OTC, New York Times, June 28, 2000.
12. *SCRIP*, 1310/18.
13. Melton, L. Chinese medicine in western packaging. *Chemistry World* May 2007, *4*, 46–50.

14. www.epa.gov/oppt/greenchemistry/pubs/pgcc/winners/gspa04.html
15. Myllylä, G. Whole blood and plasma procurement and the impact of plasmapheresis. In *Blood Separation and Plasma Fractionation* (Harris, J. R., Ed.), Wiley-Liss: New York, 1991, p. 38.
16. Reuben, B. G., Wittcoff, H. A. *Pharmaceutical Chemicals in Perspective*. Wiley: New York, 1990, pp. 337–341.
17. US International Trade Commission. *Synthetic Organic Chemicals, US Production and Sales 1992*. US International Trade Commission: Washington, DC, 1994.
18. Commission of the European Communities. *Rules Governing Medicinal Products in the European Community, Volume IV, Good Manufacturing Practice for Medicinal Products*. Office for Official Publications of the European Community: Luxembourg, 1992.
19. Last, P. E. *Chem. Br.* 1987, *23*, 1073.
20. Nathan, S. *Pharmaceutical Visions, Spring* 2001, *Spring*, 54–57.
21. Brathwaite, M. J., Ketterman, C. L. *Chem. Ind.* 1993, 1042.
22. Jackson, A., Angoh, G. *Chem. Ind.* 1993, 1046.
23. Fox, J. *Chem. Ind.* 1993, 270.
24. Rao, C. S. *The Chemistry of Process Development in Fine Chemical and Pharmaceutical Industry*, 2nd Edition. Techbooks World: East Brunswick, NJ, 2006.
25. Wittcoff, H. A., Reuben, B. S., Plotkin, J. S. *Industrial Organic Chemicals*, 2nd Edition. John Wiley and Sons: New York, 2004, 609–611.
26. Kleemann, A., Engel, J., Kutscher, B., Reichert, D. *Pharmaceutical Substances, Syntheses, Patents and Applications*, 4th Edition. Stuttgart: Thieme, 2001.
27. The sertraline story is told in *Chem. Eng. News,* April 22, 2002, p. 30 and July 1, 2002, p. 29.
28. Nature Reviews: Drug Discovery. *Nat. Rev.: Drug Discov.* 2007, *6*, 99–101.
29. DiMasi, J. A., Hansen, R. W., Grabowski, H. G. The price of innovation: new estimates of drug development costs. *J. Health Econ.* Frank J. G., New estimates of drug development costs, 2003, *22*(2003), 325–330.
30. Kettler, H. E. *Updating the Cost of a New Chemical Entity*. Office of Health Economics: London, 1999.
31. Burstall, M. L., Reuben, B. G. *Pharmaceutical R and D Productivity: The Path to Innovation*, Cambridge Healthtech Advisors, 2005; http://www.advancesreports.com/all_reports/2005_55_R and D Productivity/sample.pdf.
32. Jimenez-Andres, C. EU plan to lift R and D to begin in 07. *Drug researcher.com* 2007, *22*.
33. Burstall, M. L., Reuben, B. G. Pharmaceutical Industry Dynamics. *Spectrum* 2001, *5*. 11-1 to 11-17.
34. McKillop, T. http://www.soci.org/SCI/pressoffice/speeches/ge448.jsp
35. *The Pharmaceutical Industry Competitiveness Task Force*. Association of the British Pharmaceutical Industry: London, 2003 (www.abpi.org.uk 2003).
36. *International Trade Statistics Year Book,* United Nations, New York, published annually. See also Reuben, B. G. Pharmaceuticals. In *Encyclopedia of Physical Science and* Technology, 1991 Yearbook (Meyers, R. A., Ed.), Academic Press: New York, 1991.
37. Burstall, M. L., Reuben, B. G., Reuben, A. J. Pricing and reimbursement regulation in Europe: the industry perspective – an update. *Drug Ind. Assoc. J.* 1999, *33*, 669–688.
38. Burstall, M. L., Reuben, B. G. *Outlook for the World Pharmaceutical Industry to 2015*. Decision Resources: Waltham, MA, 2003.
39. UNCTAD, *The TRIPS Agreement and Developing Countries*. United Nations: Geneva, 1996.
40. Burstall, M. L., Reuben, B. G. *Outlook for the World Pharmaceutical Industry to 2010*. Decision Resources: Waltham, MA, 1999.
41. CMR International News, 2000, *18*(2), 1.
42. Walker, S. R. *et al. Pharmaceutical Patents*. Centre for Medicines Research: Carshalton, 1989.
43. Lis, Y., Walker, S. R., In *Medicines: Regulation, Research and Risk*. (Griffin, J. P., Ed.) Belfast: Queen's University, 1989.
44. *Chemistry World*, March 2007, p. 18; September 2007, p. 19.
45. Reymond, E. US biotech drug spending rises while Congress debates 'biogenerics' law. *DrugResearcher.com* 2007, *26*.
46. For general background, see Asbury, C. H. *Orphan Drugs: Medical versus Market value*, Lexington Books: Lexington, MA, 1985; Asbury, C. H. The Orphan Drug Act: the first seven years. *J. Am. Med. Assoc.* 1991, *265*, 893.
47. See EC No. 141/2000 and 847/2000.
48. Office of Health Economics. *Cystic Fibrosis* 1986, Office of Health Economics, London.
49. *SCRIP* Magazine, January 1994, p. 48.
50. Tobias, G., Faigen, N. *SCRIP Review* 1992, *12*.
51. Reuben, B. G., Burstall, M. L. *Generic Pharmaceuticals – The Threat*. Economists Advisory Group: London, 1989.
52. US Department of Health and Human Services. *Approved Drug Products with Therapeutic Equivalence Evaluations.* US Department of Health and Human Services, Rockville. MD, annual updated by monthly supplements. Also available electronically and updated daily.
53. Generic Pharmaceuticals Association, www/gphaonline.org
54. Health and Social Care Information Centre. Prescriptions Dispensed in the Community, Statistics for 1995–2005. Health and Social Care Information Centre: England, Bulletin: 2006/12/HSCIC.
55. Foster, P. *London Times*, March 7, 2008.
56. www.efpia.org/Objects/2/Files/infigures2007.pdf.
57. Love, J. Policy issues concerning parallel trade of pharmaceutical drugs in the United States. *FDA Prescription Drug Importation Meeting Gilman*, April 14, 2004; Price, D. J., *Property and Trade Across a Regulatory Border: Pharmaceuticals (Re)Importation from Canada and Beyond*, *U of Maryland Legal Studies Research Paper No. 2005-66, American Journal of Law and Medicine, Symposium on International Pharmaceuticals Regulation*, Spring 2006.
58. *New York Times*, February 7, 2001.
59. Burstall, M. L., Reuben, B. G. *Critics of the Pharmaceutical Industry*. REMIT Consultants: London, 1990, especially pp.19–69.
60. 15 Ways to price a pill: Medicines in the single market, *Eurohealth, 7*, 2001, 1–25.
61. *SCRIP* 1994; 1916:15
62. Reuben, B. G. http://www.eyeforpharma.com/briefing/BryanReuben FDAarticle.pdf.
63. CMR/OHE R and D Briefing, No. 33, 2001, p. 3 (http://www.cmr.org).
64. Teeling Smith, G. (Ed.). *Measuring the Benefits of Medicines: the Future Agenda*. Office of Health Economics: London, 1990.
65. Tighe, M., Goodman, S. Carboplatin versus cisplatin. *The Lancet* 1988, *ii*, 1372.
66. Olsson, G., Levin, L. A., Rehnquist, N. Economic consequences of post-infraction prophylaxis with β-blockers. *Br. Med. J.* 1987, *294*, 339.
67. Institute of Statistical Studies. *Technological Assessment of Drug Therapy-Analysis of Cost–Benefit Studies on Pharmaceuticals*. Institute of Statistical Studies: Tokyo, 1988.
68. Martens, L. L., Finn, P., in Teeling Smith, *op. cit*.
69. Rosser, R. M. A history of the development of health indicators. In *Measuring the Social Benefits of Medicine* (Teeling Smith, G., Ed.), Office of health Economics: London, 1984.
70. Devlin, N., Parkin, D. Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? *Health Econ.* 2004, *13*, 437–452.
71. Teeling Smith, G. The economics of hypertension and stroke. *Am. Heart J.* 1989, *119*, 3. pt 2, 725.
72. Reinhardt, U. E. *Accountable Health Care: Is It Compatible with Social Solidarity? Office of Health Economics Annual Lecture*. Office of Health Economics: London, 1997.
73. Reuben, B. G. A third chemotherapeutic revolution? *Chem. Ind.* 2000, *16*, 214–217.
74. Olsson, G. J. *Roy. Soc. Med.* 1988, *81*, 242; Goldman, L., Sia, B., Cook, F., Rutherford, J. D., Weinstein, M. C. *New England J. Med.* 1988, *319*, 152.

Index

(**Bold** page numbers indicate main entries; *Italic* page numbers indicate illustrations/tables)

1,2,4-oxadiazoles, 312, 314, *316*
1,2,4-thiadiazoles, 301, 302
1,3,4-oxadiazoles, 314, *316*
1,4-dihydropyridines, 519
1,5-benzodiazepine, 440
2-amino-thiazole, 137
2D autocorrelation vectors, 213, 214
$2F_o$–F_c maps, 618
2-PAM *see* N-methylpyridinium-2-carbaldoxime
3-(3-hydroxyphenyl)-*N*-*n*-propyl piperidine (3-PPP), 355, *357*
3-PPP *see* 3-(3-hydroxyphenyl)-N-n-propyl piperidine
3-pyridine-methanol, 735, *735*
3-thiaepothilone D, 174, *174*
3D autocorrelation, 215
3D descriptors, 214–16
3D MoRSE code, 215
3D pharmacophore descriptors, 215–16
3D QSAR, 491, 580
  3D molecular field descriptors calculation, 591–592
  alignment, 592–594
  applications, 594–601
  conformation analysis, 590–591
  future directions, 601
  molecular superimposition, 590–591
  statistical tools, 592
  validation, 594
  workflow, 589–590, *589*
4-hydroxycyclophosphamide, *844*
5-bromo-4-chloro-3-indolyl-(β-D-galactopyranoside (X-gal), 116
5-HT *see* serotonin
5-hydroxytryptamine (5-HT) receptors *see* serotonin receptors
5-lipoxygenase (5-LO), 396, *397*
5-thiaepothilone B, 174, *174*
6-deoxyacyclovir, 737, *737*
11β hydroxysteroid dehydrogenase (11βHSD), 201
4560 RP *see* chlorpromazine
(−) adamantaplatensimycin, 179, *179*
(−)-ephedrine, 537, *537*
(−)-pseudo-ephedrine, 537, *537*
(±)-3-(1-propyl-3-piperidinyl)phenol [(+)-3-PPP], 319, *320*
α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic-acidreceptors (AMPAR), 94–5
α-amylase, 832
α-helix, 468
α-yohimbine, 313, *314*
α_1-acid glycoprotein, 644
β-adrenoreceptor antagonists, 383, 400
β-blockers, 13, 115, 135, 348, 400–1, 898
  ring transformation, 348, *349*
β-cyclodextrin *see* cyclodextrins
β-galactosidase, 110, 113, *114*, *116*, 118, 119
β-Lactam, 458, 704, 834
β-sheets, 199, 468
β_2-adrenergic receptors, screening, 116, *116*
β_2-agonists, screening, 116, *116*
γ-aminobutyric acid *see* GABA A
γ-cyclodextrins *see* cyclodextrins
γ-glutamyldopa, 729, *729*
"λ packaging extract", 119
κ-nearest neighbors (κ-NN), 500

A

A-type prostaglandins (PGAs), 818
abacavir, 440, *441*
abbreviated new drug application (ANDA), 911, 913
ABC transporters, 698, *699*, *700*, 702–3
Abraxane®, 175, 807, 808
absolute bioavailability, 652
absolute novelty, 885
absorption, 788–9
  apparent volume of distribution, 649–50
  biovaliability, 651–2
  cyclodextrins enhancement, 826
  drugs, 637, *368*, 640–4
  gastrointestinal tract, 816, 823
  optical isomerism, 540
abstracts, Internet, 264
ACD *see* Advanced Chemistry Development
ACE *see* angiotensin-converting enzyme
acetamidoeugenol, 443
acetamino-2-fluorene, 678, *679*, 683
acetaminophen *see* paracetamol
acetaminosalol, 381
acetazolamide, 12
acetylation, 667–8, *669*, 683
acetylcholine, 283, *284*
acetylcholine receptors, 39–40, 384
acetylcholinesterase (AChE), 107, 201
  inhibitors, 301, *301*, 397, 398, *398*, 403, 594–7
acetylenic groups, 445–7
acetylsalicylic acid, 7, 8
AChE *see* acetylcholinesterase
AChE-based DMLs, **559–60**, *559*
acidic groups
  biological activity, 457–8, *457*, *458*
  NH functions, of prodrugs, 727–8, *727*, *728*
  solubility, 456, *457*
  substituents, 476
acidic ionizable moieties, 768
acids
  ionization constant, 751
  passive diffusion across living membrane, 249, 639
  salt formation, 751–2
  solubilizing moieties
    acylation of hydroxy and amino functions, 770–5
    alkylation of hydroxy and amino functions, 769–70
    direct introduction, 769
acquisition of biomass, 163–4
Actinomyces griseus, 20
active compounds, 168
active transport, 639, 698, *699*, 706
active tubular secretion, 646, *646*
activity profiles
  carboxylic acid, 457–8
  halogenation, 450
  wider selectivity optimization, **563–5**
acyclovir, 737, *737*
acyl-coA, 668–9
acylation
  amino groups with acidic chains, 770–5
  hydroxy groups with acidic functions, 770–5
  xenobiotic transformations, 668–9
acylsulfonamides, 304, 306
adalimumab, 10, 98
additional binding sites, **459–60**, *460*
additives, cyclodextrins, 829–31
adenylate cyclases, 101

ADME (absorption, distribution, metabolism, excretion), 297, 483–4, 858
computational models, 488–9
drug discovery programs, 858
lead compounds, 126
library design, 484
metabolic stability, 488–90
optical isomerism, 535
QSAR, 492
ADMET studies, 64, 211, 491, 507
administration routes, 640, *641*, 730, 860
adrenaline, 38, 653
adrenergic receptors, 93, 116, *116*, 536, *536*
adriamycin, 28, 29
Advanced Chemistry Development (ACD), 247, 262
database, 249
aflatoxin B1, 678, *678*
age factors, drug metabolism, 652
agonists, 115, 116
AIDS, 23, 24, 912
ajmalicine, 203
albifylline, 870
alcoholic groups, 724–5
Aldactone®, 34
aldehyde dehydrogenase (ALDH), 659
aldehydes, 676, 725
alglucerase, 912
alignments *see* structural alignments
Alimentation Equilibrée Commentry (AEC), *546*
aliskiren, 609–10, *609*, *610*
alkaloids, 6
alkylating stresses, *693*
alkylation, hydroxy and amino groups, 769–70
allergies, 558–9
allopregnanolone, 438
allopurinol, 28, *131*, 132, 251
allosteric modulators, 78–9
allylic constraints, 369–70, *370*
allylic series, 443–5, *444*, *445*
alprazolam, *291*
Alzheimer's disease (AD), 107, 109, 559–60
ambrisentan, 150–2
American Chemical Society (ACS), 257, 258
amide bond, 314, *316*
amidines, 458, 727
amines
acylation with acidic chains, 770
prodrugs derivatization, 726–7, *727*
solubilization, 775
aminoacridine twin drugs, *391*
aminoclonidine, 251
aminopyrine, 296, *296*
aminotetralins, 346
amiodarone, 135, *135*, 354, *356*
amoxicillin, 895
AMPA receptor antagonists, 350, *353*, 358
amphenicols, *295*
amphetamine, 89, 869
ampicillin, 20, 722–3, *723*
amyloid precursor protein (APP), 109
"analog", 291
analogical approach, 416, *416*
application strategies, 416
drug design, 126–9
analytical abstracts, 272
analytik, 267
anatomical-therapeutic-chemical (ATC) classification system, 70–1
anchor principle, 474
ANDA *see* abbreviated new drug application
anesthetics, 10, 45–6
angiotensin-I, 14–15
angiotensin-II, 14–15
angiotensin II antagonists, 257
angiotensin-converting enzyme (ACE), 608
inhibitors, 139–41, *140*, 169, *170*, *359*, 360
and sartans, 14–15
angiotensin system, for hypertension, **555–8**
angiotensine receptor ligands, *394*
anhydrides, 726, *726*
animal sources, **902**
animals
drug research, 858, 862, 864
lead compounds, 137
patentability, 883
pharmaceutical sources, 902
anionic site, 403
anological approaches, 343
conjunctive approaches, 356, 358–60, *358*, *359*
disjunctive approaches, 354–5, *355*, *356*, *357*
ring transformations, **343–54**
reorganization, 351–4, *354*, *355*, *356*
ring closure, 345–9, *346*, *347*, *348*, *349*, *350*, *351*, *352*
ring enlargement/contraction, 350–1, *352*, *353*
ring opening, 343–5, *344*, *345*
anomalies in isosterism
ether oxygen-methylene group, 329–30
fluorine–hydrogen isosterism, 328, *329*, *330*
anomalous scattering (AS), 617
ANP *see* atrial natriuretic peptide
antagonists, 76–7, 87
AMPA receptor antagonists, 350, *353*, 358
angiotensin II antagonists, 257
antiulcer H_2-receptor histamine antagonists, 303, *304*
benzazepine D_1/D_5 antagonists, 318, *319*
bradykinin B_1 antagonists, 300
calcium antagonists, 13–14, *13*
CCK antagonists, 306, 355, *357*, *309*
dopaminergic antagonists, 356
endothelin antagonists, **150–2**
H_2-receptor antagonists, 141, *141*
human histamine H_3 receptor (hH3R) antagonist, 580, 581, *581*
negative antagonists, 87–8
NMDA receptor antagonists, 318, 358
non-peptide bradykinin B_2 receptor antagonists, 301, *301*
novel cholecystokinin-2 (CCK_2) antagonist, 326, *326*
optical isomerism, 537, 539, 542
thromboxane A2 receptor antagonist, 277
anthracyclines, 28
anti-acid drugs
anti-H_2 drugs, 34
proton pump inhibitors, 34–5
antiangiogenic agents, 29–30
anti-anxiety drugs, 43–4
antibacterial, 172, 303, *303*, 401, *402*
antibacterial prodrug, 732, *733*
antibiotics, 18, 171–2
antituberculosis drugs, 20
cephalosporins, 21–2
chloramphenicol, 20–1
economic importance, 920
erythromycin, 21
misprescription, 901
pencillins, 19–20
quinolones, 22
streptomycin, 20
tetracyclines, 21
vancomycin, 21
anticancer agents
from marine organisms, 175–7
from plants, 174–5
anticancer antibiotics, 28
anticancer drugs
antiangiogenic agents, 29–30
anticancer antibiotics, 28
antimetabolites, 27–8
nitrogen mustards, 26
platinum anticancer drugs, 28–9
protein tyrosine kinases, 29
taxans, 27
vinca alkaloids, 26–7
anticholesterolemics, 169–71
antidepressants, 44–5, *45*, 327, *402*
antidiabetic drugs
biguanides, 32
insulin, 31–2
meglitinides, 32–3
sulfonylureas, 32
thiazolidinediones, 32

antiepileptic drugs, 47–9
anti-H_2 drugs, 34
antihistaminic compounds, 456, *457*
anti-HIV drugs, 23–4
antihypertensive drugs, 12, *12*, 169, 454, *455*
anti-inflammatory drugs, 346
antimalarial agents, 136, 177, *177*
antimetabolites, in cancer treatment, 27–8
antimicrobial resistance, 22–3
anti-parallel β-sheets, 828
antipodes, metabolism of, **540**
antipsychotics, 731
antipyrine, 650
antiretroviral (ARV) therapies, 24
antituberculosis drugs, 20
antiulcer H_2-receptor histamine antagonists, 303, *304*
antiviral compounds profiling, 220–2
antiviral drugs, 705
anxiolytics, 898
aplasmomycin, 332
apparent volume of distribution, 649–50, *650*
application strategies, **415–26**
  analogical approaches, 416
  biological logic rule, 423–4
  conjunctive approaches, 416
  deletion of chiral centres, 425–6
  disjunctive approaches, 416
  easy organic synthesis rule, 425
  hit optimization, 416–22
  minor modifications rule, 422–3
  pharmacological logic rule, 426
  preliminary considerations, 415–16
  rules, 422–6
  structural logic rule, 424
  substituent choice, 424–5
  topological exploration, 418–19, *418*
aqueous solubility, 494, 756–8, 788
  see also lipophilicity
  assay noise, 489
  chemistry control, **486–7**
  library design, 482
arachidonic acid, 8
  cascade, **561–3**
area under the curve (AUC), 649
arene–arene interactions, 469–71, *471*
arenologue, *284*, 286
aromatic ring equivalents, 445–6, *446*
arones, *135*
aroylpropionic precursors, 739, *739*
arpromidine, 392
arsphenamine, 145, 521
artemisin, 770, *771*
artemisinin, 89, 177, *177*
artificial intelligence, drug-likeness, 249
aryl amides, 314
arylacetic acids, 739
arylpropionic analgesics, 346
AS *see* anomalous scattering
Aspergillus terreus, 37, 170
asperlicin, 355, *357*
aspirin, 311
  and NSAIDs, 7–8
assay noise, 489
asthma drugs, 899
astrocytes, 645
atenolol, 322, *323*
atevirdine, 148
Ativan®, *793*
atorvastatin, 37
atorvastin, 170–1, *170*
ATP-binding cassette (ABC) transporters, 89
ATP-dependent membrane transport, 92
atrial natriuretic peptide (ANP), 96
AUC *see* area under the curve
autocorrelation, 214, 215
AutoDock, 259, 595
autoQSAR, 503–4
average binding energy, 474–5, *475*
avermectin, 172, *172*
avizafone, 776, *777*, *778*
avoiding chemical classes, in drug discovery and development, **252**
azathilone, 174, *174*
azathioprine, 25, 28, *131*, 132
azavinylogues, 285, *286*
azidothymidine (AZT), 23–4, 912
azo compounds reduction, 682
AZT *see* azidothymidine

B

B-factors, 620, *620*
back-propagation (BP) networks, 498–9, *499*
Bactrim, *793*
"bait" protein, 110
bapineuzumab, 109
barbiturates, 47, 48, *352*
barbiturics, 350, *352*
Barrter's syndrome, 109
bases
  ionization constant, 751
  passive diffusion across living membrane, 249, 639
  salt formation, 751–2
  solubilizing moieties, 769, **775–9**
    acidic NH function attachment, 779
    alcoholic hydroxyl attachment, 777–8
    basic NH_2 function, 779
    carboxylic acid functionality attachment, 780
    direct attachment, 775, *776*
  substituents, 458–9
basic fibroblast growth factor (bFGF), 826–7
basic groups, 458–9
basic substituents, 476
batch-mode solubility prediction, 490
Bayyesian networks (BNN), 498, 499
BCRP, 702–3
BCS *see* Biopharmaceutics Classification System
becaplermin, 96
Benazepril®, 169, *170*
benzaldehyde, 285
benzazepine D_1/D_5 antagonists, 318, *319*
benzimidazole, 35, 315, *316*
  bioisostery, 319, *320*
benziodarone, 134, 135, *135*
benzo cracking, 353, *355*
benzo-splitting, 419, *420*
benzodiazepine, 44, 447, *447*, 519
benzoxazocine, 348
benzylogues, 283–7
best selling drugs, *297–8*
bFGF *see* basic fibroblast growth factor
bifunctional molecules, *280*
biguanides, 32
bile acids, 729
biliary excretion, 646–7
binding energy, **228–9**
binding mode analysis
  identical twin drugs, 380, **401**
    interacting with adjacent binding sites, **403–5**
    interacting with different binding sites, **408–9**
    interacting with similar binding sites, **405–8**
  non-identical twin drugs, 380, **401**
    interacting with adjacent binding sites, **403–5**
    interacting with different binding sites, **408–9**
binding studies, 75–6
bioactivation, 675, **676–84**, *685*
  eliminations, 683
  mixed mechanisms, 738–9
  oxidation reactions, 676–8, *679*, *680*
  oxidative stress, 678, 680
  reduction reactions, 680–2
  substitution reactions, 682–3, *684*
bioavailability, 651–2
  carrier prologs, 724–8
bioisosterism, 290, 294, 299–303
  carbon–silicon, 330–1
  carboxamide, 313–17
  carboxylic acid, 303–10
  carboxylic esters, 310–13
  catechols, 319–20
  of heterocycles, 302–3
  imidazo[1, 2-a]pyridine, 301
  pyridazine, 301–2
  of pyridine, 299–301
  selenium, 333–4
  sulfonamides, 320
  ureas and thioureas, 317–18

biological active compounds, stereoselectivity in, **533–5**
biological activity, 492
biological assessment, 168
biological information, 134–8
biological logic rule, 423–4
biological membrane, 724–8
biological response curves
  see also dose–response curves
  homologous series, **277–83**, *277*, *278*, *279*, *280*, *281*, *282*, *283*
biological sources, **902–3**
biological space, **518–19**
biological strategy, **519–20**
Biology Oriented Synthesis (BIOS), 187–8
  futures of, 205
  protein structure similarity clustering, **199–202**
  scaffold tree, **194–9**
  SCONP and PSSC, combined application of, **202–5**
  workflow of, *195*
biomass, 163–4
BioMed Central, 268
biomembrane passage, 724–8
biopharmaceutical classification, of compounds, 861
biopharmaceutical therapies, **108–9**
Biopharmaceutics Classification System (BCS), 788
bioprecursor-prodrugs, 723, *723*
bioprecursors
  design, **721–42**
  prodrug approach, 735–40
  versus prodrugs, 740, *740*
BIOS *see* Biology Oriented Synthesis
biotin, 477, *477*
biotransformation, 647–8, **655–72**, **674–93**
  bioactivation process, 676–84
  biological factors, influence of, 671–2
  metabolic conversions, 685–93
  phase I, 647
  phase I reactions, 656–64
  phase II, 647
  phase II reactions, 664–71
biphenylylacetic acid (BPAA), 832, 832
bis(5)-tacrine, 403, *404*
bis(7)-tacrine, 403, *404*
bis-quaternary ammonium salts, 390
bismuth iodosubgallate (Airoform®), 849
bisquinolines, 390
bisubstrate inhibitors, **403–5**
bisulfite adducts, 772, *774*
bitterness, masking, 850
Black's rules, *521*
BLAST sequence alignment algorithm, 530
blood–brain barrier (BBB), 707, 708, 645
  penetration, **250**
BMS-182874, 132, *132*
BMS-192548, 130, *131*
BMS-193884, *132*, 133
BMS-207940, *132*, 133
BNCT *see* Boron Neutron Capture Therapy
boromycine, 332
Boron Neutron Capture Therapy (BNCT), 331–2
boronic chalcone, 333, *333*
Bortezomid, 333, *333*
bosentan, 150–2
Bothrops jararaca, 14, 139, 169
bovine spongiform encephalitis (BSE), 902
BPAA *see* biphenylylacetic acid
bradykinin, 368–9, *369*
bradykinin (BK) B_1 antagonists, 300
bradykinin antagonist, 385, *385*
bradykinin potentiating factor (BPF), 14
Bragg's law, *616*
brain penetration, **715–19**
  hydrogen bond interactions, 718–19
  log P, 717–178
  metabolic stability, 715–17
  pKa, 716
brain targeting, 835
brand names *see* trade names
Bristol Myers-Squibb amoxicillin, 97
Broad Institute Chemical Biology Program (BCB), 258
bromobenzene, 678, *679*
bryostatin, 1, 176–7, *176*
buserelin acetate (BLA), 822

C

C-allyl derivatives, 443–4, *444*
C. elegans, 111, 118
C2 symmetry axis, *383*
CA *see* cluster analysis
Caco-2 cell culture, 489
Caelyx®, 803
calcium antagonists, 13–14, *13*
calcium channels, 90, 91
camptothecin, 165
cancer, drugs for
  antiangiogenic agents, 29–30
  anticancer antibiotics, 28
  antimetabolites, 27–8
  nitrogen mustards, 26
  platinum anticancer drugs, 28–9
  protein tyrosine kinases, 29
  taxans, 27
  vinca alkaloids, 26–7
cannabinoid receptors, 40–1
canrenone, 344
captopril, 15, 169, *140*, *170*, *359*, 360, 465, *466*, 555, 608, *608*
carbaisosteres, 296, *296*
carbamates, *727*
carbamazepine, 48
carbocycles, *196*, 197
carbon–boron isosterism, 331–3
carbon oxidation/reduction reactions, 660–2, *660*, *661*, *662*
carbon–silicon bioisosterism, 330–1
carboplatin, 29
carboxamide bioisosteres, 313–17
carboxylic acid bioisosteres, **303–10**
  direct derivatives, 304, 306
  non-planar sulfur- or phosphorus-derived acidic functions, 306, 308–10
  planar acidic heterocycles and aryl derivatives, 306, *307*
carboxylic acids, 456, 457
  prodrugs derivatization, 725–6, *726*
carboxylic esters bioisosteres, 310–13
carboxylic functions, 308
Carcinogenic Potency Database (CPDB), 259
cardiac glycosides, 89, 819
carmofur stabilization, 821
carmustine, 145
carrier-mediated membrane transport, 639
carrier prodrugs
  application examples, 724–35
  bioavailability improvements, 724–7
  versus bioprecursors, 740, *740*
  derivatization, 724–8
  principles, 722–3, *722*
cascade prodrugs, 731–2, *733*
caspases, 98
catechol bioisosteres, 319–20
catechol *O*-methyltransferase (COMT), 665, 734
cation–π interactions, 469, *470*
cationic heads, 277, *277*
CCK antagonists *see* cholecystokinin (CCK) antagonists
Cdc25A, 201
CDE conjugates, 835, *835*
cefoxitin, 843
celecoxib, 9
cell-based screening, **113**
cell cycle
  and natural products, 162, *162*
cell membrane, 638
cell targeting, 834–5
cell therapy, 52
cellular and subcellular distribution, 699–700, *701*
cellular responses, 79–81
center modifications, 419, *420*
central nervous system (CNS), 192
cephalosporins, 21–2
Cephalosporium acremonium, 21

ceredase, 912
cetylpyridinium chloride, 734, *734*
CFTR (Cystic Fibrosis Transmembrane conductance Regulator), 89–90
CH bond oxidations, 677
charge–charge interactions, 465, *466*
charge–dipole interactions, 466, *467*, *468*
ChemAxon, 262
ChemBank, **258**
CHEMCATS, 263
ChemExper Chemical Directory, 261
ChemFinder, 261
Chemical Abstracts, 257
Chemical descriptors, 492, 495
chemical drawing and viewing software, **265**
Chemical Information Sources Wiki, 257
chemical publication
  journals, **267–8**
  open access, **268**
  theses, **268–9**
chemical software programs
  chemical drawing and viewing software, **265**
  cheminformatics software, **265–6**
  datasets, for virtual screening, **266–7**
chemical space, **521**
  search for
    biological space, **518–19**
    biological strategy, **519–20**
    size of, **518**
chemical structures representation
  3D descriptors, **214–16**
  CoMFA descriptors, **216**
  ID fingerprints, 213
  structural keys, **212–13**
  topological descriptors, **213–14**
chemical suppliers, **263–4**
chemical synthesis, **903**
ChemIDplus, 258
cheminformatics software, **265–6**
chemistry control, **487–8**
chemoinformatics, **211**, 212
chemometrics, 493
chemotherapy, *30*
ChemSources International, 263
ChemSpider, **257–8**
ChemSW, 263
Chinese hamster ovary (CHO), 834
chiral centre, deletion of, **543**
chirality
  deletion of chiral centers, 425–6, *425*
chloral hydrate, 850, *850*
chloramphenicol, 20–1
  metabolic activation, 677, *677*
chloroform, 5, 46
  metabolic activation, 677, *677*
chlorothiazide, 12
chlorpromazine (CPZ), 42, 136, 323, *324*, 327, 782, *783*, 824
chlorprothixene, *327*
chlortetracycline, 21
cholecystokinin (CCK) antagonists, 306, 355, *357*, *309*
cholinergic twin drugs, *391*
Cilazepril®, 169, *170*
cimetidine, 34, *141*
cisplatin, 28–9
claims, patents, 889
clearance, of drug elimination, 650–1
  hepatic clearance (Cl_H), 651
  renal clearance (Cl_R), 650
clemastine, 537, *538*
clinical biomarkers, **119**
clofibrate, 286, *286*
clogP software, 263
clometacin, 322, 323, *324*
clonidine, 345, 345, 434–5
clopidogrel, 100, 736
clotepin, 539
clozapine, 39, 291, *291*, 551, 555
cluster analysis (CA), 501
CMC *see* critical micelle concentration
CNG channels (cyclic-nucleotide-gated channels), 93
CNS *see* central nervous system
CNT (SLC28) transporters, 705
cocaine, 46–7, 89
  isosterism, 312, *314*
cocaine-derived local anesthetics, 355, *356*
co-crystallization, 614
codrugs, 734, *734*
cognate groups
  halogenation, **451–2**, *451*, *452*, *453*
  methyl groups, 444
colon targeting, 832–4
colonic damage score (CDS), 833
combinatorial chemistry, 177
combinatorial libraries, 482–4
CoMFA *see* comparative molecular field analysis
common ion effect, 758
Common names *see* nonproprietary names
compactin, *131*, 346, *347*
comparative molecular field analysis (CoMFA), 216, 575, 588, 601
comparative molecular similarity (CoMSIA), 601
composition claims, patents, 889
compound collections, **146**
compound properties, **481–90**
  aqueous solubility, 486–7
  chemistry control, 487–8
  combinatorial libraries, 482–4
  intestinal permeability, 484–6
  in vitro potency, 487–8
  metabolic stability, 488–90
  permeability screens, 490
conformation analysis
  3D QSAR, 590–1
conformational analysis, 367–8
  pharmacophores, 575–6
conformational energy, 472
conformational entropy, 473
conformational restriction, **363–77**
  analysis, 367–8
  bioavailability, 368
  case studies, 368–77
  constrained analogs, 366–7
  methyl groups, 434–5
  rigid compounds, 368
  steric effects, 368
  theoretical points, 364–6, *364*, 365, *366*
conjugation reactions *see* Phase II reactions
conjunctive approaches, 343, 356, 358–60, *358*, *359*, 416, *416*, *417*, *418*
consumption and production, of pharmaceuticals, **894**
  drugs, **895**
    manufacture of, **903–9**
    national differences, in prescribing, **899–902**
    prescribed drugs, **895–9**
    sources of, **902–3**
    top-earning drugs, **895**
  future of pharmaceutical industry, **918–20**
  social and economic factors, **909**
contraceptive "pill", 33–4
contraceptive steroids, 730
controlled release, drug absorption, 640
cooperative binding, 478–9
corticosteroids, 33
cosolvent, drug solubilization, *793*
cost containment measures, **914–16**
co-transporters, 698
covalent derivatives, **846–7**
COX-2, 8–9
  inhibitors *see* cyclooxygenase-2 inhibitors
"coxibs", 8–9
creams, 764–5
Cremophor EL®, 797
Creutzfeld-Jakob disease (CJD), 902
critical micelle concentration (CMC), 246, 247, 248, 794
cromakalim, 134, *135*, 345, *345*
crystal lattice cohesion, 433–4, *434*
crystal structures, 618–21
crystalline salts, 755
  aqueous solubility, 756–8
  chemical stability, 761
  co-solvents, 759
  dissolution rate, 759–60
  electrolyte effects, 758
  hygroscopicity, 758

crystalline salts (*Continued*)
melting point, 756
morphology, 760
polymorphism, 760–1
pseudopolymorphism, 760–1
crystallization, 611–15
Current Drugs' IDdb, 260
cyanosafracin, 175, *176*
cyclanol carbamates, 276, *276*
cyclenes, 447
cyclic hemi-orthoester, 771, *772*
cyclization, 345–9
cyclo-oxygenase (COX), 8
cyclodextrins (CyDs), **813–35**, 843
characteristics, 814
drug property improvement, 816–31
combined use, with additives, 829–31
drug absorption, enhancement of, 822–4
peptide and protein drugs, usage in, 826–9
release control, 821–2
side-effects, reduction of, 824–6
solid properties, control of, 819–21
solubilization, 817–18
stabilization, in solution, 818–19
pharmaceutical uses, 813–16
biological profiles, 814–16
physicochemical profiles, 814
site-specific drug delivery, 831–5
brain targeting, 835
cell targeting, 834–5
colon targeting, 832–4
structure, 814
cyclooxygenase-2 inhibitors, 302
cyclopentyl group, 440–1
cyclophosphamide, 26, 737, *737*
cyclopolymethylenic compounds, **276**
cyclopropyl group, 440, 443
cyclopropyl peptides, 376–7, *376*, *377*
cyclopropylamino acid amide, 300, *301*
cyclosporine, 25, 842
cyclosporine A, 145
cyclovinylogues, 286, *286*
CYP enzyme *see* cytochrome P450 (CYP) enzyme family
CYP monooxygenase, 676–7, *676*
Cys-loop receptor superfamily, 95
cystein delivery systems, 740, *740*
Cystic fibrosis, 912
cytochrome P450 (CYP), 674–5, 657–8, 658, 676
enzyme family, 647
cytosol, 647

D

Danio rerio, 193
Danishefsky diene, 197
DARA, 7, 393
data collection, protein crystallography, 615
data mining methods, **212**, **216–17**
database searches, **217–18**
datasets, for virtual screening, **266–7**
daunomycin, 28
Daylight Chemical Information Systems, 258
decision tree approach, 499, 526
decision trees, 499, *500*
definitions and objectives
drug development, stages of, 64–6
drug substances, 64
drugs, 64
medicinal chemistry and related disciplines, 63–4
delavirdine, **148–9**
Delphion, 269
derivatization, 724–7
Derwent, 270
Derwent World Drug Index, 522
design
natural-products-based drug discovery
active compounds isolation, 168
biomass acquisition, 163–4
extraction, 167
large-scale supplies procurement, 168–9
microbial diversity, unexplored potential of, 164–7
screening methods, 167–8
structure–activity relationships determination, 169
structure elucidation, 168
prodrugs, **721–42**
design templates, 178
designing libraries
for drug-like compound, **251–3**
Developmental Therapeutics Program Division of Cancer Treatment, 258–9
dexapanthenol, 735, *735*
"DFG-out" pocket, 610, *611*
diacylcyclopropane, 314, *315*
diamino-cyclobutene-diones, 308, *311*
diastereoisomers, **537**
diazaborine, 333, *333*
diazepam, 128, 291, *291*, *292*, 447, 776–7, *777*
dibenzoyl-2-amino-6, 7-dihydroxytetra-hydronapthalene (DB-ADTN), 725
dibotermin alfa (rhBMP2), 96
Dictionary of Natural Products (DNP), 194, 195
dicumarol, 383, *383*
diethyl dithio*iso*phtalate (Etisul®), 849, *849*
diethylstilbestrol, 344, *344*
difference maps, 618
differential scanning calorimetry (DSC), 240
differential titration calorimetry, **240**
diflunisal, *667*
digitalis, 10–11
digoxin, 646, 650, 819
dihydroartemisinin ethers, 770
dihydrofolate, 574
diltiazem, 13, 822
dioxolanes, 739, *739*
diphenhydramine, 348–9, *349*, 433, 434, *434*
dipivaloyl-epinephrine, 724–5, *725*
dipole–dipole interactions, 466, *468*
dipyrone, 773
direct derivatives
of carboxylic acid bioisosteres, 304, 306
direct ligand-gated ion channels, *93*, 93–5, *94*
Directory of Open Access Journals (DOAJ), 268
discolsure, 884–5
disease etiology, **107–8**
disease-modifying antirheumatic drugs (DMARDs), 9
disease state, 653
disintegration, drug release, 640
disjunctive approaches, **343**, 354–5, *355*, *356*, *357*, 416, *416*
dispersion forces, 471
dissociation constants, 330
dissolution rate, 759–60, *759*, *760*
distomer counteracts, the eutomer, **542**
distomer metabolization, **542–3**
distribution phase, 648
distributions, of transporters, 699–702, *701*
disulfide thiamine prodrugs, 740–1, *741*
diuretics, 12–13, 89
divalent atoms, 294
Diversity Oriented Synthesis (DOS), 187, 371, *371*, *372*
applications of, **192–3**
of small molecule libraries, **188–92**
principles, **188**
DML discovery, **553–68**
DNA-binding domain (DBD), 110
DNA ligands, **390**, *390*
DNA technology and genetics, 49–52
DNP *see* Dictionary of Natural Products
docetaxel, 27
Document Object Identifier (DOI), 259
donepezil, 40
l-Dopa (1-3,4-dihydroxyphenylalanine), 38, **138–9**
dopamine, 346, *348*
dopamine D_2-plus DMLs, for schizophrenia, **555**
dopamine receptors, 38–9
dopaminergic antagonists, 356, *358*

dopaminergic and serotonergic dual acting drug, *392*
doramectin, 172, *172*
dorzolamide, 52, 608, *608*
DOS *see* Diversity Oriented Synthesis
dosage forms, 762–5
dose–response curves, 117
double prodrugs, 732, *733*
Doxil®, 808
Doxorubicin, 28
DPI products, 764
Drosophila, 111, 118
drug, sources of, **902**
animal sources, **902**
biological sources, **902–3**
chemical synthesis, **903**
fermentation, **903**
vegetable sources, **902**
drug absorption, 637, *638*, **640–4**
dosage form, 640
enhancement of, 822
first-pass effect, *643*, 643–4
gastric emptying, 640, 642
gastrointestinal absorption, 823
GI motility, 640, 642
GI permeability, 642–3
nasal absorption, 824
perfusion of GI tract, 643
rectal absorption, 824
transdermal absorption, 823–4
drug accumulation, 644, 645
drug activity phases
pharmaceutical phase, 66
pharmacodynamic phase, 67
pharmacokinetic phase, 66, *67*
successful drug development, 67
drug administration routes, 640, *641*
drug binding, 87
drug classification systems, 67, 70–1
mechanism of action, 68–70, *69*, *70*
targets, 68
drug delivery systems (DDS), 813
drug development
stages, 64–6, *863*
drug discovery, 3–6, **144**, **857–9**
early discovery program, 858
evolution
anti-acid drugs, 34–5
chemists' contribution, against cancer, 26–30
drugs of mind, 41–9
endocrine disorders, drugs for, 30–4
giving back, 10–15
immunosuppressive drugs, 24–6
lipid lowering drugs, 35–7
microbes and viruses, fight against, 15–24
neurotransmitters and receptors, 37–41
pain killers, 6–10
exploratory research, 858
learning experiences, 859
mature discovery program, 858
protein crystallography, **605–29**
recent trends, considerations on
drug hunting, hopes and limits for, 52–5
genetics to DNA technology, 49–52
research–development interface, 859
drug disposition, **702–5**
drug distribution, 637, *638*, 644–5, 653
blood–brain barrier, 645
drug accumulation, 644, 645
plasma protein binding, 644
drug effects, measurement and expression of, 73–5, *75*, 79
ex vivo experiments, 81–2
in vitro experiments
allosteric interaction, 78–9
binding studies, 75–6
cellular and tissular functional responses, 79–81
ligand–receptor interaction-induced functional effects, 76–8
in vivo experiments, 82–3
drug elimination, 637, *638*, 645–8
biliary excretion, 646–7
biotransformation, 647–8
renal excretion, 645, *646*
active tubular secretion, 646
glomerular filtration, 646
tubular reabsorption, 646
drug hunting
hopes and limits for, 52–5
drug latentiation, 722
drug-like space, **521**
drug-likeness
assessment of, **245**
mimicking known drugs, **247–9**
prediction of, **249–50**
QSAR, 493, 509
threats avoidance, **245–7**
drug metabolism, **118**
drug nomenclature, **867–75**
INNs, for pharmaceutical substances, **868–74**
developments, 872–4
history, 868–9
procedure, 871
publication, 871–2
selection process and criteria, 869–71
ISO, common names by, **874**
national nomenclature, 874
principles, 874
proprietary names, **874–5**
protection of, 874–5
use of, 874
trade names and nonproprietary names, **867–8**
drug property improvement, **816–31**
combined use, with additives, 829
morphine, transmucosal delivery of, 831
nifedipine, modified release of, 829–30
prostaglandin, transdermal delivery of, 830–1
drug absorption, enhancement of, 822
gastrointestinal absorption, 823
nasal absorption, 824
rectal absorption, 824
transdermal absorption, 823–4
peptide and protein drugs, usage in, 826
absorption enhancement, 826
insulin, release of, 828–9
self-association, prevention of, 827–8
stabilization, 826–7
release control, 821–2
side-effects, reduction of, 824
local irritancy, 824–5
systemic detoxication, 825–6
solid properties, control of, 819
carmofur, stabilization of, 821
metastable polymorphs, crystallization of, 819–20
nifedipine, modification of, 820–1
solubilization, 817–18
stabilization, in solution, 818
cardiac glycosides, 819
prostanoids, 818–19
drug quality, **481–90**
drug–receptor interactions, **464–79**
cooperative binding, 478–9
electrostatic interactions, 465–71
enthalpy/entropy compensation, 472–3
functional group contributions, 473–8
principles, 464–5
steric interactions, 471–2
drug release
site-specific, 729–30
drug residence times, 640
drug targets
biopharmaceutical therapies, **108**
passive immunotherapy, **108–9**
disease etiology, **107–8**
identification, validation and screening
cell-based screening, **113**
clinical biomarkers, **119**
drug metabolism, **118**
G-protein-coupled receptors (GPCRs), **115–17**
Hit-to-Lead, **113**
intracellular enzymes, **115**
intracellular receptors, **113–14**
proteome mining, **109–10**
rare mutations, **109**
RNA interference (RNAi), **111–13**
toxicology, **118–19**
transgenic animals, **117–18**
yeast two-hybrid system, **110–11**, *110*

drug transport mechanisms, **697–709**
DrugBank, **260–1**
druggability, **525–6**
  degrees, of, **527–9**
druggable genome, **529–31**
drugs of mind
  anti-anxiety drugs, 43–4
  antidepressants, 44–5
  antiepileptic drugs, 47–9
  general and local anesthetics, 45–7
  lithium, 42
  neuroleptics, 42–3
  psychotrophic drugs, 41–2
duodenum, 640, *642*
dysidiolide, 201, 203
dysiherabine, 192

E

E-type prostaglandins (PGEs), 818
E7389, 169, 176, *176*
early discovery program, 858
east side modifications, 418, *419*, *420*
easy organic synthesis (EOS) rule, 425
ebselen, *333*, 334
ECO-0501, *179*, 180
ecteinascidin, 743, 175, *176*
efavirenz, **148–9**, 440
efflux transporters, 639
electron density maps, 618, *622*
electronic aspects
  isosterism, 328, *329*
electronic effects
  acetylenic groups, 445, *446*
  halogenation, 448–9, *449*
  methyl groups, 435–6, *435*, *436*
electronic Medicines Compendium, 261
electronic parameters, 327, *328*
electronic theses and dissertations (ETDs), 268, 269
electrostatic interactions, **465–71**
electrostatic similitude
  halogenation, 449, *450*
elimination
  drugs, 637, *368*
elimination half-life ($T_{1/2}$), 651
elimination phase, 648, 649
elimination reactions, 683
ELISA *see* enzyme-linked immunosorbent assays
emulsions, 764–5, 798–802
enalapril, *140*, 141, 276, 460
enalaprilat, 276
enantiomers, 541–2
  differences in, 538–9
  pharmacological profile of, 539
Enbrel, 107
"end-of-pipe" technology, 905
endocrine disorders, drugs for, 30
  antidiabetic drugs, 31–3
  contraceptive "pill", 33–4
  corticosteroids, 33
endocytosis, 639
endomorphin-2, 190
endonentan, 150–2
endophytes, *165*, 165–6
endoplasmic reticulum, 647
endothelin antagonists, **150–2**
endothelin-converting enzyme (ECE-1), 298
England, patent history prior to, 1883, 878–9
Enovid®, 34
ENT (SLC29) transporters, 705
enteral routes, drug administration, 640, *641*
enterohepatic cycling, 638, 647
enthalpy–entropy compensation, 365, *365*
enthalpy- versus entropy-driven binding, 365–6, *365*, *366*, *367*
entropy–enthalpy compensation, 365, *365*, 472–3
enzymatic inversion, **540–1**
enzyme inhibitors, 76, 383, 653
  as drugs, **88–9**
  angiotensin-converting, 169, *332*
  homodimers as, 387–90
enzyme ligands, 87
enzyme-linked immunosorbent assays (ELISA), 192
enzymes, 74
  see also metabolism
  as drug targets, 88–9
  xenobiotic transformations, **655–72**
EOS rule *see* easy organic synthesis rule
EPC *see* European Patent Convention
ephedrine, 89, 537
epibatidine (EPI), 371–2, *373*
epinephrine, 13
Epirubicin, 28
episulfonium ion, 683, *684*
epothilones, 172–4, *173*
ER β *see* estrogen receptor β
Erlenmeyer, H., 293
errors, in crystal structures, 619–20
erythromycin, 21
Estimation Program Interface (EOI) Suite, 262
estrogen receptor β (ER β), 202
estrogen-related receptors (ERR), 104
etanercept, 10, 98
ether oxygen-methylene group paradox *see* Friedman's ether oxygen-methylene group paradox
ethnopharmacology, 134
ethyl 4-biphenylylacetate (EBA), 823
ethynologues, 285, *285*, *286*
etiology *see* disease etiology
Eudismic Index, *534–5*
eudismic ratio, 537
European Patent Convention (EPC), 880, 881
European Patent Office (EPO), 269, 880–1
ex vivo experiments, for drug effects evaluation, 81–2
excitatory neurotransmitters, 93
exenatide, 32
Exo–endo amidine, 313, *314*, 318
exocytosis, 639–40
Expereact™ WEB, 263
experiential formulations, **859–63**
  biopharmaceutical classification, of compounds, 861
  oral formulations, 859, *860*
  principles, 862–3
  procedures, 860–1
  proceeding, at practical level, 861–2
  purpose, 859–60
Experimental Data Checker, 267–8
experimental design, 504–5
experimental setup, X-ray crystallography, 615–18
experimental synthesis success rate, 483
exploratory research, 858
extensive screening, **129–30**
extraction ratio, 651
extremophiles, 164–5, *165*

F

facilitated diffusion, 639
factorial design (FD), 505
false negatives
  in fragment-based screening, 239–40
  protein crystallography, 621
false positives
  in fragment-based screening, 239
  protein crystallography, 621
farnesoid X receptor (FXR), 104, 201–2, *202*
fat emulsions, 799
fast-release portion (FRP), 830
fatty alcohol propylene glycol (FAPG), 823
FBDD libraries
  creation and analysis, 234–5
FD *see* factorial design
feature trees, **213–14**
FeedDemon, 257
fenbendazole, 287, *287*
fermentation, 903
fexofenadin, 100
FFD *see* fractional factorial design
Fick's law, 639
first-pass effect, *643*, 643–4
fixed dose combination (FDC), 550
flexibility factor, 620–1
FlexX, 259
Flo, 579
flosequinan, *129*
fludelone, 174, *174*
flufenisal, 328, *329*
flumazenil, 44
flunarizine, 328, *329*

fluorescence resonance energy transfer (FRET) assays, 192
fluorine–hydrogen isosterism, 328, *329*, *330*
fluoro-nitrogen, 312, *314*
fluozetine, 45
folding, proteins, 199
Food and Drug Administration (FDA), 144, 260, 524, 866
formulation, purpose of, 859–60
formulation and packaging-sterile products, **905–6**
formulation problems solving, by chemical and physicochemical approaches
  chemical stability, increasing of, **841–4**
  gastrointestinal irritability, **847–8**
  melting point, increasing, **846**
    covalent derivatives, **846–7**
    salt/complex formation, **846**
    symmetry, introduction of, **847**
  mesomorphic crystalline forms, dealing with, **845–6**
  organoleptic properties, **849**
    odor, **849–50**
    taste, **850–2**
  painful injections, **848–9**
  peptides and proteins, improved formulation of, **844–5**
formylacetone, *283*
Fortovase®, 24
forward synthetic analysis, 188, *189*
Fourier transformation, 616
fraction bound, 644
fractional factorial design (FFD), 505
fragment, definition, 231
fragment-based drug discovery (FBDD)
  concept, **231–5**
  differential titration calorimetry, **240**
  follow-up, **239–40**
  future, **240–1**
  in late 1990s, **230–1**
  in situ click chemistry, **239**
  isothermal calorimetry, **240**
  ligand–protein interactions, **228–30**
  mass spectroscopy, **239**
  nuclear magnetic resonance, **235–7**
  SPR spectroscopy, **239**
  substrate activity screening, **238–9**
  X-ray crystallography, **237–8**
fragment-based screening (FBS)
  X-ray crystallography, 625–7, *627*
framework combination, 551, 552, *552*
France, patent history prior to 1883, 879
free R-factor, 619
Free-Wilson (FW) method, 492–3
FreePatentsOnline, 269
frequent hitters, 252
FRET *see* fluorescence resonance energy transfer
Friedman, H.L., 294
Friedman's ether oxygen-methylene group paradox, 329–30
functional analogs, 128
functional groups, **464–79**
  cooperative binding, 478–9
  drug–receptor interactions, 473–8
  electrostatic interactions, 465–71
  enthalpy/entropy compensation, 472–3
  steric interactions, 471–2
functionalization reactions *see* Phase I reactions
Furanodictin, 197
furans, 684, *685*
furosemide, 89
FXR *see* farnesoid X receptor

G

GABA A (γ-aminobutyric acid), 291
GABA site ligands, 95
GABA-A ligands, 324, *325*
GABA-A receptors, 95
GABAergic agonists, 303, *306*, 346, *348*
gabapentin, 353, *354*
galantamine, 40
ganaxolone, 438
gastric emptying, 640, 642
gastrointestinal absorption, 823
gastrointestinal irritability, **847–8**
Gaucher's disease, 912
Gaussian processes (GP), 500
gem-dimethyl groups, 440, *441*
gene therapy, 52, 834
general solubility equation, 750
generic names *see* nonproprietary names
generic pharmaceuticals, **912–14**
genes, in drug transport, 698–9
genetic algorithms, 499
genetic factor, drug metabolism, 652
genetic polymorphism, 652
genetics and DNA technology, 49–52
genistein, 202
gentamicin, 825
germs, 15–17
GI absorption, 653
GI motility, 640, 642
GI permeability, drug absorption, 642–3
GI tract
  absorption, 640, *642*
  perfusion, 643
Gleevec®, 578
glomerular filtration, 646
glucocorticoids, 25, 833
glucosides, 781–2, *781*, *782*
glucuronic acid conjugation, 683
glucuronidation, 665, *666*, 667, *668*
glucuronides, 647
glutamate-activated receptors, 94–5
glutathione conjugation, 669–71, 683
glutathione S-transferases (GSTs), **406–7**
glutathione transferases, 671
glycerol trinitrate, 11
glyceryl side chains, 780–1, *780*
glycyrrhetinic acid, 203, *203*
good manufacturing practice (GMP), **904**
gossypol, 383, *383*
G-protein-coupled receptors (GPCRs), 73, 99–103, *99*, *100*–1, **115–17**
  heterodimer ligands, 408–9
  ligands, 407–8
"green" chemistry, **906–7**
GRID, 259, 599
Grimm's hydride displacement law, 293
GRIND, 593
GSH conjugation, 669–71
guanethidine, 351, *354*
guanobenz, 251

H

H_2-receptor antagonists, 141, *141*
halichondrin B, 175–6, *176*
halogenated phenols, 328, *330*
halogenation, **448–52**
  cognate groups, 451–2
  electronic effects, 448–9
  electrostatic similitude, 449
  hydrophobic effects, 449–50
  steric effects, 448
haloperidol, 137, 555
halothane, 687–89, *688*, *690*
hanging/sitting drop crystallization set-up, 612, *613*
Hansch–Fujita approach, 492
hard gelatine capsules, 763
Harvard mouse, 883
HCFU *see* 1-hexylcarbamoyl-5-flurouracil
HCN channels (hyperpolarization-activated cyclic nucleated-gated cation channels), 93
Health Management Organizations (HMOs), 918
heavy metals, chelation, 454
Helicibacter pylori, 35
Henderson–Hasselbalch equations, 639
hepatic biotransformation, 647
hepatic clearance (Cl_H), 651, 706–7, *707*
hepatitis C virus (HCV)
  ligands for, 372–4
hepatocarcinogenicity, 678
Herceptin, 107
hereroatom oxidations, **678**
heterocycle bioisosteres, 302–3
heterocycles, 425
hetero-Diels-Alder reaction, 197
heterodimer, and dual acting ligands, 382, **391**
  dual acting drugs, examples of, **400–1**
  hybrid molecules
    acting, on receptor and enzyme, **398–400**

heterodimer, and dual acting ligands (*Continued*)
of different receptors, **391–4**
as enzymes inhibitors, **394–8**
hexane, 683, *685*
1-hexylcarbamoyl-5-flurouracil (HCFU), 821
high speed chemistry libraries, drug-likeness, 245
high-throughput screening (HTS), 129, 130–1, 244, 253, 482, 509, 522, 627
ambrisentan, 150–2
bosentan, 150–2
delavirdine, 148–9
efavirenz, 148–9
endonentan, 150–2
endothelin antagonists, 150–2
library design, 482–3
nevirapine, 148–50
Raf kinase inhibitor, 152, *153*
reverse transcriptase inhibitors, 148, *149*
screen to lead, 146–7
sitaxsentan, 150–2
sorafenib, 152
histamine H_1-plus DMLs, **558–9**
histamine H_3 ligands, 370, *370*
histaminergic hybrid drugs, *393*
histamines, 435–6, *435*, *436*
histaprodifen, 459
historical material
anaesthesia, 45–7
antidepressants, 44–5
antiepileptics, 47–9
drug discovery, 909
endocrine disorders, 30–4
immunosuppressants, 24–6
organic chemistry, 4
patents, 878
psychotrophic, 41–2
sulfonamides, 17–18
hit finding *see* in silico screening
hit optimization strategies, 416–22
hit-to-lead, **113**, 146, 147
hit/lead generation, 623–7
hits, 211–12
and leads, 125–6
HIV PR, 387
HIV protease inhibitors profiling, **222**
HIV reverse transcriptase (HIV RT), 387HIV-1 reverse transcriptase inhibitors, 148, *149*
HLB *see* hydrophile/lipophile balance system
HMOs *see* Health Management Organizations
homodimer, 383
as DNA ligands, **390**, *390*
as enzyme inhibitors, **386–90**
of pharmacological interest, **390–1**
as receptors ligands, **383–6**
symmetric nature, **383**
homologous series
benzologues, **275**
biological response curves, **277–83**
classification, **275–7**
vinylogues, **275**
Hospice movement, 7
HR22C16, 192
HRV coat protein, 326, *327*
HTS *see* high-throughput screening
HTS/HTL approaches versus FBDD, **233–4**
human genome project
database, 260
lead compounds, **125**
target identification, **109–19**
human histamine H_3 receptor (hH_3R) antagonist, 580, 581, *581*
human immunodeficiency virus (HIV), 23–24
human papilloma virus (HPV), 902
Human Rhinovirus (HRV) coat protein, 326, *327*
human thrombin crystals, *612*
humors, 3
Huntingdon's chorea, 911
hybrids *see* non-identical twin drugs
hydantoins, 663–4
isosterism, 294, *294*
ring transformations, **343**
hydride displacement law, 293
hydrogels, 805–6
hydrogen bond interactions, 713–15, 718–19
hydrogen bonding
functional groups, 468–9, *468*, *469*
intestinal permeability, 485
hydrolases, 659–60, *659*
hydrolysis, 663–4, *664*
hydrophile/lipophile balance (HLB), 79
hydrophilic CyDs, 813
hydrophilic solvents, *790*
hydrophobic CyDs, 813–14
hydrophobic interactions, 432–3, *433*, 449–50, *450*, 472–3
hydroxamic acid vinylogues, 284–5
hydroxamic acids, 284, 304, *306*
hydroxyethylureas, 314, *315*
hydroxy-isoxazoles, 306
hydroxyl group
drug–receptor interactions, 476, *476*
hydroxylation, **452–4**
ligand–receptor interactions, 453, *454*
metabolism, 453–4
solubility, 453, *453*
hydroxypropylcellulose (HPC), 830
hygroscopicity, salt selection criteria, 758
hypochlorites, 16
hypothermia, 653
hystaminergic hybrid drugs, *393*

I

ibritumomab, 98
ibufenac, 869
ID fingerprints, 213
ifosfamide, 26
iGluRs *see* ionotropic glutamate receptors
IKK inhibitors, 152, *154*
ileum, 640, *642*
Ilotycin®, 21
imatinib, 88, 560–1, 610, *610*
imidazolinediones, 432, *433*
imidazo[1,2-a]pyridine bioisosteres, 301
imidazo[1,2-a]pyridines, 345, *346*
imipramine, 44, 89, 129, 323, *324*, 327, *327*
immunoglobulins, 903
immunosuppressive drugs, 24–6, 171, *171*
"important" drugs, **895–902**
in silico screening, **210**, 624–5
chemical structures representation, **212–16**
chemoinformatics, **211**
data mining, **212**, **216–17**
database searches, **217–18**
future directions, **222**
hit versus lead structures, **211–12**
ligand-based *in silico* screening, **218–19**
parallel *in silico* screening, **219**
parallel pharmacophore-based virtual screening, **219–22**
structure-based *in silico* screening, **219**
in situ click chemistry, **239**
in vitro experiments, for drug effects evaluation
allosteric interaction, 78–9
binding studies, 75–6
cellular and tissular functional responses, 79–81
ligand–receptor interaction-induced functional effects, 76–8
in vitro false positives, *252*
in vitro potency, **487–8**
in vivo experiments, for drug effects evaluation, 82–3
InChI, 257
inclusion complex, 817
incremental cost-effectiveness ratio (ICERs), 917
indazolones, 315, *316*
indenoisoquinoline NSC, 314622, 136, *136*
indifferent electrolyte effect, 758
indigenous medicines *see* ethnopharmacology
indirect toxicity, 675
indole ring, *316*, 317
indolylacetic acid, 137, 138
indomethacin, 322, 323, *324*, *739*
inductive interactions, 466–8
industrial applications, inventions, **887–8**

industrial chemical products, 136–7
infliximab, 10, 98
influx transporters, 639
information, patents as source, 891
infringement of patents, 890–1
inglicretin, 869
inhibitors of cyclooxygenase (COX-2), 396, *397*
inhibitors of factor Xa (FXa), 387
inhibitory neurotransmitters, 93–4, *94*
Initiative for Chemical Genetics (ICG), 258
INNM *see* modified International Nonproprietory Names
innovation, drug discovery, 857–9
INNs *see* International Nonproprietory Names
inogatran, 350, *352*
inotropic drugs, 392
InstantMT™, 270
insulin, 31–2
  nasal absorption, 826
  release of, 828–9
  self-association prevention, 827–8
intellectual property digital library (IPDL), 270
interaction capacities, 215, 536
Interactive Laboratory (I-Lab), 262
intercellular fluid *see* interstitial fluid
intercellular recognition events, 834
International Chemical Identifier (InChI), 212
International Committee of Harmonization (ICH), 866
International Nonproprietory Names (INNs), 867
  developments, 872–4
  history, 868–9
  procedure, 871
  publication, 871–2
  selection process and criteria, 869–71
International Organization for Standardization (ISO), 874
International Patent Classification (IFC), 270
International Proprietary Name (INN), 245
International Union for Pure and Applied Chemistry (IUPAC), 257, 868
International Union of Biochemistry (IUB), 257, 868
Internet Drug Index, 260
internet *see* web alert
InterPro domain annotation, 529
interstitial fluid, 649–50
intestinal absorption, 705–6, *706*
intestinal permeability, 484–6
intracellular enzymes, **115**
intracellular receptors, **113–14**
intramolecular hydrogen bonding, 485
intrinsic binding energy, 473–4, *474*
intrinsic clearance (Cl'_{int}), 651
intrinsic enzyme activity, 95–6
inventions, 882
  see also patents
  categories, 888
  industrial application, 887–8
  novelty, 884–6
  patentability, 886–7
inventive step concept, 886–7
inverse agonists *see* negative antagonists
Invirase®, 24
invitro potency
  chemistry control, 487–8
iodine-free chlorquinaldol (Sterosan®), 849
iodine "tincture", 16
iodochlorhydroxyquin (Vioform®), 849
iodoform, 849
ion channels *see* direct ligand-gated ion channels; non-selective cation channels
  ion-gated, 74, **93–5**
  voltage-gated, **90–2**
ion trapping, 639
ionic bonds, 465
ionizable cyclodextrins, 814
ionizable CyDs, 814
ionization constant, 494, 751, *752*
ionotropic glutamate receptors (iGluRs), 192
ionotropic receptors *see* glutamate-activated receptors
IP News Flash, 270
iproniazid, 44
IRAK (interleukin receptor-associated kinase), 98
irritancy, 824–5
  cyclodextrins, 816, *816*, **824–5**
  injection site, **824–5**, **847–9**
ISO, common names by, **874**
  national nomenclature, 874
  principles, 874
isochondrodendrine, 383, *383*
isoenzymes, 88, 387, 394
isoguvacine, 303, *306*
isolation of active compounds, **168**
isomer effects
  on absorption and distribution, **540**
  on excretion, **541**
  on metabolism, **540–1**
  on uptake, **541**
isoniazid, 683, *684*
isoniazide, 134
isonicotinic acid hydrazide, 20
isopropyl group, 440
isopropyl hydrazide, *728*
isosterism, **290–334**
  see also bioisosterism
  analysis, 326–30
  anomalies, 328–30
  carboxylic acid, 306, *306*, *307*
  concepts, 291–4
  definition, 290
  electronic parameters, 328
  ether oxygen and methylene group, 329
  functional groups, reversal of, 320–3
  metalloids, 330–4
  modifications, 294–323
  peptides, 320
  phenols, 318
  ring equivalents, 297–303, *299*
  scaffold hopping, 323–6
  solubility, 327–8
  steric effects, 328
  structural parameters, 327
isothermal calorimetry, **240**
iterative loops, *862*
IUB *see* International Union of Biochemistry
IUPAC *see* International Union for Pure and Applied Chemistry
ivermectin, 172, *172*

J

JAK/STAT-coupled receptors, 97, *97*
JChem, 267
jejunum, 640, *642*

K

K. bulgaricus, 834
K^+ channels *see* potassium channels
Kaletra®, 24
kallikrein–kinin pathway, 139
ketones
  prodrugs derivatization, 725, *726*
  solubilization, 775, *776*
ketoprofene, 346, *347*
khellin, 354, *356*
kininase-II, 14
"Kinome", 109
Kohonen mapping (KM), 501–2

L

labetalol, 537
language problems, International Nonproprietory Names, 869
large-scale fermentation, 168–9
latentiation, drug, 722
law of mass action, 644
LDA *see* linear discriminant analysis
lead compounds, 476–7
  analog design, 126–9
  biological information, 134–8
  chemistry control, 487–8
  clinical observations, 134–5
  complexity, 487–8
  computer-assisted drug design (CADD), 417
  drug discovery, **858**

lead compounds (*Continued*)
extensive screening, **129–30**
generation, **551–3**
high-throughput screening, **130–1**
and hits, 125–6
natural products, 159–80
old drugs, 135–6
planned research and rational approaches, 138–41
privileged structures, 247–8
random screening, **130**
rule-based filters, 246–7
SOSA approach, **132–3**
strategies, **125–42**
synthethic intermediates, screening of, **131–2**
systematic screening, 129–33
undesirable functionalities, 245
lead optimization (LO), 144, 627–9, *629*
setting, 241
lead structure, 148, 211–12
lead-likeness
designing libraries for, **251–3**
quality of, **250–1**
learning curve, 145
learning experiences, 859
leave-one-out (LOO) cross validation method, 594
legal issues
patents, **878–93**
representatives, 889
leinamycin, 844
leucine-enkephalin, 438
leuprorelin, 100
LFER (linear free-energy relationship) *see* Hansch–Fujita approach
LHRH *see* luteinizing hormone-releasing hormone
library design, 508–9
Librium®, 43, 44, *793*
lifetime of patents, 890
ligand efficiency (LE), 230, **250**, 478
ligand–protein interactions, **228–30**
hydroxylation, 453, *454*
ligand–receptor interaction-induced functional effects, 76–8
see also drug–receptor interactions; protein–ligand complexes
ligand superposition techniques, 576
ligand-based *in silico* screening, **218–19**
ligand-based NMR techniques, **235–7**
ligand-binding domains, 96, 202
ligand-fishing protocol, 614–15
ligand-gated ion channels, 74, **93–5**
LigandScout, 578–9
light microscopy, 615–16
linear discriminant analysis (LDA), 495, 498
linear free-energy relationship (LFER) *see* Hansch–Fujita approach
LinkOut, 270
Lipid Library, 802
lipid lowering drugs
lipoproteins, 35–6
statins, 36
Lipid Metabolites and Pathways Strategy (Lipid MAPS), 802
lipid vehicles, **798–803**
"Lipinski's Rule of Five", 788
lipitor, 170–1
lipophilic lead molecules, 145
lipophilicity, 459, 494
methyl groups, 430, *432*, *433*
lipoproteins, 35–6
liposomes, 802–3, 807–8
lithium therapy, 42
liver, 643, 646–7
liver X receptors (LXR), 103–4
lobelanine, 383
LOC *see* US National Cancer Institute's Letter of Collection
local anesthesia, 45–7
log *D*, 246
log P, 245, 246, 717–18
LOO method, 502
lopinavir, 24
loratidine, 555, 558
losartan, 15, 458, 555, 735, *735*
lovastatin, *131*, 36
Lovastatin® *see* mevinolin
lozapine succinate, *291*
lucensimycin A, 179, *179*
luteinizing hormone-releasing hormone (LHRH), 822

M

M. vastus lateralis, 816
Macherey-Nagel, 267
machine learning, 217
magic bullet, 519
"magic methyl", 526, 527
magic shotgun, 519
manufacture of drugs, **903**
choice of, regents, **906**
downstream processing, **907**
formulation and packaging-sterile products, **905–6**
good manufacturing practice (GMP), **904**
"green" chemistry, **906–7**
outsourcing, **907–9**
plant design, **904–5**
MAOIs *see* monoamine oxidase inhibitors
maprotiline, 327, *327*
marine microbes, 166
marinomycins, 166, *166*
Markush searching, 218
mass spectroscopy (MS), **239**, 251
MATE genes, 698
MATE transporters, 705
Materia Medica, 4
matrix metalloproteinases (MMPs), 387, *389*
mature discovery program, 858
MatWeb, 261
MDDR database, drug likeness, 247
MDI products, 763
MDL Drug Data Report (MDDR), 259
me-too compounds, 126, 127, *127*
me-too copies, 349
me-too products, 128
mean residence time (MRT), 821
mechanism of action (MOA), 68–70, *70*, *71*
medicinal chemist, **234**
mefenamic acid, 847
meglitinides, 32–3
melagatran, 350, *352*
melatonin, *359*, 360, 385, *386*
melting point, 756
membrane transport, 724–8
membrane transporters, as drug targets, 87, 89–90
menopause treatment, 730–1, *731*
meperidines, 295, 322, *323*
meprobamate, 43
MEPs *see* molecular electrostatic potentials
mercaptopurine, *131*, 132
mesomorphic crystalline forms, dealing with, **845–6**
met-enkephalin, 438, *438*
META software, 250
MetabolExpert software, 250
metabolic conversions, 674–5, **685–93**
metabolic databases, **260**
metabolic diseases, 560
metabolic stability, 488–90
brain penetration, 715–17
oral bioavailability, 711–12
"metabolic syndrome", 560
metabolism, 653
hydroxylation, 453–4
methyl groups, 437–8, *437*, *438*, *439*
metalloids-toxic isosters, 330–4
metasites and databases, **272**
metastable polymorphs
crystallization of, 819–20
metformin, 32
methicillin resistant *Staphyloccus aureus* (MRSA), 919
methoctramine, 385
methotrexate, 24–5, 131–2, *131*, 574, 895
methoxy-6-amino-8-quinoline, 280, *281*
methyl group, **432–41**
conformational effects, 434–5, *435*
drug–receptor interactions, 475–6, *476*
electronics effects, 435–6, *435*, *436*
extensions, 440–1
metabolism, 437–8, *437*, *438*, *439*

oxidation, 437, *437*
solubility, 432–4, *432*, *433*, *434*
methylation, 683
xenobiotic transformations, 665, *667*
methylene blue, 17
methylenedioxyl derivatives, 735, *735*
mevalonic acid, 170
mevastatin, 36
mevinolin, *130*, 169–71, 346, *347*, 425–6, 543
micelles, 794, 795, 803
micro- and nano drug delivery systems, 803–6, *804*
microbial anticancer agents, 172–4
microbial diversity
combinatorial biosynthesis, 167
culturing procedures, 164
endophytes, *165*, 165–6
extremophiles, 164–5, *165*
marine microbes, 166
microbial symbionts, 167
microbial symbionts, 167
microbiology, 137–8
microemulsions, 798–802
microparticles, 805
MicroPatent, 269
microsomes, 647
microvilli *see* villi
Midazolam, 44, *291*
MIFs *see* molecular interaction fields
milnacipran, 370, *370*
minaprine, 133, *133*
mini-receptor model *see* pseudo-receptor model
minor modifications rule, 422–3, *423*
minoxidil, 665, *667*
MIR *see* multiple isomorphous replacement
mixed bioactivation mechanisms, 738–9
mixed-type prodrugs, 740–1
MLR *see* multiple linear regression
MOA *see* mechanism of action
model building, protein crystallography, 618
modified International Nonproprietory Names (INNM), 869
molecular datasets, **259–60**
molecular docking, 578
molecular drug targets
direct ligand-gated ion channels, 93–5
enzymes, 88–9
ligands for, 87–8
membrane transporters, 89–90
non-selective cation channels, 92–3
quantity, 86
receptors
cytosolic proteins, coupled to, 97–9
G-protein-coupled receptors, 99–103
nuclear receptors, 103–4
with intrinsic enzyme activity, 95–6
response of organism, 86–7
target–drug interaction, 87
voltage-gated ion channels, 90–2
molecular dynamics (MD) simulations, 576
molecular electrostatic potentials, 591
molecular interaction fields (MIFs), 590
molecular number, 291–2
molecular property prediction, 261–3
molecular replacement, 617
molecular similarities, 500–2
determination, 597, 599–601, *599*
molecular variations
benzologues, 286–7, *287*
homologous series, **275–83**
isosterism, **290–334**
vinylogues, **283–7**
molecular viewers, 265
"molecule from hell", 482, *482*
MOLPRINT 2D, 266
Monascus rubber, 170
monoalkylated derivatives, **275**
monoamine oxidase, 88, 443
monoamine oxidase inhibitors (MAOIs), 44, 542
monoamine transporter, 555
monoclonal antibodies, 51, 903
monooxygenases, 657–8, *657*, *658*
morphine, 6–7, 355, *357*, *439*, 645
transmucosal delivery of, 831
MS techniques, 168
Mu-opioid-plus DMLs, **563**
multi-drug resistance (MDR), 89, 807
multiple emulsions, 799
multiple isomorphous replacement (MIR), 617
multiple kinases, **560–1**
multiple linear regression (MLR), 493, 498, 502
Multiple wavelength Anomalous Dispersion (MAD), 617
multi-target drugs, *550*
activity profile and wider selectivity, optimization of, **563–5**
designed multiple ligands (DML) discovery, **553**
AChE-based DMLs, for Alzheimer's disease, **559–60**, *559*
angiotensin system, for hypertension, **555–8**
archidonic acid cascade, **561–3**
dopamine D_2-plus DMLs, for schizophrenia, **555**
histamine H_1-plus DMLs, for allergies, **558–9**
Mu-opioid-plus DMLs, for treating pain, **563**
multiple kinases, for treating cancer, **560–1**
PPAR-based DMLs, for metabolic disease, **560**
SERT-plus DMLs, for depression, **554–5**
lead generation, **551–3**
physicochemical challenge, **565–8**
Muromonab-CD3, 109
muscarinic ligands, 312, *312*
muscarinic M1 agonists, *133*
mutagenecity, chain length, 422, *423*
Mycobacterium tuberculosis, 205
Mycocet®, 809
myeloperoxidase (MPO), 833

N

N-acetyl-*p*-benzoquinone imine (NAPQI), 685–7
N-benzylimidazole, 15
N-heterocycles, *196*, 197
N-methylation, 665, *667*
N-methylpyridinium-2-carbaldoxime, 737
N-type calcium channel, 111
nafadotride, *133*
nanocolloids, 806–7
cancer treatment, **807–8**
drug delivery and clearance mechanisms, **806–7**
drug toxicity modification, **808–9**
nanomedicines, 803
nanoparticles, 804–5
Nanoxel®, 807
NAPAP, 345, *346*
NAPQI *see* *N*-acetyl-*p*-benzoquinone imine
narcotics, 442
nasal absorption, 824
natalizumab, 99
National Cancer Institute (NCI), 145, 259
national differences, in prescribing, **899–902**
National Institute for Clinical Excellence (NICE), 916, 917, 918
National Institute of Health (NIH) Roadmap, **258**
national nomenclature, 874
natural product scaffold tree, 197
natural products as drugs, 159
design, 163–9
examples, 169–77
future directions, 177–80
importance, 160–3
Nature Chemical Biology, 258
NCI DIS 3D database, 258
NCI *see* National Cancer Institute
NCMS, 270
nefopam, 348, *349*
negative antagonists, 87–8
Nilevar®, 34
neostigmine, 416, *416*
nephron, 645, *646*
NETSci, 266

Networked Digital Library of Theses and Dissertations (NDLTD), 268
neural networks, 498–9
neuraminidse inhibitors, *275*
neurokinin NK_1, 555
neuroleptics, 42–3
neutral endopeptidase (NEP), 396, *397*
nevirapine, **148–50**
  evolution of, *150*
new drug applications (NDA), 160, 173
new molecular entity (NME), 909
nicotine, 299, 346, 348, *348*
nicotinic acetylcholine receptors (nAChRs), 95, 371, 385
nifedipine, 400
  modification of, 820–1
  modified release of, 829–30
nilotinib, 610, *610*
NIST Chemistry webBook, 261
nitrendipine, 385
nitric oxide (NO), 11
nitro compounds reduction, 681–2, *682*
nitrogen mustards, 26, 737–8, *738*
nitroglycerine, **11**
nitroimidazolylmethyluracils, 738
NLM *see* non-linear mapping
NMDA receptor antagonists, 318, 358
NME *see* new molecular entity
N-methyl-d-aspartate receptors (NMDAR), 94–5
NMR *see* nuclear magnetic resonance
NMRShiftDB, 258, 267
NNRTIs *see* non-nucleoside reverse transcriptase inhibitors
NO *see* nitric oxide
nociceptin, 374, *374*
noncovalent interactions, *467*
non-identical twin drugs, 381, **401**, **403**, **408**
non-ionizable moieties, 769
non-ionizable side chains
  solubilization, **780–2**
    glucosides and related compounds, 781–2
    glycolyl and glyceryl side chains, **780–1**
    polyethylene glycol derivatives, **781**
non-linear mapping (NLM), 501
non-nucleoside reverse transcriptase inhibitors (NNRTIs), 24, 219, 396
non-peptide bradykinin B_2 receptor antagonists, 301, *301*
non-planar sulfur- or phosphorus-derived acidic functions, 306, 308–10
nonproprietary names
  and trade names, **867–8**
non-responders, 901
non-selective cation channels, 92–3
non-steroidal antiinflammatory drugs (NSAIDs), 324
  and aspirin, 7–8
NOP receptors, 374
Nordic Pharmacopoeia Names (NFN), 874
norepinephrine, 13
norfloxacin, *129*, *358*, 359, *359*
north side modifications, 418–19, *419*
noscapine, 715
novel cholecystokinin-2 (CCK_2) antagonist, 326, *326*
novel lead structures, 144, 145, 180
novelty, inventions, 884–6
NRTIs *see* Nucleoside reverse transcriptase inhibitors
NSAIDs *see* non-steroidal antiinflammatory drugs
NSC 314622, 136, *136*
nuclear magnetic resonance (NMR), 168, 251
  ligand-based screening, **235–6**
  protein-based screening, **237**
nuclear receptors, 103–4
Nucleoside reverse transcriptase inhibitors (NRTIs), 396

O

O-glucuronidation, 665, 667
O-heterocycles, *196*, 197
O-methylation, 665, *667*
OAT (SLC22) transporters, 704
OATP transporters *see* organic anion-transporting proteins transporters
observation exploitation, 134–8
OCT (SLC22) transporters, 703–4
octanol–water partition coefficient, 261
oil-in-water emulsions, 798
ointments, 764–5
olanzapine, 43, 100
old drugs, new uses of, 135–6, *136*
omalizumab, 98
omeprazole, 89, 738, *738*, 843
omit maps, 618
omits over-the-counter (OTC), 895, 901
one disease–one target–one ligand approach, 382
One-Cell Systems, 164
one-electron oxidation, 677, *678*
Online Mendelian Inheritance in Man (OMIM), 259
open access approach, **268**
open-chain analogs, **343–5**, *344*, *345*
open difunctional polymethylenic series, **276–7**, *276*
OpenBabel, 266
opioid motif, 7
opioid receptor ligands, 374–5, *375*
optical isomerism, in drugs
  experimental facts, interpretation
    biological active compounds, stereoselectivity in, **533–5**
    diastereoisomers, **537**
      Pfeiffer's rule, **538**
      stereoselectivity ratios, **537–8**
      three-point contact model, **535–6**
  and pharmacodynamic aspects, **538**
    enantiomers, differences in, **538–9**
    pharmacological profile, of enantiomers, **539**
  pharmacokinetic effects, **539**
    isomer effects, **540–1**
  practical considerations
    chiral centre, deletion of, **543**
    distomer counteracts, the eutomer, **542**
    distomer metabolization, **542–3**
    enantiomers, **541–2**
    racemates, **541–2**
    racemic mixtures, usefulness of, **543–6**
    racemic switches, **542**
oral bioavailability, **711–15**
  hydrogen bond interactions, 713–15
  metabolic stability, 711–12
  structural rigidity, 712–13
oral formulations, 859, *860*
oral solutions, 763
organic anion-transporting proteins (OATP) transporters, 703
Organic Compounds Database, 261
organic solvents, **789–94**
Organic Synthesis (OS), 264
organoborons, 210, 333
organoleptic properties, **849**
  odor, **849–50**
  taste, **850–2**
organosilicon compounds, 330–1
orphan drugs, **911–12**
orphan receptor, 103, 113
orphenadrine, 348, *349*
OS *see* Organic Synthesis
OSCAR toolkit, 267–8
oseltamivir, *608*, 609
OSIRIS Property Explorer, 262, 272
"Ostwald's Rule of Stages", 820
overlap mode, 381, *381*
ownership of patents, 890
oxacillin, 869
oxazaborolidin, 333, *333*
oxazolidine-diones, 294, *294*
oxfenicine, 790, *791*, 792
oxidation, **676–8**, *679*, *680*
  bioactivation, 735–7
  CH bond oxidations, 677, *677*, *678*
  heteroatom oxidations, 678, *680*
  methyl groups, 437
  N-oxidations, 678, *679*
  unsaturated bond oxidations, 678
  xenobiotic transformations, 660–3, *660*, 662, *663*
oxidative attack, *437*, 438
oxidative stress, **678–80**, *693*

oxidoreductase, 657, *657*, 659
oxime ethers, *286*
l-2-oxothiazolidine-4-carboxylate, 740, *740*
l-2-oxothiozolidine-4-carboxylate, 740, *740*
oxotremorine, 285, *285*, 350–1, *353*
oxytetracycline, 21
OZ277, 177, *177*

P

P-glycoprotein (P-gp) multidrug transporter, 643, 646, 702
p-halogenated phenols, 328, *330*
P13Ks, 102–3
P2X-ATP receptors, 94
paclitaxel, 27, 165, 168, 175, *175*
pain killers
  best-sellers and controversies
    aspirin and NSAIDs, 7–8
    "coxibs", 8–9
    poppy extracts led to receptors, 6–7
    rheumatoid arthritis, new strategies for, 9–10
painful injections, **848–9**
p-aminosalicylic acid (PAS), 20
paracellular drug transport, 640
paracetamol, 685–7, *686*, *687*, 868
parallel artificial membrane permeation assay (PAMPA), 485
parallel chemistry, 253
parallel *in silico* screening, **219**
parallel pharmacophore-based virtual screening, **219–22**
parallel trade, **914**
parenteral routes, of drug administration, 640, *641*
parenteral solutions, 763
Paris Convention, 879, 886
parkinsonism, **138–9**, 555
paroxetine, 89
partial least squares (PLS), 493, 498, 502, 592
passive diffusion, 638–9
passive immunotherapy, **108–9**
Patent and Trademark Office (US), 269
Patent Cooperation Treaty (PCT), 880–1
patent information, **269**
  Japanese patent, **270**
patents, **878–93**, **910–11**
  basic requirements, 882–8
  definition, 882
  formal requirements, 888–90
  history prior to, 1883, 878–9
  infringement, 890–1
  kinds of inventions, 882
  lifetime, 890
  ownership, 890
  pharmaceutical industries, 891–2
  source of information, 891
Paterra, 270
pathways, *518*
pattern and cost, of innovation, **909–10**
PCA *see* principal component analysis
PCT *see* Patent Cooperation Treaty
PD 118440, 319, *320*
PDB *see* Protein Data Bank
PDE-deficient yeast (PMY), *115*
pegaptanib, 96
pencillins, **19–20**
penicillamine, 454
Penicillium brevicompactum, *170*
Penicillium citrinum, 170
pentoxifylline, 870
PEPT1 (SLC15A1) transporters, 704
PEPT2 (SLC15A2)transporters, 704
peptide and protein drugs, usage in, 826
  absorption enhancement, 826
  insulin, release of, 828–9
  self-association, prevention of, 827–8
  stabilization, 826–7
peptide GPCRs, 566–7
peptides
  isosterism, 313, *315*
  reversal of functional groups, 320, 322–3
peptides and proteins, improved formulation of, **844–5**
peptidomimetics, for SH2 domains, 375–7
  cyclopropyl peptides, 376–7
  metathesis, rigidification by, 375–6
percentage bound, 644
perfusion limited distribution, 644
peripheral binding site, 403
permeability limited distribution, 644
permeability screens, 490
permeability testing, 485–6
PERMUTE, 259
peroxisome proliferators-activated receptor (PPAR), 32, 385–6, *387*, 394, *396*
peroxisome proliferation-activated receptor γ (PPAR γ), 202
pethidine, 283, *284*
PFam, 529
Pfeiffer's rule, **538**
"Pfizer's rule of five", 859
pH
  salt former selection, 754–5
pharmaceutical development, in industry, **863–5**
pharmaceutical industry, 127, 211, 222, 903, 904, 907, 909, 911
  drug discovery, **230–1**
  patenting, 891–2
  patents, 891–2, **919–20**
  trend, 542
pharmacodynamics, 67, 705
pharmacoeconomics, **916–18**
pharmacokinetic (PK) profile, 147
pharmacokinetic effects, **539–46**
pharmacokinetics, physiological aspects determining, **637–54**
  age, 652
  bioavailability, 651–2
  disease state, 653
  drug absorption, 640–4
  drug distribution, 644–5
  drug elimination, 637, *638*, 645–8
  drug interactions, 653
  drug passage through biological barriers, 638–40
  elimination half-life ($T_{1/2}$), 651
  genetic factors, 652
  hepatic clearance (Cl_H), 651
  plasma concentration–time curve, 648–9
  pregnancy, 653–4
  renal clearance (Cl_R), 650
  volume of distribution, 649–50
pharmacological logic rule, 426
pharmacological profile, 147
pharmacophore, **572–84**, 590
  applications, 574–5
  definitions, 573–4
  determination, 581–2, *583*
  identification, 424
  importance, 574
  modeling, 575–7
  pseudo-receptor models, 579–80
  screening, 580, 582, 584, *584*
  structure-based pharmacophores, 577–9
Pharmacophore Point Filter, 247
pharmalogical space, **521**
  chemical space, **521**
    drug-like space, **521**
  target space, **524–31**
    degrees, of druggability, **527–9**
    druggability, **525–6**
    druggable genome, **529–31**
    structure-based druggability, **526–7**
PharmGKB, 260
phase I reactions, 656–64
  bioprecursor-prodrugs principles, *723*
  enzymes, 657–60
  oxidation, 660–4
  reduction, 660–4
phase II reactions, 664–71
  acetylation, 667–8
  acyl coA intermediates, 668–9
  glucuronidation, 665, *666*, 667, *668*
  glutathione conjugation, 669–71
  methylation, 665, *667*
  sulfonation, 665
phase problem, X-ray crystallography, 616–17, *617*
phenformin, 32
phenol bioisosteres, 318–19
phenols, 724–5
phenylbutazone, 283, *284*, 349, *352*
phenylpropanolamine, 348, *349*
phenytoin, 647, 848

phodophyllotoxine, 165
phomallenic acid C, 179, *179*
phosphate esters, 772, *773*
phosphodiesterases (PDEs), 102, 115
phospholipases C (PLC), 102
phosphonates, 295
phosphonates, 306, *308*
Physarum plycephalum, 292
physostigmine, 296, *296*, 416, *416*
picenadol (LY 150720), 542
pilocarpine, 311, 734, *734*, 842
piperaquine, 390
pirenzepine, 148, 418, *418*, *419*
pivampicillin, 647
PJB Publications' Pharmaprojects®, 260
pKa, 716
  see also ionization constant
PKC *see* Protein kinase C
placental barrier, 653
planar acidic heterocycles and aryl derivatives, 306, *307*
plant design, **904–5**
plant observations, **137–8**
plasma concentration–time curve, 648–9
plasma membrane *see* cell membrane
plasma protein binding, of drugs, 644
Plasmodium falciparum, 770
Plasmodium gallinaceum, 131
platencin, 179, *179*
platensimycin, 179, *179*
platinum anticancer drugs, 28–9
PLS *see* partial least squares
polar surface area (PSA), 246, *246*, 247, 485
polarized expression, of transporters, 700, *701*
poling, 576
poly(ethylene glycol) (PEG), 828
polyethylene glycol (PEG), 781, *781*
polyhalogenated compounds reduction, 681, *682*
polymeric micelles, 803–4
polymorphism
  salt solid-state properties, 760–1
polyspecific transport, 701
poor solubility, 486, 489
  library design, 483
positional isomery, *323*
potassium canrenoate, 344, *344*
potassium channels, *91*, 91–2
potential rings, 344, *344*
PPAR γ *see* Peroxisome proliferation-activated receptor γ
PPAR-based DMLs, **560**
PPARs *see* peroxisome proliferators-activated receptor
PPs *see* principle properties
practolol, 322, *323*
pramipexole, 319, *320*
pravastatin®, *131*, 170–1, *170*
pregabalin, 49
pregnancy, variability in pharmacokinetics, 653
prescribed drugs, **895–9**
"prey" protein, 110
PrGen, 579
Prialt®, 161
PriceWaterhouseCoopers, 863
primary transporters, 698
principal component analysis (PCA), 500–1, 592
principle properties (PPs), 505
Priority Right, 880, 885
privileged structures, **247**, *247*
probucol, 137
procainamide, 286, *286*
process claim, patents, 889
prodrugs, 647
prodrugs designing, **721–42**
  bioavailability, 734
  biomembrane passage, 724–8
  carrier prodrugs, 722–3, 724–35
  cascades, 731–3
  classification, 721–2
  codrugs, 734
  definitions, 721–2
  limitations/difficulties, 741–2
  mixed bioactivation mechanisms, 738–9
  mixed-type prodrugs, 740–1
  oxidative bioactivations, 735–7
  practical applications, 724, *724*
  prolonged duration of action, 730–1
  reductive bioactivations, 737–8
  site-specific delivery, 728–30
  soft drugs, 734
product protection, legal aspects, **878–93**
products claim, patents, 889
proguanil, 344, *344*
prolonged duration of action, 730–1
promethazine, 323, *324*
Prontosil®, 18
propranolol, 13, 542, *656*
proprietary names, **874–5**
  protection of, 874–5
  use of, 874
prostacyclin (PGI_2), 818
prostaglandin
  transdermal delivery of, 830–1
prostaglandin E_1 (PGE_1), 817
prostaglandins (PGs), 818
prostanoids, 818–19
protease inhibitors (PIs), 24
protein crystallography, **605–29**
  crystallization, 611–15
  data collection, 615
  experimental setup, 615–18
  hit/lead generation, 623–7
  information content, 618–21
  lead optimization, 627–9
  target, 621–2
    identification, 622–3
    selection, 623
    validation, 623
Protein Data Bank (PDB), 199
protein kinase A (PKA), 101–2
protein kinase C (PKC), 197, 387
protein kinase inhibitors, *388*
protein–ligand complexes, 614
protein structure similarity clustering (PSSC), 194, **199–202**
protein tyrosine kinases (TKs), 29
protein tyrosine Phosphatase Shp-2, 197
protein tyrosine Phosphatase-1B, 197
protein-based NMR techniques, **237**, 240
proteinuria, 653
proteome mining, **109–10**
proton pump inhibitors, 34–5
Protys, 270
Prous' Ensemble, 260
pseudoatoms, 293
pseudocycles, **344–5**
pseudo-receptor models, 579–80
PSSC *see* Protein structure similarity clustering
psychotrophic drugs, 41–2
PubChem, **258**
Public Library of Science (PLoS), 268
PubMed, 260
purchasing power parties (PPPs), *900*
pyridazine bioisosteres, 301–2
pyridine bioactivation, 683, *684*
pyridine bioisosteres, 299–301, *300*
pyrroline-3-ones, 314, *315*

Q

QikProp, 249
QSAR *see* quantitative structure–activity relationships
QSAR softwares, *509–10*
Quality Adjusted Life Year (QALY), 917
quality criteria, 618–19
quality fixing, and product development, **865–6**
quantitative structure–activity relationships (QSAR), 216, 250, **491–510**, 580
  applications, 505–10
  autoQSAR, 503–4
  consensus modeling, 503
  descriptors, 493–6
  experimental design, 504–5
  global models, 503
  inverse QSAR, 505
  local models, 503
  predictive models, 496–503
  time-series behavior, 503–4
Quasar, 579
Quetiapine fumarate, *291*
quinidine, 539
quinine, 539

quinolones, 22
quorum sensing, 161, *161*

R

R(−)epinephrine, 535, 536, *536*
R-factor, 619
R-sym, 619
racemates, **541–2**
racemic mixtures, usefulness of, **543–6**
racemic switches, **542**
radial distribution function code (RDF code), **215**
Raf kinase inhibitor, **152**, *153*
raloxifene, 103
Ramachandran plot, 619, *620*
random screening, **130**, *130*
Rapamune®, 171
rapamycin, 25, 171, *171*
Raptor, 579
rare mutations, **109**
rate-determining steps, 483–4
rational drug design, 28, **138–41**
Rauwolfia serpentina, *134*
rauwolscine *see* α-yohimbine
RCM *see* ring-closing metathesis
RDF code *see* radial distribution function code
receptor mapping, 424
receptors, 37
  acetylcholine receptors, 39–40
  cannabinoid receptors, 40–1
  caspases and NFκB, 98–9
  cellular adhesion, 99
  cytosolic serine/threonine kinase IRAK, 98
  cytosolic Src, Zap70/Syk and Btk tyrosine kinase, coupling, 97–8
  cytosolic tyrosine kinase JAK, coupling, 97
  dopamine receptors, 38–9
  of mediators, 87–8
  serotonin receptors, 39
  with guanylate cyclase activity, 95–6
  with intrinsic ion channels *see* direct ligand-gated ion channels
  with serine/threonine kinase activity, 96
  with tyrosine kinase activity, 96
Reciprocal Net project, 259
rectal absorption, 824
reduction
  xenobiotic transformations, 660–3, *661*, *662*, *663*
reductive bioactivations, 737–8
reductive biotransformations, 680–2
relative bioavailability, 652
release control, 821–2
renal clearance, 650, 707–9, *708*
renal excretion, of drugs, 645, *646*, 653
  active tubular secretion, 646
  glomerular filtration, 646
  tubular reabsorption, 646
rennin-angiotensin system (SAS), 555
research–development interface, 859
reserpine, 89
residence times, 640
responders, 901
restrictive elimination, 651
restructured ring systems, 353–4, *355*, *356*
retinoic acid receptor-related orphan receptor α (RARα), 103
retinoic acid receptors (RAR), 103
retinoic X receptors (RXR), 103
retro-thiorphan, 320, *322*
retropeptides, 320
reversal of functional groups
  isosterism, 320, 322–3
reverse docking, 623
reverse transcriptase inhibitors, **148**, *149*
rexinoid receptors *see* retinoic X receptors
rheumatoid arthritis (RA)
  new strategies for, 9–10
Ridogrel, 300
rifampicin, 703
right substituent choice, 424–5, *425*
rigid compounds, 368
rigidification, 375–6, *376*
rimonabant analogs, 349, *350*, *351*
ring equivalents, 297–303
ring transformations
  anological approaches, 343
    reorganization, 351–4, *354*, *355*, *356*
    ring closure, 345–9, *346*, *347*, *348*, *349*, *350*, *351*, *352*
    ring enlargement/contraction, 350–1, *352*, *353*
    ring opening, 343–5, *344*, *345*
  conjunctive approaches, 356, 358–60, *358*, *359*
  disjunctive approaches, 354–5, *355*, *356*, *357*
ring-closing metathesis (RCM), 189, 197
ritanserin, 392
ritonavir, 24, 846
rituxizumab, 98
rivastigmine, 40, 416, *416*
RNA chip technology (REF), 112
RNA-induced silencing complex (RISC), 111
RNA interference (RNAi), **111–13**
rofecoxib, 9
rolipram, 440, *441*
rotational entropy, 473
Royal Society journals, 268
RTECS database, 262
rule-based filters, **246**
Rule of, 3, **251**
"Rule-of-5" (RO5), 145, **245**, 565
rule-of-five, 509
RxList, 260

S

Saccharomyces cerevisiae, 108, 113, 116
salicin, 7
salicylic acid, 7
salt/complex formation, **846**
salts
  water-soluble compounds preparation, **749–65**
sartans, 115
  angiotensin-converting enzyme inhibitors, 14–15
saturation transfer difference (STD) NMR, 235–6
scaffold hopping, 323–6
  and virtual screning, 325–6
scaffold tree, **194–9**
Scholarly Publishing and Academic Resource Coalition (SPARC), 268
SciDev.Net, 268
SCONP *see* Structural Classification of Natural Products
SCOP (Structural Classification of Proteins), 199, 526
screen to lead
  assays, **146–7**
  compound collections, **146**
  hit-to-lead process, **147**, *148*
screening, 144, 145
  natural products, 162–3, 167–8
  of leads, 552
    drugs derived from, *145*, 147–52
SEDDS *see* self emulsifying delivery system
selective serotonin re-uptake inhibitor (SSRI), 45
selenium bioisosteres, 333–4
self-association, prevention of, 827–8
self emulsifying delivery system (SEDDS), 799–800
self micro-emulsifying delivery system (SMEDDS), 799
serotonergic drug, *392*
serotonin
  isosterism, 323, *324*
serotonin receptors, 39
serotonin re-uptake inhibitor (SSRI), 899
serotonin transporters (SERT) inhibitors, 398, *398*
serotoninergic 5-HT3 receptors, 95
SERT-plus DMLs, for depression, **554–5**
sertraline synthesis, *908*
short-range repulsive forces, 471–2
shRNA, 112
side-effects, 134–5
  reduction of, 824
    local irritancy, 824–5
    systemic detoxication, 825–6

sildenafil, 136, *137*
silicon–carbon bioisosterism, 330–1
simplified molecular input line system (SMILES), 212
simulated annealing *see* molecular dynamics
simvastatin®, 37, *131*, 170–1, *170*
sirtuin inhibitors, **406**
sitaxsentan, **150–2**
site-directed drug delivery, 728–9, *728*, *729*
site-specific delivery, 728–30
site-specific drug delivery, **831–5**
  brain targeting, 835
  cell targeting, 834–5
  colon targeting, 832–4
site-specific drug release, 729–30, *729*, *730*
skeletal diversity, 191, *192*
SLC genes, 698
SLC transporters, 699, *700*, 703–5
slow-release portion (SRP), 830
SMEDDS *see* self emulsifying delivery system
SMILES, 257, 262
SMILES *see* simplified molecular input line system
SmPCs *see* Summaries of Product Characteristics
social and economic factors, **909**
  cost containment measures, **914–16**
  generic pharmaceuticals, **912–14**
  orphan drugs, **911–12**
  parallel trade, **914**
  patents, **910–11**
  pattern and cost, of innovation, **909–10**
  pharmacoeconomics, **916–18**
sodium channels, *90*, 90–1
"sodium–potassium pump", 10
soft drugs, 734, *734*
soft gelatine capsules, 764
soft independent modeling of class analogy (SIMCA), 498, *499*
solbilization
  lipid vehicles, **798–803**
  surfactants, **793–8**
solid properties, control of, 819
  carmofur, stabilization of, 821
  metastable polymorphs, crystallization of, 819–20
  nifedipine, modification of, 820–1
solubility, 858
  acidic groups, 456–8
  hydroxylation, 453
  methyl groups, 432–4
  parameters, 327–8, *329*
  prediction, 750–1
solubilization, 817–18
  acidic side chain, **769–75**
  artemisin, 770, *771*
  basic side chain, **775–9**
  covalent attachment of side chains, **767–81**
  dissolution rate, 759–60
  hydroxy and amino group alkylation, 769–70
  non-ionizible side chains, **780–2**
  strategies, 768–9
"soluseptazine principle", 772
Solv-DB, 261
solvents, 789–94
sorafenib, **152**
SOSA (selective optimization of side activities), **132–3**, *132*
south side modifications, 419, *420*
sp-carbon, 661–2, *662*
sp^2 carbon atoms, 661–2, *662*
sp^3-carbon atoms, 660–1, *660*, *661*
sparteine, 383, *383*
SPCs *see* Summaries of Product Characteristics
Spectra Online, 267
spiro derivatives, 352, *354*
spiro-cyclopropyl groups, 440
splanchnic circulation, 643, *643*
SSRI *see* selective serotonin re-uptake inhibitor
stability
  salt form, 761
stabilization, 826–7
staphylococcus aureus, 20
"state of art", patents, 884, 886
statins, 36–7
statistical tools, 592
stereochemical diversity, 190, *191*
stereoselectivity ratios, **537–8**
steric aspects
  isosterism, 328, *329*
steric hindrance, **363–77**
  methyl groups, 434–5
steric interactions, 471–2
STI-571 *see* Gleevec®
stomach, drug residence time in, 640
Streptomyces aureofaciens, 21
Streptomyces erythreus, 21
streptomycin, 20
Streptozan®, 18
structural–activity relationship (SAR), 113
structural alignments, 592–4
structural analogs, 128
Structural Classification of Natural Products (SCONP), 194
structural constraints, 446–7, *447*
structural keys, **212–13**
structural logic's rule, 424, *424*
structural parameters, 327, *328*
structural rigidity, 712–13
structure–activity relationship (SAR), **239**, 300, 482
  application strategies, **415–26**
  natural products, 169
  substituent groups, **431–60**
structure-based drug design (SBDD), *606*
structure-based druggability, **526–7**
structure-based *in silico* screening, **219**
structure-based pharmacophores, 577–9
structure elucidation
  natural products, 168
strychnine-sensitive glycine receptors (GlyRs), 95
substituent groups, **431–60**
  acidic functions, 456–8
  additional binding sites, 459–60
  basic groups, 458–9
  halogenation, 448–52
  hydroxylation, 452–4
  methyl groups, 432–41
  sulfur-containing groups, 454–5, *456*
  unsaturated groups, 441–7
substituents
  effectiveness assessment, 477–8
substitution reactions
  conjugation reactions, 683
  hydrolysis, 682
substrate activity screening, **238–9**
sulfamethoxazole, 729–30, *730*
sulfamide, 718
sulfamidothiodiazol, 32
sulfanilamide, 12
sulfate esters, 772, *774*
sulfathiazole, 132, *132*, 137, 152
sulfation, 683
sulfisoxazole, 132, *132*, 150, 152
sulfonamides, 17–18, 433
  acidic groups, 458
  isosterism, 298, *299*
sulfonamides bioisosters, 320
sulfonates, *775*
sulfonation, 665
sulfonic acids, *769*, 772
sulfonylureas, 32
sulfur, isosterism, *308*
sulfur-containing groups, 454–5, *456*
sulindac, 723, *724*, 847
sulpiride, 133, *133*
sulpiride side-chain, 346, *347*
sumatriptan, 100
Summaries of Product Characteristics (SPCs), 261
Sunitinib, 565
support vector machines (MVMs), 500
surface plasmon resonance (SPR) spectroscopy, **239**, 240
surfactants, **793–8**, *794*, *795–7*
suspension formulations, 763–4
symbiotic approach, 381, *382*
synthesis intermediates, **131–2**
synthetic dyes, 6
SyracuseResearchCorporation (SRC), 261
systamatic screening

extensive screening, **129–30**
high-throughput screening, **130–1**
random screening, **130**
SOSA approach, **132–3**
synthesis intermediates, **131–2**
systemic clearance, 650
systemic detoxication, 825–6

T

tablet products, 762–3
tacrine, 389
tacrolimus, 25
talipexole, 319, *320*
tamoxifen, 103, 113
tarenflurbil, 117, 118
target, 68, *69*, 417–18, *417*, *418*
protein crystallography, 621–3
target–drug interaction, 87
target oriented synthesis (TOS), 188
target space, **524–31**
taxane anticancer agents, 175, *175*
Taxol®, 27, *130*, 797
Taxotere®, 27
see also paclitaxel
Taxus brevifolia, 168
tazonermin, 98
temperature factors, protein crystallography, 620
temsirolimus, 171, *171*
Tenconi, 283
teprotide, 15, 139
tertiary amines oxidation, 677, 678, *678*
tetracycline, 21, *131*, 645
tetraethylthiurame disulfide, 137
tetrazepam, 447
tetrazole, 306
Thalidomide, 135
theophylline, *667*, 822
thermolysin inhibitors, 295, *296*
thiatriazole dioxide, 318
thiazolidinediones, 32, 306
thienylurea, 152
thiocyanates, 454–5, *456*
thiols, 454, *455*
thiopentone, 645
thiophene, 678, *680*
thiorphan, 320, *322*
thiourea bioisosteres, 317–18
thioureas, 455, *456*, *457*
THIP, 303, *306*
Thornber, C.W., 294
three-dimensional quantitative structure–activity relationships *see* 3D QSAR
three-point contact model, **535–6**
thromboxane A2 receptor antagonist, 277
thromboxane A_2 (TXA_2), 392, 399, 400
thyroid receptors (TR), 103
tienilic acid, 687, *688*
time-series behavior, 503–4
tissue distribution, 707
tissular responses, 79–81
tolcapone, 283–4, *284*
top-earning drugs, **895**
topological descriptors, **213–14**
Torpedo Californica, 403
torsemide, 317
TOS *see* Target oriented synthesis
total body clearance, 650
toxic isosters, 330
toxic metabolites, **685–93**
toxicology, **118–19**
toxicology databases, **270–2**
Toxicology Tutor, 270
TOXNER, 270
toxophoric groups, *693*
trade names
and nonproprietary names, **867–8**
Trade Related Aspects of Intellectual Property Rights (TRIPS), 881–2
Trade Related Intellectual Property (TRIPS), 911
trademarks *see* trade names
transcellular drug transport, **638**
carrier-mediated processes, 639
passive diffusion, 638–9
vesicular transport, 639–40
transcytosis, 640
transdermal absorption, 823–4
transferases, *666*
transgenic animals, **117–18**
translational entropy, 473
transport kinetics, 701–2
transporters, 74, 639, **697–709**
ABC transporters, 702–3
active transport, 698, *699*
basic structure, 699, *700*
blood barriers, 707
distributions, 699–702
efflux transporters, 639
genes, 698–9
influx transporters, 639
intestinal absorption, 705–6, *706*
kidney and renal clearance, 707–9
liver and hepatic clearance, 706–7
properties, 699–702
SLC transporters, 703–5
tissue distribution, 707
trifluoperazine, 43, *472*
trigonelline esters/amines, 741, *741*
triplatin tetranitrate, 29
Tripos, 599
Tritrichomonas fetus, 219
trivalent atoms, isosterism, 296
troglitazone, 202, 690–2, *691*, *692*
Trypanosoma brucei, 238
Tsg, 101, 112
tubular reabsorption, 646
tumor necrosis factor-alpha (TNF-α), 9–10
twin drug approach, *382*
combination modes for, *381*
heterodimer ligands, **391–401**
homodimer, **383–91**
for opoid receptors, *386*
tyrosine kinase receptors, 73
tyrothricin, 19

U

UK Patent Office, 269
ultrahigh-throughput screening (uHTS) *see* high-throughput screening
undesirable functionalities, **245**
Union Catalog Project, 268
United States Adopted Name (USAN), 245
unity of invention, 888
univalent atoms, 294
University of Minnesota Biocatalysis/Biodegradation database, 260
University of Minnesota Biocatalysis/Biodegradation Database (UM-BBD), 264
unnatural natural products, 177
unsaturated groups, **441–7**
acetylenic series, 445–7
allylic series, 443–5
cyclenes, 447
vinyl series, 442–3
urea, 5
urea bioisosteres, 317–18
Uretupamine B, 192
urokinase, *627*
US National Cancer Institute's Letter of Collection, 163
US Patent and Trademark Office (USPTO), 269
US Patent Bibliographic Database, 269
"useless" drugs, 901

V

validation, 866
of 3D QSAR models, 594
valinomycin, 383
valproic acid, 48, 690, *690*
valsartan, 100
vancomycin, 21, 468–9
vancomycin resistant *Enterocci* (VRE), 919
vapor diffusion method, 612, *613*
variation of activity
of antihistaminic drugs, 328, 330
vascular endothelial growth factor (VEGF), 29–30
vascular endothelial protein tyrosine phosphatase (VEPTP), 197
vectorial transport, 700
vegetable sources, **902**
VEGF *see* vascular endothelial growth factor
Velcade *see* bortezomid

VEPTP *see* Vascular endothelial protein tyrosine phosphatase
vesicular transport, 639–40
Viagra® *see* sildenafil
villi, 640
vinblastine, 26, 137, 165
vinca alkaloids, 26–7
Vinca rosea, 137
vincristine, 26, 165
vinyl fluorides, 314, *315*
vinyl series, 442–3, *442*, *443*
vinylogoues, **283–7**
virtual receptor site, 591
virtual scaffolds, 197
virtual screens, 508–9
  and scaffold hopping, 325–6
vitamin C, *283*
vitamin D receptors (VDR), 103
VolSurff, 593
voltage-gated ion channels
  calcium channels, 91
  potassium channels, 91–2
  sodium channels, 90–1
voltage-operated cation channels, 74
volume of distribution, 649–50
voucher specimens, 163
Vytorin®, 171

W
water–cosolvent systems, **789–94**, *789*, *790*, *791*, *793*
water-in-oil emulsions, 799
water-miscible solvents, 759
water molecules, 620
water-soluble compounds
  acids and bases for salt formation, 751–2
  covalent attachment of solubilizing moieties, **767–81**
  preparation through salt formation, **749–65**
  solubility, 750–1
water-soluble taxol prodrugs, 731–2, *732*
WDI *see* World Drug Index
web alert
  blogs, **256–7**
  chemical analysis, **267**
  chemical publication
    journals, **267–8**
    open access, **268**
    theses, **268–9**
  chemical software programs
    chemical drawing and viewing software, **265**
    cheminformatics software, **265–6**
    datasets, for virtual screening, **266–7**
  chemical suppliers, **263–4**
  chemical synthesis, **264**
  compound information
    biological properties of, **258–60**
    ChemBank, **258**
    ChemSpider, **257–8**
    NIH Roadmap, **258**
    PubChem, **258**
  drug information
    DrugBank, **260–1**
    molecular property prediction and calculation of, **261–3**
    physico-chemical properties of, **261**
  metasites and databases, **272**
  patent information, **269**
    Japanese patent, **270**
  toxicology, databases, **270–2**
  wikis
    RSS information feeds, **257**
WebReactions, 264
Website *see* web alert
west side modifications, 419, *420*
WHO Drug Information, 871
WHO *see* World Health Organization
WIPO *see* World Intellectual Property Organization
Wiswesser Line Notation (WLN), 212
WLN *see* Wiswesser Line Notation
Wolf-Parkinson-White syndrome, 135
World Drug Index (WDI), 197, 245
World Health Organization (WHO), 867
World Intellectual Property Organization, 881

X
X-ray crystallography, 168, **237–8**, 251
  drug discovery, **605–29**
xenobiotics
  phase I reactions, 656–64
  phase II reactions, 664–71

Y
Yak, 579
yeast-based screening, *114*, *115*, *116*
yeast two-hybrid system, **110–11**, *110*
yohimbine, 203
Yondelis®, 169, 175, *176*

Z
zalcitabine (ddC), 396
zaleplon, 291, *292*, *468*
zanamivir, *608*, 609
zanoterone, 303, *304*
zaprinast, 287, *287*
zebra fish *see Danio rerio*
Ziconotide, 161
zidovudine (AZT), 396
ZINC database, 266
zinc proteases, 320, *322*
zolpidem, *128*, 291, *292*
zonisamide, 299
zopiclone, *128*, *130*, 291, *292*
Zovirax®, 28
zwitterions, 756