Biochemical Targets of Plant Bioactive Compounds

A pharmacological reference guide to sites of action and biological effects

Gideon Polya

Taylor & Francis
Taylor & Francis Group
LONDON AND NEW YORK

Biochemical Targets of Plant Bioactive Compounds

Biochemical Targets of Plant Bioactive Compounds

A pharmacological reference guide to sites of action and biological effects

Gideon Polya

Taylor & Francis
Taylor & Francis Group
LONDON AND NEW YORK

First published 2003
by Taylor & Francis
11 New Fetter Lane, London EC4P 4EE

Simultaneously published in the USA and Canada
by Taylor & Francis Inc.
29 West 35th Street, New York, NY 10001

Taylor & Francis is an imprint of the Taylor & Francis Group

This edition published in the Taylor & Francis e-Library, 2004.

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
Polya, Gideon Maxwell.
Biochemical targets of plant bioactive compounds : a pharmacological reference guide to sites of action and biological effects / Gideon Polya.
p. cm.
Includes bibliographical references and index.
1. Materia medica, Vegetable – Handbooks, manuals, etc. 2. Botanical chemistry – Handbooks, manuals, etc. 3. Plant products – Handbooks, manuals, etc. 4. Pharmacognosy – Handbooks, manuals, etc.
5. Plants – Metabolism – Handbooks, manuals, etc. I. Title.

RS164 .P766 2003
615′.32–dc21 2002155281

ISBN 0-203-01371-9 Master e-book ISBN

ISBN 0-203-34161-9 (Adobe eReader Format)
ISBN 0–415–30829–1 (Print Edition)

Contents

Tables

Preface

Plants defend themselves from other organisms by elaborating bioactive chemical defences. This is the essential basis of the use of herbal medicines that still represents a major therapeutic resort for much of humanity. However, at the outset, it must be stated that any plant that is not part of our evolved dietary cultures is potentially dangerous. Commercial herbal medicinal preparations approved by expert regulatory authorities have a significant place in mainstream conventional medicine and in complementary medicine. The first and last message of this book on the biochemical targets of bioactive plant constituents is that use of herbal preparations for medicinal purposes should only occur subject to expert medical advice. In the language of popular culture, **DO NOT TRY THIS AT HOME!**

This book arose from 40 years as a student, researcher and academic teacher in biochemistry, a discipline fundamentally informed by both chemistry and physiology. This book is aimed at a very wide readership from biomedical researchers and practitioners to a wide range of scientifically literate lay persons. Lay readers (notably high school and university students and graduates) would range from everyone following public media reports and discussions on health, environmental and other scientific matters to potential readers of popular generalist scientific journals such as *Scientific American* or *New Scientist*. The scientific readership would include researchers, professionals, practitioners, teachers and industry specialists in a wide range of disciplines including the life sciences, ecology, nursing, naturopathy, psychology, veterinary science, paramedical disciplines, medicine, complementary medicine, chemistry, biochemistry, molecular biology, toxicology and pharmacology.

This book condenses a huge body of information in a succinct and user-friendly way. Ready access to a goldmine of key chemical structure/plant source/biochemical target/physiological effect data from a huge scientific literature is via a Plant Common names index, a Plant genus index and a Compound index. Such information is obviously useful for biomedical and other science specialists. The introductory chemical and biochemical summaries will be very useful to students in these and allied disciplines. However, at a universal, everyday level, one can also use the book to readily find out about the nature and targets of bioactive substances in what you are eating at a dinner party. Further, plants and their constituents play an important part in human culture and the bed-time or aeroplane reader will find a wealth of interesting snippets on the historical, literary, artistic and general cultural impact of plant bioactive substances.

Many people have variously helped and encouraged me in this project, most notably my wife, Zareena, my children Daniel, Michael and Susannah, my mother and siblings, recent

research collaborators, colleagues who have given computing and scientific advice and further colleagues and other professionals who have read specific chapters. I must gratefully acknowledge the profound influence of my late father, Dr John Polya. Any deficiencies of this book are simply due to me and have occurred despite such helpful interactions.

Dr Gideon Polya
Department of Biochemistry, La Trobe University
Bundoora, Melbourne, Australia
August 2002

1 Plant defensive compounds and their molecular targets

1.1 Introduction

Higher plants are sessile and are consumed by motile organisms, namely other eukaryotes and prokaryotes. Plants defend themselves by physical barriers including cell walls at the cellular level, by the waxy cuticle of leaves and by bark and thorns at the macroscopic level. Plants also defend themselves from fungal and bacterial pathogens and animal herbivores by elaborating a variety of bioactive secondary metabolites and defensive proteins. There may be as many as 100,000 different kinds of plant defensive compounds of which about 30,000 have been isolated and structurally characterized. Biochemical targets have been determined *in vitro* or *in vivo* for some thousands of the defensive compounds isolated to date.

The word "target" is being used rather broadly and loosely here to encompass the molecular sites of interaction demonstrated for such compounds. However, the demonstrated binding of a plant compound to a protein *in vitro* or *in vivo* does not necessarily mean that this particular interaction is actually the critical site of action of the defensive compound. Further, a particular defensive compound may have multiple molecular sites of action and may well have synergistic effects with other such compounds. This book is concerned with the biochemical targets of plant defensive compounds.

This treatise has been designed to address a very wide audience ranging from scientifically literate lay people to researchers in many disciplines and health professionals. Plant products have had a huge impact on the way in which different human societies have developed, especially over the last twelve thousand years since the advent of agriculture. Thus, the evolution of specific day-length and temperature requirements for plant development meant adaptation of specific plants to particular latitudes. Accordingly, exploitation of "useful" plants (and of domesticatable animals feeding upon them) would have spread rapidly on an East–West axis. This contributed to the technological and military dominance of cultures of the Eurasian axis in the colonial era (as opposed to those of the North–South long axis continents of Africa and the Americas) (Diamond, 1997).

Particular plant products have had a massive impact on human populations and cultures in recent centuries as evidenced by the slave trade to the Americas (for the purposes of coffee, sugar and cotton production), colonial conquest in the East (opium, indigo, tea, cotton and preservative spices), African subjugation (slavery, cocoa, rubber and timber) and temperate colonization (grain, cotton, timber and herbivore production). Notwithstanding the European "Enlightenment", these economic expansions and social reorganizations (both domestic and colonial) were accompanied by horrendous abuses connected with war and famine (problems that are continuing today in the "New World Order").

Plants provide a bulk supply of carbohydrate (typically as seed or tuber starch) to support the global human population that now totals 6 billion as compared to an estimated 1 million

hunter-gatherers before the advent of agriculture-based civilization twelve thousand years ago. However, plants also provide humanity with a variety of bioactive constituents used for their taste, preservative, psychotropic or medicinal properties. Notwithstanding synthetic sweeteners, non-plant preservatives and an explosion of psychotropic drugs and other pharmaceuticals, plants are still major sources of such ameliorative and protective agents. While the "Western" pharmaceutical global market reached a value of US$354 billion in 2000, the total global herbal medicine market is currently about US$30 billion. Herbal medicine remains a major core recourse for the impoverished majority of the world's population.

Herbal medicinal traditions can be traced back to our primate forebears. Thus, parasite-infected chimpanzees make recourse to particular plants, which they evidently associate with symptomatic relief. Human cultures in general have accumulated medicinal protocols based on use of plants, major traditions including Chinese medicine and Indian Ayurvedic herbal medicine. As detailed in this book, in some instances, specific bioactive substances from medicinal plants (or derivatives of such compounds) have found application in conventional medicine. Thus, the cardiotonic cardiac glycoside sodium pump (Na^+, K^+-ATPase) inhibitors derived from the initial use for cardiac insufficiency of digitalis (dried leaves of the foxglove, *Digitalis purpureum*).

Determining the molecular sites of action of bioactive medicinal plant constituents is clearly important for establishing the chemical and physiological basis for herbal medicinal efficacy, for quality control of commercial herbal preparations and for the discovery of "lead compounds" for synthetic (or semi-synthetic) pharmaceutical development. Of course, it must be recognized that medicinal plant efficacy may derive from complex synergistic effects or even from quasi-placebo effects connected with the taste, mild effects and appearance of the preparation. While recognizing these possible "holistic" complications, in order to find out how such preparations work, it is clearly important to initially isolate, structurally characterize and define the biochemical targets of plant bioactive substances.

1.2 Organization and scope of the book

The book has been devised and organized so that it can be used by a wide range of people as (a) a textbook, (b) a user-friendly reference and (c) as a comprehensive summary of the biochemical pharmacology of plant compounds. This book focuses specifically on purified plant compounds (secondary metabolites and proteins) and the molecular entities (principally proteins) with which they interact in the target microbial pathogens and animal herbivores. In contrast, there are many essentially ethnobotanical books that variously deal with medicinal and psychotropic plants, detailing the nature, distribution, physiological effects, chemical components (where known) and cultural significance of such plants. In addition, there are many books that deal with purified and characterized plant defensive components from a chemical structure perspective. The Merck Index (Budavari, 2001) and the Phytochemical Dictionary (Harborne and Baxter, 1993) are notable examples of such chemical compendia that were particularly useful in the writing of this book and indeed are very useful adjuncts to the present work (especially for the chemical structures of plant compounds).

This first chapter deals with the structural diversity of plant defensive compounds. Chapter 2 provides a succinct but comprehensive summary of the essentials of biochemistry (the chemistry of living things). This biochemical review provides a detailed background for understanding the nature and function of the targets of plant defensive metabolites and proteins. The remainder of the book summarizes (mainly in table form) a wealth of information

about the molecular targets which are mainly proteins (such as receptors and enzymes) but also include polynucleotides (RNA and DNA), phospholipids and reactive oxygen species (ROS).

It will be apparent from a preliminary scan of this book that most of the biochemical targets are directly or indirectly concerned with cellular signalling, that is, the machinery enabling cells to perceive and respond to extracellular signals. Obvious major differences aside (e.g. the occurrence of chloroplasts in plants), the fundamental biochemical processes of metabolism and replication in plants and the organisms that consume plants are very similar. Accordingly, plants must be protected from compounds they produce that poison metabolism and replication. Such protection is achieved, for example, by defensive compounds being deposited extracellularly, being temporarily inactivated by chemical modification (e.g. glycosylation) and being highly specific for the non-plant targets. However, a major "strategy" that has evidently evolved in the defence of sessile plants against their mobile enemies has been to impair signalling processes, that is, it is energetically more efficient for plants to discourage rather than kill plant-consuming organisms.

1.3 Description of the tables

Most of the book is comprised of tables dedicated to specific targets or groups of targets of plant defensive compounds. Target-related tables are grouped into specific chapters that are prefaced by succinct summaries of the biochemistry of the targets. The tables in general have three columns that are dedicated respectively to (a) compound name, synonym and general chemical class, (b) plant sources of the compound together with common plant names of well-known plants, plant family and the plant part involved and (c) the biochemical target being considered, a measure of the affinity of the compound for the target, other biochemical targets and *in vivo* cellular and physiological effects of the compound. The information provided for any compound entry has been pared to a minimum and extensive use is necessarily made of abbreviations that are defined within the text and at the end of the book.

It should be noted that the literature covered for this book was enormous and varied. Accordingly, plant parts, numerous plant sources and compound affinities are not given in all entries. Measures of the affinity of a compound for its target are given in various ways. IC_{50} value (concentration for 50% inhibition of an enzyme, 50% displacement of a known ligand from the target molecule or 50% inhibition of an *in vivo* process) is routinely presented in round brackets in micromolar units (μM; micromoles per litre; 10^{-6} moles per litre). Compound-target dissociation constant (K_d) or inhibitor-target dissociation constant (inhibitor constant, K_i) (another measure of tightness of association) is presented in square brackets in micromolar units. For simplicity, the IC_{50}, K_d or K_i values (when provided) are given as a simple number with the unit (μM) being assumed because most of these values are indeed in the range of 1–100 μM. However, in cases when these values are much less than 1 μM, the value is given with the appropriate unit explicitly specified, for example, nM (nanomolar; nanomoles per litre; 10^{-9} moles per litre) and pM (picomolar; picomoles per litre; 10^{-12} moles per litre). Of course, the quantitation of such affinities depends upon the conditions of measurement and the source of the biochemical target entity. However, it was felt that provision of such values in many cases would give a useful "ball park" figure for comparative purposes and for indicating concentrations required for *in vitro* or *in vivo* effects. Thus (1 pM) would indicate that the compound binds very tightly to the target or causes *in vitro* or *in vivo* effects at extremely low concentrations. Conversely, (100) (i.e. 100 μM) would indicate a low affinity of the compound for the target and a relatively high concentration being required for *in vitro* or *in vivo* effects.

A selection of major plant sources has been provided in the tables but space limitations precluded an exhaustive listing of plant sources. Thus, the triterpene bioactive betulinic acid has so far been found in some 460 plant species and the flavonol kaempferol has been isolated from over 150 plant species. Conversely, some 600 bioactive secondary metabolites have been isolated from plants of the *Piper* genus alone. Most of the information on the plant bioactives and their sources have been derived from Web searching (e.g. using Alta Vista, Google and the PubMed system of the National Library of Medicine of the National Institutes of Health, USA), Biological Abstracts, review journals, a huge body of primary research papers and key compendia such as the Phytochemical Dictionary (Harborne and Baxter, 1993), the Merck Index (Budavari, 2001) and the Bioactive Natural Products series (Atta-ur-Rahman, 2001). Of especial use in surveying and checking bioactive compounds, plant sources and compound biological effects were the Merck Index (Budavari, 2001), the Phytochemical Dictionary (Harborne and Baxter, 1993) and a key Web-accessible compendium, namely Dr Duke's Phytochemical and Ethnobotanical Databases (the US Department of Agriculture (USDA) Agricultural Research Service, Beltsville, Maryland, USA).

Scientific and common names are provided for the compounds described. Obviously in some cases, the chemical structure can be rigorously defined in words understandable to readers with a modest chemistry background (e.g. the amino acid neurotransmitter GABA = γ-aminobutyric acid = gamma-aminobutyric acid = 4-aminobutyric acid = $H_2N{-}CH_2{-}CH_2{-}CH_2{-}COOH$). In other cases, a similar rigorous specification is based on the structure of a parent nucleus that is substituted (e.g. the flavonol phenolic quercetin = 3,5,7,3′,4′-pentahydroxyflavone) and indeed the structures of a variety of such "parent compounds" (e.g. flavone) are described later in this chapter and in the Appendix. For the lay reader, typical covalent chemical bonding can be summarized "Lego"-style by saying that hydrogen (H), oxygen (O), nitrogen (N), carbon (C) and phosphorus (P), respectively, have 1, 2, 3, 4 and 5 "friends" (i.e. single bond or equivalent single/double/triple bond combination connections). Reduced sulfur (S) is bivalent in hydrogen sulfide (H–S–H) but is hexavalent in the highly oxidized sulfate ion $[O{-}S(=O)_2{-}O]^{2-}$.

In many cases the compound structure is very complex but the name(s) and general chemical class description (provided for all compounds) provide a reasonable structural definition given the space limitations. However, the information provided will generally enable rapid sourcing of the chemical structure via the Web, the Merck Index (Budavari, 2001), the Phytochemical Dictionary (Harborne and Baxter, 1993), Chemical Abstracts and other chemical compendia and chemical and biochemical textbooks listed in the Bibliography. In this chapter and Chapter 2, the structures of a large number of bioactive compounds are given precisely in the text where this is readily possible. However, more complex structures are efficiently dealt with in a way to be described later that succinctly conveys the essential "skeletal" structure of a compound without confusing the reader with lengthy descriptions of additional structural details.

It must be appreciated that compounds with a carbon (C) atom having four different substituents (A, B, C and D) can exist as stereoisomers (mirror image configurations) that can only be interconverted by breaking and re-forming bonds (this interconversion being called racemization). You can readily establish this for yourself using matches tetrahedrally disposed on a piece of fruit representing the C atom (or by inspecting your "mirror image" left and right hands). Such isomerism can be of major importance for biological activity. Thus the α-amino acids that are constituents of proteins (poly-amino acids, polypeptides) can, in general, exist as mirror-image stereoisomers referred to as the so-called L- and

D-configurational isomers – however, only L-amino acids are found in proteins. The reader must be aware that such stereoisomerism is indicated in some key examples but not in all cases for reasons of space and didactic effectiveness.

In tables dealing specifically with proteins, a convention has been followed that the genus name of the protein source is generally given prior to naming the protein because particular types of defensive proteins (e.g. lectins, lipid transfer proteins and Bowman–Birk protease inhibitors) have been isolated from a variety of plants. Further, a brief description of the protein involving selected bits of information is provided in parentheses, for example, how many amino acids constitute the polypeptide (*x* aa); the molecular mass (*x* kDa = *x* kilodaltons, where 12 Da = the mass of a carbon-12 atom); the number of cysteine residues in the protein (*x* Cys); the number of disulfide bonds formed between cysteine residues (*x*/2 S–S); whether the protein is a glycoprotein and is glycosylated, that is, has sugar residues attached.

Because some compounds have been found to interact with a variety of targets, it was necessary to make a large number of abbreviations that are comprehensively listed at the end of the book. Thus, for example, an "Acetylcholine receptor of the nicotinic kind" is abbreviated as "nACh-R". The abbreviations for the particular targets that are the subject of specific tables are also defined within those tables.

For some particular targets (such as particular hormone receptors that have only recently been detected, purified or expressed), very few interacting plant compounds have as yet been identified and accordingly the tabulation process has been simple. However, in many cases a large number of compounds belonging to different chemical classes have been found to interact with particular targets. These compounds have been grouped into various categories, namely alkaloids, phenolics, terpenes, other compounds and non-plant reference compounds (the latter category being introduced to link the plant compounds with notable non-plant compounds of pharmacological and medical interest). Within such groupings the compounds are listed alphabetically and indeed throughout the tables compounds, compound synonyms, plant families and physiological properties of compounds are all consistently listed in alphabetical order for convenience.

Non-plant reference compounds are provided (listed within square brackets) for many targets (notably in the tables concerned with compounds binding to hormone or neurotransmitter receptors). Some of these non-plant compounds derive from fungi and indeed in some cases from pathogenic fungi growing on plants. Others are well-known bioactive compounds derived from other organisms or synthetic compounds of pharmacological and/or clinical importance. In some cases the affinities of plant substances for particular targets have been determined from the ability of the plant compound to displace a radioactively labelled non-plant ligand from the target protein or the plant compound and the non-plant compound compete or antagonize each other in bioassays. The *in vivo* physiological effects of the various bioactive compounds are very briefly described in square brackets at the end of each entry.

Finally, it was recognized that plants and their constituents have an intimate place in human cultures for a variety of reasons connected with food, hunting, medicine, war, religious practice, poisoning and psychotropic properties. Accordingly, in entries scattered throughout the tables, brief mention is made of historical, medicinal and toxicological properties of well-known plants and their products. In particular, the tables have been leavened by reference to notable interactions of famous people (including scientists) with particular plants or plant defensive compounds.

1.4 Using the tables

Because of the comprehensiveness of this book and the need to update entries in the future, the tables have been organized rationally in relation to groups of biochemical targets. In short, if you know the name of the compound or the plant genus from which it has been isolated, then you can rapidly turn to table-specific entries (as opposed to page-specific entries). If you know the common name of the plant, you can find the "genus" part of the binomial scientific name of the plant by consulting the Common Plant Name Index at the end of the book. Knowing the genus name of the plant species, you can look up the Plant Genus Index and find the relevant entries successively specifying genus name, table number, specific target section (a capital letter) and subsection (a lower case letter – a for alkaloid, p for phenolic, t for terpene and o for other; n specifies a non-plant compound). In tables dealing specifically with plant proteins, the name of the protein is preceded by the genus name. One can also look up the separate Compound Index listing all chemical compounds referred to in the tables and also obtain table references as described above.

By way of example, you can quickly find from the Plant Genus Index what has been found in *Coffea arabica* (family Rubiaceae) (coffee), the entry being:

Coffea 4.3Aa, 4.3Ba, 4.3Ca, 4.4D, 4.4E, 5.1Aa, 7.4a, 9.2p, 10.2a, 10.2t, 10.4a, 10.4o, 10.4p, 10.4t, 13.8ZOp, 14.1Ap, 14.2p, 14.5p

It is "common knowledge" that coffee contains caffeine (a methylxanthine compound) and inspection of the Compound Index yields the following entry:

Caffeine 4.3Aa, 4.3Ba, 4.3Ca, 4.4D, 4.4E, 5.1Aa, 7.4a, 10.2a

These entries succinctly describe coffee constituents that have been isolated, structurally characterized and shown to interact with particular biochemical targets.

1.5 The structural diversity of plant defensive compounds

As previously indicated, some 30,000 plant defensive compounds (either secondary metabolites or proteins) have so far been purified and characterized. This huge diversity has been reviewed in major monographs and monograph series listed in the Bibliography at the end of the book. A huge literature was examined in preparing this book, this amounting to tens of thousands of individual primary scientific papers and reviews describing the isolation, structural characterization, pharmacological effects and biochemical targets of thousands of plant-derived and other chemical compounds. Because of limitations of space it was simply not possible to reference each entry (such documentation would have required thousands of pages in itself). For the primary literature, for each entry the reader is referred to Web search vehicles (notably Google and PubMed) and the abstracting compendia, monographs and monograph series listed in the Bibliography.

Because of the need for user-friendly tables, the chemical complexity of plant-derived natural products has been simplified in this book into four categories, namely the alkaloids (a), phenolics (p), terpenes (t) and "other compounds" (o). These categories have been used flexibly so that the "alkaloids" category includes nitrogen-containing, heterocyclic pseudo alkaloids and the "phenolics" category includes some compounds that are phenolic derivatives. The chemical complexity of these various groups of compounds is briefly reviewed below. The chemical complexity increases through covalent modification of many of these compounds through processes such as glycosylation, hydroxylation, methylation and epoxide and *N*-oxide formation. Further, new bioactive entities may be generated after ingestion of plant material through hydrolysis of peptide, ester and glycoside linkages.

As indicated previously, space simply does not permit comprehensive presentation of the chemical structures of the thousands of plant defensive compounds dealt with in this book, although the structures of particular representative compounds or their related "parent" compounds are shown in the Appendix. Indeed there are clear advantages in attempting to "distil" molecular complexity down to readily comprehended groupings of covalently linked moieties that can be described by succinct text. Thus, this approach enables common structural patterns of pharmacological interest to become more evident and reduces molecular complexity to a kind of functional "Lego" that can be appreciated by chemist and non-chemist readers alike. The conventions for the simplified skeletal structural presentations used in this chapter are summarized below.

Carbon chain length of alkyl groups or the total number of carbons in a molecule is represented as C_n, for example, ethane (C_2; CH_3–CH_3). When a C has four different substituents, as for example the α-C of α-amino acids, parentheses are used to define the substituents. Thus, the general structure of an α-amino acid is $^-$OOC–CH(R)–NH_3^+ and the structure of the α-amino acid alanine (R=CH_3) is $^-$OOC–CH(CH_3)–NH_3^+.

In describing ring structures, the total number of C atoms is given as C_n and the other atoms (typically O, S and N) are also indicated. Thus, tetrahydropyrrole (a fully reduced or saturated five-membered ring with four Cs and one N) is C4N. In order to keep the descriptions as simple as possible the number of double bonds will not be specified but some attempt is made to address this by specifying particular structures (e.g. phenyl or benzene (Phe); isoquinoline (IQ); methylene dioxy (–O–CH_2–O–) (MD); and epoxy (–O–), pyrrole, pyridine, furan and pyran as themselves) and by blanket statements about groups of compounds (e.g. the sterols are polycyclics largely involving unsaturated, alicyclic ring structures).

Dihydro-, tetrahydro- and hexahydro- simplify to DH, TH and HH, respectively, as in dihydrofuran (DHfuran), tetrahydrofuran (THfuran; C4O, a cyclic ether), tetrahydropyran (THpyran; C5O, a cyclic ether) and hexahydropyridine (HHpyridine) (C5N). Note that hexahydropyridine is completely reduced, that is, fully saturated. Cyclic esters (lactones) have a –C–CO–O–C–moiety and are specified as C*n*OL. Cyclic hemiacetals have a –C–O–CH(OH)–C– grouping and are specified as C*n*OH. Again, to keep structural representations simple, aliphatic side chains will be represented explicitly if they are small (e.g. ethyl, –CH_2–CH_3) or simply represented as C_n if large and complex.

In some cases, a group cross-links across a ring and hence creates two further rings; however, clarity dictates that in this case the cross-link is indicated simply in square brackets. Thus, a compound with a ring cross-linked with a N-methyl group would be denoted X[–CH_3–N<], the epoxy analogue as X[–O–] (or X[epoxy]) and the dimethylene cross-link analogue as X[–CH_2–CH_2–].

In polycyclic structures, rings joined by C–C bonds are simply indicated thus: C*n*–C*n* or C*n*–C_n–C*n*. Thus the stilbene "skeleton" (Section 2, Appendix) could be "loosely" presented as Phe–C_2–Phe or, precisely, as Phe–CH=CH–Phe. Where rings are fused and share two Cs, the fusion is indicated thus: C*n*|C*n*, for example, fully reduced naphthalene is precisely C6|C6. When three Cs are shared in a polycyclic fusion, the symbol || is employed. When only one C is shared, the notation is C*n*·C*n*. When more than two rings are fused, the structure could be "linear" or "angular" and it is assumed (unless stated otherwise) that the angular "foetal" orientation is the default situation. Thus, anthracene is Phe|Phe|Phe (linear), phenanthrene is Phe|Phe|Phe (angular) and the fully reduced entities are C6|C6|C6 (linear) and C6|C6|C6 (angular), respectively (see Appendix, Section 4).

Further complexity arises when, for example, three rings are all fused with each other (as opposed to the linear and angular arrangements indicated above) and share a common C.

A simple example is the tricyclic aromatic phenalene, this arrangement being indicated by an asterisk: Phe*|Phe*|Phe* (or C6*|C6*|C6* in the case of the fully hydrogenated entity). In very few and very complicated structures multiple "shared Cs" are indicated by * and *′ superscripts (or, in the most complicated example to be encountered here, by 3*, 3*′ and 4* superscripts to indicate two Cs each shared by three rings and another C shared by four rings).

Unsaturated heterocyclic ring compounds to be encountered include thiophene (C4S), pyrrole (C4N), furan (C4O), pyran (C5O), pyrylium ($C5O^+$) and pyridine (C5N). When alkaloid rings are fused and share a N, a similar system is used of a vertical line to indicate sharing of two C atoms, * to indicate a C shared with three rings and N# to indicate sharing of a N (thus a pyrrolizidine ring involving two fused five-membered rings sharing a C and a N is represented as C4N#|C4N#). Just as we describe 2-hydroxy, 3-hydroxy and 4-hydroxy benzoic acid as *ortho* (*o*)-, *meta* (*m*)- and *para* (*p*)-benzoic acid, we can conveniently apply the same nomenclature to rings containing more than one N. Thus the unsaturated six-membered ring compounds 2-azapyridine, pyrimidine and pyrazine are denoted here as *o*C4N2, *m*C4N2 and *p*C4N2, respectively. The frequently encountered five-membered ring compound imidazole can be simplistically denoted as C3N2, the Ns being separated by a C. The important heterocyclic "parent" compound purine found in RNA and DNA is pyrimidine|imidazole (or *m*C4N2|C3N2).

The "rules" outlined above conveniently provide simple, succinct representations of complex polycyclic compounds and avoid the problem of the reader being "unable to see the wood for the trees". The structures of key "parent" ring compounds to be encountered in this book are presented in the Appendix together with the structures of some representative alkaloids, phenolics, terpenes and other compounds. Before sketching the complexity of plant bioactive compounds and their modes of action, it should be noted that many such compounds act as "agonists" by mimicking the action of particular hormones or neurotransmitters at specific receptors whereas others may act as "antagonists" by simply competing for binding to the receptor and thus blocking the normal receptor-mediated response.

1.6 Plant alkaloids

The alkaloids are basic compounds in which an N atom is typically part of a heterocyclic ring but in some cases is merely a substituent of an alicyclic or aromatic ring system (as for example with colchicine, some peptide alkaloids and some Amaryllidaceae alkaloids). Various N-based heterocyclics such as the purine and pyrimidine bases of DNA and RNA (see Chapter 2) and the methylxanthine purine derivatives variously found in tea and coffee (caffeine, theobromine and theophylline) are sometimes referred to as pseudoalkaloids and for consistency will be included as alkaloids in this classification. Indeed all plant heterocyclics with a ring N will be conveniently lumped in with the alkaloids in the tables for didactic simplicity and consistency.

Alkaloids are widespread in plants and include some very well-known poisons (notably coniine and strychnine), hallucinogens (morphine, cocaine and muscimol) and other potentially lethal compounds that are nevertheless used in medical practice (e.g. atropine, codeine, colchicine and morphine). As indicated by the preliminary snap-shot above, alkaloids typically have names ending in -ine and which are often related to the plant source or properties. Thus, morphine was named after Morpheus (the God of sleep) and coniine derives from *Conium maculatum* (hemlock), the plant used in the judicial murder of Socrates (399 BC). Various chemical tests for alkaloids are used as preliminary indicators of alkaloid presence in crude plant extracts. Finally, it should be noted that alkaloids can also exist as *N*-oxides of the alkaloid base.

i. Monoterpene alkaloids are formed from iridoid monoterpene lactone glycoside precursors (with ten carbon chain (C_{10}) deglycosylated aglycones) such as loganin (C5|C5O, C5|pyran) and seco-loganin (C5O, DHpyran) by condensation with ammonia (NH_3). Indeed such reactions may occur during isolation in the presence of ammonium hydroxide (NH_4OH). Monoterpenes in turn derive biosynthetically from two isoprene (C_5) ($2 \times C_5 = C_{10}$) precursors. Examples include the bicyclic monoterpenes tecomine (a hypoglycaemic antidiabetic) from *Tecoma stans* (Bignoniaceae) and the anti-inflammatory compounds gentianamine, gentianadine and gentianine (pyridine|C5L) (from *Gentiana* species (Gentianaceae)). The tricyclic *N*-(*p*-hydroxyphenethyl)actinidine (*p*–OH–Phe–CH_2–CH_2–N–pyridine|C5) from *Valerian officinalis* (valerian) (Valerianaceae) is an acetylcholinesterase (AChE) inhibitor.

ii. Sesquiterpene alkaloids deriving from the sesquiterpene farnesol ($3 \times C_5$ isoprene units = C_{15}) include α-nupharidine (furan–C5N#|C5N#) and thiobinupharidine (furan–C5N#|C5N#·C4S·C5N#|C5N#–furan) from *Nuphar* species (Nymphaeaceae) rhizomes used for sedative and narcotic extracts.

iii. Diterpene alkaloids derive from diterpene ($4 \times C_5$ isoprene units = C_{20}) precursors and include some very toxic compounds, for example, heart-slowing, blood pressure-lowering, voltage-gated Na^+ channel activators from *Aconitum* (wolfsbane) species (Ranunculaceae) (aconitine, aconifine, delphinine, falaconitine, hypaconitine, indaconitine, jesaconitine, lappaconitine, lycoctonine, mesaconitine and pseudoaconitine) and neuromuscular blockers with curare-like effects from *Delphinium* species (Ranunculaceae) (condelphine, elatine and methylaconitine), the representative compound of this group being aconitine ([–CH_2–N(CH_2CH_3)–CH<]C6|C7|C5|C6–O–CO–Phe]). Further diterpene alkaloids include the cardiotonic, digitalis-like Na^+, K^+-ATPase inhibitors from *Erythrophleum guineense* (Fabaceae) (cassaine, cassaidine and erythrophleguine) (C6|C6|C6-alkylamine); and ryanodine (methylene–[pyrrole–CO–O–C5*|C4O*,*′|C5*,*′|C6*′]) from *Ryonia speciosa* (Flacourtiaceae) (a ligand that modulates the endoplasmic reticulum "ryanodine receptor" Ca^{2+} channel that is variously opened in excited skeletal muscle, cardiac and neuronal cells).

iv. Steroid alkaloids derive from triterpene ($6 \times C_5$ isoprene units = C_{30}) precursors. These generally toxic compounds include some AChE inhibitors from *Lycopersicon* (tomato) and *Solanum* (potato) species (Solanaceae) such as demissidine (C6|C6|C6|C5|C4N#|C5N#) and tomatidine (C6|C6|C6|C5|C4O·C5N) and their glycosylated derivatives (demissine and tomatine, respectively). A number of steroid alkaloids are teratogenic (cause embryological defects) including some from *Veratrum* species (Liliaceae) namely 3-*O*-acetyljervine, *N*-butyl-3-*O*-acetyl-12β, 13α-dihydrojervine, cyclopamine, cycloposine, *O*-diacetyljervine, 12β,13α-dihydrojervine, jervine (C6|C6|C5|C6·C4O|C5N), *N*-formyljervine, *N*-methyljervine and protoverine (C6|C6|C5|C6|C5N#|C5N#). Related teratogens from *Solanum* tubers include the glycosides α-chaconine, α-solanine and solasonine and their aglycones (deglycosylated entities) α-chaconidine (C6|C6|C6|C5|C4N#|C5N#), solanidine (C6|C6|C6|C5|C4N#|C5N#) and solasodine (C6|C6|C6|C5|C4O·C5N), respectively.

v. Peptide alkaloids or cyclopeptides have macrocyclic 13–15-membered rings involving several peptide (–CO–NH–) links. Cyclopeptides have been isolated from various sources, notably *Ceanothus* and *Zizyphus* species (Rhamnaceae) (e.g. Zizyphine A). These 0.6 kDa cyclopeptides are synthesized by a non-ribosomal mechanism in contrast to the much larger 2–3 kDa protease inhibitory cyclotides that are cyclic peptides synthesized as proproteins on

ribosomes (see Chapter 13) (and as such are considered under "other" plant defensive compounds in Section 1.9).

vi. Betalain alkaloids are non-toxic, water soluble, purple or yellow coloured plant pigments deriving from the amino acid derivative 3,4-dihydroxyphenylalanine (dopa, 3-hydroxytyrosine). Dopa rearranges to yield betalamic acid (a tetrahydropyridine, C5N) and can form a further derivative cyclodopa (a dihydroindole, Phe|C4N). Betalamic acid condensation with cyclodopa yields purple betacyanins that can be further modified by glycosylation. Betalamic acid condensation with aliphatic amino acids yields yellow betaxanthins. *Beta vulgaris* (beetroot) (Chenopodiaceae) contains betalamic acid, purple betacyanins (namely betanidin, DHpyridine=CH–CH=(N)–indole) and glycosylated betanidin derivatives (betanin and betanin sulfate) and yellow betaxanthins (vulgaxanthins I and II, DHpyridines). A relatively common inability to degrade these compounds gives rise to the coloured urine of "beeturia". The gorgeous purple of *Bougainvillea* species (Nyctaginaceae) bracts derives from betalains such as the glycosylated betanidin bougainvillein-r-1.

vii. Indole alkaloids include a variety of polycyclic compounds involving the bicyclic basic compound indole (2,3-benzopyrrole, Phe|pyrrole, Phe|C4N) and hence related to the amino acid tryptophan (Trp, 2-amino-3-indolylpropionic acid). Tryptophan decarboxylates to tryptamine (3-(2-aminoethyl)indole) which is thence converted to a variety of neuroactive compounds acting as agonists for serotonin receptors (5HT-Rs) including: bufotenine (*N*,*N*-dimethyl-5-hydroxytryptamine) (hallucinogenic); *N*,*N*-dimethyltryptamine (hallucinogenic); 5-hydroxytryptamine (5HT) (the excitatory neurotransmitter serotonin); 5-methoxy-*N*,*N*-dimethyltryptamine and gramine (3-(dimethylaminomethyl)indole) (agents causing *Phalaris* staggers in sheep); and the hallucinogens psilocin (3-dimethylaminoethyl-6-hydroxyindole) and psilocybin (6-phosphopsilocin) (from the *Psilocybe* "magic mushroom" species).

Further "simple" indoles include the faecal-smelling 3-methylindole and indole; and the cell wall-expanding plant hormone indole 3-acetic acid (IAA, auxin) and its precursors indole-3-acetonitrile and indole-3-carboxaldehyde. Tricyclic indoles include: harman (a DNA intercalator) (Phe|pyrrole|pyridine), the related hallucinogens harmine and harmaline (3,4-dihydroharmine) and chanoclavine (Phe*|pyrrole*|C6*); the narcotic mesembrine (saturated indole-Phe); and the Fabaceae tricyclic AChE inhibitors eseramine (Phe|DHpyrrole|THpyrrole), eserine (physostigmine) (Phe|DHpyrrole|THpyrrole) and eseridine (Phe|DHpyrrole|C4NO). Indican (3-(β-glucoside)indole) from *Indigofera* species (Fabaceae) and *Polygonum tinctorum* (Polygonaceae) oxidizes to yield the dark blue dye indigo. Similarly isotan B (a 3-hydroxyindole sugar ester) from *Isatis tinctoria* (Brassicaceae) (the woad used for body painting by the ancient Britons) is oxidized to yield indigo. A sulfur-containing *N*-methoxyindole derivative methoxybrassinin is a phytoalexin produced by *Brassica* species (Brassicaceae) in response to fungal infection.

A variety of more complex indole compounds derive from condensation of an indole precursor (deriving from tryptophan) and the aglycone of the C_{10} monoterpene-based iridoid glycoside secologanin. These indole derivatives range from tetracyclics to compounds with as many as eleven rings. Some of these indole alkaloids include the nicotinic acetylcholine receptor (nACh-R) antagonists *C*-curarine (quaternary amine, eleven-ring, epoxy structure), sarpagine (Phe|pyrrole|C5N#|C5N#[methylene]) and toxiferine (eleven-ring quaternary amine); the glycine receptor antagonist strychnine (seven compactly fused Phe, C4N#, C5N#, C6O, C6, C4N# and C5N# rings); the muscarinic acetylcholine receptor antagonist usambarensine (Phe|pyrrole|C5N#|C5N#–CH_2–|pyridine|pyrrole|Phe); the anti-addictive and hallucinogenic glutamate receptor antagonist ibogaine (Phe|pyrrole|C6N|C6 N-methylene); the α-adrenergic and 5HT receptor antagonist yohimbine

(Phe|pyrrole|C5N#|C5N#|C6); the *Rauwolfia* species (Apocynaceae) antipsychotic and neurotransmitter transport inhibitor reserpine (Phe|pyrrole|C5N#|C5N#|C6–O–CO–Phe); and the anti-mitotic, tubulin-binding antitumour agents vinblastine and vincristine (Phe|pyrrole|C8N#|C5N#–Phe|pyrrole|C6*|C4N*#|C5N*#).

The hallucinogenic tetracyclic ergine (lysergic acid amide) (Phe*|pyrrole*|C6*|DHpyridine carboxamide) is found (like chanoclavine) in *Rivea corumbosa* and *Ipomoea* species (ololiuqui) (Convolvulaceae). Ergine is also found in the fungal ergot (*Claviceps purpurea*) that infects Poaceae (such as rye) as are a variety of hallucinogenic ergine derivatives namely the tetracyclics elymoclavine (a teratogen) and ergometrine and hallucinogenic compounds involving ergine substituted with polycyclic substituents namely ergocornine, ergocristine, ergocryptine, ergosine and ergotamine. The ergot alkaloids are hallucinogens that act as serotonin receptor (5HT-R) agonists and block prolactin release in herbivores. Ergot consumption has had a tragic history in susceptible regions of Western Europe and North America because consequent behavioural alteration was construed as "devil possession" leading to appalling torture and execution of as many as 100,000 victims as "witches".

viii. Isoquinoline (IQ) alkaloids include a variety of bioactive compounds variously deriving from the amino acids phenylalanine and tyrosine and including IQ (benzo[c]pyridine) (Phe|pyridine; Phe|C5N) or its derivatives as part of their structure. In many cases the pyridine moiety is reduced to give tetrahydroisoquinoline and the benzo moiety is often substituted with a MD (–O–CH_2–O–) to form an additional ring. This very large group of alkaloids includes many compounds which are psychoactive and/or which affect muscle function. Chemically the IQ alkaloids are classified into structural subgroups named for key members (e.g. morphine-related morphinans) or structural complexity (e.g. simple IQs, ring-opened IQs and benzylisoquinolines).

Many opium-derived and other IQs are psychoactive, the best known being the analgesic, addictive, narcotic, opium-derived morphinan alkaloids codeine and morphine (heroin being the semi-synthetic diacetate of morphine). The tertiary or quaternary amine structural component is important for the activity of some *Erythrina* alkaloids and bisbenzylisoquinolines (notably the major curare component (+)-tubocurarine) as antagonists of the nACh-R involved in neuronal excitation of skeletal muscle. The planar disposition of some polycyclic benzophenanthridines enables intercalation (parallel interleaving) between the base pairs of DNA. A variety of naturally occurring and synthetic IQ compounds are protein kinase inhibitors.

The chemical and pharmacological complexity of the various IQ alkaloid sub-groups is sketched below with pharmacological and other attributes for each compound given in parentheses. Some of the better-known IQ alkaloids derive from opium, the dried milky latex from the unripe seed pods of *Papaver somniferum* (opium poppy) (Papaveraceae) and accordingly whether a substance is opium-derived is also indicated. Selected representative examples are given for each IQ alkaloid subgroup.

Simple isoquinolines (IQs) (−)-pellotine (IQ) (*Lophophora williamsii* (peyote) (Cactaceae) paralytic convulsant); (−)-salsolinol (IQ) (*Musa paradisiaca* (banana) (Musaceae) and *Theobroma cacao* (cocoa) (Sterculiaceae) dopamine antagonist linked to chocolate craving).

Ring-opened isoquinolines Narceine (MD–Phe–CH_2–CO–Phe amine) (opium-derived antitussive).

Aporphines Magnoflorine (IQ*|C6*|Phe) (a weak neuromuscular blocker of widespread occurrence); xylopine (MD–IQ*|C6*|Phe) and xylopinine (Phe|C5N*|C5N*|Phe) (*Xylopia* spp. (Annonaceae) α-adrenergic antagonists).

Cularines Cularicine, cularidine, cularimine and cularine (Fumariaceae cytotoxics) (IQ*|C6O*|Phe–MD).

Morphinans (compactly fused Phe, C6, C5N, C6 and C4O rings) Codeine (opium-derived addictive, analgesic, antitussive, spasmolytic narcotic); morphine (opium-derived addictive, analgesic, antitussive, sedative, spasmolytic narcotic; heroin is the semi-synthetic diacetate); thebaine (non-analgesic, toxic, convulsant narcotic and semi-synthesis precursor of the anti-addiction drug naltrexone).

Phthalideisoquinolines α-narcotine and narcotoline (MD–IQ–C4L|Phe) (opium-derived spasmolytics); (+)-bicucculine (MD–IQ–C4L|Phe–MD) (*Corydalis* species (Papaveraceae) GABA receptor antagonist).

Rhoedans Rhoeadine (MD–Phe|C9ON|Phe–MD) (*Papaver rhoeas* (red poppy) (Papaveraceae) narcotic).

Pavines (−)-argemonine (Phe|C8[$CH_3N<$]|Phe) (*Argemone* species (Papaveraceae) weak analgesic).

Benzylisoquinolines (IQ–CH_2–Phe) Ethaverine and laudanosine (L-type Ca^{2+} channel blockers from opium); papaverine (cAMP phosphodiesterase inhibitor and smooth muscle relaxant derived from opium and *Rauwolfia serpentina* (Apocynaceae)); protopine (MD–Phe|C9N|Phe–MD); opium-derived smooth muscle relaxant); (+)-reticuline (opium-derived adrenergic receptor ligand and hair growth accelerant).

Emetines (Phe|C6N#|C6N#–CH_2–C5N|Phe) Emetine, emetamine and psychotrine (from *Cephaelis ipecacuanha* (Rubiaceae), ipecacuanha being used as an emetic and expectorant due principally to its content of emetine, a DNA-binding compound).

Protoberberines Berberine (umbellatine) (MD–Phe|C5N#|C5N#|Phe) (DNA-binding cytotoxic, adrenergic receptor antagonist and AChE inhibitor from *Berberis vulgaris* (Berberidaceae) and other plants).

Benzophenanthridines (IQ|Phe|Phe) Fagaronine (*Fagara xanthoxylum* (Rutaceae) DNA-binding antibacterial); palmatine (calystigine) (Berberidaceae and Papaveraceae adrenergic ligand and AChE inhibitor); sanguinarine (pseudochelerythrine) (MD–IQ|Phe|Phe–MD) (antibacterial, DNA-binding protein kinase inhibitor derived from *Chelidonium majus* (Papaveraceae) and opium); chelerythrine (MD–IQ|Phe|Phe) (*C. majus* (Papaveraceae) protein kinase inhibitor).

Bisbenzylisoquinolines (macrocyclic or linear, formed by 2 benzylisoquinolines) (+)-tubocurarine (macrocyclic) (acetylcholine (nicotinic) receptor antagonist and skeletal muscle relaxant; major component of *Chondrodendron* species (Menispermaceae) pareira bark-derived "curare" arrow poison); dauricine (linear) (Menispermaceae curare-like anaesthetic); rodiasine (macrocyclic) (*Ocotoea venenosa* (Lauraceae) curare-like skeletal muscle relaxant); cepharanthine (macrocyclic) (*Stephania* species (Menispermaceae) anti-mycobacterial active against leprosy and tuberculosis).

Erythrina isoquinolines (Phe|C5N*#|C4N*#|C6*) Erysonine, erysotrine, erythratidine, α-erythroidine and β-erythroidine (*Erythrina* species (Fabaceae) curare-like neuromuscular blockers).

ix. Pyrrolidine alkaloids are based on tetrahydropyrrole (pyrrolidine, C4N), a five-membered ring containing one N atom, that is, the fully reduced derivative of pyrrole (Section 1, Appendix). Examples include cuscohygrine, hygrine and hygroline from *Erythroxylum coca* (coca) (Erythroxylaceae); the anti-schistosomiatic cucurbitine from *Cucurbita moschata* (Cucurbitaceae); the antimicrobial tricyclic gerrardine from *Cassipourea* species (Rhizophoraceae); the renal osmoprotectant stachydrine (proline betaine) and 3-hydroxystachydrine from *Capparis* species (Capparidaceae); and the anti-inflammatory (−)-betonicine

(achillein or 4-hydroxyproline betaine) from *Betonica officinalis* (Lamiaceae) and *Achillea* species (Asteraceae).

DMDP (2,5-dihydroxymethyl-3,4-dihydroxypyrrolidine) from *Derris elliptica* and *Lonchocarpus sericeus* (Fabaceae) and the related homoDMDP and several homoDMDP glycosides from *Scilla campanulata* and *Hyacinthoides non-scripta* (Hyacinthaceae) are variously active as inhibitors of particular glycosidases (enzymes cleaving glycosidic linkages in sugar oligosaccharides and polysaccharides). These polyhydroxypyrrolidine compounds are structurally similar to so-called furanose sugars (see Section 1.9 and Chapter 2).

Myosmine (3[2-pyrrolidinyl]pyridine) and nicotine (3[1-methyl-2-pyrrolidinyl]pyridine) and a variety of related pyrrolidinylpyridine compounds notably occur in *Nicotiana tabacum* (tobacco) (Solanaceae) and are discussed in Section xii under pyridine alkaloids.

x. Pyrrolizidine alkaloids (C4N#|C4N#) have an N atom shared between two fused five-membered rings. Some pyrrolizidine alkaloids are α-glycosidase inhibitors, namely (sources in parentheses) alexine (*Alexa leiopetala* (Fabaceae)), australine (*Castanospermum australe* (Fabaceae)) and casuarine (*Casuarina equisitefolia* (Casuarinaceae)). 1,2-Dihydroxy-3, 5-dihydroxymethylpyrrolizidine (hyacinthacine B2) from *Scilla campanulata* (Hyacinthaceae), its C-5 epimer (hyacinthacine B1) from *Scilla campanulata* and *Hyacinthoides non-scripta* (Hyacinthaceae) and 3-hydroxymethyl-5-methyl-1,2,6,7-tetrahydroxyquinolizidine (hyacynthacine C1) from *Hyacinthoides non-scripta* all inhibit various glycosidases.

The highly poisonous *Senecio* species (ragworts) (Asteraceae) have a major role in global livestock poisoning through the elaboration of hepatotoxic pyrrolizidines including the angelic acid ester O^7-angelylheliotridine and a variety of related compounds having a lactone (cyclic ester) ring (angularine, isatidine, jacobine, retrorsine, riddelline, senecionine, senecivernine). Senecionine is a teratogen as are other pyrrolizidines (namely fulvine and heliotrine), these compounds having unwanted developmental effects connected with mutagenicity and toxicity. Other variously hepatotoxic and carcinogenic pyrrolizidines derive from *Crotalaria* species (Fabaceae) (including the lactones fulvine (a teratogen), monocrotaline, riddelline and usaramine); *Heliotropium* species (Boraginaceae) (heliosupine, heliotridine, heliotrine (a teratogen), indicine, intermedine, lasiocarpine, lycopsamine and supinine); and from *Symphytum* (comfrey) species (Boraginaceae) (echimidine, heliosupine, lasiocarpine, lycopsamine and symlandine). The diester echimidine also occurs in *Echium plantagineum* (Paterson's curse or Salvation Jane) (Boraginaceae), a pretty plant that covers 33 million hectares of Southern Australia from Western Australia to northern New South Wales and costs the Australian livestock industry US$125 million per annum.

xi. Indolizidine alkaloids (C5N#|C4N#) have an N atom shared between a five-membered ring and a six-membered ring. Castanospermine from *Castanospermum australe* (Fabaceae) inhibits α- and β-glucosidases and swainsonine from *Swainsona* species (Fabaceae) inhibits α-mannosidase. The indolizidine slaframine (produced on *Trifolium repens* (red clover) (Fabaceae) by the fungal pathogen *Rhizoctonia leguminicola*) is a muscarinic acetylcholine receptor (mACh-R) agonist (i.e. an acetylcholine "mimic" on such receptors) and is hence a parasympathetic stimulant causing salivation and diarrhoea in livestock.

xii. Pyridine and piperidine alkaloids. Piperidine alkaloids are based on piperidine (hexahydropyridine) which has a six-membered saturated ring including an N atom (C5N). An example of a simple pyridine compound is trigonelline (*N*-methylpyridine 3-carboxylic acid), a hypoglycaemic compound from *Trigonella foenum-graecum* (fenugreek), *Medicago sativa* (alfalfa) (Fabaceae) and *Coffea* species (Rubiaceae). Piperidine- and pyridine-based alkaloids often have more than one ring and the degree of saturation can vary. Thus, (−)-anabasine (3-(2-piperidinyl)-pyridine) involves a piperidine (six-membered ring) linked to

pyridine and is an analogue of nicotine (3[1-methyl-2-pyrrolidinyl]pyridine) which involves a pyrrolidine (five-membered ring) linked to pyridine.

Myosmine (3[2-pyrrolidinyl]pyridine) and nicotine (3[1-methyl-2-pyrrolidinyl]pyridine) (Section 1, Appendix) and a number of related bioactive alkaloids occur in *Nicotiana tabacum* (tobacco) (Solanaceae) and variously in other Solanaceae such as *Duboisia* species. Nicotine and the related tobacco compounds nicotyrine and (−)-nornicotine are agonists (neurotransmitter "mimics") of the so-called (nicotine binding) nACh-R involved in neurotransmission and in neuromuscular transmission for skeletal muscle. The extraordinary addictiveness of nicotine derives from nACh-R agonists causing dopamine release and activating the mesolimbic dopamine system yielding "reward" effects. The antidepressant (−)-cotinine is the major nicotine metabolite in humans and a nicotinic agonist.

(−)-Anabasine (3-(2-piperidinyl)-pyridine) from *Nicotiana* and *Duboisia* species (Solanaceae) is an nACh-R agonist used to discourage tobacco smoking as is the *N*-methylated tricyclic piperidine (−)-lobeline from *Lobelia* species (Campanulaceae). Lobeline-related compounds from *Lobelia* species include the bicyclic *N*-methyltetrahydropyridines isolobinine and lobinine and the tricyclic *N*-methylpiperidines lobelanine and lobelanidine. Anabasine-related compounds include anatabine (2-(3-pyridyl)-1,2,3,6-tetrahydropyridine) from *N. tabacum* and (+)-ammodendrine (*N*-acetyltetrahydroanabasine) from *Ammodendron* and *Sophora* species (Fabaceae).

Apart from nicotine, the best-known piperidine alkaloid is (+)-coniine (2-propylpiperidine) from *C. maculatum* (hemlock) (Apiaceae) and *Sarracenia flava* (carnivorous pitcher plant) (Sarraceniaceae). Hemlock was drunk in the judicial murder of Socrates (Athens, 399 BC). Coniine is a paralysis-inducing nACh-R agonist as are (+)-*N*-methylconiine and γ-coniceine from the same source, the latter also deriving from *Aloe* species (Liliaceae). Coniine and γ-coniceine are teratogenic as well as being highly toxic. Other piperidine-related teratogens include (−)-anabasine from *Nicotiana* species, mimosine from *Leucaena leucocephala* and *Mimosa pudica* (Fabaceae) and (+)-ammodendrine, *N*-methylammodendrine and *N*-acetylhystrine from toxic *Lupinus* (lupine) species (Fabaceae) that can give rise to "crooked calf disease".

Seeds of *Areca catechu* (betel nut) (Palmae) contain the simple *N*-methyltetrahydropyridine 3-carboxylic acid (*N*-methyl-Δ^3-tetrahydronicotinic acid) arecaidine and arecoline (arecaidine methyl ester) (Section 1, Appendix) that are mACh-R agonists and accordingly parasympathetic stimulants. Betel nut also yields guvacine (Δ^3-tetrahydronicotinic acid) that is an anti-epileptic GABA transport inhibitor. Conversely the *N*-methyl dihydropyridone derivative ricinine from seeds of *Ricinus communis* (castor seed) (Euphorbiaceae) is a stimulatory agonist acting at the benzodiazepine site of the GABA(A) receptor.

The simple piperidine pelletierine from *Punica granatum* (pomegranate) (Punicaceae) and *Duboisia myoporoides* (Solanaceae) is an anthelmintic. The simple piperidine derivatives deoxymannojirimycin (DMJ) and deoxynojirimycin (DNJ) from *Lonchocarpus* species (Fabaceae) are glycosidase inhibitors because they are structurally similar to the pyranose (six-membered ring) sugar moieties of the glycosidase disaccharide substrates.

xiii. Quinoline alkaloids are based on a benzo[b]pyridine (quinoline) nucleus (Phe|pyridine) and are biosynthetically derived from 2-aminobenzoic acid (anthranilic acid), a key intermediate in the biosynthesis of the indole-containing amino acid tryptophan. Quinoline alkaloids can be simple or composed of a quinoline nucleus fused with other moieties to yield polycyclic derivatives. Thus, quinoline fused with benzene is acridine (dibenzo[b,e]pyridine) (Phe|pyridine|Phe); furoquinolines have a fused furan ring (a five-membered ring with an O) (Phe|pyridine|C4O); and pyranoquinolines have a fused pyran ring (a six-membered ring with an O) (Phe|pyridine|C5O). Quinazolines have two N atoms

in the same ring. The anticancer quinoline-based compound camptothecin has a structure involving fused quinoline, indolizidine and pyran lactone rings. Simple and more complex quinolines can have an additional ring formed by an MD substituent. The structural and pharmacological complexity of quinoline alkaloids is sketched below.

Simple quinolines (Phe|pyridine) include the *Cinchona* and *Remijia* species (Rubiaceae) antimalarials cinchonidine (α-quinidine), cinchonine (a stereoisomer of cinchonidine), hydroquinidine (quinotidine), quinine and quinidine (β-quinine), these compounds all having a quinuclidinemethanol (1,4-ethylpiperidinylmethanol) substituent. Quinine is also an extremely bitter tasting compound. Of a range of other simple quinolines, eduline and its *O*-methyl derivative japonine, both from *Orixa japonica* (Rutaceae), are notable for being intestinal smooth muscle relaxants and echinopsine from *Echinops* species (Asteraceae) is psychotropic.

Furoquinolines (Phe|pyridine|C4O) notably derive from the Rutaceae and include a variety of antibacterial and antifungal compounds. Thus, *O*-methylptelefolonium and pteleatine from *Ptelea trifoliata* (Rutaceae) and veprisinium from *Vepris louisii* (Rutaceae) are antimicrobial. Ribalinium from *Ruta graveolens* (Rutaceae) is anti-mycobacterial. The Rutaceae furoquinolines dictamnine(dictamine), γ-fagarine, haplopine, isodictamnine, kokusaginine, maculosidine and skimmianine (β-fagarine) are phototoxic antimicrobials. Dictamnine, γ-fagarine (8-methoxydictamnine) and skimmianine (7,8-dimethoxydictamnine) from *Ruta graveolens* (rue) (Rutaceae) are photomutagenic, forming DNA monoadducts in a light-dependent process and thus contributing to the phototoxic phytodermatitis of rue. Confusameline, kokusaginine and skimmianine (β-fagarine) are 5-hydroxytryptamine (5HT, serotonin) receptor (5HT-R) antagonists and platelet aggregation inhibitors. Haplophyllidine and robustine are psychoactive.

Pyranoquinolines (Phe|pyridine|C5O) include the antimicrobials flindersine and *N*-methylflindersine from *Flindersia* and *Glycosmis* species (Rutaceae).

Acridines (Phe|pyridine|Phe) include arborinine from *Ruta graveolens* and other Rutaceae (a spasmolytic and A1 adenosine receptor antagonist) and the pyranoquinoline acronycine (with cytotoxic and antitumour activity) from *Acronychia* species and *Melicope leptococca* (Rutaceae) and which has become a useful lead compound for the synthesis of other anticancer compounds. A variety of synthetic acridines are DNA binding anticancer compounds.

Quinazoline alkaloids (Phe|C4N2) include a variety of bioactive compounds from a number of plant families. Febrifugine (Phe|C4N2−C_3−HHpyridine) and the hemiacetal isofebrifugine (Phe|C4N2−CH_2−C4OH−HHpyridine) are potent antimalarials from *Dichroa febrifuga* and *Hydrangea* species (Saxifragaceae). The quinazolines deoxypeganine, deoxyvasicinone and peganine (Phe|C4NN#|C4N#) from *Peganum* species (Zygophyllaceae) are AChE inhibitors. The structurally related vasicinol (7-hydroxypeganine) from *Adhatoda vasica* (Acanthaceae) and *Sida cordifolia* (Malvaceae) is also an AChE inhibitor and the related vasicinone from the same sources is bronchodilatory. Tryptanthrine (couroupitine A) (Phe|C4NN#|C4N#|Phe) from *Strobilanthes cusia* (Acanthaceae), *Isatis tinctoria* (woad) (Brassicaceae) and *Polygonum tinctorum* (Polygonaceae) is a potent inhibitor of inducible cyclooxygenase (COX) 2, inhibits inducible nitric oxide synthase (iNOS) expression and is an agonist of the xenobiotic-responsive element-interacting aryl hydrocarbon receptor (dioxin receptor).

Camptothecins. The alkaloid camptothecin from *Camptotheca acuminata* (Nyssaceae) and *Mappia foetida* (Icacinaceae) has a pyranoindolizoquinoline structure (Phe|pyridine|C4N#|C5N#|C5L) involving the fusion of quinoline (Phe|pyridine), indolizidine (C4N#|C5N#) and C5 lactone (C5L) rings. Camptothecin is a topoisomerase I inhibitor and is a potent cytotoxic and antitumour compound that is used clinically as an anticancer

compound and has been the "lead compound" for the synthesis of a variety of anticancer compounds such as irinotecan, topotecan and 9-aminocamptothecin.

xiv. Tropane alkaloids are alicyclic compounds containing an N atom and structurally based on the bicyclic aliphatic tropine (8-methyl-8-azabicyclo[3.2.1]octan-3-α-ol) (C7[CH_3–N<]). which can be simply viewed as a cycloheptane (C7) cross-linked by a methylamino (CH_3–N<) group. Pseudotropine is the corresponding 3-β-ol isomer, nortropine lacks the *N*-methyl and tropane lacks the 3-hydroxy. Ecgonine (tropine 2-carboxylic acid) is the precursor of the important narcotic cocaine (ecgonine benzoate methyl ester). The highly toxic anticholinergic atropine (tropine tropate), a potent antagonist of mACh-Rs, is an ester of tropine and tropic acid (α-(hydroxymethyl)phenylacetic acid) (Section 1, Appendix). The tropine moiety derives biosynthetically from ornithine and the tropic acid from the amino acid phenylalanine.

Tropine derivatives are typically found in certain highly poisonous Solanaceae species, most notably *Atropa belladonna* (deadly nightshade), *Datura stramonium* (thornapple), other *Datura* species, *Duboisia myoporoides* (corkwood elm), *Hyoscyamus niger* (henbane) and other *Hyoscyamus* species. Other sources include *Convolvulus* species (Convolvulaceae), *Erythroxylum coca* (coca), other *Erythroxylum* species (Erythroxylaceae) and *Bruguiera species* (Rhizophoraceae).

Hyoscyamine (duboisine) and the racemate atropine are mACh-R antagonists and a number of atropine derivatives also have this property, namely anisodamine (6β-hydroxyhyoscyamine), 7β-hydroxyhyoscyamine, hyoscine (6,7-epoxyhyoscyamine or scopolamine), benzoyltropein (tropine benzoate), littorine (tropine α-hydroxyphenylpropionate), tigloidine (pseudotropane tiglate) and tropacocaine (pseudotropine benzoate). The further derivatives apoatropine (α-dehydrohyoscyamine) and tropine are very toxic.

The stimulant narcotic cocaine (benzoylmethylecgonine) from *Erythroxylum coca* (coca) and other *Erythroxylum* species (Erythroxylaceae) inhibits serotonin (5HT) and dopamine reuptake. Related bioactive tropane alkaloids from *Erythroxylum* species include benzoylecgonine, benzoyltropeine (tropine benzoate), cinnamoylcocaine (cinnamoylmethylecgonine) and ecgonine.

A variety of other tropane alkaloids have been isolated of which the most important is anatoxin-A, a highly toxic nACh-R agonist and depolarizing neuromuscular blocking agent deriving from *Anabaena* cyanobacterium species that can contaminate inland waters.

xv. Quinolizidine and *Lycopodium* alkaloids. Quinolizidine alkaloids have two fused six-membered rings sharing an N atom, the simplest such entity being the saturated two-ring compound quinolizidine (C5N#|C5N#). More complex entities are formed by the addition of further N-containing rings through addition of substituents such as –CH_2–NH–CH_2–, –$(CH_2)_3$–NH– and –$(CH_2)_4$–NH– as well as other ring and "side chain" substituents. The major source of quinolizidine alkaloids are the legumes (Fabaceae). However, various quinolizidine and related alkaloids have been isolated from *Lycopodium* species (club mosses) (Lycopodiaceae).

Legume quinolizidines. The simplest legume quinolizidine is the toxic lupinine (quinolizidine-1-methanol) from *Lupinus* (lupine) species as well as from *Anabasis aphylla* (Chenopodiaceae). Quinolizidine-based legume toxicity is a significant agricultural problem. Other toxic legume quinolizidines (other attributes in parentheses) include anagyrine (C5N#|C5N#||C5N#|CN5# i.e. quinolizidine||quinolizidine) (teratogen), cytisine (C5N#|C5N#||C5N) (nACh-R agonist, hallucinogen and teratogen), *N*-methylcytisine (nACh-R agonist and teratogen), (−)-sparteine (lupinidine) (quinolizidine||quinolizidine)

(which blocks voltage-gated Na^+ channels and ATP-regulated K^+ channels), lupanine (2-oxo-11α-sparteine) (weak sedative and Na^+ channel blocker), and 13-hydroxylupanine (anti-arrhythmic and hypoglycaemic). Sophoramine (C5N*#|C5N*#||C5N*#|CN5#) is also anti-arrhythmic. (+)-Matrine (C5N*#|C5N*#||C5N*#|CN5#) inhibits lipopolysaccharide-induced cytokine expression in immune cells and is anti-nociceptive by acting through μ and κ opiate receptors. (+)-Allomatrine (the C-6 epimer of (+)-matrine) is anti-nociceptive, acting through κ opiate receptors.

***Lycopodium* alkaloids.** The *Lycopodium* (or club moss) alkaloids include quinolizidine alkaloids in which N atoms are variously shared between two or three six-membered rings. The toxic alkaloid lycopodine (C5N*#|C5N*#||C6*[isobutyl<]) is a tetracyclic alkaloid with an N shared between two six-membered rings. The toxic alkaloid carolinianine (C5N*#|C5N*#||C5N*#N#|CN5#)) is a tetracyclic with two Ns shared between three and two six-membered rings, respectively. Other such alkaloids, such as lycodine (C5N|C6[isobutyl<]|C5N), have Ns that are associated with only one ring.

xvi. Amaryllidaceae alkaloids derive from the bulbs of plants such as amaryllis or belladonna lily (*Amarillus belladonna*), daffodil and narcissus (*Narcissus* species) and snowdrop (*Galanthus nivalis*). These alkaloids are typically tetracyclic with a five- or six-membered N-containing ring as a common feature, many having a further ring created by an MD bridge (–O–CH_2–O–).

Many Amaryllidaceae alkaloids are toxic and are of interest as anticancer and selective anti-protozoal agents because of their cytotoxicity. Examples (some source genera in parentheses) include: the cytotoxic antimalarials augustine (MD–Phe|C5N[OH–CH–CH_2<]|C6) (*Crinum*), crinamine (MD–Phe|C5N[OH–CH–CH_2<]|C6) (*Crinum*), lycorine (MD–Phe|C5N|C6) (*Brunsvigia*, *Lycoris*), 1,2-di-*O*-acetyllycorine (*Brunsvigia*); the related antineoplastic cytotoxic alkaloids ambelline (MD–Phe|C5N[OH–CH–CH_2<]|C6), acetylcaranine and anhydrolycorinium (*Amaryllis*); the cytotoxics tazettine, hippeastrine (MD– Phe|C5L|C6|C4N) and haemanthidine (*Hymenocallis*); the specific anti-microsporidium (*Encephalitozoon intestinalis*) antimitotics pancratistatin (MD–Phe|C5N|C6) (*Pancratium*) and 7-deoxynarciclasine (*Narcissus*); and the further toxic alkaloids 3-acetylnerbowdine (*Nerine*), candimine (MD–Phe|C5L|C6|C4N) (*Hippeastrum*) and caranine (MD–Phe|C5N*#|C4N*#|C6*) (*Amaryllis*).

The phenanthridine alkaloid lycorine (narcissine, galanthidine) (MD–Phe|C5N|C6) has a widespread occurrence and inhibits protein synthesis. Like lycorine, the structurally similar alkaloids dihydrolycorinine, haemanthamine, narciclasine, pretazettine and pseudolycorine also inhibit protein synthesis at the level of peptide bond formation. Galanthamine (lycorimine) (Phe*|C6N*·*′|C4O*·*′|C6*′), from daffodil bulbs but also of widespread occurrence, is both a nACh-R allosteric modulator and an inhibitor of AChE. Galanthamine is clinically employed in the treatment of Alzheimer's disease (dementia linked to deficiency in acetylcholine-mediated signalling in the central nervous system).

xvii. Other polycyclic alkaloids not covered above include the following groups of alkaloids:

Benzofuranone tetrahydropyrrole alkaloids. Shihunidine (Phe|C4OL·C4N) and shihunine (Phe|C4OL·C4N) from *Dendrodium* species are inhibitors of the Na^+, K^+-ATPase (sodium pump).

Benzoxazolinone alkaloids include some types of phytoalexins (compounds produced by plants in response to microbial infection), examples including *Avena sativa* (oats) (Poaceae) avenalumin I (*p*OH-Phe|C4NOL–CH=CH–Phe-*p*OH), *Triticum aestivum* (wheat)

and *Zea mays* (maize) (Poaceae) 2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one (DIMBOA) (Phe|C4NO) and DIMBOA glucoside and *Dianthus caryophyllus* (carnation) (Caryophyllaceae) dianthalexin (Phe|C4NOL-Phe).

Cepahalotaxine alkaloids are based on cephalotaxine which has a pentacyclic system including a seven-membered ring and a five-membered ring sharing an N atom (MD-Phe|C6N*#|C4*N#|C5*). Cephalotaxine alkaloids include the cytotoxic, anticancer protein synthesis inhibitors cephalotaxine, harringtonine and homoharringtonine.

Imidazole-containing alkaloids related to the amino acid histidine include histamine (imidazole-4-ethanamine) (C3N2) (from numerous plant sources) and casimiroedine (an *N*-glycoside), *N*-methylhistamine and *N*,*N*-dimethylhistamine from *Casimiroa edulis* (Rutaceae) that are hypotensive through interaction with histamine receptors.

Imidazoloylmethylfuranones include the parasympathetic agonist pilocarpine (C4OL-CH_2-C3N2) and pilosine (carpidine) (Phe-CH_2-C4OL-CH_2-C3N2) from *Pilocarpus* species (Rutaceae), narcotic compounds that are agonists of muscarinic acetylcholine receptors (mACh-Rs) and accordingly stimulate salivation and tear secretion.

Isoxazole alkaloids involve a five-membered unsaturated ring having an O and an N atom (C3NO). Isoxazole alkaloids notably include ibotenic acid (C3NO–$CH(NH_3^+)COO^-$) and muscimol (OH–C3NO–CH_2–NH_2) from the reputedly aphrodisiac, hallucinogenic and extremely toxic *Amanita* species mushrooms. Ibotenic acid (= α-amino-3-hydroxy-5-isoxazoleacetic) is neurotoxic and an agonist of excitatory NMDA- and non-NMDA ionotropic glutamate receptors and of inhibitory ionotropic glutamate receptors. Muscimol (3-hydroxy-5-aminomethyl-isoxazole) is an hallucinogenic GABA(A) receptor agonist.

Phenanthroindolizidine and phenanthroquinolizidine alkaloids involve a phenanthrene (Phe|Phe|Phe (angular)) fused with an indolizidine or quinolizidine, respectively. The phenanthroindolizidines tylophorine (phenanthrene|C5N#|C4N#) and tylocrebrine (phenanthrene|C5N#|C4N#) and the phenanthroquinolizidine cryptopleurine (phenanthrene|C5N#|C5N#) are toxic, cytotoxic protein synthesis inhibitors. The phenanthroindolizidines tylophorine and pergularinine are thymidylate synthase inhibitors.

Taxine alkaloids are complex polycyclic compounds in which N is present but not as an integral part of a ring. The taxines are found in *Taxus* (yew) species (Taxaceae). Taxine A (C6|C10|C6–O–CO–CH(OH)–CH($N(CH_3)_2$)–Phe) is substantially responsible for yew toxicity. The related polycyclic amide taxol (paclitaxel) and the closely related docetaxel are tubulin-binding, antimitotic cytotoxics that are used clinically as anticancer drugs. A variety of taxines have been isolated from *Taxus* species.

Other alkaloids include: the **quinine-like chloroalkaloids** (C5·ChloroC5*,*′ (–CH_2*CH_2–NH*′–)|C6*,*′) acutumine, acutumidine, dauricumine and dauricumidine from *Menispermum dauricum* (Menispermaceae); **tricyclic pyrazole alkaloids** (THpyrrole#|C3NN#–Phe) from *Newbouldia laevis* and *Withania somnifera* (Solanaceae) including withasomnine, newbouldine and the 4′-hydroxy and 4′-methoxy derivatives of these alkaloids; **pyrrolidinoquinolines** variously from *Calycanthus* species (Calycanthaceae) and *Psychotria* species (Rubiaceae) including calycanthine (Phe|C5N*,*′(||C4N*)|C5N*,*′(||C4N*′)|Phe), isocalycanthine, and tetrahydroisocalycanthine; **pyrazine alkaloids** (*p*C4N2), namely the antibiotic mycotoxin aspergillic acids from *Aspergillus* species (fungi); **polycyclic quinolizidine lactones** include the anti-inflammatory prostaglandin synthetase inhibitors cryogenine (Phe|C11OL(Phe|)|C5N#|C5N#) and nesodine from *Heimia* species (Lythraceae); **various diverse peptide macrocyclic alkaloids** including the DNA-binding RNA- and DNA-polymerase inhibitor pithecolobine from *Pithecolobium*

saman (Fabaceae) and the potent cytotoxic, antitumour, antitubulin compounds maytansine (from *Maytenus* species (Celastraceae)) and cryptophycin A (a cyclic depsipeptide from the cyanobacterium (blue–green alga) *Nostoc*); **colchicine-related** antimitotic alkaloids variously from *Androcymbium*, *Colchicum* and *Gloriosa* species (Liliaceae) and including androcymbine, *O*-methylandrocymbine, colchicine (Phe|C7(NH−CO−CH_3)|C7) and demecolcine (colchicine being used to treat gout); and **securinine** (in which piperidine shares an N with a pyrrolidine (five-membered ring) and a seven-membered ring) (C5N#|C6N#(−CH_2−)|C4OL); securinine derives from *Securinega suffruticosa* (Euphorbiaceae) and *Securidaca longepedunculata* (Fabaceae) and is a GABA(A) receptor antagonist.

xvii. Pseudoalkaloids. As indicated previously, for the sake of consistency and simplicity, all heterocyclics with a ring N have been included here in the category of "alkaloids" including a variety of "universal" biochemically important derivatives of pyrimidine (a six-membered ring with two Ns) and purine (pyrimidine fused with a five-membered ring with two Ns). Unsaturated pyrimidine (*m*C4N2) and purine (*m*C4N2|C3N2; pyrimidine|imidazole) derivatives are involved in RNA and DNA structure and biosynthesis as well as related compounds used in signalling and for "defensive" purposes.

The bases found in RNA (ribonucleic acid) are the purine heterocyclics adenine (6-aminopurine) and guanine (2-amino-6-oxypurine) and their "complementary" pyrimidine bases uracil (2,4-dioxypyrimidine) and cytosine (2-oxy-4-aminopyrimidine), respectively (Section 1, Appendix). In RNA double-stranded duplexes adenine (A) base-pairs with uracil (U) via two hydrogen bonds (A=U) and guanine base-pairs with cytosine (C) via 3 hydrogen bonds (G≡C). Adenine forms the nucleoside adenosine by an N-glycosidic link with the 5-carbon (C5) sugar ribose. Adenosine can be successively modified by phosphorylation to yield the nucleotides adenosine 5′-monophosphate (5′-AMP), adenosine 5′-diphosphate (5′-ADP) and adenosine 5′-triphosphate (5′-ATP). The other bases form the corresponding nucleosides (and nucleotides) guanosine (5′-GMP, 5′-GDP and 5′-GTP), uridine (5′-UMP, 5′-UDP and 5′-UTP) and cytidine (5′-CMP, 5′-CDP and 5′-CTP).

The bases found in DNA (deoxyribonucleic acid) are adenine and guanine and the corresponding base-pairing complements thymine (T) (5-methyluracil, 2,4-dioxy-5-methylpyrimidine) and cytosine (C) that hydrogen bond in double-stranded (duplex) DNA thus: A=T and G≡C. The corresponding nucleosides (deoxyribonucleosides) are formed via *N*-glycosidic links with 2′-deoxyribose (2′-deoxyadenosine, 2′-deoxyguanosine, 2′-deoxythymidine and 2′-deoxyuridine) and thence the corresponding deoxyribonucleotides (5′-dAMP, 5′-dADP, 5′-dATP, 5′-dGMP, 5′-dGDP, 5′-dGTP, 5′-dTMP, 5′-dTDP, 5′-dTTP, 5′-dCMP, 5′-dCDP and 5′-dCTP).

The 3′,5′-cyclic nucleoside monophosphates 3′,5′-cyclic AMP (cAMP) and 3′,5′-cyclic GMP (cGMP) are so-called "second messengers", the cytosolic levels of which rise in response to binding of particular "primary messengers" (such as hormones or neurotransmitters) to plasma membrane receptors (Chapters 5 and 7). Both cGMP and cAMP have been found in plants. ATP is the so-called "energy currency" of cells. UDPglucose is involved in protein glycosylation and in synthesis of sucrose, cellulose (a β-1,4-glucan), callose (a β-1,3-glucan) and glycogen (an α-1,4-glucose polymer). Synthesis of starch (an α-1,4-glucose polymer) involves ADP-glucose, CDP-glucose and GDP-glucose as precursors (Chapter 2).

In addition to the bases outlined above, transfer RNA (tRNA) (involved in amino acid-specific codon recognition in protein synthesis) contains unusual chemically modified bases (e.g. 6-methylaminopurine). DNA can be modified by methylation yielding 5-methylcytosine. A number of other adenine (6-aminopurine) derivatives are plant growth regulator "cytokinins" having mitogenic and anti-senescent activity in plants including plant-derived

dihydrozeatin (N^6-isopentanoladenine), N^6-(Δ^2-isopentenyl)adenine and zeatin (N^6-(Δ^2-isopentenol)adenine) and the semi-synthetics N^6-furfuryladenine (kinetin) and N^6-benzyladenine.

Critical N-containing heterocyclics are chlorophyll a and chlorophyll b, Mg^{2+}-chelated cyclic tetrapyrroles that are involved in light harvesting in the chloroplast photosystems. The Fe^{3+}(Fe^{2+})-complexed tetrapyrrole haems are involved as the prosthetic groups of cytochromes in mitochondrial and chloroplast electron transport chains and of cytochrome P450 of the endoplasmic reticulum (ER)-associated xenobiotic detoxification system. The non-cyclic tetrapyrrole phytochrome is the key chromophore in red/far red light perception and signalling in plants. Haem is the prosthetic group of the oxygen-binding protein haemoglobin.

Vitamins are plant-derived compounds that we cannot synthesize ourselves and which accordingly must be ingested for survival. Vitamins are typically ring structures involving one or more ring Ns. **Thiamine (vitamin B_1)** (pyrimidine-CH_2-(N)-thiazole) involves a pyrimidinylmethyl (*m*C4N2) linked to a thiazole (C3NS) ring and as the thiamine pyrophosphate (TPP) coenzyme derivative is involved in pyruvate dehydrogenase, α-ketoglutarate dehydrogenase and transketolase function. Good vitamin B_1 sources are leafy vegetables, grain and legumes and deficiency causes beri beri (diarrhoea and fatigue). **Riboflavin (vitamin B_2)** is a riboside of isoalloxazine (Phe|pyrazine|pyrimidine) (Phe|*p*C4N2|*m*C4N2) (Section 1, Appendix) and is part of the redox coenzymes flavin adenine dinucleotide (FAD/$FADH_2$) and flavin mononucleotide (FMN/$FMNH_2$) (oxidized/reduced forms). Riboflavin is present in leafy vegetables and cereals and deficiency is associated with growth retardation. **Pyridoxine (vitamin B_6)** (1-methyl-3-hydroxy-4,5-dicarboxymethylpyridine) is the precursor of pyridoxal phosphate, a coenzyme involved in transaminase and lysyl oxidase. Vitamin B_6 is found in cereals and legumes and deficiency is associated with dermatitis, depression and particular infantile convulsions. **Biotin (vitamin H or coenzyme R)** (C4S|C3N2) involves fused, fully reduced (saturated) thiophene and imidazole rings and is involved in carboxylation reactions (e.g. fatty acid synthesis). **Folic acid** (pteroylglutamate) has a pteridine (*m*C4N2|*p*C4N2) (pyrimidine|pyrazine) heterocyclic ring and is involved in methylation reactions crucial for DNA precursor (thymine) synthesis. Folate is present in green leafy dietary vegetables and maternal folate deficiency is associated with occurrence of spina bifida. **Cyanocobalamin (vitamin B_{12})** (5,6-dimethylbenzimidazolyl cyanocobamide), produced by colonic bacteria, is a cobalt ion-chelated tetrapyrrole, the coenzyme derivatives of which are involved in C–C bond breakage and re-formation in methionine (C_5) and succinyl-CoA (C_4) formation from homocysteine (C_4) and methylmalonyl-CoA (C_4), respectively. Vitamin B_{12} deficiency is associated with pernicious anaemia. **Niacin** (nicotinic acid, pyridine 3-carboxylic acid) is the precursor of nicotinamide which is part of the nicotinamide adenine dinucleotide redox coenzymes NAD^+/NADH and $NADP^+$/NADPH (oxidized/reduced forms). Niacin is found in grain and legumes and niacin deficiency is associated with pellagra (involving mental and physical weakness).

Methyl derivatives of xanthine (2,3-dioxypurine) namely caffeine (1,3,7-trimethylxanthine), theobromine (3,7-dimethylxanthine) and theophylline (1,3-dimethylxanthine) (Section 1, Appendix) are variously found in plants used for stimulatory drinks such as *Ilex paraguayensis* (maté) (Aquifoliaceae), *Coffea* species (coffee) (Rubiaceae), *Paullinia cupana* (guarana) (Sapindaceae), *Cola acuminata* (cola) and *Theabroma cacao* (cocoa) (Sterculiaceae) and *Camellia sinensis* (tea) (Theaceae). These methylxanthines are variously active as inhibitors of

cAMP phosphodiesterase or as adenosine receptor antagonists. Caffeine also activates the ryanodine receptor Ca^{2+} channel.

The pyrimidine nucleosides convicine (3,6-diamino-2,4,5-trihydroxypyrimidine 5-*O*-β-glucoside) and vicine (divicine-β-glucoside, 2,6-diamino-4,5-dihydroxypyrimidine 5-*O*-β-glucoside) derive from *Vicia fava* (fava beans) (Fabaceae) and give rise to Favism in people with glucose-6-phosphate dehydrogenase (G6PDH) deficiency (typically in Mediterranean countries in which this deficiency was selected for as a protectant against malaria). The aglycones (non-glycosylated pyrimidines) are involved in oxidative reactions resulting in glutathione deficiency, red blood cell haemolysis and anaemia in G6PDH-deficient individuals.

1.7 Plant phenolics

Plant phenolics represent a very large group of defensive compounds defined here as having a phenol (hydroxybenzene) moiety. In some instances substances having a phenolic precursor (e.g. methoxybenzene derivatives) have conveniently also been included in this category. Phenolics derive biosynthetically from hydroxycinnamoyl coenzyme A (yielding a phenylpropanoid moiety).

The phenolics range in complexity from simple phenolics and quinones (with one ring), through chalcones and stilbenes (with two rings) to a range of phenolics with three rings namely anthocyanins, anthochlors, benzofurans, chromones, chromenes, coumarins, flavonoids, isoflavonoids, neoflavonoids, stilbenoids and xanthones (see Section 2, Appendix). More complex polycyclic phenolics exist, notably the hydrolysable tannins (gallotannins and ellagitannins) and the condensed tannins.

The phenolic ring system (Phenyl-OH, or for aromatics in general, Aryl-OH) is planar and electron-rich. The planar benzene ring is hydrophobic but the phenolic OH confers polarity and water-solubility and the capacity for hydrogen bonding, for example, Phenyl–OH$\cdots{}^{-}$OOC–X and Phenyl–OH$\cdots H_2N$–X (these properties permitting phenolic–protein interactions that are stronger, the greater the number of interactions involved). The phenolic group can be deprotonated (to form the phenolate (Phenyl-O^-) and can be oxidized yielding a quinone (Aryl=O) and the radical Aryl–$O^{\bullet}$. Accordingly, phenolics have antioxidant properties that are biologically important. Because of the extensive conjugated double bond systems found in the more complex phenolics (e.g. Aryl–$(CH_2–CH=CH)_n$), such compounds absorb light well in the visible part of the spectrum, that is, they are coloured.

The above properties of phenolics provide molecular rationales for phenolic compound functions. Thus, coloured phenolics act as pollinator-attractants and complex polyphenolics (tannins) bind tightly to proteins and act as herbivore deterrents through being bitter tastants. The planar ring systems of flavonoids and related compounds can mimic key enzyme substrates such as ATP and the key redox coenzymes NADPH, NADH, $FMNH_2$ and $FADH_2$. Many phenolics can act as anti-inflammatory antioxidants through covalent reaction with free radicals, notably ROS such as superoxide (O_2^-). Conversely, many phenolics have antimicrobial (antibacterial or antifungal) properties. The complex structure and function features of the various groups of phenolics are sketched below. The structures of a variety of simple and more complex polycyclic phenolics are presented in the order of increasing complexity in the Appendix (Section 2).

i. Simple phenols include a variety of compounds noted because of their antimicrobial, topical antimicrobial, antiseptic, dermatitic and odorant properties. The denaturant, irritant, odorant and antiseptic properties of the parent compound phenol are familiar.

Antiseptic plant-derived phenols include phenol (Phe–OH, hydroxybenzene, carbolic acid), *p*-cresol (4-methylphenol), catechol (1,2-dihydroxybenzene), resorcinol (1,3-dihydroxybenzene) and pyrogallol (1,2,3-trihydroxybenzene). Other simple phenols with antimicrobial properties include some related to benzoic acid (benzenecarboxylic acid), namely salicylic acid (2-hydroxybenzoic acid), ginkgoic acid (2-hydroxy-6-(pentadec-8-enyl)benzoic acid), gentisic acid (2,5-dihydroxybenzoic acid), pyrocatechuic acid (3,4-dihydroxybenzoic acid) and gallic acid (3,4,5-trihydroxybenzoic acid). Other plant-derived phenol-related compounds include 4-methylcatechol, 1,3-dihydroxy-5-(heptadec-12-enyl)benzene, hydroquinone (1,4-dihydroxybenzene), 1,4-dihydroxy-2-geranyl (di-isoprenyl)benzene and 4-methoxybenzaldehyde (*p*-anisealdehyde).

The non-specific biocidal properties of phenols give rise to dermatitic properties. Noted plant phenol dermatitics include anacardic acids (2-hydroxy-6-(long chain alkyl)-benzoic acids), catechol (1,2-dihydroxybenzene), ginkgol (3-(pentadec-8-enyl)phenol), *Grevillea robusta* (Proteaceae) grevillol (1,3-dihydroxy-5-tridecylbenzene), salicylic acid (2-hydroxybenzoic acid), sesamol (3,4-methylene dioxyphenol), *Turricula parryi* (poodle dog bush) (Hydrophyllaceae) turricolol E (1,4-dihydroxy-2-(tri-isoprenyl)benzene) and the *Toxicodendron radicans* (poison ivy) (Anacardiaceae) 3-(long chain alkenyl)-catechols.

Phenols have distinct odours. Notable simple phenol-related odorants/tastants include 4-methoxybenzaldehyde (*p*-anisealdehyde), guaiacol (2-methoxyphenol), 4-hydroxybenzaldehyde, phenethyl alcohol, piperonal (heliotropin, 3,4-methylenedioxybenzoic acid) and *Vanilla planifolia* (vanilla) (Orchidaceae) pod vanillin (3-methoxy-4-hydroxybenzaldehyde) (Chapter 10).

Some simple phenolics inhibit COX (prostaglandin synthetase) and/or 5-lipoxygenase (5-LOX). COX inhibitors include the anacardic acids, 2,6-dimethoxyphenol and *Ginkgo biloba* (Ginkgoaceae) ginkgoic acid (2-hydroxy-5-pentadec-8-enyl)benzoic acid) and ginkgol (3-(pentadec-8-enyl)phenol). Simple phenol 5-LOX inhibitors include ginkgol and grevillol. The acetyl ester of salicylic acid (2-hydroxybenzoic acid) is the synthetic COX-inhibitory anti-inflammatory aspirin (Chapter 14).

ii. Phenolic ketones. Phenolic ketones typically have a phenol-related benzene (unsaturated C6) ring with a 2-carbon (C_2) sidechain as exemplified by the phenolic precursor acetophenone (Phe–CO–CH_3). Such compounds derive from phenylpropanoids (Phe–C_3). A variety of such phenolic ketones are based upon phloroglucinol (1,3,5-trihydroxybenene) including: the COX and 5-LOX inhibitors, 2,6-dimethoxy-4-hydroxyacetophenone and xanthoxylin (4,6-dimethoxy-2-hydroxyacetophenone; phloroacetophenone 4,6-dimethyl ether) and the *Humulus lupulus* (hops) (Cannabaceae) bitter-tasting, isoprenylated antibacterials humulone (α-lupulic acid) and lupulone (β-lupulic acid). The non-aromatic, hops-derived, tricyclic ketone tricyclodehydrohumulone is also a bitter tastant. Other phenolic ketones include acetosyringone (3′,5′-dimethoxy-4′-hydroxyacetophenone) (the tobacco inducer of *Agrobacterium tumefaciens* virulence gene expression required for infection), the phloroglucinol benzophenone maclurin, the benzophenone tubulin-binding anti-mitotic xanthochymol and the oestrogenic macrocyclic mycotoxin zearalenone from the fungus *Gibberella zeae*.

iii. Phenylpropanoids. The phenylpropanoids derive biosynthetically from phenylalanine (Phenyl–CH_2–$CH(NH_2)$–COOH) through deamination. The phenylpropanoids (Phe–C_3) in turn give rise to lignans in which benzene rings are linked by a C–C bond (Phe–Phe) and coumarins in which ring closure by a lactone grouping (–O–CO–) creates a benzopyran-2-one (Phe|C5OL).

Major simple phenylpropanoids include cinnamic acid (Phe–CH=CH–COOH), *p*-coumaric acid (*p*-hydroxycinnamic acid), *o*-coumaric acid (*o*-hydroxycinnamic acid), caffeic

acid (3,4-dihydroxycinnamic acid), ferulic acid (3-methoxy-4-hydroxycinnamic acid) and isoferulic acid (3-hydroxy-4-methoxycinnamic acid). These parent compounds can in turn be altered through reduction of the sidechain double bond or of the carboxyl (to yield aldehydes and alcohols); formation of glycosides with sugars; formation of carboxylic acid esters with sugars and other compounds (notably quinic acid and shikimic acid); formation of amides; decarboxylation (to yield phenylpropenes and phenylpropanes); methylation of phenolic hydroxyls; and formation of an MD ring from phenolic hydroxyls.

Some non-polar phenylprop-2-ene (allylbenzene (AB); Phe–CH_2–CH=CH_2) derivatives can form 2,3-epoxides and thence covalent adducts with DNA, such genotoxic (and potentially mutagenic and carcinogenic) compounds including elemicin (3,4,5-trimethoxyAB), estragole (3-methoxyAB), methyleugenol (4,5-dimethoxyAB) and safrole (4,5-methylenedioxyAB), noting that such compounds occur in plant material ingested by humans. While the phenylprop-1-ene (prop-1-enebenzene; PB) compounds *trans*- and *cis*-asarone (2,4,5-trimethoxyPB) form DNA adducts, a range of other plant-derived PB or AB compounds are not genotoxic including eugenol (4-hydroxy-5-methoxyAB), isosafrole (4,5-methylenedioxyPB), methylisoeugenol (4-hydroxy-5-methoxyPB) and myristicin (3-methoxysafrole) (which forms such adducts poorly). Epoxide hydrolases provide some protection from genotoxic phenylpropenes.

A variety of **phenylpropanoid ketones** are anti-inflammatory inhibitors of COX and 5-LOX, enzymes that are involved in the formation of prostaglandins and leukotrienes, respectively. Thus, the dihydroferulic acid-derived ketone [6]-Gingerol (4′-hydroxy-5′-methoxyphenylpropane–CO–CH_2–CH(OH)–$(CH_2)_4$–CH_3) (Phe-alkyl ketone) inhibits both COX and 5-LOX as variously do the corresponding [2]-, [4]-, [8]-, [10]-, [12]-, [14]- and [16]-gingerols and the diketones [6]- and [8]-gingerdione, all of these compounds deriving from the rhizome of *Zingiber officinale* (ginger) (Zingiberaceae). The structurally related **diarylheptanoids** are ketones (R–CO–R′) from *Alpinia* species (Zingiberaceae) rhizomes in which the aryl R–CO– and R′– groups are phenylpropanoid (Phe–C_3) and phenylpropanoid-related (Phe–C_4), respectively. The diarylheptanoids are variously COX and 5-LOX inhibitors.

A variety of other phenylpropanoids have been shown to inhibit particular enzymes including (target enzyme in parentheses): coniferyl aldehyde and the amide fagaramide (COX); the biphenylpropanoid glycosides forsythiaside, hellicoside and suspensaside (5-LOX and cAMP phosphodiesterase); the allylbenzene myristicin (monoamine oxidase); the tricaffeic acid salvianolic acid A (gastric H^+ secreting H^+-ATPase); the caffeic acid esters vanicosides A and B and the diferuloyl curcumin (protein kinases); curcumin and caffeic phenethyl ester (HIV-1 integrase); caffeic acid (xanthine oxidase); and ferulic acid, curcumin, the diarylheptanoid yakuchinone B and 4-hydroxy-3-methoxy cinnamaldehyde (tyrosinase).

iv. Lignans. Simple lignans derive from dimerization of phenylpropanoids (Phe–C_3), typically through a sidechain (C_3) C–C link, that is, Phe–C_3 + Phe–C_3 → Phe–C_3–C_3–Phe (typically Phe–CH_2–$CH(CH_3)$–$CH(CH_3)$–CH_2–Phe). However, alternative linkages could be phenyl C–C links (i.e. Phe–C_3 + Phe–C_3 → C_3–Phe–Phe–C_3). In monoepoxylignans, a tetrahydrofuran (THF) (C4O) is formed linking the two phenyls, that is, Phe–CH_2–$CH(CH_3)$–$CH(CH_3)$–CH_2–Phe + O → Phe–CH_2–C4O–CH_2–Phe or Phe|C4O–Phe (in which the THF moiety is fused with one of the phenyls). Further oxidation yields lignanolides in which there is a central tetrahydrofuranone (C4OL) lactone ring (Phe–CH_2–C4OL–CH_2–Phe) and bisepoxylignans in which phenyl (Phe-) moieties are linked by two fused THF rings (Phe–C4O|C4O–Phe). In the more complex podophyllotoxin-related cyclolignans, there is sidechain cyclization to form a ring system fused with one of the

phenyl groups and further cyclolignan possibilities exist. These various structural types are further varied by substitutions with hydroxyl, methoxy, methylenedioxy and *O*-glycosyl groups. Lignans are mostly found in wood and many have cytotoxic properties.

Simple lignans involving a Phe–C_3–C_3–Phe structure are illustrated by the antioxidant and Ca^{2+} channel blocker nordihydroguaiaretic acid (NDGA) (3,4-dihydroxyphenyl–CH_2–$CH(CH_3)$–$CH(CH_3)$–CH_2–(3′,4′-dihydroxyphenyl)), the bitter-tasting phyllanthin and the cAMP phosphodiesterase inhibitor *cis*-hinokiresinol. Simple lignans of the C_3–Phe–Phe–C_3 kind are illustrated by the antibacterials honokiol and the protein kinase inhibitor magnolol.

Lignanolides (Phe–CH_2–C4OL–CH_2–Phe) include the Ca^{2+} channel blocker trachelogenin, the cytochrome P450-linked oxygenase inhibitor cubebin, the cAMP phosphodiesterase inhibitor (−)-arctigenin and the antimitotic glycoside podorhizol-β-D-glucoside from *Podophyllum* species (Podophyllaceae).

Monoepoxylignans include the Ca^{2+} channel blockers fargesone A and fargesone B (Phe|C4O–Phe(MD)); the antitumour compound burseran ((MD)Phe–CH_2–C4O–CH_2–Phe); the platelet activating factor (PAF) receptor antagonists grandisin, magnosalicin, saucernetin and (+)-veraguensin (Phe–CH_2–C4O–CH_2–Phe); and the PAF antagonists kadsurene and kadsurin A (DHPhe|C4O–Phe).

Bisepoxylignans (Phe–C4O|C4O–Phe) include the 1-acetoxypinoresinol and pinoresinol (cAMP PDE inhibitors), (−)-eudesmin (Ca^{2+} channel blocker), sesamolinol (antioxidant) and sesartemin (an inhibitor of cytochrome P450-linked oxygenase).

Podophyllotoxin-related cyclolignans include the important antitumour antimitotic podophyllotoxin ((MD)Phe(Phe)|C4OL) from *Podophyllum* species (Podophyllaceae) that inhibits topoisomerase and binds to tubulin. Podophyllotoxin-related compounds with antimitotic, cytotoxic and antitumour activity include 4′-demethylpodophyllotoxin, 4′-demethyldeoxypodophyllotoxin and deoxypodophyllotoxin. A variety of other kinds of cyclolignans and polycyclic neolignans have been characterized.

v. Benzoquinones, naphthoquinones and anthraquinones. The benzoquinone parent compound quinone (*p*O=Phe=O) (Q) is an oxidant which is readily reduced to *p*-hydroxyphenol (hydroquinone) (HO–Phe–OH). Quinone is a cytotoxic antimicrobial found in plants. A variety of simple antimicrobial hydroquinone-based phenolics are elaborated by plants as also outlined in Section i above. The reactivity of quinones in terms of redox reactions, hydrogen bonding (–C=O···H–X–) and hydrophobic binding in relation to proteins in general contributes to their irritant, cytotoxic and antimicrobial effects.

The napthoquinones are fused benzene and quinone rings (Phe|Q) and the anthraquinones involve a quinone ring fused with two benzene rings (Phe|Q|Phe). Furanobenzoquinones and furanonaphthoquinones involve a furan ring (C4O) fused with a benzoquinone or naphthoquinone ring, respectively. Similarly, pyranoquinones involve fusion of quinones with a pyran (C5O) ring. Binaphthoquinones and bianthraquinones derive from C–C links between the monomeric precursors. Substituents include hydroxy, hydroxymethyl methoxy, alkyl (notably isoprenyl), *C*-glycosyl and *O*-glycosyl groups. The compounds with more extensive conjugated systems (e.g. the anthraquinones) are coloured.

Benzoquinones (Q) include the bicyclic COX inhibitor arnebinone (DHPhe|Q) and the leukotriene receptor antagonists ardisianone and cornudentanone, which are 6′-methoxy-2′-alkylbenzoquinones (Q-alkyl) where the long chain alkyl substituents are 3-acetoxypentadecyl and 3-acetoxytridecyl, respectively. A number of benzoquinones are allergens including acamelin, 2,6-dimethoxybenzoquinone, geranylbenzoquinone,

prenylbenzoquinone and primin. The universal isoprenylated benzoquinone ubiquinones (e.g. coenzyme Q_{10}; benzoquinone-2-methyl-5,6-dimethoxy-3-(isoprenyl)$_{10}$) are key redox components in the mitochondrial electron transport chain and coenzyme Q_{10} is used as an anti-aging nutriceutical. The plastoquinones are analogous 3-isoprenylated 5,6-dimethyl-benzoquinone redox components in the chloroplast photosynthetic electron transport chain.

Naphthoquinones (Phe|Q). The benign isoprenylated naphthoquinones alkannin and shikonin are used for red lipstick and lawsone (1-hydroxynaphthoquinone) is the henna principle used to dye hair and for painting hands in Indian ceremonies. A variety of naphthoquinones are antimicrobials. Juglone, naphthazarin and plumbagin are protein kinase inhibitors. The widespread isoprenylated naphthoquinone vitamin K_1 (phylloquinone) is required for the formation of γ-carboxyglutamate residues in prothrombin, this permitting Ca^{2+} binding, prothrombin activation and subsequent blood clotting.

Anthraquinones (Phe|Q|Phe). Alizarin (1,2-dihydroxyanthraquinone) is the orange–red compound of *Rubia tinctorum* (madder) (Rubiaceae), a longstanding dyestuff in human history. A range of anthraquinones are variously cathartic, antimicrobial and cytotoxic. A variety of anthraquinones are protein kinase inhibitors including alizarin, chrysazin, damnacanthal, emodin and purpurin.

Binapthoquinones include the phototoxic phytotoxin cercosporin from the fungus *Cercospora* (two Phe|Q moieties linked by two Phe–Phe links and an MD link). Hypericin (two anthraquinones linked by three Phe–Phe linkages) is a **bianthraquinone** from *Hypericum* species (Hypericaceae). Hypericin is a phototoxic protein kinase inhibitor that causes light-dependent ovine facial eczema. **Benzonaphthoquinones** include the dermatitic cypripedin (Phe|Phe|Q). Lichen 7-chloroemodin is a novel **chloroanthraquinone** and the fused tricyclic **pyrano-α-naphthoquinone** β-lapachone (Phe|*o*Q|C5O) is a reverse transcriptase inhibitor with antimicrobial and cytotoxic activity.

vi. Stilbenes, bisbenzyls and phenanthrenes. Stilbenes (Phe–CH=CH–Phe) derive from the phenylpropanoid *p*-hydroxycinnamic acid (Phe–C_3; *p*OH–Phe–CH=CH–CO_2^-) and malonylCoA ($C_3^{\bullet}$; ^-O_2C–CH_2–CO–S–CoA) with loss of CO_2 (C_1): Phe–C_3 + 3 C_3 → Phe–C_2–Phe + 4C_1). A further C–C link between the phenyl rings yields the three fused benzene rings of phenanthrene (the non-linear isomer of the linear anthracene, Phe|Phe|Phe). Stilbene reduction yields bisbenzyls (Phe–CH_2–CH_2–Phe). Stilbenoid compounds can be modified by reduction and by hydroxyl, methoxy, isoprenyl and glycosyl ring substituents. Stilbenes are often found as antifungal agents in wood.

Simple stilbenes (Phe–CH=CH–Phe) include the *Vitis vinifera* (grape) (Vitaceae) cytotoxic resveratrol (4,3′,5′-trihydroxystilbene), the mitochondrial electron transport inhibitor oxyresveratrol (3,5,2′,4′-tetrahydroxystilbene) and the protein kinase inhibitor piceatannol (3,4,3′,5′-tetrahydroxystilbene), all these compounds having antifungal activity. The isoprenylated stilbene chlorophorin (4-geranyl-3,5,2′,4′-tetrahydroxystilbene) is an antioxidant free radical scavenger (AO/FRS).

Bisbenzyl (Phe–CH_2–CH_2–Phe) compounds include dihydroresveratrol (4,3′,5′-trihydroxybisbenzyl) and the allergenic benzopyranone hydrangenol from *Hydrangea macrophylla* (Saxifragaceae).

Phenanthrenes (angular Phe|Phe|Phe) include the antifungal methoxyphenanthrenes batatasin I and isobatatasin I from bulbs of *Dioscorea* species (Dioscoraceae). The pyranophenanthrenes have a tetracyclic structure (involving linkage of the outer phenanthrene rings with an –O–CH_2– group), examples including the spasmolytic compounds coelogin and flavidin from *Coelogyne* species (Orchidaceae).

vii. Anthochlors (chalcones and aurones), anthocyanidins and anthocyanins. Anthochlors (chalcones and aurones), anthocyanidins and anthocyanins provide colour to flowers that is required for attracting pollinating herbivores. The anthochlors are yellow but the anthocyanins (and the corresponding aglycone anthocyanidins) have colours ranging from blue to red.

Chalcones. The parent compound is chalcone (1,3-diphenyl-2-propen-1-one or benzylideneacetophenone; Phe–CH=CH–CO–Phe), the ring numbering being 1–6 (benzylidene phenyl) and 1′–6′ (acetophenone phenyl). Chalcone variants derive from hydroxy, prenyl (isopentenyl) and glycosyl substituents. Phenols are weak acids and as such can act as "protonophores" to increase the proton (H^+) permeability of the mitochondrial inner membrane and hence act as "uncoupling" inhibitors of the key ATP-providing process of oxidative phosphorylation. Butein (2′,4′,3,4-tetrahydroxychalcone), isoliquiritigenin (2′,4′,4-trihydroxychalcone) and okanin (2′,3′,4′,3,4-pentahydroxychalcone) are uncouplers of oxidative phosphorylation. Various chalcones inhibit the following particular enzymes (in parentheses): abyssinone VI (3,5-isoprenyl-2′,3′,4-trihydroxychalcone) (steroid aromatase); buteine (receptor tyrosine kinase and NADH and succinate dehydrogenases); liquiritigenin and isoliquiritigenin (monoamine oxidase); and chalconaringenin (2′,4′,6′,4-tetrahydroxychalcone) (iodothyronine deiodinase).

Dihydrochalcones. The parent compound is dihydrochalcone (1,3-diphenylpropan-2-one). Phloretin (4,2′,4′,6′-tetrahydroxydihydrochalcone) is an uncoupler and an inhibitor of iodothyronine deiodinase and protein kinase. Phloridzin (phloretin 2′-*O*-glucoside) is a bitter tastant and an inhibitor of glucose transport. Odoratol (α-hydroxy-4,4′-dimethoxy-6′-hydroxydihydrochalcone) is a *Lathyrus odoratus* (sweet pea) (Fabaceae) phytoalexin. Various methylated dihydrochalcones including loureirins B and D from *Dracaena loureiri* (Agavaceae) are oestrogen receptor agonists.

Aurones (Phe|C4O(=O)=CH–Phe). Aurones (2-benzylidenebenzofuranones) derive from oxidation and cyclization of chalcone precursors to yield the corresponding benzofuranone (benzene fused with a five-membered furanone ring): Phenyl–CO– CH=CH–Phenyl + O_2 → Benzofuranone = CH–Phenyl. Various aurones inhibit iodothyronine deiodinase, namely (numbering 1–9 in the bicyclic benzofuranone and 1′–6′ in the benzylidene phenyl) aureusidin (4,6,3′,4′-tetrahydroxyaurone), bracteatin (4,6,3′,4′,5′- pentahydroxyaurone), maritimetin (6,7,3′,4′-tetrahydroxyaurone) and sulfuretin (6,3′,4′-trihydroxyaurone).

Anthocyanins and anthocyanidins. Anthocyanidins are the aglycones of the corresponding anthocyanins, the parent compound being 2-phenylbenzopyrylium (flavylium) (Phe|pyrylium$^+$–Phe). The benzopyrylium moiety is benzene fused with an unsaturated six-membered pyrylium ring containing five Cs and a positively charged O. Cyanidin (ring numbering 1–10 in the benzopyrylium ring and 1′–6′ in the phenyl ring) is 3,5,7,3′,4′-pentahydroxyflavylium and is very widespread, particularly as the anthocyanin cyanidin 3-*O*-glucoside. Other anthocyanidins include apigeninidin, delphinidin, hirsutidin, luteolinidin, malvidin, pelargonidin, peonidin and petunidin, the structural variations arising from differing patterns of hydroxy and methoxy substitution (and thence of differing glycosylation in the corresponding anthocyanins).

Cyanidin inhibits epidermal growth factor receptor tyrosine kinase (EGF-RTK), α-glycosidase and COX-1 and COX-2. Delphinidin (3,5,7,3′,4′,5′-hexahydroxyflavylium) also inhibits EGF-RTK. Anthocyanidins and anthocyanins can be anti-inflammatory antioxidants by acting as free radical scavengers. Thus, nasunin (delphinidin-3-(*p*-coumaroylrutinoside)-5-glucoside) scavenges OH (hydroxyl), O_2^- (superoxide) and lipid peroxyl radicals and inhibits lipid peroxidation.

viii. Benzofurans. The parent compound benzofuran (Phe|furan) involves a fused benzene (unsaturated C6 ring) and furan (unsaturated five-membered ring including four Cs and one O). In addition to simple benzofurans there are dibenzofurans (Phe|furan|Phe) in which the furan ring is fused with two benzenes to make a tricyclic nucleus. The simple benzofurans and dibenzofurans are generally toxic with antimicrobial and notably antifungal activity.

Simple benzofurans (Phe|furan) involve benzofuran variously having acetoxy, hydroxy, methoxy or more complex substituents on the benzo moiety and typically a 2-phenyl or 2-(2-propenyl) substituent on the furan moiety. Asteraceae benzofurans with a 2-propenyl substituent include toxol and toxyl angelate (from *Haplopappus heterophyllus*) and dehydrotremetone and tremetone (from *Eupatorium* (snakeroot) species); ingestion of these plants by cows gives rise to "milk sickness". Snakeroot "milk sickness" involves blockage of glucose-supplying gluconeogenesis (see Chapter 2) and was responsible for the death of Abraham Lincoln's mother Nancy. The *Penicillium*-derived tricyclic chlorobenzofuran metabolite griseofulvin (Phe|C4O(=O)·C6) is an antifungal drug that interferes with microtubule tubulin and is used against tinea capitis (cradle cap) in children. The 2-phenylbenzofurans include *Morus* species (mulberry) (Moraceae) albanol A (mulberrofuran G) (Phe|furan-polycyclic) and mulberrofuran A (Phe|furan-Phe-isoprenyl) (COX inhibitors); lithospermic acid (aryl-Phe|furan-Phe) (from Boraginaceae) (a free radical scavenger and inhibitor of prolyl hydroxylase and collagen hydroxylation); and *Morus alba* (mulberry) (Moraceae) antifungal phytoalexins moracins A–Z and chalcomoracin (Phe|furan-Phe) (superoxide scavengers).

Dibenzofurans (Phe|furan|Phe) include various fungal infection-induced plant antifungal compounds (phytoalexins) such as the Rosaceae-derived cotonefuran (from *Cotoneaster lactea*) and α-pyrofurans (from *Pyrus communis*). Usnic acid from lichens (notably *Usnea* species) is anti-mycobacterial, anti-mitotic, an uncoupler and a potent inhibitor of plant protoporphyrinogen synthetase and 4-hydroxphenylpyruvate dioxygenase.

ix. Chromones and chromenes. Chromones and chromenes involve a benzene ring fused with pyran (an unsaturated six-membered ring containing five Cs and one O). In chromenes (Phe|α-pyran), the heterocyclic ring is an unsaturated α-pyran (1,2-pyran) moiety (C5O, two asymmetric double bonds) and in chromones (Phe|γ-pyran-4-one), the O-containing ring is an unsaturated γ-pyran-4-one (1,4-pyran-4-one) moiety (C5, O, two symmetrically placed double bonds and a keto O). The flavonoids (2-phenylchromones), isoflavonoids (3-phenylchromones) and xanthones (Phe|γ-pyran-4-one|Phe) will be dealt with in Sections xi–xvi. The chromones and chromenes are variously condensed with other ring systems and substituted with hydroxy, methoxy, alkyl and aryl groups. A number of these compounds are variously antimicrobial and cytotoxic.

Simple chromones (Phe|γ-pyran-4-one) include the glucoside biflorin (a cAMP phosphodiesterase inhibitor and free radical scavenger) and the 2-phenoxychromone capillarisin (an aldose reductase inhibitor) as well as a number of variously cytotoxic and antimicrobial compounds.

Furanochromones (furan|Phe|γ-pyran-4-one) have a furan ring fused with the benzene moiety of the chromone. Khellin, the related khellol glucoside and visnagin (dehydrokhellin) derive from seeds of *Ammi visnaga* (Apiaceae), both khellin and visnagin being phototoxic and vasorelaxant cAMP phosphodiesterase inhibitors.

Pyranochromones (α-pyran| Phe|γ-pyran-4-one) have an α-pyran ring fused with the benzene ring of the chromone and include the *Cneorum* species (Cneoraceae) antibacterial and cytotoxic compounds pulverochromenol (having an α-pyran fused with

a benzochromone) and spatheliabischromene (α-pyran|Phe(α-pyran)|γ-pyran-4-one) (having two α-pyran rings condensed with a benzochromone).

Chromenes (Phe|α-pyran) include encecalin (a phototoxic antimicrobial from various Asteraceae) and the phloroglucinol derivative mallotochromene (cytotoxic and an HIV-1 reverse transcriptase inhibitor). Precocene 1 (7-methoxy-2,2-dimethylchromene) and precocene 2 (6,7-dimethoxy-2,2-dimethylchromene) produced by *Ageratum* species (Asteraceae) inhibit the production of insect juvenile hormone (JH) as a result of "suicidal" conversion of these "pro-toxins" to cytotoxic derivatives by the JH-producing insect cells.

x. Coumarins. The parent compound coumarin (benzopyran-2-one; 1,2-benzopyrone) (Phe|pyran-2-one) involves the fusion of benzene (Phe–H) and pyran-2-one (C5, O, two double bonds and a 2-keto; unsaturated C5OL). Coumarin is responsible for the smell of newly cut grass. In addition to simple coumarins, there are furanocoumarins (in which a five-membered furan ring is fused with the benzo moiety of coumarin in either an angular or linear fashion) and pyranocoumarins (in which a six-membered pyran ring is fused with the benzo moiety of coumarin in either an angular or linear fashion). These coumarins are variously substituted with hydroxy, methoxy, methyl, acetoxy, glycosyl and other groups.

Simple coumarins (Phe|pyran-2-one) include coumarin and a variety of antibacterial derivatives including ammoresinol (7-hydroxy-3-geranylgeranylcoumarin), daphnetin (7,8-dihydroxycoumarin), esculetin (6,7-dihydroxycoumarin), esculin (esculetin 6-*O*-glucoside), herniarin (7-methoxycoumarin) and umbelliferone (7-hydroxycoumarin). Fraxetin and 4-methyldaphnetin (6,7-dimethoxycoumarin) are antioxidant ROS scavengers and 5-LOX inhibitors. Esculetin, 7-hydroxy-4-methylcoumarin and umbelliferone are xanthine oxidase inhibitors. Coumarins inhibiting other enzymes (enzyme target in parentheses) include: osthol (7-methoxy-8-isopentenylcoumarin) (cAMP phosphodiesterase) and the antioxidant scoparone (6,7-dimethoxycoumarin) (tyrosine kinase). Dicoumarol (3,3′-methylenebis (4-hydroxycoumarin); dicumarol) is a haemorrhagic anticoagulant from *Melilotus alba* (sweet clover) (Fabaceae) hay. Dicoumarol acts by being an antagonist of vitamin K_1 (a quinone that is required for prothrombin carboxylation and consequent Ca^{2+} binding and activation leading to blood clotting).

Furanocoumarins (furan|Phe|pyran-2-one) include a variety of angular and linear furanocoumarins as exemplified by the respective parent compounds isopsoralen and psoralen. Many furanocoumarins and the parent compounds themselves bind to DNA and form covalent adducts with DNA in a light-activated process involving alkylation of pyrimidine bases. Such photoactivatable compounds include the angular furanocoumarin isopsoralen (angelicin) and the linear furanocoumarins psoralen, bergapten (5-methoxypsoralen), 4,5′,8-trimethoxypsoralen and xanthotoxin (8-methoxypsoralen). Xanthotoxol (8-hydroxypsoralen) is an antioxidant ROS scavenger. A variety of angular and linear furanocoumarins inhibit inducible NO synthase expression, including isopsoralen, pimpinellin, sphondin, byakangelicol, oxypeucedanin, cnidilin and xanthotoxin. Isopsoralen and psoralen inhibit both monoamine oxidases A and B.

Pyranocoumarins (C5O|Phe|pyran-2-one) include a variety of angular and linear compounds. A number of angular pyranocoumarins are spasmolytic and vasodilatory, notably the Ca^{2+} channel blocker visnadin. The inophyllums B and P from *Calophyllum ionophyllum* (Guttiferae) are inhibitors of HIV-1 reverse transcriptase.

xi. Flavones and flavonols. Flavones, biflavones and flavone-3-ols (flavonols) are derivatives of the parent 2-phenylchromone, flavone (2-phenyl-1-benzopyran-4-one;

2-phenyl-γ-benzopyrone), the ring numbering system being 1 (pyrone ring O), 4 (pyrone ring keto C), 5–8 (benzo ring Cs) and 1′–6′ (2-phenyl ring Cs). Flavones and flavonols (3-hydroxyflavones) contribute to petal colour (especially as perceived by insects) together with anthocyanins and also function in UV protection and defence against herbivores.

Flavones. Flavone structural variation derives from hydroxylation, *O*-methylation and *O*-glycosylation. In addition, there can be C6- and C8-linked *C*-glycosides, isoprenyl (isopentenyl, C_5) substituents and C–C or C–O–C links to form biflavones. Methylation of the phenolic OHs decreases polarity to permit an external location such as in the waxy leaf or fruit surface.

Flavones with a widespread occurrence include apigenin (5,7,4′-trihydroxyflavone), luteolin (5,7,3′,4′-tetrahydroxyflavone) and the corresponding derivatives apigenin 7,4′-dimethylether, apigenin 7-*O*-glucoside (cosmosiin), apigenin 8-*C*-glucoside (vitexin), apigenin 6,8-*C*-diglucoside (vicenin-2), luteolin 7-*O*-glucoside, luteolin 6-*C*-glucoside (isoorientin), luteolin 6-*C*-glucoside (orientin) and luteolin 6,8-*C*-diglucoside (lucenin-2).

Some bioactive flavones include: aldose reductase inhibitors (apigenin 4′-methyl ether (acacetin), apigenin 7-*O*-apioside (apiin), 5,7-dihydroxyflavone (chrysin) and luteolin); anti-inflammatory 5-LOX inhibitors (5,6,7-trihydroxyflavone (baicalein), 5,6,3′,4′-tetrahydroxy 7-methoxyflavone (pedalitin), 5,3′,4′-trihydroxy 6,7-dimethoxyflavone (cirsiliol, 6-*O*-methylpedalitin) and flavone); a COX inhibitor (flavone); iodothyronine deiodinase inhibitors (acacetin, chrysin and luteolin); a NADH and succinate dehydrogenase inhibitor (luteolin); millet-derived, goitrogenic inhibitors of thyroid peroxidase (flavone *C*-glycosides orientin and vitexin); and protein kinase inhibitors (acacetin, apigenin, baicalein, flavone, luteolin, 5,7,3′,4′,5′-pentahydroxyflavone (tricetin) and tricetin 3′,4′,5′-trimethyl ether).

A variety of flavones are anti-inflammatory (apigenin, apigenin 7,4′-dimethylether, baicalein, 8-hydroxyluteolin and luteolin); insect feeding attractants (notably the C-glycosides carlinoside, isoorientin, isoscoparin, neocarlinoside, schaftoside and neoschaftoside); oestrogenic (wogonin); and oviposition stimulants (luteolin 7-(6′′-malonylglucoside) and vicenin-2).

Biflavones. A number of biflavones are formed via C–C linkages, notably the cAMP phosphodiesterase (cAMP PDE) inhibitory biapigenins agathisflavone (6,8′′-biapigenin), amentoflavone (3′,8′′-biapigenin), cupressiflavone (8,8′′-biapigenin) and robustaflavone (3′,6′′-biapigenin). The 4′-C-O-6′′-C-linked biapigenin hinokiflavone is also a cAMP PDE inhibitor.

Flavonols. The most common flavonols (3-hydroxyflavones) include kaempferol (3,5,7,4′-tetrahydroxyflavone), quercetin (3,5,7,3′,4′-pentahydroxyflavone), myricetin (3,5,7,3′,4′,5′-hexahydroxyflavone), quercitrin (quercetin 3-*O*-rhamnoside), isoquercitrin (quercetin 3-*O*-glucoside), isorhamnetin (quercetin 3′-methyl ether) and rutin (quercetin 3-rutinoside). A large number of other flavonols variously have hydroxy, methoxy, isoprenyl, *O*-glycoside and other substituents.

Some bioactive flavonols include: aldose reductase inhibitors (axillarin (5,7,3′,4′-tetrahydroxy-6-methoxyflavone), 2,3-dihydroquercetin (taxifolin), 6-hydroxykaempferol (galetin), hyperin (quercetin 3-*O*-galactoside), isoquercetrin, morin (3,5,7,2′,4′-pentahydroxyflavone), quercetin, quercitrin and rutin); anti-inflammatory 5-LOX inhibitors (fisetin (3,7,3′,4′-tetrahydroxyflavone), kaempferol, morin, myricetin, quercetin and rutin); a COX inhibitor (galangin (3,5,7-trihydroxyflavone)); iodothyronine deiodinase inhibitors (fisetin, kaempferol and morin); NADH and succinate dehydrogenase inhibitors (fisetin and myricetin); and protein kinase inhibitors (fisetin, galangin, isorhamnetin, kaempferide

(kaempferol 4′-methyl ether), morin, quercetagetin (6-hydroxyquercetin), quercetin, quercitrin and rutin).

Flavonols are variously good ROS scavengers (e.g. kaempferol and quercetin). Particular flavonols are insect feeding attractants or stimulants (quercetin, quercetin 7-*O*-glucoside (quercimeritrin), isoquercitrin, quercitrin and rutin).

xii. Dihydroflavonoids. Dihydroflavonoids are flavonoids in which the 2,3 double bond of the chromene ring has been reduced. Such compounds include the flavanones (2,3-dihydroflavones such as naringenin or 2,3-dihydroapigenin) and 2,3-dihydroflavonols (such as taxifolin or 2,3-dihydroquercetin). Related compounds include flavan-3-ols, 2,3-dihydrochalcones (1,3-diphenylpropan-1-ones) and flavans. Further, more complex flavan-based compounds include the biflavans and biflavanones. The basic skeleton in each case can be modified with hydroxyl, methoxy, glycosyl, isopentenyl (isoprenyl) and other groups. The condensed tannins derive from C–C-linkage of flavan-3-ols such as afzelechin, (+)-catechin and (−)-epicatechin and are considered separately in Section xiii.

Flavanones. Widespread flavanones (2,3-dihydroflavones) include the 2,3-dihydroflavones eriodictyol (5,7,3′,4′-tetrahydroxyflavanone; 2,3-dihydroluteolin), naringenin (5,7,4′-trihydroxyflavanone; 2,3-dihydroapigenin) and pinocembrin (5,7-dihydroxyflavanone; 2,3-dihydrochrysin). Eriodictyol and eriodictyol 4′-methyl ether (hesperetin) induce *Rhizobium* nodulation gene expression; hesperetin and eriodictyol 3′-methyl ether are insect feeding deterrents; and several hesperetin glycosides are oviposition stimulants. The 7-*O*-neohesperidosides of naringenin, eriodictyol and hesperetin are bitter tasting. The flavanolignan flavanone derivatives silandrin, silybin and silychristin from *Silibum marianum* (Asteraceae) are antihepatotoxic. Sanggenon C and sanggenon D bind to the phorbol ester binding site on protein kinase C (PKC).

Dihydroflavonols. Widely distributed 2,3-dihydroflavon-3-ols include the antioxidant 2,3-dihydroflavonols aromadendrin (3,5,7,4′-tetrahydroxyflavanone; 2,3-dihydrokaempferol), ampelopsin (3,5,7,3′,4′,5′-hexahydroxyflavanone; 2,3-dihydromyricetin), fustin (3,7,3′,4′-tetrahydroxyflavanone; 2,3-dihydrofisetin) and taxifolin (3,5,7,3′,4′-pentahydroxyflavanone; 2,3-dihydroquercetin). Some flavanols are sweet-tasting, notably 6-methoxyaromadendrin 3-*O*-aceate, 6-methoxytaxifolin and taxifolin 3-*O*-acetate. Taxifolin and fustin inhibit NADH and succinate dehydrogenases and taxifolin inhibits 5-LOX.

Flavans. A number of flavans are variously antimicrobial or dermatitic. The isoprenyl flavans kazinols A, Q and R from *Broussonetia* species (Moraceae) are cytotoxic.

Biflavanoids. Biflavanoids are linked by C–C bonds. Biflavanones include isochamaejasmin (3,3′-binaringenin), kolaflavanone (3′,8′′-binaringenin) and the aldose reductase inhibitor manniflavanone (3′,8′-bieriodictyol). The *Camellia sinensis* (tea) (Theaceae) biflavanol theasinensin A (6′,6′′-bi(5,7,3′,4′,5′-pentahydroxyflavan 3-*O*-galloyl ester), a theaflavin precursor, is apoptotic, cancer chemopreventative and an inhibitor of squalene epoxidase.

Flavan-3-ols. Flavan-3-ols include afzelechin (3,5,7,4′-tetrahydroxyflavan), (+)-catechin, (−)-epicatechin, (–)-epicatechin, (+)-catechin ((+)-3,5,7,3′,4′-pentahydroxyflavan), (−)-epicatechin ((+)-3,5,7,3′,4′-pentahydroxy-flavan), epigallocatechin (EGC; 5′-hydroxyepicatechin), epicatechin-3-*O*-gallate (ECG), (−)-epigallocatechin-3-*O*-gallate (EGCG) and (−)-gallocatechin-3-*O*-gallate (GCG). These polyphenols are variously antioxidant ROS scavengers and the monomeric units of condensed tannins (see Section xiii). Enzymes variously inhibited by these flavan-3-ols include squalene epoxidase, protein kinase, aldose reductase, COX and 5-LOX.

xiii. Tannins. The tannins are widely distributed defensive compounds in plants and fall into two major categories, the condensed tannins and the hydrolysable tannins. The condensed tannins essentially derive from the polymerization of the flavan-3-ols (+)-catechin, (−)-epicatechin and their derivatives via C–C links, thus generating flavan oligomers (flavolans). The hydrolysable tannins, defensive compounds confined to dicots, involve a glucose esterified to gallic acid (gallotannins) or ellagic acid-derived hexahydroxydiphenic acid (ellagitannins). The multiplicity of phenolic hydroxy groups enables tannins to hydrogen bond extensively with protein peptide links (–CO–NH–) and protonatable R groups, this property being the basis of "tanning" animal skins to generate leather. Tannins have antioxidant activity as ROS scavengers. Various condensed and hydrolysable tannins are cytotoxic with antitumour activity. Notwithstanding the general avidity of tannins for polypeptides, there are many examples of specificity in tannin-protein interactions.

Condensed tannins derive from the polymerization of flavan-3-ols such as (+)-catechin (C; (+)-3,5,7,3′,4′-pentahydroxyflavan), (−)-epicatechin (E; (+)-3,5,7,3′,4′-pentahydroxyflavan), EGC; 5′-hydroxyepicatechin) and ECG, this typically involving 4→8 and 6→8 C–C links. The condensed tannins are classified on the basis of the mauve- to red-coloured monomeric anthocyanidin products produced by heating the tannin in acid, for example, as indicated in parentheses as follows: procyanidins (product cyanidin, 3,5,7,3′,4′-pentahydroxyflavylium), prodelphinidins (delphinidin, 3,5,7,3′,4′,5′-hexahydroxyflavylium), propelargonidins (pelargonidin, 3,5,7,4′-tetrahydroxyflavylium) and proluteolinidins (luteolinidin, 5,7,3′,4′-tetrahydroxyflavylium).

A variety of condensed tannins are antagonists of particular hormone receptors or inhibitors of particular enzymes, most notably protein kinases.

Hydrolysable tannins involve a glucose esterified to gallic acid (gallotannins) or ellagic acid-derived hexahydroxydiphenic acid (ellagitannins). These complex structures can be described simply in essence by representing galloyl (Phe), galloyl variants and hexahydroxydiphenoyl (Phe–Phe) as G, G′ and H, respectively, noting that glucose has five hydroxy groups that can potentially form ester linkages (X–CO–O–Y) with these acids, diesters being formed with H and monoesters with G and G′. Thus, pentagalloyl-β-D-glucose can be represented as G_5-glucose and casuarinin as H_2-glucose-G. The more complex coriariin A can be represented as HG_2–glucose–G′–G′–glucose–HG.

A variety of hydrolysable tannins have been shown to act as hormone receptor antagonists or inhibitors of particular enzymes. The inhibition of protein kinases by various hydrolysable tannins becomes more potent as the number of phenolic groups increases.

xiv. Isoflavonoids. Isoflavonoids have a common structural element of a 3-phenyl chromane which is thence modified by oxidation and substitution to yield the different classes within this group, namely the isoflavones, isoflavanones, isoflavans, pterocarpans, rotenoids and coumestans. The isoflavonoids are very largely confined to the legumes (Fabaceae) and many such compounds have antifungal activity. Many isoflavonoids are phytoalexins, that is, antifungal compounds synthesized in response to fungal infection of the plant. The structural and functional complexity of the isoflavonoids is briefly sketched below.

Isoflavones are derivatives of the parent compound isoflavone (3-phenylchromone; 3-phenylbenzopyran-4-one). The dietary isoflavone phytoestrogens (notably from soya bean) that bind to the oestrogen receptor are the best known, namely: daidzein (7,4′-dihydroxyisoflavone), genistein (5,7,4′-trihydroxyisoflavone) and glycitein (7-hydroxy-6-methoxyisoflavone) and their respective "pro-phytoestrogen" 7-*O*-glucoside precursors daidzin,

glycitin and genistin, respectively, that are inactive or poor as ligands for the oestrogen receptor but which are hydrolysed to the active aglycones after ingestion. Formonetin (daidzein 4′-methylether) is also a "pro-phytoestrogen". The further postprandial 2,3-dihydro products dihydrodaidzein (equol), dihydroglycitein and dihydrogenistein are also active as oestrogen receptor ligands. Isoflavone *C*-glycosides include the anti-atherosclerotic genistein 8-*C*-glycoside, and daidzein 8-*C*-glycoside (puerarin). The isoprenylated isoflavones licoisoflavone A (5,7,2′,4′-tetrahydroxy-3′-isopentenylisoflavone), luteone (5,7,2′,4′-tetrahydroxy-6-isopentenylisoflavone) and wighteone (5,7,4′-trihydroxy-6-isopentenylisoflavone) are phytoalexins.

Isoflavanones are 2,3-dihydroisoflavones and a number of such compounds are antifungal phytoalexins. Thus, kievitone (2′,4′,5,7-tetrahydroxy-8-isopentenylflavanone) and the related compounds cyclokievitone, dalbergioidin and 5-deoxykievitone are antifungal phytoalexins induced in *Phaseolus vulgaris* (bean) and other Fabaceae species by fungal infection.

Isoflavans are analogues of the isoflavanones that lack the 4-keto, that is, they are 3-phenylchromanes. The simple isoflavans sativan and vestitol from *Lotus* species (Fabaceae) are antifungal phytoalexins. The pyranoisoflavans glabridin and hispaglabridin are antimicrobials from *Glycyrrhiza glabra* (liquorice) (Fabaceae) roots.

xv. Polycyclic isoflavan-related compounds (neoflavonoids). Ptercocarpans, pterocarpenes, coumestans and rotenoids are polycyclic compounds related to isoflavans and coumarins through the formation of an additional fused furan or pyran ring as a result of introduction of an ether (C–O–C) link between the chromane ring and a 3-phenyl substituent.

Pterocarpans and pterocarpenes. Pterocarpans (Phe|C5O|C4O|Phe) are isoflavonoids involving a fusion of chromane and benzofuran rings, that is, they are isoflavans in which a furan ring is formed through generation of an ether link between the chromane and the 3-phenyl. Pterocarpenes are 2,3-dehydropterocarpanes. The phytoalexins anhydroglycinol and phaseollidin are examples of a pterocarpene and a pterocarpan, respectively. Glyceollins I and II (C5O|Phe|C5O|C4O|Phe) from *Glycine* species (Fabaceae) and phaseolin (Phe|C5O|C4O|Phe|C5O) from *Phaseolus* species are pyranopterocarpan phytoalexins.

Coumestans are benzofuranocoumarins. Coumestrol (Phe|C5OL|furan|Phe; coumarin|furan|Phe) is a phytoalexin in *Glycine max* and *Phaseolus* species (Fabaceae). Coumestrol is also oestrogenic as is the pyranocoumestan phytoalexin sojagol (coumarin|furan|Phe|C5O) from *Glycine max*.

Rotenoids have a basic structural element involving fused chromone and chromane rings. The best-known rotenoid is the furanorotenoid rotenone (C5O|Phe|pyran-4-one|C5O|Phe) from *Derris* and *Lonchocarpus* species (Fabaceae), a potent inhibitor of the mitochondrial electron transport chain NADH dehydrogenase (complex I).

xvi. Xanthones. Xanthones have a basic parent tricyclic ring structure, namely that of xanthone (dibenzo-γ-pyrone) (Phe|(4-keto)C4O|Phe). This structure arises from phenylpropanoid (Phe–C_3) and malonyl-coenzyme A (C_3–CoA, $^{-}O_2CCH_2CO$–S–X) precursors (Phe–C_3 + $2C_3O_3$ → Phe–CO_2–Phe + $2CO_2$). Xanthones are grouped below into simple xanthones, prenylated xanthones, xanthone-*O*-glycosides, xanthone-*C*-glycosides and pyranoxanthones. In addition, these compounds differ in hydroxy, methoxy, glucosyl, methyl and alkyl substituents.

Simple xanthones include various mutagenic and antibacterial compounds such as bellidifolin (3-methoxy-1,5,8-trihydroxyxanthone). A number of simple xanthones are inhibitors of monamine oxidase A (bellidifolin, demethylbellidifolin, gentiacaulin, isogentisin and swerchirin), protein kinase (norathyriol) and of xanthine oxidase (athyriol, isoathyriol and norathyriol).

Prenylated xanthones include α-mangostin (2,8-di-isoprenyl-1,3,6-trihydroxy-7-methoxyxanthone) and γ-mangostin from the fruit of *Garcinia mangostana* (Guttiferae). α-Mangostin inhibits various protein kinases, Ca^{2+} ATPase and HIV-1 protease and binds to the oestrogen receptor and the histamine receptor. γ-Mangostin inhibits HIV-1 protease and various protein kinases.

Glycosylated xanthones include xanthone-*O*-glycosides such as the antibacterial bellidifolin 8-*O*-glucose (swertianolin) and the widely distributed xanthone-*C*-glycoside mangiferin (1,3,6,7-tetrahydroxyxanthone 2-*C*-glucoside).

Pyranoxanthones have a pyran ring fused with a xanthone, an example being the antimicrobial isomangostin (C5O|Phe|(4-keto)C4O|Phe) that is structurally related to the prenylated xanthone α-mangostin (2,8-di-isoprenyl-1,3,6-trihydroxy-7-methoxyxanthone) through cyclization involving the 1-hydroxy and 2-isoprenyl. The furanoxanthone psorospermin (C4O|Phe|(4-keto)C4O|Phe) derives from cyclizing involving adjacent C5 side chain and hydroxy substituents yielding a fused furan ring.

1.8 Plant terpenes

Terpenes are composed of isoprenyl (C_5) units and are conveniently grouped as monoterpenes (skeletal basis $C_{10} = 2\times C_5$), sesquiterpenes ($C_{15} = 3\times C_5$), diterpenes ($C_{20} = 4\times C_5$), triterpenes ($C_{30} = 6\times C_5$) and tetraterpenes ($C_{40} = 8\times C_5$). The structures of some representative terpenes are shown in the Appendix (Section 3). Terpenes ultimately derive biosynthetically from acetate (C_2) via the activated acetyl thioester (CH_3–CO–S–X) acetyl-coenzyme A (acetylCoA; CH_3–CO–S–CoA) as outlined below (enzymes catalysing key steps being indicated in parentheses).

Acetate (C_2) is generated as a result of primary "catabolic" "energy metabolism" involving glucose (C_6) oxidation to pyruvate (C_3) (by the enzymes of the ATP-yielding glycolysis pathway) and subsequent pyruvate decarboxylation (loss of CO_2, C_1) and oxidation to yield acetylCoA and reduced coenzymes. AcetylCoA (C_2) condenses with oxaloacetate (C_4) to yield the tricarboxylic acid citrate (C_6) which is ultimately oxidized via a succession of C_6, C_5 and C_4 intermediates to yield oxaloacetate (C_4), CO_2 (C_1), ATP and reduced coenzymes (catalysed by the enzymes of the tricarboxylic acid (or citric acid, Krebs) cycle). The reduced coenzymes (NADH and $FMNH_2$) are oxidized via the mitochondrial respiratory chain (oxygen being the ultimate electron acceptor), this being coupled to the formation of the "energy-rich" "cellular energy currency" ATP (catalysed by the ATP synthase (F_0–F_1) complex of oxidative phosphorylation).

Excess acetate (C_2) can be converted to the "mobile" ketone body energy source acetoacetate (C_4) and thence its reduced form hydroxybutyrate (C_4) for transport throughout the body. Excess acetate can be carboxylated (via acetylCoA carboxylase) to form malonylCoA (C_3), the donor for further C_2 additions (with CO_2 elimination) in the "anabolic" synthesis of long chain fatty acids. Fatty acids are components of the phospholipids of cellular membranes and are also stored as triacylglycerols (triglycerides) for subsequent hydrolysis and "catabolic" fatty acid oxidation to yield reduced coenzymes and thence ATP (see Chapter 2).

AcetylCoA (C_2) can also react with acetoacetylCoA (C_4) to generate hydroxymethylglutarylCoA (HMGCoA) (C_6) and thence the isoprenoid precursor mevalonate (C_6). Mevalonate (C_6) ultimately yields the key C_5 isoprenoids isopentenylpyrophosphate ($CH_3C(=CH_2)$–CH_2–CH_2–O–PO_3–PO_3) (IP-PP) and dimethylallylpyrophosphate (CH_3–$C(CH_3)$=CH–CH_2–O–PO_3–PO_3) (DMA-PP), the immediate precursors of cholesterol and

steroid hormones in animals and of a wide range of terpenes in plants. These reactions are summarized below:

$CH_3CO{-}S{-}CoA\ (C_2) + CH_3COCH_2CO{-}S{-}CoA\ (C_4)$
$\rightarrow {}^{-}OOC{-}CH_2{-}C(CH_3,OH){-}CH_2{-}CO{-}S{-}CoA$ (HMGCoA) (C_6)
$\rightarrow {}^{-}OOC{-}CH_2{-}C(CH_3,OH){-}CH_2{-}CH_2OH$ (mevalonate) (C_6)
[via HMGCoA reductase + NADPH]
$\rightarrow$ 3-phospho-5-pyrophosphomevalonate (C_6)
$\rightarrow CO_2\ (C_1) + PO_4^{2-}$ (inorganic phosphate, P_i) +
$CH_3C(=CH_2){-}CH_2{-}CH_2{-}O{-}PO_3{-}PO_3$ (IP-PP) (C_5)
$\rightarrow (CH_3{-}C(CH_3)=CH{-}CH_2{-}O{-}PO_3{-}PO_3)$ (DMA-PP).

IP-PP and DMA-PP can yield volatile C_5 hemiterpenes. At the other extreme, extensive polymerization of the C_5-pyrophosphates (with release of pyrophosphate, PP_i) yields the formation of the plant latex polymers such as *cis*-polyisoprenes (rubber) and *trans*-polyisoprenes (gutta-percha). In between these extremes, a variety of monoterpenes, sesquiterpenes, triterpenes and C_{40} carotenes derive from these C_5-pyrophosphate precursors.

Head to tail condensation of IP-PP (C_5) and DMA-PP (C_5) with release of PP_i forms geranylpyrophosphate: $CH_3{-}C(CH_3)=CH{-}CH_2{-}CH_2{-}C(CH_3)=CH{-}CH_2{-}O{-}PO_3{-}PO_3$, that is, $H(CH_2{-}C(CH_3)=CH{-}CH_2)_2{-}O{-}PO_3{-}PO_3$ (C_{10}–PP), the starting point for plant monoterpenes. Further, head-to-tail reaction of geranylpyrophosphate (C_{10}–PP) with isopentenylpyrophosphate (C_5–PP) yields farnesylpyrophosphate $H(CH_2{-}C(CH_3)=CH{-}CH_2)_3{-}O{-}PO_3{-}PO_3$ (C_{15}–PP), the parent of plant sesquiterpenes. Head to tail condensation of farnesylpyrophosphate (C_{15}–PP) with IP-PP (C_5–PP) yields geranylgeranylpyrophosphate $H(CH_2{-}C(CH_3)=CH{-}CH_2)_4{-}O{-}PO_3{-}PO_3$ (C_{20}–PP), the parent of plant diterpenes.

Representing the PP-end as the "head", head-to-head condensation of two geranylgeranylpyrophosphate (C_{20}–PP) molecules ultimately yields phytoene (C_{40}), that is, if one represents the isoprenylpyrophosphate polarities as IP-PP and PP-PI, one could represent phytoene as $(IP)_4$-$(PI)_4$.

Head-to-head condensation of two farnesylpyrophosphate (C_{15}–PP) molecules yields a C_{13}-cyclopropane (C3)–C_{14} intermediate which is then reduced to yield squalene: $H(CH_2{-}C(CH_3)=CH{-}CH_2)_3{-}(CH_2{-}CH=C(CH_3)CH_2)_3$ (C_{30}), that is, if one represents the isoprene polarities as IP and PI, one could represent squalene as $(IP)_3{-}(PI)_3$. Squalene is subsequently oxidized [via a squalene monooxygenase] to yield squalene 2,3-epoxide which is cyclized to the tetracyclic sterol terpene lanosterol (C_{30}) [via squalene cyclase].

If as above we simply represent alicyclic rings sharing two Cs by a vertical line, then we can represent the basic tetracyclic structure of lanosterol as C6|C6|C6|C5 (noting that there are two double bonds and various alkyl substituents and also a 3-hydroxyl on the first of the alicyclic rings). Many subsequent reactions yield cholesterol, a major triterpene membrane component that modifies the fluidity of animal cell membranes and is a precursor for synthesis of animal bile acids (fat solubilizing amphipathic detergents); plant triterpenes; and steroid hormones such as the corticosteroids cortisol and cortisone, the mineralocorticoid aldosterone and the sex hormones testosterone and 17-β-oestradiol. The structure and bioactivity of the plant terpenes is sketched below.

i. Monoterpenes. The monoterpenes (di-isoprenes) are typically strong smelling oils and part of the so-called "essential oils" of odoriferous plants.

Non-cyclic monoterpenes are unsaturated, pleasant-smelling, C_{10} aliphatic compounds including aldehydes such as citronellal and citral (lemon-scented); the sweet-rose

scented alcohols geraniol and nerol; esters such as geranyl acetate and linalyl acetate (bergamol); and alkenes such as myrcene and β-ocimene.

Monocyclic monoterpenes include the fully saturated menthol (5-methyl-2-isopropyl-cyclohexanol) (C6) (peppermint smell), the fully unsaturated analogue thymol (5-methyl-2-isopropylphenol) (C6) (smell of thyme) and the partially unsaturated α-terpinene (5,6-dihydro-4-isopropyltoluene) (C6) (lemon odour). Variants derive from different degrees of unsaturation and substitution and from different functional groups (e.g. alkyl, hydroxyl, aldehyde, peroxy and keto groups).

Bornane monoterpenes are exemplified by camphene (2,2-dimethyl-3-methylene-bicyclo[2,2,1]heptane), a structure in which two fused cyclopentane rings share three Cs. We can simply represent the camphene skeleton as a cyclohexane with a methylene (–CH_2–) cross-link (C6(–CH_2–)). The keto derivative camphor (camphor smell), the ether eucalyptol (eucalyptus smell) and the simple bornene α-pinene (pine smell) are familiar examples.

Tropolone monoterpenes include the antifungals β- and γ-thujaplicin (4- and 5-isopropyltropolones, respectively, tropolone being 2-hydroxycycloheptatrienone (C7)). The antioxidant β- thujaplicin (hinokitiol) is an inhibitor of 5-, 12- and 15-LOXs.

Thujane monoterpenes are based on the bicyclic (C3|C5) monoterpene thujane and include umbellone (thujan-2-one) and the neuroactives α-thujone and β-thujone (thujan-3-one isomers) that can cause convulsions. Thujones are GABA(A) receptor antagonists and are the active constituents in oil of wormwood from *Artemisia absinthium* (Asteraceae) used in the alcoholic drink absinthe that was eventually banned because of its deleterious neurotoxic effects.

Chrysanthemum carboxylic acid esters. Chrysanthemum monocarboxylic acid (CMC) and dicarboxylic acid (CDC) esters include the toxic cinerins and pyrethrins from *Pyrethrum* (*Chrysanthemum*) *cinerarifolium* (Asteraceae) namely cinerin I (CMC cineralone ester), cinerin II (CDC monomethyl ester cinerolone ester), pyrethrin I (CMC pyrethrolone ester) and pyrethrin II (CDC monomethyl ester pyrethrolone ester). The chrysanthemum carboxylic acids are cyclopropane-based monoterpenes and cineralone and pyrethrolone are cyclopentanone monoterpene alcohols. The pyrethrins (and their insecticidal synthetic derivatives) are toxic to insects through keeping cell membrane voltage-gated Na^+ channels open and thus impairing neurotransmission.

ii. Iridoids. Iridoids are monoterpenes deriving biosynthetically from geranyl-pyrophosphate (C_{10}) and are typically bicyclic hemiacetals (C5|C5OH) or lactones (C5|C5OL). The heterocyclic ring is typically a hemiacetal, ring closure deriving from intramolecular reaction between an aldehyde (–CHO) and another aldehyde or a hydroxyl (–OH) to yield a –CH(OH)–O-linkage in the iridoid (i.e. C5|C5OH). Alternatively, heterocyclic ring closure involves lactone formation involving reaction of a carboxyl (–COOH) with an hydroxyl (–OH) to form an intracyclic ester or lactone linkage (–CO–O–) as in nepetalactone (C5|C5OL). The lactone and hemiacetal rings are denoted below as C*n*OL and C*n*OH, respectively, (where *n* is the number of C atoms in the ring). The hemiacetal hydroxyl can be glycosylated. The hemiacetal structure is unstable and ring opening of the aglycone (e.g. generated by acid hydrolysis) yields an aldehyde that is very reactive (yielding coloured polymeric forms). In the seco-iridoids, the alicyclic C5 ring is opened or expanded by oxygen insertion.

Simple iridoids are volatile iridoids of which the best known is the cat-exciting nepetalactone (C5|C5OL) from *Nepeta cataria* (catnip) (Lamiaceae). The lactone nepetalactone, the hemiacetal neomatatabiol (C5|C5OH), iridodiol (in the ring opened bi-aldehyde

OHC−C5−CH(CH_3)−CHO) and ring-closed (C5|C5OH) forms) and the ring-opened bi-aldehyde dolichodial are volatile simple iridoids variously having insect repellent and attractant activity. The iridoid hemiacetal valtratum from the roots of *Valeriana* (valerian) and *Centranthus* is an anxiolytic psycholeptic and related valepotriates such as isovaltrate may also contribute to the tranquillizing, anxiolytic and anti-insomnia effects of valerian (noting that the baldrinal and homobaldrinal aldehyde products are mutagenic).

Iridoid glycosides include the bitter hemiacetal glucosides (C3|C5|C4OH−O−glucoside) catalpol, harpagoside and loganin (loganoside). The hemiacetal glucoside aucubin (aucuboside) is toxic because the aglycone C5OH ring can open and thence react with proteins to form imine adducts.

Seco-iridoids involve opening of the C5 ring and include the glucoside swertiamarin (C5OL|C5OH−O−glucoside) (the aglycone of which, erythrocentaurin, is very bitter) and oleuropein (aryl−C5OH−O−glucoside) which can form covalent adducts with proteins through reaction with a readily oxidized alkene side chain. Secologanin is similarly reactive through the aldehyde and ethylenyl substituents on the residual heterocyclic ring and is a precursor for particular alkaloids through reaction with amines.

Non-glycoside iridoids include aglycones stabilized through formation of a cyclic ether ring involving the hemiacetal hydroxy, examples including the antimicrobials plumericin and isoplumericin (C4OL*|C4O*,*′|C5*,*′|C5O*,*′) (where the superscripts * and *′ indicate that three Cs are respectively shared by the three rings thus denoted). The tranquillizing iridoid hemiacetals didrovaltratum and valtratum from *Valeriana officinalis* (valerian) (Valerianaceae) have isobutyric acid esterified on the hemiacetal hydroxy and another hydroxy.

iii. Sesquiterpenes. Sesquiterpenes derive from farnesylpyrophosphate (C_{15}) having three isoprene units (C_5) linked head-to-tail and occur in plant essential oils. Sesquiterpenes include a huge variety of cyclic compounds as well as simple non-cyclic farnesyl derivatives. The cyclic sesquiterpenes include monocyclic, bicyclic and tricyclic compounds and the sesquiterpene lactones. The sesquiterpene lactones are a particularly large group and are dealt with separately in Section iv.

Non-cyclic sesquiterpenes include the volatiles α- and β-farnesene (which have alarm pheromone activity) and pleasant odorants from *Citrus sinsensis* (orange) (Rutaceae), namely α- and β-sinensal (mandarin peel odour) and nerolidol from orange flower oil (oil of neroli). The epoxide JH III is produced by *Cyperus iria* (Cyperaceae) and acts critically on insect development.

Monocyclic sesquiterpenes typically have an alkylated C6 ring but macrocyclic examples include the insect antifeedant shiromodiol diacetate (C10) and the insect attractant odorants α-humulene from *Humulus lupulus* (hops) (Cannabaceae) (C11 ring) and germacrene B from *Citrus* peel (C10 ring). Monocyclic sesquiterpenes with a C6 ring include: juvabione (that has insect JH activity); the important plant growth regulator abscisic acid (that regulates stomatal opening, bud dormancy and leaf abscission); the odorants curcumene and zingiberene from the oil of both *Curcuma aromatica* (turmeric) and *Zingiber officinale* (ginger); the sweet compounds hernandulcin and 4β-hydroxyhernandulcin; the anti-inflammatory bisabolol (that promotes wound healing); and the *Viola* (Violaceae) violet scents, the α- and β-irone isomers.

Bicyclic sesquiterpenes include a variety of bioactive compounds including: the Solanaceae phytoalexins capsidiol (C6|C6) from *Capsicum frutescens* (pepper) and the *Solanum tuberosum* (potato) antifungals rishitin (C6|C6) and solavetivone (C6·C5); the *Ipomoea batatus*

(sweet potato) (Convolvulaceae) furanoid phytoalexin ipomeamarone (THfuran-furan); the hepatotoxic furanoid sesquiterpenes dehydromyodesmone (C6–C5) and dehydrongainone (THfuran-furan) from the toxic shrub *Myoporum deserti* (Myoporaceae); the piscicidal 5-LOX inhibitors buddledin A, B and C (C4|C6); and the spasmolytic cAMP PDE inhibitor petasin (C6|C6).

Neuroactive bicyclic sesquiterpenes include the antifeedant cinnamodial (C6|C6) (a vanilloid (capsaicin) receptor agonist), α-eudesmol (C6|C6) (a Ca^{2+} channel blocker) and valerenic acid (C5|C6) (which inhibits GABA breakdown). Odorant bicyclic sesquiterpenes include α-vetivone (C6|C6) and β-vetivone (C6·C5) from *Vetiveria zizanoides* (vetiver grass) (Poaceae) roots. The anti-inflammatory antioxidants chamazulene (C5|C7) and guaiazulene (C5|C7) derive post-extraction from steam distillation of leaves of *Matricaria chamomilla* (chamomile) (Asteraceae). The dialdehyde warburganal (C6|C6) is toxic because of its reactivity with thiols and the amino groups of proteins. The dimeric, bicyclic, sesquiterpene phenolic aldehyde gossypol (Phe|Phe–Phe|Phe) from *Gossypium hirsutum* (cotton) (Malvaceae) seed oil is a potent inhibitor of various protein kinases and of the Ca^{2+}-dependent protein phosphatase calcineurin.

Non-lactone tricyclic sesquiterpenes include the *Juniper* (Cupressaceae) odorants α-cedrol and α-cedrene (C5|C5|C6); the *Piper cubeba* (cubeb fruit) (Piperaceae) flavours α- and β-cubebene (C5|C3|C6); and the fragrant patchouli alcohol (tetramethyl-1,6-methano-octahydronaphthalene) (C6||C6(–CH(OH)–$(CH_3)_2$–)) from *Pogostemon patchouly* (Lamiaceae) patchouli oil. The carcinogenic DNA alkylating and breaking norsesquiterpene pterosin B (C3·C6|C5) occurs as a glucoside ptaquiloside in *Pteridium aquilinum* (bracken fern) (Dennstaedtiaceae) (the fern "fiddlehead" sprouts are eaten in New Brunswick, Canada and elsewhere in the region and are toxic if insufficiently cooked).

iv. Sesquiterpene lactones. The sesquiterpene lactones are a large class of C_{15}-based terpenes having a common five-membered γ-lactone ring system (a tetrahydrofuranone) involving cyclization through esterification of a carboxy with a γ-hydroxy of the precursor HO–CH(X)–CH(Y)–CH(Z)–COO^-. Sesquiterpene lactones are typically unsaturated di- or tri-cyclics and many have a reactive methylene (=CH_2) substituent. Many of these terpenes derive from Asteraceae (Compositae) plants and are variously bitter tasting, insect antifeedants, cytotoxic and antineoplastic. These compounds can be grouped based on the various fused ring structure arrangements. In summarizing, the sesquiterpene lactone structures below the common C4 lactone ring element will be represented as C4OL and the corresponding C5 lactone as C5OL. Similarly, the commonly occurring C4 and C5 cyclic ethers are represented as C4O and C5O, respectively. Where condensed rings share three Cs, a double vertical line (‖) is used and remember that an asterisk (*) indicates that a ring is part of a tricyclic structure in which all rings share one C.

Elemanolides. The elemanolide sesquiterpene lactones (C5OL|C6|C4OL) are exemplified by the antifeedants vernodalin (C5OL|C6|C4OL) and vernodalol (C5OL|C6) (in which the C4OL lactone ring is opened and the carboxy methylated). Vernodalin contributes to the bitter taste of *Vernonia amygdalina* (Asteraceae) ingested by parasite-infected chimpanzees.

Eudesmanolide sesquiterpene lactones (C6|C6|C4OL) include a variety of variously cytotoxic compounds. The antifeedants alantolactone and isoalantolactone (helenin being a mixture of the two) (C6|C5|C4OL=CH_2) are antimicrobial. The pro-apoptotic activity of these compounds may derive from the reactivity of the α-methylene-γ-butyrolactone ring

(C4OL=CH_2). Santamarin may form a covalent adduct in inhibiting transcription factor NFκB binding to DNA.

Guaianolides (C5|C7|C4OL) include many cytotoxic and antineoplastic compounds. Various guaianolides are bitter tasting and insect antifeedants. Zaluzanin inhibits bacterial lipopolysaccharide-induced NFκB-mediated expression of iNOS by immune cells and cynaropicrin inhibits similar induction of TNF-α expression through formation of a covalent protein adduct. Costunolide, 7-hydroxycostunolide and 3,4-epoxydehydroleucodin act in a similar manner to inhibit NFκB binding to DNA.

Other notable guaianolides include achillin, artabsin and matricin (that can be converted on heating to the anti-inflammatory radical scavenger and COX inhibitor chamazulene); the cytotoxic and antineoplastic chloroguaianolides eupachlorin, eupachlorin acetate and eupachloroxin; the cytochrome P450 aromatase inhibitors 10-epi-8-deoxycumambrin, dehydroleucodin and ludartin; and thapsigargin, thapsivillosin and trilobolide (inhibitors of the transmembrane Ca^{2+} pumping Ca^{2+} ATPase).

Pseudoguaianolides (C5|C7|C4OL) differ from the guaianolides in having a 5-methyl substituent (at the junction of the C5 and C7 rings) and this group includes many insect antifeedants and compounds with cytotoxic and antineoplastic activity. Ambrosin and hymenin trigger apoptosis in leukaemia cells. The anti-inflammatory helenalin alkylates the p65 subunit of NFκB, thereby inhibiting the function of this inflammation-related transcription factor. Glutathione adducts of helenalin and 11α,13-dihydrohelenalin acetate inhibit glutathione S-transferase and helenalin inhibits 5-LOX. 2,3-Dihydrohelenalin and *bis*-helenalinyl malonate inhibit IMP dehydrogenase.

Germacranolides have a common bicyclic (C10|C4OL) structure but some (e.g. budlein A) have an additional fused furan ring through an ether linkage across the larger ring. Many germacranolides are cytotoxic and antineoplastic. Parthenolide (C10|C4OL) from the antimigraine herb *Tanacetum* (*Chrysanthemum*) *parthenium* (feverfew) (Asteraceae) is a serotonin receptor (5HT-R) antagonist. Parthenolide and costunolide inhibit the phosphorylation of IκB that is required for pro-inflammatory activation of NFκB.

Tutinanolide sesquiterpene lactones are epoxides having a common C6|C5epoxide bicyclic element to which is appended a five-membered lactone structure involving three Cs of the C6 ring through an ester (−CO−O−) cross-link across the ring (this is denoted as a (C4OL*|C6*(−CO−O−)|C5*epoxide) structure). Important compounds of this kind include the *Menispermum occulus* (Menispermaceae) GABA(A) receptor and glycine receptor antagonists picrotin and picrotoxinin (C4OL*|C6*(−CO−O−)|C5*epoxide). Other excitatory tutinanolides that are GABA(A) receptor antagonists include coriamyrtin (C6(−CO−O−)|C5epoxide) and tutin (2-hydroxycoriamyrtin) from *Coriaria* species (Coriariaceae) and the Euphorbiaceae mellitoxin (C6(−CO−O−)|C5epoxide) that also derives from the honey of bees feeding on *Coriaria* species.

Other sesquiterpene lactones include the toxic Asteraceae seco-pseudoguianolides hymenoxyon (P|C7|C4OL) (from the toxic *Hymenoxys* and *Helenium* species) that alkylates DNA and vermeerin (C5OL|C7|C4OL) (from the toxic *Geigera* species) that forms adducts with protein cysteines; the seco-guaianolides (xantholides) (C7|C4OL) xanthinin and xanthumin that have auxin antagonist and antifeedant activity, respectively; and the GABA(A) receptor antagonist anisatin (C5|C6(·C3OL)|C5OL).

Artemisinin (quinghaosu) (3,12-peroxyC6O*|C5OL*|C6*) from *Artemisia annua* (Asteraceae) is of major importance as an antimalarial because of extensive resistance of *Plasmodium falciparum* to antimalarials such as chloroquine. Artemisinin in has a 3,12-peroxy (−O−O−) substituent spanning the C6O ring. Artemisinin alkylates and inhibits glutathione S-transferase.

v. Diterpenes. Diterpenes derive from the C_{20} isoprenoid geranylgeranylpyrophosphate (four head-to-tail-linked isoprenes). Geranylgeraniol and the chlorophyll moiety phytol are acyclic diterpenes (Section 3, Appendix). Cyclic diterpenes vary in the number, nature and disposition of the ring structures. The core alicyclic skeleton of the various diterpenes usually involves fused C3, C4, C5, C6 and C7 alicyclic rings that are typically (but not always) completely saturated. Carboxy and hydroxymethyl groups (or carboxymethyl and hydroxy groups) on adjacent Cs can cyclize to form a fused five-membered lactone ring (C4OL). A fused, reduced six-membered lactone ring (C5OL) derives from lactone formation from, for example, carboxymethyl and hydroxymethyl groups on adjacent Cs. In addition, epoxides can be formed from oxidation of double bonds.

The diterpenes are a structurally diverse group of natural products, of which many are toxic or otherwise bioactive. In the following sketch, diterpene structural complexity has been simplified as before by representing fused rings sharing two Cs as C*n*|C*n*. Ring systems with more than two fused rings are mostly angular (cf. fully reduced phenanthrene) rather than linear (cf. fully reduced anthracene). In some cases, furan and pyran rings are involved that have different degrees of saturation. The different classes of diterpenes are dealt with below in alphabetical order for ease of reference.

Abietane diterpenes (C6|C6|C6 with varying degrees of unsaturation) include the 5-LOX inhibitor abietane, the GABA(A) receptor antagonist taxodione and the bitter tastant carnosol.

Clerodanes involve a (C6|C6) group variously linked to furan (unsaturated C4O), pyran (unsaturated C5O), methyleneoxy and methylenedioxy rings. Clerodanes include bitter tastants and antifeedants as exemplified by the extremely bitter component columbin of the bitter tonic made from roots of *Jateorhiza columba* (columba root) (Menispermaceae).

Daphnanes (C5|C7|C6) include a variety of cytotoxic, irritant, inflammatory and toxic compounds from the Thymelaeaceae and the Euphorbiaceae. Of particular note are the highly inflammatory PKC activators resiniferatoxin and tinyatoxin from *Euphorbia* species (Euphorbiaceae) and thymeleatoxin from *Thymelea hirsuta* (Thymelaeaceae). While the non-ester resiniferonol is inactive, the ester (X–CO–O–Y) resiniferatoxin is both an anti-nociceptive vanilloid receptor (capsaicin receptor) agonist and a PKC activator, as is the ester tinyatoxin.

Gibbanes have a complex (C6*|C4OL*|C5*′|C6*′|C5*′) structure (noting that the central C5*′ shares a common C with both C6*|C4OL* rings and the C6*′|C5*′ rings, respectively). The gibbanes include a large number of plant growth regulators called gibberellins of which the best known is gibberellic acid (gibberellin A_3 or GA_3) which controls growth and seed dormancy. Gibberellic acid is produced by the rice pathogen *Gibberella fujikuroi* (*Fusarium moniliforme*) and causes greatly increased, spindly rice stalk growth. Gibberellic acid is used to induce barley seed aleurone α-amylase production in malting prior to brewing beer.

Ginkgolides (C4OL|$C5^{3*,4*}$|$C4OL^{3*,4*}$(|$C4O^{3*,3'*,4*}$|$C4OL^{3'*}$)|$C5^{3'*,4*}$) (noting – in the worst case you will encounter in this chapter – that the superscripts 3*, 3′* and 4* indicate three separate C atoms shared by 3, 3 and 4 rings, respectively, as denoted). Ginkgolides are anti-inflammatory, bitter antifeedants from *Ginkgo biloba* (Ginkgoaceae). Ginkgolide A is a PAF antagonist and consequently anti-inflammatory.

Grayanotoxins (C5|C7*|C6*||C5*) are highly toxic Ericaceae compounds of which grayanotoxin I is the best known. Grayanotoxin I opens voltage-gated Na^+ channels from the inside of the cell thus causing depolarization, impairment of neurotransmission and interference with proper cellular signalling.

Ingenanes (C5*|C7*||C7*|C3) include irritants and secondary tumour promoters (co-carcinogens) from *Euphorbia* species (Euphorbiaceae) that activate PKC. While the non-ester

precursor ingenol is active, 17-hydroxyingenol 20-hexadecanoate, ingenol 3-benzoate, ingenol 3,20-dibenzoate and ingenol 20-hexadecanoate all interact with PKC.

Jatrophanes (C5|C12) are cytotoxic, antitumour compounds from *Jatropha* species (Euphorbiaceae). Jatrophone binds to DNA and also has activity as a glutamate receptor antagonist.

Kauranes (common structural element C6|C6*|C6*||C5*) and related compounds include natural products that are variously antifeedants, bitter and otherwise bioactive. Thus, the glycoside toxin atractyloside inhibits the mitochondrial ADP/ATP translocator, the sweet glycoside stevioside blocks Ca^{2+} channels and the antifeedant inflexin inhibits aromatase. Some kauranes having a fused furan ring (i.e. skeletal structure furan|C6|C6*|C6*||C5*)) include the very bitter glycoside mascaroside and the coffee components cafestrol (cafesterol) and kahweol ($\Delta^{1,2}$-cafestrol). Cafestrol and kahweol (present in boiled (unfiltered) coffee) raise plasma low density lipoprotein (LDL)-associated cholesterol (by decreasing expression of LDL receptors) but are also chemopreventative by decreasing expression of the cytochrome P450-linked CYP oxygenases that generate genotoxic metabolites from precursors (e.g. genotoxic aflatoxin B1-8,9-epoxide from the consumed *Aspergillus flavus* (fungal) coumarin procarcinogen aflatoxin B1).

Labdanes (core C6|C6 linked to variously reduced furan or pyran moieties) include the Lamiaceae (Labiatae) diterpenes forskolin (C6|C6|pyran) (that activates adenylyl cyclase, the enzyme catalysing cyclic AMP formation from ATP) and premarrubiin (C4OL*|C6*|C6*·furan·furan) that converts to the bitter non-opiate antinociceptive marrubiin (C4L*|C6*|C6*–$(CH_2)_2$-furan).

Tigliane (C5|C7|C6|C3) diterpenes include the highly irritant, toxic, co-carcinogenic, PKC activating phorbol esters from Euphorbiaceae plants. While not being activated by the parent compound phorbol, PKC is activated by plant-derived esters of phorbol, 4-deoxyphorbol and 12-deoxyphorbol (e.g. 12-*O*-palmitoyl-16-hydroxy-phorbol 13-acetate, 12-deoxyphorbol 13-benzoate, 12-deoxyphorbol 13-phenylacetate, 12-deoxyphorbol 13-phenylacetate-20-acetate, sapintoxin A (4-deoxyphorbol 12-(2-methylamino)benzoate-13-acetate) and 12-tetradecanoylphorbol 13-acetate (TPA)), as well as by synthetic phorbol esters.

Other diterpenes include the antifungal phytoalexin casbene (C3|C14); the PKC-binding dihydroxyatisanone and trihydroxyatisane (C6|C6|C6|C6); the toxic totarane diterpenes hallactones A and B (C4OL*|C6*|C6*|C5OL); the pimarane pimaric acid (C6|C6|C6); the labour-inducing oxepane macrocyclic ethers montanol and zoapatanol (C7O); portulal (C7|C5); lathyrol (C5|C11|C3) and the macrocyclic insect trail pheromone neocembrene (C14) that is also found in certain plants.

vi. Triterpenes. Triterpenes derive from cyclization of the linear C_{30} precursor squalene, remembering that if we denote isoprene (C_5) as IP (to indicate structural head and tail polarity) then squalene has the structure $(IP)_3$–$(PI)_3$. Triterpenes are polycyclic and often glycosylated. The non-glycosylated aglycones usually have about thirty Cs, but some have more or fewer C atoms. Thus, one can distinguish between the C_{30} triterpenoid sapogenins (saponin aglycones), the cucurbitacin aglycones and other C_{30} triterpenes as opposed to the C_{27} spirastane-based steroid sapogenins, C_{24} bufadienolides, C_{23} cardenolides, nortriterpenoid C_{26} limonoids and C_{19}–C_{20} quassinoids. Further, the phytosterols are structurally very similar to cholesterol (C_{27}) but the major phytosterols have 1–2 more Cs in the side chain furthest removed from the 3-hydroxy. The structural and functional complexity of the terpenes is briefly sketched below.

C_{30} triterpenoid saponins and sapogenins. Saponins are terpenoid amphipathic compounds having water-soluble sugar residues linked (via glycosidic links formed between the sugar hemiacetal and terpenoid OHs) and a relatively hydrophobic (water repelling) triterpenoid aglycone part. Amphipathic compounds (i.e. compounds having both hydrophilic and hydrophobic regions) typically foam (by accumulating at an air–water interface) and can act as detergents (by solubilizing hydrophobic compounds such as phospholipids). Accordingly, saponins have detergent properties and can be haemolytic through solubilizing the cell membrane of red blood cells.

The C_{30} sapogenins (the aglycone moieties) typically have a non-linear (i.e. reduced phenanthrene-like) 3-OH–C6|C6|C6|C6|C6 skeletal structure with glycosylation often at the 3-OH but also occurring at OHs more distal to the 3-OH. Unless indicated otherwise, the compounds discussed below have this skeletal structure. An illustrative exception is the nucleoside transport inhibitor cimicifugoside which has a 3-*O*-glycosyl (Glyc)-C6|C6|C6|C5|pyran·furan-epoxide structure.

Some triterpenoid saponins are bitter tastants (e.g. helianthoside A) and others are sweet tasting, most notably the 3-*O*-glycosides abrusosides A–D (C6|C6|C6|C5–CH(CH_3)–C5L 3-*O*-glycosides) from *Abrus* species (Fabaceae) and glycyrrhizin (glycyrrhizic acid) from the rhizomes and roots of *Glycyrrhiza glabra* (licorice) (Fabaceae). In contrast, the 3-*O*-glycoside gymnemic acid, the 3-*O*-glycoside of barringtogenol (from tea) and jegosaponins A–D have antisweet activity (i.e. abolish a sweet tastant response) (Chapter 10).

Glycyrrhetinic acid (glycyrrhetic acid) (the aglycone of glycyrrhizic acid) inhibits 11β-hydroxysteroid dehydrogenase (thus impairing cortisol oxidation to cortisone and causing hyper-mineralocorticosteroidism when licorice is taken in excess). Glycyrrhetinic acid, oleanolic acid and ursolic acid inhibit protein kinases. Gypenosides (C6|C6|C6|C5 glycosides) and saikosaponin A inhibit the Na^+ pump (Na^+, K^+-ATPase). α-Hederin (sapindoside A) and oleanolic acid inhibit chitin synthetase II. Oleanolic acid and ursolic acid inhibit DNA polymerase.

Other C_{30} triterpenoids have been shown to interact with specific biochemical targets. Such triterpenes are classified in terms of their skeletal arrangement, for example, cycloartanes, friedelanes, oleane, taraxanes and ursanes (C6|C6|C6|C6|C6), fernanes, hopanes and lupanes (C6|C6|C6|C6|C5), dammaranes and euphanes (C6|C6|C6|C5–C_8), and protolimonoids (C6|C6|C6|C5–P). The friedelane tingenone binds to DNA and inhibits DNA-dependent RNA and DNA synthesis. The ursane α-amyrin, the lupane lupeol and fatty acid esters of these triterpenes inhibit cyclic AMP-dependent protein kinase (PKA) and the serine proteases trypsin and chymotrypsin. A range of Asteraceae cycloartane, dammarane, oleane and taraxane triterpenoids that inhibit phorbol ester-induced inflammation are also inhibitors of trypsin and chymotrypsin.

Steroid saponins are glycosides of spirostane triterpenoid sapogenins that have a basic 3-OH–C6|C6|C6|C5|THfuran·THpyran skeleton. Steroid saponins are in general non-toxic but have a foaming and detergent propensity. The steroid glycoside digitonin and its aglycone digitogenin derive from seeds of *Digitalis purpurea* (foxglove) (Scrophulariaceae). Digitonin is widely used in biochemical investigations as a "gentle" non-ionic detergent to solubilize membranes, for example, to prepare submitochondrial particles from the mitochondrial inner membrane. The steroid glycoside officinalisnin I from the roots of *Asparagus officinalis* (asparagus) (Liliaceae) is bitter whereas the glycoside osladin is sweet. The steroid glycoside gitonin from foxglove leaves is a cyclic AMP phosphodiesterase inhibitor.

Cucurbitacins are oxygenated triterpenes (C_{30}; typical skeletal structure C6|C6|C6|C5–C_8) that can be glycosylated. Cucurbitacins are typically bitter tastants and

antifeedants present in plants of the Cucurbitaceae in particular as well as having been found in some other plant families. Cucurbitacins are in general bitter tastants and antifeedants. However, some cucurbitacins such as the glycosides bryodulcoside and carnosifloside VI are sweet tastants. The aglycones cucurbitacins B and D are ecdysone antagonists and can act both as antifeedants and as insect attractants. Cucurbitacins can be toxic and cytotoxic. The aglycone cucurbitacin E disrupts the cellular actin cytoskeleton. Some cucurbitacin glycosides from *Picria fel-terrae* (Scrophulariaceae) inhibit both the classical and alternative pathways of the complement system.

Phytosterols are structurally very similar to cholesterol and the major phytosterols (campesterol, sitosterol and stigmasterol) have the same kind of membrane viscosity modulating function in plants that cholesterol (C_{27}; 3-OH–C6|C6|C6|C5–C_8) has in animals. Campesterol (24-methylcholesterol), sitosterol (24-ethylcholesterol) and stigmasterol (Δ^{22}, 24-ethylcholesterol) are widespread phytosterols. The "animal" sterols lanosterol and cholesterol are present in particular plants. Phytosterol esters reduce cholesterol absorption and lower LDL-cholesterol.

The insect moulting hormones ecdysone and 20-hydroxyecdysone (2,3-OH–C6|C6|C6|C5–C_8) are elaborated by particular plants as are a large number of structurally very similar phytoecdysones that mimic ecdysone action in insect metamorphosis. The C_{19} animal androgens androstenedione and testosterone (3-keto-C6|C6|C6|C5) are present in *Pinus sylvestris* (Pinaceae) and the C_{18} oestrogens 17β-oestradiol, oestriol and oestrone are elaborated by particular plants. The elaboration of phytoecdysones and testosterone and oestrogen receptor agonists would potentially perturb the development of herbivore pests. The plant growth regulator brassinolide (2,3-OH–C6|C6OL|C6|C5–C_8) is also active as an ecdysone antagonist.

Cardenolides, cyclic bridged cardiac glycosides and bufadienolides are extremely toxic triterpenoids that are C_{23} and C_{24}, respectively, as aglycones and derive from a C_{30} triterpene precursor. The cardenolides (3-OH–C6|C6|C6|C5–C4OL) (fused rings successively denoted A, B, C and D) can have a *cis*-configuration at the junction of the A and B rings (5β-cardenolides such as digoxigenin) or a *trans*-configuration (5α-cardenolides such as aspeciogenin from *Asclepias* species (Asclepiadaceae)). The cardenolides are typically glycosylated and the cardiac-active compounds are referred to as cardiac glycosides. The cardiac glycosides inhibit the Na^+ pump (Na^+, K^+-ATPase) that is responsible for maintaining a low cytosolic Na^+ and high cytosolic K^+ concentration critical for cell excitability, maintenance of low cytosolic Ca^{2+} concentration and for neurotransmission.

Among the best-known cardenolide glycosides (aglycones in parenthesis) are digitalin (gitoxigenin), gitoxin (gitoxigenin), digitoxin (digitoxigenin) and digoxin (digoxigenin) from *Digitalis* species (Scrophulariaceae), notably *Digitalis purpurea* (foxglove). The foxglove leaf extract (digitalis) has been used for several centuries for cardiac insufficiency, inhibition of the Na^+ pump successively lowering the Na^+ gradient across the cell membrane, decreasing Na^+-dependent Ca^{2+} pumping out of the cell, increasing cytosolic Ca^{2+} concentration and increasing cardiac muscle contraction. Other important cardiotonic cardiac glycosides (aglycones in parenthesis) are ouabain (ouabagenin) and strophanthin-K (strophanthidin) from *Strophanthus* species (Apocynaceae). Ouabain has been found to be an endogeous Na^+ pump regulator and signalling compound in animals.

Various Asclepiadaceae 5α-cardenolides (C_{23}; 2,3-di-OH–C6|C6|C6|C5–C4OL) form cyclic bridged glycosides linking the sugar via the 2- and 3-hydroxyls of the aglycone, an example being asclepin from *Asclepias* species. Bufadienolide (C_{24}; 3-OH–C6|C6|C6|C5–C4OL) glycosides include (aglycone in parenthesis) scillaren A (scillarenin) from *Scilla maritima* (Liliaceae) and hellebrin (hellebrigenin) from *Helleborus niger* (Ranunculaceae).

Limonoids are C_{26} nortriterpenoids deriving from a C_{30} triterpene precursor. The best known limonoids are the *Azadirachta indica* (neem tree) antifeedant azadirachtin (C5OL*|C4O*|C6*–C6O(epoxide; methylene cross-link)|furan) and the *Citrus* species (Rutaceae) bitter antifeedant limonin (C5OL*|C4O*|C6*|C6|C5OL(epoxide)-furan). Limonin gives a delayed bitter taste to *Citrus* fruit. The limonoids are typically bitter compounds with insect antifeedant activity.

Quassinoids are typically C_{19} and C_{20} nortriterpenoids deriving from processing of a C_{30} triterpene precursor. These compounds typically have a basic C6|C6*|C6*||C4O*|C5OL*| skeleton as typified by the cytotoxic bruceines from *Brucea* species (Simaroubaceae). Other quassionoids include the very bitter tastants quassin (C6|C6*|C6*|C5OL*) from *Quassia amara* (Simaroubaceae) and nigakihemiacetal A (C6|C6*|C6*|C5OH*). Many quassinoids are bitter tastants and cytotoxic. Chaparrinone and related quassinoids from *Hannoa* species (Simaroubaceae) are antimalarials.

vii. Carotenes. Carotenes derive from geranylgeranylpyrophosphate $H(CH_2–C(CH_3)=CH–CH_2)_4–O–PO_3–PO_3$ (C_{20}–PP). Representing the PP-end as the "head", head-to-head condensation of two geranylgeranylpyrophosphate (C_{20}–PP) molecules ultimately yields phytoene (C_{40}), that is, if one represents the isoprene polarities in isopentylpyrophosphate (IP-PP) as IP (tail-head) and PI (head-tail), one could represent phytoene as $(IP)_4–(PI)_4$. Because of the extensive conjugated double bond systems (i.e. $(–C=C–C=C–)_n$) the carotenes are coloured, the colour ranging from yellow to red. Accordingly, carotenes are important for pollination (attracting insects to flowers) and seed dispersal (attracting herbivores to fruit).

The most abundant carotene is β-carotene which after ingestion gives rise to vitamin A (all-*trans*-retinol) (C_{20}) that is involved in proper development (via the cytoplasmic retinoid receptors that switch on expression of specific sets of genes). The aldehyde derivative retinal is involved in vision as the chromophore covalently linked to opsin proteins and which initiates a G protein-linked signalling pathway after undergoing light-dependent isomerization. The signalling pathway in vision successively involves conformational change of the opsin–retinal complex (rhodopsin), release of Gαt-GTP from the G protein complex, activation of cyclic GMP phosphodiesterase by Gαt-GTP, decreased cyclic GMP, closure of cyclic GMP-gated Na^+ channels and transmembrane potential hyperpolarization (see Chapters 3 and 5). Vitamin A ((all *E*)-2,3,7-dimethyl-9-(2,6,6,-trimethyl)-1-cyclohex-1-yl)-2-4,6,8-nonatetraen-1-ol) can be simply represented as cyclic $(IP)_2–(IP)_2–OH$ and β-carotene as cyclic$(IP)_2–(IP)_2–(PI)_2$-cyclic$(PI)_2$. Accordingly, oxidation of β-carotene yields two molecules of vitamin A.

In addition to β-carotene there are a variety of other C_{40} pro-vitamin A carotenes that differ from β-carotene in the nature of the terminal cyclic moieties. Thus, representing the "right" cyclic moiety as X, we can represent β-carotene as $X–(IP)_2–(PI)_2–X$ that yields two molecules of vitamin A or $X–(IP)_2–OH$. Carotenes can have different cyclic moieties X′ (where $X' \neq X$) or no cyclic isoprene dimer moieties. Other C_{40} pro-vitamin A carotenes that yield only one vitamin A molecule on oxidation include α-carotene, β-cryptoxanthin, β-carotene epoxide, echinenone and mutachrome (generalized structure $X–(IP)_2–(PI)_2–X'$) and γ-carotene and torulene ($X–(IP)_2–(PI)_4$).

A variety of C_{40} carotenes do not yield vitamin A on oxidation and these variously have altered cyclic groups or no cyclic groups at all and can be variously oxidized or reduced. Good examples are the widespread lutein ($X'–(IP)_2–(PI)_2–X'$) (yellow) and the non-cyclic carotenes lycopene (ψ,ψ-carotene; the orange–red colour of tomatoes and other fruits), ζ-carotene (7,8,7′,8′-tetrahydro-ψ,ψ-carotene; yellow) and lycoxanthin (ψ,ψ-caroten-16-ol; yellow).

Crocetin (C_{20}) is a yellow $(IP)_2$–$(PI)_2$-derived dicarboxylic acid (generalized structure –OOC–C_{18}–COO–) from the styles of *Crocus sativus* (Iridaceae) (the saffron of Indian cooking and Buddhist robes). Crocin, the digentiobiose ester of crocetin, is water soluble, unlike other carotenoids which are lipophilic (fat soluble). Crocetin is a protein kinase inhibitor. Excess vitamin A (or excess pro-vitamin A) ingestion is toxic (dog liver consumption having caused the death of Sir Douglas Mawson's explorer companions in the Antarctic by this mechanism).

1.9 Other plant compounds

A variety of other plant compounds are bioactive as toxins, pro-toxins, sweet or bitter tastants, odorants, semiochemicals, enzyme inhibitors, receptor agonists, receptor antagonists or psychoactive agents. The structure and bioactivity of non-alkaloid, non-phenolic and non-terpenoid plant compounds is briefly reviewed below. Some selected structures of cyclic compounds in this category are shown in the Appendix (Section 4).

i. Sugars. Sugars such as monosaccharides (e.g. glucose and fructose) and disaccharides (e.g. sucrose) are typically sweet tastants, this pleasant animal perception having been selected evolutionarily because of the energy-rich, catabolizable nature of sugars. However, sugars are often linked to toxic defensive compounds as glycosides and such compounds can be bitter. Sugars can have a general structure of $HOCH_2$–$(CH(OH))_n$–CHO (aldoses) or $HOCH_2$–$(CH(OH))_n$–CO–CH_2OH (ketoses). A C atom having four different substituents can give rise to two possible mirror image isomers (stereoisomers) that as configurational isomers can only be interconverted by breaking and re-forming bonds. The stereoisomers of sugars due to these C atom "chiral centres" were detected by differential "optical activity" (rotation of the plane of polarization of plane polarized light in a polarimeter) and the absolute configurations have been established. Most of the sugars of living organisms have a so-called D configuration (as with the key metabolite D-glucose) as opposed to an L-configuration (as with the 5-carbon sugar L-arabinose).

Sugars are further classified by the number of carbons. Thus, aldoses include aldotrioses (C_3; D-glyceraldehyde, HO–CH_2–CHO), aldotetroses (C_4; D-erythrose); aldopentoses (C_5; D-ribose, D-arabinose, D-xylose) and aldohexoses (C_6; D-glucose, D-mannose, D-gulose, D-galactose). Ketoses include ketotrioses (C_3; dihydroxyacetonephosphate, HO–CH_2–CO–CH_2OH), ketotetroses (C_4; D-erythrulose), ketopentoses (C_5; D-ribulose, D-xylulose) and ketohexoses (C_6; D-fructose).

Aldose sugars (such as glucose) can exist in an open chain form as described above but in aqueous solution largely condense to form cyclic hemiacetals, the ring closure linkage being: –$CH(CH_2OH)$–O–CH(OH)–. Similarly, ketose sugars (such as fructose) condense to form a hemiketal, the ring closure linkage being: –$CH(CH_2OH)$–O–C(OH, CH_2OH)–. Glucose forms a six-membered ring containing five Cs and one O and is called a glucopyranose form after the cyclic ether tetrahydropyran (C5O). Fructose forms a five-membered ring containing four Cs and one O and is called a fructofuranose after the cyclic ether tetrahydrofuran (C4O).

If D-glucopyranose is drawn with the hemiacetal O going into the plane of the paper and the C-6 CH_2OH group pointing above the chain, then the C-1 hemiacetal OH can either point up (in the β anomer) or point down (in the α-anomer). Hence, we can either have β-D-glucopyranose or α-D-glucopyranose and the same anomeric possibilities exist for other sugars, for example, β-D-fructofuranose or α-D-fructofuranose (Section 4, Appendix). Hemiacetal and hemiketal OHs can react with OH groups on other molecules (HO–X) with the elimination of H_2O to form a glycosidic link: C-1-O–X, noting that this is either an α- or β-glycosidic link, for example, quercimeritrin (in which a glucoside is formed

via reaction of the hemiacetal glucose with the 7-OH of the flavonol quercetin) is quercetin 7-*O*-β-D-glucoside.

Monosaccharides can link together through glycosidic links and thence to form oligosaccharides and ultimately polysaccharides (such as starch, glycogen, cellulose and callose) (see Chapter 2). Thus, maltose (α-D-glucopyranosyl-(1 → 4)-D-glucopyranose or α-D-Glc-(1 → 4)-D-Glc derives from the reaction of two glucopyranoses to form an "α(1 → 4) bond" with the elimination of H_2O (HO–H, with OH coming from the hemiacetal C-1 OH of one glucose and H from the alcohol 4-OH of the second glucose). Note that mutarotation means that the second glucose moiety in maltose still has a hemiacetal OH (a "reducing end" because it can react with an oxidant) and could exist in either an α- or β-form.

Lactose (β-D-galactopyranosyl-(1 → 4)-β-D-glucopyranoside; β-D-Gal-(1 → 4)-D-Glc) involves a "β(1 → 4) bond", noting that lactose has a "reducing end", that is, C-1 of the glucose part. Sucrose (α-D-glucopyranosyl-(1 → 2)-β-D-fructofuranoside; α-D-Glc-(1 → 2)-β-D-Fru) does not have a reducing end, the reducing ends of both the constituent monosaccharides being involved in glycosidic bond formation (Section 4, Appendix). Maltose, lactose and sucrose are sweet tasting. Other sweet tasting sugars include: melibiose (β-D-Gal-(1 → 6)-D-Glc); the sulfated fucose polymer fucoidan from brown algae; the sugar alcohols ($H(CH–OH)_nCH_2OH$) glycerol (C_3), erythritol (C_4), D-arabitol (C_5), dulcitol (C_6), D-mannitol (C_6), D-sorbitol (C_6) and sedoheptitol (C_7); and the cyclohexanehexols (cyclitols; $C6(OH)_6$) inositol and quercitol. Gentiobiose (β-D-Glc-(1 → 6)-D-Glc) is a bitter tastant and the sulfated galactose polymer carageenan from red seaweed can induce gastric inflammation and oedema in mammals.

ii. Other aliphatics. In addition to the compounds outlined above, plants variously produce numerous aliphatic carboxylic acids, alcohols, aldehydes, fatty acids, sulfides and hydrocarbons that are variously bioactive. The structure and bioactivity of these other aliphatic plant natural products are outlined below.

Carboxylic acids. Aliphatic carboxylic acids (R–COOH) are deprotonated at physiological pH (pH 7) and are therefore represented as $R–COO^-$. Thus, acetic acid ($CH_3–COOH$) exists as acetate (CH_3COO^-) at pH 7. A variety of short chain mono-, di- and tricarboxylic acids are important intermediates in metabolism and may be present at low concentrations in all cells either as the acid or as a covalent adduct. Thus, acetate (C_2) and malonate (C_3) can exist as the key acyl-coenzyme A thioester intermediates acetylCoA and malonylCoA, respectively. Phosphoenolpyruvate (C_3), 1,3-bisphosphoglyceric acid (C_3) and 3-phosphoglycerate (C_3) are key metabolic intermediates.

Major monocarboxylic acids include formate (C_1; $HCOO^-$), glycolate (C_2; $HO–CH_2–COO^-$), glyoxylate (C_2; $OHC–COO^-$), acetate (C_2; $CH_3–COO^-$), pyruvate (C_3; $CH_3–CO–COO^-$), lactate (C_3; $CH_3–C(H,OH)–COO^-$), mevalonate (C_6), shikimate (C_7; 3,4,5-trihydroxycyclohexenecarboxylate, an intermediate in aromatic compound biosynthesis) and quinate (C_7; tetrahydroxycyclohexanecarboxylate). Dicarboxylic acids include oxalate (C_2; $^-OOC–COO^-$), malonate (C_3; $^-OOC–CH_2–COO^-$), tartrate (C_4; $^-OOC–C(H,OH)–C(H,OH)–COO^-$) and the successive TCA cycle intermediates α-ketoglutarate (C_5; $^-OOC–CO–CH_2–CH_2–COO^-$), succinate (C_4; $^-OOC–CH_2–CH_2–COO^-$), fumarate (C_4;$^-OOC–$ $CH=CH–COO^-$), malate (C4; $^-OOC–CH_2–CH(OH)–COO^-$) and oxaloacetate (C_4; $^-OOC–CH_2–CO–COO^-$). Tricarboxylic acids include the successive TCA cycle intermediates citrate (C_6; $^-OOC–CH_2–C(H,COO^-)–CH_2–COO^-$), *cis*-aconitate ($C_6$; $^-OOC–CH=C(H,COO^-)–CH_2–COO^-$) and isocitrate ($C_6$; $^-OOC–C(H,OH)–C(H,COO^-)–$ $CH_2–COO^-$).

Some organic acids (notably citrate and malate) may be present at high concentrations in the acid vacuoles of plant cells. Thus, in particular, desert plants having so-called Crassulacean acid metabolism (CAM plants) there is a water-saving adaptation that involves fixing CO_2 as malate [via PEP carboxylase] during the night (with leaf stomata open) and then releasing the CO_2 intracellularly for photosynthetic CO_2 fixation during the day (with leaf stomata closed and thereby minimizing H_2O loss). Thus, malate concentration increases during the night. Some fruits have a high level of organic acids such as malate, notably *Malus* species (apples) (Rosaceae).

Organic acids are sour tastants and particular organic acids accumulated in acid vacuoles contribute to the sourness of fruit, including malic acid and quinic acid (apple, apricot, pear, peach and banana fruit), citric acid (citrus fruits) and tartaric acid (grapes). Isovaleric acid (isopropylacetic acid) has a rancid smell but organic acid esters can have very pleasant smells (e.g. that of ethylbutyrate, the smell of apples). Oxalic acid is neurotoxic by chelating Ca^{2+} and malonate is a competitive inhibitor of succinate dehydrogenase. Fluoroacetic acid ($FCH_2–COO^-$) is toxic because of its conversion to fluorocitrate, a potent inhibitor of the enzyme aconitase that catalyses the conversion of citrate to isocitrate in the TCA cycle (Krebs cycle, Citric acid cycle). Ascorbic acid (vitamin C) is a tetrahydroxylactone co-enzyme for collagen hydroxylation and is readily oxidized to dehydroascorbic acid. Vitamin C must be derived from plants by humans and its absence causes scurvy.

Long chain fatty acids (general structure of a saturated fatty acid: $CH_3–(CH_2)_n–COOH$) are key components of all living things. Fatty acids form esters (X–CO–O–Y) with glycerol (trihydroxypropane; $CH_2(OH)–CH(OH)–CH_2OH$) to form monacyl-, diacyl- and triacylglycerides as high density "energy stores" in animal fatty tissue and in plants, notably in the oil-rich seeds of cotton, sunflower, linseed, coconut, peanut, soya bean and canola (rapeseed). 3-Phosphodiacylglycerol (phosphatidic acid) is the parent compound for phosphodiester phospholipids (e.g. phosphatidylinositol, phosphatidylcholine, phosphatidylserine and phosphatidylethanolamine) that are the bulk components of the molecular bilayers making up biological membranes. Plant fatty acids are typically unsaturated and membranes having a higher degree of unsaturated fatty acids in the phospholipids are more fluid (i.e. less viscous) and "freeze" at lower temperatures than membranes with more saturated fatty acyl components.

Unsaturated fatty acids generally have a *cis*-configuration of the double bonds, an exception being vaccenic acid (*trans*-11-octadecenoic acid). The most common plant saturated fatty acids are palmitic acid (C_{16}; $CH_3–(CH_2)_{14}–COOH$; *n*-hexadecanoic acid) and stearic acid (C_{18}; $CH_3–(CH_2)_{16}–COOH$; *n*-octadecanoic acid). Common plant unsaturated C_{18} fatty acids include oleic acid (*cis*-9-octadecenoic acid; *cis*-Δ^9-octadecenoic acid), linoleic acid (*cis*-$\Delta^{9,12}$-octadecadienoic acid), α-linolenic acid (*cis*-$\Delta^{9,12,15}$-octadecatrienoic acid) and γ-linolenic acid (*cis*-$\Delta^{6,9,12}$-octadecatrienoic acid). Ricinoleic acid (12-hydroxyoleic acid) is abundant in castor oil. Arachidonic acid (C_{20}; *cis*-$\Delta^{5,8,11,14}$-eicosatetraenoic acid) is absent in higher plants but is the precursor for the pro-inflammatory oxidized prostaglandins, thromboxanes and leukotrienes in animals.

Some plant fatty acids are notably bioactive such as erucic acid (C_{22}; *cis*-Δ^{13}-docosenoic acid) which was greatly reduced by breeding in canola rapeseed because of indications of negative effects in animals (e.g. myocardial fibrosis in long-term dietary experiments with rats). "Lorenzo's oil" (a 4 : 1 mixture of glyceroltrioleate and glyceroltrierucate) apparently does not assist X-linked adrenoleukodystrophy progression in symptomatic patients but may help pre-symptomatic patients. Chaulmoogric acid ((*S*)-13-(cyclopent-2-enyl)tridecanoic

acid) inhibits the growth of the leprosy-causing *Mycobacterium leprae*. The cotton seed oil fatty acids sterculic acid (8-(2-octylcyclopropenyl)octanoic acid; C_8–C3–C_7–COOH) and malvalic acid (7-(2-octylcyclopropenyl)heptanoic acid; C_8–C3–C_6–COOH) inhibit fatty acid desaturase.

Acetylenes. Plants elaborate various acetylenes having the general structure R–(C≡C–)$_n$, where R is an alkyl, aryl or heterocyclic group (e.g. pyran, furan, thiophene or cyclic disulfide) and other functional groups include carboxyl, alcohol, amide, ester, aryl and keto groups. The plant alkynes are often toxic and antifungal. Thus, the alkynes dehydrosafynol (C_{13}), safynol (C_{13}), mycosinol (C_{13}), falcarindiol (C_{17}), falcarinone (C_{17}), wyerone acid (C_{14}) and wyerone acid methyl ester (C_{15}) are antifungal phytoalexins the synthesis of which is induced by fungal pathogen infection. The cytotoxic, antineoplastic toxins virol A, virol B and cicutoxin (C_{17}) from *Cicuta virosa* (water hemlock) (Apiaceae) are acutely toxic through binding to the GABA(A) receptor chloride (Cl^-) channel. Crepenynic acid (C_{18}) is a COX inhibitor and the phytoalexins falcarindiol (C_{17}) and falcarinone (C_{17}) inhibit pro-inflammatory iNOS induction.

The arylacetylene phenylheptatriyne (Phe–C≡C–C≡C–C≡C–CH_3) from *Bidens*, *Dahlia* and *Coreopsis* species (Asteraceae) has phototoxic antimicrobial activity as have 5-(3-buten-1-ynyl)-2,2′-bithienyl (thiophene–thiophene–C≡C–C=CH_2) and the cyclic disulfide acetylenes thiarubrine A (C_3–(C4,S–S)–C_6) and thiarubrine B (C_5–(C4,S–S)–C_4). The photoactivation of acetylenes derives from light absorption by these conjugated systems and ready reaction with oxygen to form reactive intermediates.

Alkyl sulfides and thiols. Some alkyl thiols and sulfides, notably those from commonly ingested *Allium sativum* (garlic) and *Allium cepa* (onion) (Alliaceae), are variously bioactive as odorants and antimicrobials. Propanethial S-oxide (CH_3–CH_2–CH=S=O) is a lachrymatory irritant principle of onion. Allicin (*S*-oxodiallydisulfide; CH_2=CH–CH_2–SO–S–CH_2–CH=CH_2), diallyldisulfide (CH_2=CH–CH_2–S–S–CH_2–CH=CH_2) and diallylsulfide (CH_2=CH–CH_2–S–CH_2–CH=CH_2) are major odorants of garlic that are reactive and irritant because of the allyl groups. Dimethyl disulfide (CH_3–S–S–CH_3), dipropyl disulfide (CH_3–CH_2–CH_2–S–S–CH_2–CH_2–CH_3), methyl allyl disulfide (CH_3–S–S–CH_2–CH=CH_2) and propane-1-thiol (CH_3–CH_2–CH_2–SH) are further *Allium* odorants. Methane thiol (methyl mercaptan; CH_3–SH) is a widespread plant volatile and notably derives from anaerobic bacterial degradation of cysteine as in human flatus and bad mouth odour. The aliphatic disulfides allicin and ajoene inhibit proinflammatory expression of iNOS.

Other aliphatics. In addition to the compounds described above, plants generate a variety of hydrocarbons and other aliphatic compounds ranging from low molecular weight volatiles to high molecular weight alcohols, acids, ketones and esters found in the waxy external cuticle of leaves and fruit.

In addition to the monoterpene and sesquiterpene volatiles described earlier and the thiols and sulfides outlined above, many other low molecular weight volatiles are produced by plants that variously have attractant, repellant or other signalling functions. Cucurbic acid (C_5–C5–CH_2–COOH) is a volatile plant growth regulator as is jasmonic acid (C_5–C5–CH_2–COOH), a major volatile that signals tissue wounding in plants. Volatile plant wounding signals enable herbivore damage to one plant to be communicated to an otherwise untouched plant. Leaf alcohol (*cis*-hex-3-en-1-ol; *cis*-CH_3–CH_2–CH=CH–CH_2–CH_2OH) and leaf aldehyde (*trans*-hex-2-enal; *trans*-CH_3–CH_2–CH_2–CH=CH–CHO) are major green leaf odorants. Nona-2,6-dienal (CH_3–CH_2–CH=CH–CH_2–CH_2–CH=CH–CHO) and jasmone (C_1–C5–C_5) are the characteristic odorants of cucumber

and jasmine, respectively, and octan-1-ol is a major orchid flower (Orchidaceae) bee attractant (Chapter 10).

Higher molecular weight tastants include the peachy flavour γ-undecalactone (C_{11}; 4-hydroxyundecanoic acid lactone; C_7-C4OL) and the coconut flavour γ-nonalactone (C_9; 4-hydroxynonanoic acid lactone; C_5-C4OL). Very high molecular weight aliphatics include long chain fatty acids, alcohols, esters, ketones and hydrocarbons, for example, the plant growth regulator triacont-1-ol (C_{30}; CH_3–$(CH_2)_{28}$–CH_2OH), the Crassulaceae cuticle wax component tritriacontane (C_{33}; CH_3–$(CH_2)_{31}$–CH_3 and Eucalyptus wax (C_{33}; tritriacontane-16,18-dione).

iii. Amino acids and other non-alkaloid amines. The structures of the twenty L-amino acids found in proteins are dealt with in detail in Chapter 2. The diversity of L-amino acids and structurally related non-alkaloid plant amines is briefly outlined below.

α-Amino acids have the general structure $^-OOC–C(H,R)–NH_3^+$. The C carrying the so-called "R group" is the α-C and is a chiral centre (optical activity centre) in all amino acids in which its four substituents are different (glycine in which R=H is not optically active). The other amino acids found in proteins are exclusively L-stereoisomers but D-amino acids can be generated (i.e. through racemization) by heating plant material. D-amino acids are also present in various toxic microbial peptides. The presence of D-amino acid oxidase in animal peroxisomes indicates a need for detoxification of D-amino acid xenobiotics. D-histidine, D-asparagine, D-glutamine and D-phenylalanine are sweet tastants. *N*-Malonyl-D-alanine is present in pea seedlings.

L-Aminoacid analogues such as azetidine 2-carboxylic acid (the C4 ring analogue of the C5 ring L-proline) and L-canavanine (2-amino-4-(guanidinoxy)butyric acid, an analogue of L-arginine) are plant defensive amino acids that are incorporated into protein by the pathogen or herbivore with resultant toxic or debilitating protein mis-folding. L-Homoarginine and γ-hydroxyarginine are also L-arginine analogues.

L-amino acid analogues elaborated by plants inhibit particular enzymes. Thus, L-albizziine (a L-glutamine analogue) inhibits glutamine-dependent asparagine synthase. γ-Hydroxyarginine (a L-arginine analogue) inhibits arginase (the enzyme that catalyses the critical urea cycle detoxifying reaction: arginine ($^-OOC–CH(NH_3^+)–(CH_2)_3–NH–C(=NH)–NH_3^+$) + $H_2O \rightarrow$ ornithine ($^-OOC–CH(NH_3^+)–(CH_2)_3–NH_3^+$) + urea ($H_2N–CO–NH_2$)). L-Canaline ($^-OOC–CH(NH_3^+)–CH_2–CH_2–O–NH_3^+$), an analogue of the non-protein-derived L-amino acid ornithine, inhibits ornthine transcarbamoylase, a key enzyme involved in the ammonia detoxifying urea cycle.

A variety of plant amino acids are neuroactive or neurotoxic including: GABA (γ-aminobutyric acid=4-aminobutyric acid; GABA receptor agonist); β-alanine (3-aminopropionic acid; GABA receptor agonist); glutamate receptor agonists, including glutamate, isowillardiine, willardiine, the Fabaceae neurotoxic, neurolathyrism-inducing compounds L-α-amino-γ-oxalylaminobutyric acid, L-α-amino-γ-oxalylaminopropionic acid and 3-cyano-L-alanine and the cycad neurotoxin L-β-methylaminoalanine; L-dopa (3,4-dihydroxy-L-phenylalanine) (the dopamine precursor used to treat Parkinsonism); L-tryptophan and 5-hydroxytryptophan (antidepressive serotonin precursors); and L-α,γ-diaminobutyric acid (a GABA transport inhibitor).

Some further toxic plant amino acids include the *N*-methylpyridinone mimosine (DNA binding and damaging) and 2-methylenecyclopropylalanine (hypoglycin) and 2-methylenecyclopropylglycine that, respectively, yield 2-methylenecyclopropylacetylCoA and 2-methylenecyclopropylformylCoA (inhibitors of acylCoA dehydrogenases). The cancer

chemopreventative, pro-apoptotic and selenosis-inducing toxic seleno-amino acids Se-methylselenocysteine, L-selenocysteine and L-selenomethionine (from selenium accumulating plants growing on seleniferous soils) yield antimitotic methylseleninic acid (CH_3–Se(=O)–OH), dimethyldiselenide (CH_3–Se–Se–CH_3) and methylselenol (CH_3–SeH) (which generate apoptotic superoxide O_2^-) and SeO_2 (a pro-apoptotic inhibitor of PKC).

Other plant bioactive amines include a variety of neuroactive compounds and polyamines. Notable polyamines include cadaverine (1,5-diaminopentane), putrescine (1,4-diaminobutane), spermidine (NH_2–$(CH_2)_4$–NH–$(CH_2)_3$–NH_2), spermine (NH_2–$(CH_2)_3$–NH–$(CH_2)_4$–NH–$(CH_2)_3$–NH_2) and agmatine (NH_2–C(=NH)–$(CH_2)_4$–NH_2).

The following phenethylamine (Phe–CH_2–CH_2–NH_3^+) derivatives are neuroactive (hormone/neurotransmitter receptor interaction in parenthesis): dopamine (dopamine receptor); norepinephrine, phenethylamine, *Catha edulis* (khat) (Celastraceae) D-cathine and D-cathinone and *Ephedra* species (Ephedraceae) ephedrine and pseudoephedrine (β-adrenergic receptor agonists); *Lophophora williamsii* (Cactaceae) (peyote) hallucinogens mescaline and *N*-methylmescaline (serotonin (5-hydroxytryptamine) 5HT2 receptor agonists).

iv. Cyanogenic and other toxic glycosides. Cyanogenic glycosides have the general structure glycosyl–O–C(X,Y)–CN and are inactive in themselves but break down (either spontaneously in acid conditions or in hydrolytic reactions catalysed by β-glycosidases) to generate cyanide (CN^-). CN^- is a potent inhibitor of cytochrome oxidase that catalyses the final transfer of electrons to molecular oxygen in the mitochondrial respiratory (electron transport) chain. Many cyanogenic glycosides derive biosynthetically from amino acids which have the general structure ^-OOC–C(H,R)–NH_3^+ where R is an alkyl, aromatic or heterocyclic group (see Chapter 2).

The best known cyanogenic glycosides are those occurring in plants of economic importance including: amygdalin (gentiobiosyl–O–C(H,Phe)–CN) from *Prunus amygdalis* (almond) (Rosaceae) seeds; dhurrin (*p*-hydroxymandelonitrile glucoside; glucosyl–O–C(H, *p*–OH–Phe)–CN) from *Sorghum* species (Poaceae); linamarin (manihotoxine) (glucosyl–O–C(CH_3,CH_3)–CN) from *Linum usitatissimum* (flax) (Linaceae) seedlings and in *Manihot esculentum* (cassava) (Euphorbiaceae); linustatin (gentiobiosyl–O–C(CH_3, CH_3)–CN) and neolinustatin (gentiobiosyl–O–C(CH_3, CH_2CH_3)–CN) from flax seeds; prunasin (glucosyl–O–C(H,Phe)–CN) from bark of *Prunus* species (Rosaceae); lucumin (xylosyl-(1→6)-glucosyl−O−C(H,Phe)−CN) from seeds of *Calocarpum sapota* (sapote) (Sapotaceae); lotaustralin (glucosyl−O−C(CH_3,CH_2CH_3)–CN) from *Lotus australis* and *Trifolium repens* (clover) (Fabaceae), flax and *Triticum* species (Poaceae); and vicianin (vicianosyl−O−C(H,Phe)−CN) from seeds of *Vicia* species (vetches) (Fabaceae).

Variants on the above theme are provided by cyanogenic glycosides in which the nitrile (CN) group is attached to an *O*-glycosylated C within a cyclic structure, for example, a cyclopentene as in gynocardin from *Gynocardia odorata* (Flacourtiaceae) seeds and a dihydropyridone as in acalyphin from *Acalypha indica* (Euphorbiaceae) seeds. An interesting exception to the above structural generality is *p*-glucosyloxymandelonitrile (glucosyl–O–Phe–C(H, OH)–CN) from *Goodia latifolia* (Fabaceae) which can generate CN^- without cleavage of the glycosidic link.

Other toxic glycosides include the 3-nitropropanoyl glucosides cibarian and coronarian from *Astragalus* species (Fabaceae) and the *Cycas* species (cycad sago palm) (Cycadaceae) cycasin (methylazoxymethanol-β-D-glucoside; CH_3–$N^+(O^-)$=N–CH_2–O–glucose). Deglycosylation of cycasin and related *Cycas* azoxyglycosides yields methylazoxymethanol

(CH_3–$N^+(O^-)$=N–CH_2–OH), a DNA-damaging, genotoxic, mutagenic, toxic and teratogenic compound.

v. Glucosinolates. Glucosinolates are thioglucosides having the general structure β-D-glucosyl–S–C(R)=N–O–SO_3^-. Thus, R=Phe–CH_2– in benzylglucosinolate. The glucosinolates derive biosynthetically from amino acids (general structure: $^-$OOC–C(H,R)–NH_3^+) as can be seen by comparing the structure of benzylglucosinolate (glucosyl–S–C(CH_2–Phe)=N–O–SO_3^-) with that of the amino acid phenylalanine (R=CH_2–Phe) ($^-$OOC–C(H,CH_2–Phe)–NH_3^+). Myrosinase (thioglucosidase) present in the glucosinolate-producing plant catalyses R-glucosinolate hydrolysis when the plant material is crushed (e.g. by herbivores) with resultant production of the corresponding isothiocyanate R–N=C=S, together with minor by-products, namely R–S=C=N^- (R thiocyanate) and R–CN (R nitrile). The Brassicaceae are a major source of glucosinolates which function as insect deterrents and antifeedants. Isothiocyanates (R–N=C=S) are chemically reactive and can react with thiol and amino groups of proteins.

Glucosinolates are found in familiar *Brassica* species (broccoli, Brussel's sprouts, cabbage, chinese cabbage, cauliflower, mustard, rape cress, swede) as well as in other familiar Brassicaceae species such as *Rapahanus sativus* (radish), *Armoracia lapathifolia* (horseradish) and *Lepidium sativum* (garden cress). Glucosinolate breakdown during cooking and ingestion gives rise to isothiocyanates with characteristic flavours and properties. Thus, methylglucosinolate (glucocapparin) yields methylisothiocyanate that is responsible for the pungent flavour of horseradish and various glucosinolate breakdown products give rise to the characteristic odour of boiled cabbage so intimately redolent of British establishments.

The various substituents (R) of glucosinolate (R-glucosinolate) compounds include alkyl, hydroxyalkyl, aryl (e.g. Phe–CH_2, *p*-HO–Phe, Phe–$(CH_2)_2$), indol-3-yl (Phe|pyrrole), methylsulfonyl alkyl (CH_3–SO_2–$(CH_2)_n$), methylsulfinylalkyl (CH_3–SO–$(CH_2)_n$) and methylthioalkyl (CH_3–S–$(CH_2)_n$) groups. These give rise to the corresponding isothiocyanates (R–N=C=S) that can have particular bioactivities such as insect attractant, insect deterrent, cytotoxic, lachrymatory, tastant and odorant activities.

Of particular note are goitrogenic glucosinolates such as benzylglucosinolate (glucotropaeolin), 3-(methylsulfinyl)propylglucosinolate (glucocheirolin) and progoitrin (2-hydroxybut-3-enylglucosinolate) that yield goitrogenic products that impair thyroid hormone production. Goitrin ((*R*)-5-vinyl-2-oxazolidinethione) is a potent goitrogen and decreases thyroid hormones T3 and T4. Goitrin also induces glutathione S-transferase activity and increases aflatoxin detoxification. Accordingly, moderate *Brassica* consumption is advocated because of the chemopreventative, anticarcinogenic effects of glycosinolate decomposition products.

Other examples of *Brassica* species glucosinolate (R–G) compounds include prop-2-enylG (prop-2-enylglucosinolate) (sinigrin), 4-(methylsulfinyl)butylG (glucoraphanin), 3-(methylsulfinyl)propylG (glucoiberin), 4-(methylsulfinyl)pentylG (glucoalyssin), 4-(methylsulfonyl)butylG (glucoerysolin), 5-(methylthio)butylG (glucoerucin), 5-(methylthio)pentylG (glucoberteroin), indol-3-ylmethylglucosinolate (glucobrassicin), *N*-methoxybrassicin (neoglucobrassicin) and *p*-hydroxybenzylG (sinalbin).

vi. Proteins. Plants produce a number of different kinds of defensive proteins. The most complex of these are polysaccharide hydrolases such as glycan hydrolases (that can hydrolyse the cell walls of invading plant pathogenic fungi), chitinases (that can damage the chitin of the insect digestive system), monosaccharide/oligosaccharide-binding proteins called lectins (that can be potent mitogens), *c.* 40 kDa polygalacturonase-inhibiting proteins

(Chapter 12) and *c.* 20 kDa Kunitz serine protease inhibitor proteins (Chapter 13). Ribosome-inactivating proteins having purine aminoglycosidase activity can be extraordinarily toxic when associated with a lectin subunit enabling entry into the target cell, ricin from seeds of *Ricinus communis* (Euphorbiaceae) being the best known example of such toxic proteins (Chapter 9). Plant thiaminase in ingested plant material degrades thiamine (vitamin B_1) and can consequently cause beriberi from vitamin B_1 deficiency. Thiaminase in insufficiently leached nardoo seed flour (flour made from the sporocarps of the nardoo fern *Marsilea drummondii*) caused peripheral neuropathy in the starving members of the Burke and Wills expedition that crossed Australia from south to north in 1860–1861. Robert O'Hara Burke, William John Wills and Charles Gray died but the sole survivor John King had permanent peripheral neuropathy. Thiamine deficiency disease is also exhibited by livestock feeding on nardoo in "outback" western New South Wales.

Plants also produced a variety of relatively small (3–15 kDa), disulfide-rich, stable defensive proteins that are variously protease and α-amylase inhibitors (thereby inhibiting herbivore digestion and feeding activity) (Chapter 13) or membrane-active entities (such as lipid transfer proteins, defensins, thionins, napins, osmotins and thaumatins) that can damage the cell membranes of pathogenic fungi. The squash family protease inhibitor proteins are among the most potent protease inhibitors known with dissociation constants for the target enzyme-inhibitor complexes of about 10 pM (Chapter 13).

Not dealt with specifically in this book are the plant proteins of importance to humans because of their immunogenicity. Various seed proteins have been shown to cause immunological hypersensitivity after ingestion or inhalation. Thus, a napin protein from rapeseed flour (Chapter 12) causes allergic reactions. The gliadins of wheat flour gluten and the prolamins of barley and rye flour are immunogenic and resultant inflammatory responses affecting the small intestinal mucosa of genetically susceptible people give rise to coeliac disease. Grass pollen is a major outdoor cause of hay fever and allergic asthma and the culprits are protein allergens associated with pollen starch grains (allergenic starch grains released from hydrated pollen being responsible for thunderstorm-associated asthma epidemics). Hevein, a defensive chitin-binding protein present in rubber tree latex, causes allergy to rubber products (Chapter 13).

Chapter 2 deals in part with the structure and function of proteins, including plant defensive proteins and the proteins that are the principal targets of plant defensive compounds.