

What is pharmacology?

OVERVIEW

In this introductory chapter we explain how pharmacology came into being and evolved as a scientific discipline, and describe the present-day structure of the subject and its links to other biomedical sciences. The structure that has emerged forms the basis of the organisation of the rest of the book. Readers in a hurry to get to the here-and-now of pharmacology can safely skip this chapter.

WHAT IS A DRUG?

For the purposes of this book, a drug can be defined as *a chemical substance of known structure, other than a nutrient or an essential dietary ingredient,¹ which, when administered to a living organism, produces a biological effect*.

A few points are worth noting. Drugs may be synthetic chemicals, chemicals obtained from plants or animals, or products of biotechnology (biopharmaceuticals). A medicine is a chemical preparation, which usually, but not necessarily, contains one or more drugs, administered with the intention of producing a therapeutic effect. Medicines usually contain other substances (excipients, stabilisers, solvents, etc.) besides the active drug, to make them more convenient to use. To count as a drug, the substance must be administered as such, rather than released by physiological mechanisms. Many substances, such as insulin or thyroxine, are endogenous hormones but are also drugs when they are administered intentionally. Many drugs are not used commonly in medicine but are nevertheless useful research tools. The definition of drug also covers toxins, which again are not usually administered in the clinic but nonetheless are critical pharmacological tools. In everyday parlance, the word *drug* is often associated with psychoactive substances and addiction – unfortunate negative connotations that tend to bias uninformed opinion against any form of chemical therapy. In this book we focus mainly on drugs used for therapeutic purposes but also describe psychoactive drugs and provide important examples of drugs used as experimental tools. Poisons fall strictly within the definition of drugs, and indeed ‘all drugs are poisons... it is only the dose which makes a thing poison’ (an aphorism credited to Paracelsus, a 16th century Swiss physician); conversely, poisons may be effective therapeutic agents when administered in sub-toxic

doses. Botulinum toxin (Ch. 14) provides a striking example: it is the most potent poison known in terms of its lethal dose, but is widely used both medically and cosmetically. General aspects of harmful effects of drugs are considered in Chapter 58. Toxicology is the study of toxic effects of chemical substances (including drugs), and toxicological testing is undertaken on new chemical entities during their development as potential medicinal products (Ch. 60), but the subject is not otherwise covered in this book.

ORIGINS AND ANTECEDENTS

Pharmacology can be defined as the study of the effects of drugs on the function of living systems. As a science, it was born in the mid-19th century, one of a host of new biomedical sciences based on principles of experimentation rather than dogma that came into being in that remarkable period. Long before that – indeed from the dawn of civilisation – herbal remedies were widely used, pharmacopoeias were written, and the apothecaries’ trade flourished. However, nothing resembling scientific principles was applied to therapeutics, which was known at that time as *materia medica*.² Even Robert Boyle, who laid the scientific foundations of chemistry in the middle of the 17th century, was content, when dealing with therapeutics (*A Collection of Choice Remedies*, 1692), to recommend concoctions of worms, dung, urine and the moss from a dead man’s skull. The impetus for pharmacology came from the need to improve the outcome of therapeutic intervention by doctors, who were at that time skilled at clinical observation and diagnosis but broadly ineffectual when it came to treatment.³ Until the late 19th century, knowledge of the normal and abnormal functioning of the body was too rudimentary to provide even a rough basis for understanding drug effects; at the same time, disease and death were regarded as semi-sacred subjects, appropriately dealt with by authoritarian, rather than scientific, doctrines. Clinical practice often displayed an obedience to authority and ignored what appear to be easily ascertainable facts. For example, cinchona bark was recognised as a specific and effective treatment for malaria, and a sound protocol for its use was laid down by Lind in 1765. In 1804, however, Johnson declared it to be unsafe until the fever had subsided, and he recommended instead the use of large doses of calomel (mercurous chloride) in the early stages – a murderous piece of advice that was slavishly followed for the next 40 years.

¹Like most definitions, this one has its limits. For example, there are a number of essential dietary constituents, such as iron and various vitamins, that are used as medicines. Furthermore, some biological products (e.g. epoietin) show batch-to-batch variation in their chemical constitution that significantly affects their properties. There is also the study of pharmaceutical-grade nutrients or ‘nutraceuticals’.

²The name persists today in some ancient universities, being attached to chairs of what we would call clinical pharmacology.

³Oliver Wendell Holmes, an eminent physician, wrote in 1860: ‘[I] firmly believe that if the whole *materia medica*, as now used, could be sunk to the bottom of the sea, it would be all the better for mankind and the worse for the fishes’ (see Porter, 1997).

The motivation for understanding what drugs can and cannot do came from clinical practice, but the science could be built only on the basis of secure foundations in physiology, pathology and chemistry. It was not until 1858 that Virchow proposed the cell theory. The first use of a structural formula to describe a chemical compound was in 1868. Bacteria as a cause of disease were discovered by Pasteur in 1878. Previously, pharmacology hardly had the legs to stand on, and we may wonder at the bold vision of Rudolf Buchheim, who created the first pharmacology institute (in his own house) in Estonia in 1847.

In its beginnings, before the advent of synthetic organic chemistry, pharmacology concerned itself exclusively with understanding the effects of natural substances, mainly plant extracts – and a few (mainly toxic) chemicals such as mercury and arsenic. An early development in chemistry was the purification of active compounds from plants. Friedrich Sertürner, a young German apothecary, purified morphine from opium in 1805. Other substances quickly followed, and, even though their structures were unknown, these compounds showed that chemicals, not magic or vital forces, were responsible for the effects that plant extracts produced on living organisms. Early pharmacologists focused most of their attention on such plant-derived drugs as quinine, digitalis, atropine, ephedrine, strychnine and others (many of which are still used today and will have become old friends by the time you have finished reading this book).⁴

PHARMACOLOGY IN THE 20TH AND 21ST CENTURIES

Beginning in the 20th century, the fresh wind of synthetic chemistry began to revolutionise the pharmaceutical industry, and with it the science of pharmacology. New synthetic drugs, such as barbiturates and local anaesthetics, began to appear, and the era of antimicrobial chemotherapy began with the discovery by Paul Ehrlich in 1909 of arsenical compounds for treating syphilis. Around the same time, William Blair-Bell was world renowned for his pioneering work at Liverpool in the treatment of breast cancers with another relatively poisonous agent, lead colloid mixtures. The thinking was that yes, drugs were toxic, but they were slightly more toxic to a microbe or cancer cell. This early chemotherapy has laid the foundations for much of the antimicrobial and anticancer therapies still used today. Further breakthroughs came when the sulfonamides, the first antibacterial drugs, were discovered by Gerhard Domagk in 1935, and with the development of penicillin

by Chain and Florey during the Second World War, based on the earlier work of Fleming.

These few well-known examples show how the growth of synthetic chemistry, and the resurgence of natural product chemistry, caused a dramatic revitalisation of therapeutics in the first half of the 20th century. Each new drug class that emerged gave pharmacologists a new challenge, and it was then that pharmacology really established its identity and its status among the biomedical sciences.

In parallel with the exuberant proliferation of therapeutic molecules – driven mainly by chemistry – which gave pharmacologists so much to think about, physiology was also making rapid progress, particularly in relation to chemical mediators, which are discussed in depth throughout this book. Many hormones, neurotransmitters and inflammatory mediators were discovered in this period, and the realisation that chemical communication plays a central role in almost every regulatory mechanism that our bodies possess immediately established a large area of common ground between physiology and pharmacology, for interactions between chemical substances and living systems were exactly what pharmacologists had been preoccupied with from the outset. Indeed, these fields have developed hand-in-hand as wherever there is either a physiological or pathological mechanism, pharmacology could be there to exploit it with a drug. The concept of ‘receptors’ for chemical mediators, first proposed by Langley in 1905, was quickly taken up by pharmacologists such as Clark, Gaddum, Schild and others, and is a constant theme in present-day pharmacology (as you will soon discover as you plough through the next two chapters). The receptor concept, and the technologies developed from it, have had a massive impact on drug discovery and therapeutics. Biochemistry also emerged as a distinct science early in the 20th century, and the discovery of enzymes and the delineation of biochemical pathways provided yet another framework for understanding drug effects. The picture of pharmacology that emerges from this brief glance at history (Fig. 1.1) is of a subject evolved from ancient prescientific therapeutics, involved in commerce from the 17th century onwards, and which gained respectability by donning the trappings of science as soon as this became possible in the mid-19th century. Pharmacology grew rapidly in partnership with the evolution of organic chemistry and other biomedical sciences, and was quick to assimilate the dramatic advances in molecular and cell biology in the late 20th century. Signs of its carpetbagger past still cling to pharmacology, for the pharmaceutical industry has become very big business and much pharmacological research nowadays takes place in a commercial environment, a rougher and more pragmatic place than academia.⁵ No other biomedical ‘ology’ is so close to Mammon.

ALTERNATIVE THERAPEUTIC PRINCIPLES

Modern medicine relies heavily on drugs as the main tool of therapeutics. Other therapeutic procedures, such

⁴A handful of synthetic substances achieved pharmacological prominence long before the era of synthetic chemistry began. Diethyl ether, first prepared as ‘sweet oil of vitriol’ in the 16th century, and nitrous oxide, prepared by Humphrey Davy in 1799, were used to liven up parties before being introduced as anaesthetic agents in the mid-19th century (see Ch. 42). Amyl nitrite (see Ch. 21) was made in 1859 and can claim to be the first ‘rational’ therapeutic drug; its therapeutic effect in angina was predicted on the basis of its physiological effects – a true ‘pharmacologist’s drug’ and the smelly forerunner of the nitrovasodilators that are widely used today. Aspirin (Ch. 27), the most widely used therapeutic drug in history, was first synthesised in 1853, with no therapeutic application in mind. It was rediscovered in 1897 in the laboratories of the German company Bayer, who were seeking a less toxic derivative of salicylic acid. Bayer commercialised aspirin in 1899 and made a fortune.

⁵Some of our most distinguished pharmacological pioneers made their careers in industry: for example, Henry Dale, who laid the foundations of our knowledge of chemical transmission and the autonomic nervous system (Ch. 13); George Hitchings and Gertrude Elion, who described the antimetabolite principle and produced the first effective anticancer drugs (Ch. 57); and James Black, who introduced the first β -adrenoceptor and histamine H₂-receptor antagonists (Chs 15 and 31). It is no accident that in this book, where we focus on the scientific principles of pharmacology, most of our examples are products of industry, not of nature.

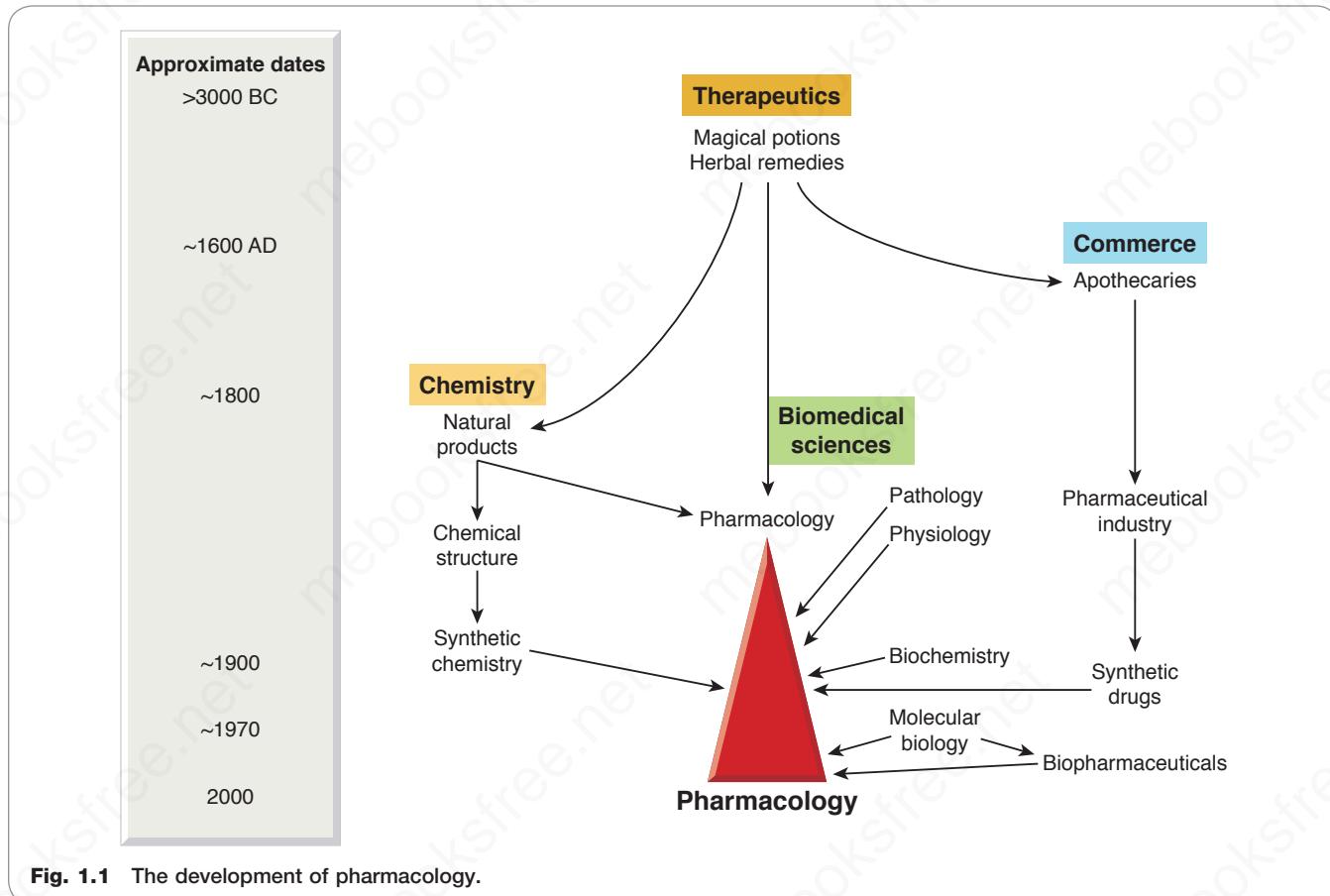


Fig. 1.1 The development of pharmacology.

as surgery, diet, exercise, psychological treatments etc., are also important, of course, as is deliberate non-intervention, but none is so widely applied as drug-based therapeutics.

Before the advent of science-based approaches, repeated attempts were made to construct systems of therapeutics, many of which produced even worse results than pure empiricism. One of these was *allopathy*, espoused by James Gregory (1735–1821). The favoured remedies included bloodletting, emetics and purgatives, which were used until the dominant symptoms of the disease were suppressed. Many patients died from such treatment, and it was in reaction against it that Hahnemann introduced the practice of *homeopathy* in the early 19th century. The implausible guiding principles of homeopathy are:

- like cures like
- activity can be enhanced by dilution

The system rapidly drifted into absurdity: for example, Hahnemann recommended the use of drugs at dilutions of 1:10⁶⁰, equivalent to one molecule in a sphere the size of the orbit of Neptune.

Many other systems of therapeutics have come and gone, and the variety of dogmatic principles that they embodied have tended to hinder rather than advance scientific progress. Currently, therapeutic systems that have a basis that lies outside the domain of science remain popular under the general banner of ‘alternative’ or ‘complementary’ medicine. Mostly, they reject the ‘medical model’, which attributes disease to an underlying derangement of normal function that can be defined in physiological or structural

terms, detected by objective means, and influenced beneficially by appropriate chemical or physical interventions. They focus instead mainly on subjective malaise, which may be disease-associated or not. Abandoning objectivity in defining and measuring disease goes along with a similar departure from scientific principles in assessing therapeutic efficacy and risk, with the result that principles and practices can gain acceptance without satisfying any of the criteria of validity that would convince a critical scientist, and that are required by law to be satisfied before a new drug can be introduced into therapy. Demand for ‘alternative’ therapies by the general public, alas, has little to do with demonstrable efficacy.⁶

THE EMERGENCE OF BIOTECHNOLOGY

Since the 1980s, biotechnology has emerged as a major source of new therapeutic agents in the form of antibodies, enzymes and various regulatory proteins, including hormones, growth factors and cytokines (see Clark & Pazdernik, 2015). Although such products (known as *biopharmaceuticals*, *biologics* or *biologics*) are generally produced by genetic engineering rather than by synthetic chemistry, the pharmacological principles are essentially the same as for conventional drugs, although the details of absorption,

⁶The UK Medicines and Healthcare Regulatory Agency (MHRA) requires detailed evidence of therapeutic efficacy based on controlled clinical trials before a new drug is registered, but no clinical trials data for homeopathic products or for the many herbal medicines that were on sale before the Medicines Act of 1968.

distribution and elimination, specificity, harmful effects and clinical effectiveness all differ markedly between high molecular-weight biopharmaceuticals and low molecular-weight drugs – as does their cost! Looking further ahead, gene- and cell-based therapies (Ch. 5), although still in their infancy, are beginning to take therapeutics into a new domain. The principles governing gene suppression, the design, delivery and control of functioning artificial genes introduced into cells, or of engineered cells introduced into the body, are very different from those of drug-based therapeutics and will require a different conceptual framework, which texts such as this will increasingly need to embrace if they are to stay abreast of modern medical treatment.

PHARMACOLOGY TODAY

As with other biomedical disciplines, the boundaries of pharmacology are not sharply defined, nor are they constant. Its exponents are, as befits pragmatists, ever ready to poach on the territory and techniques of other disciplines. If it ever had a conceptual and technical core that it could really call its own, this has now dwindled almost to the point of extinction, and the subject is defined by its purpose – to understand what drugs do to living organisms, and more particularly how their effects can be applied to therapeutics – rather than by its scientific coherence.

Fig. 1.2 shows the structure of pharmacology as it appears today. Within the main subject fall a number of compartments (neuropharmacology, immunopharmacology,

pharmacokinetics, etc.), which are convenient, if not watertight, subdivisions. These topics form the main subject matter of this book. Around the edges are several interface disciplines, not covered in this book, which form important two-way bridges between pharmacology and other fields of biomedicine. Pharmacology tends to have more of these than other disciplines. Recent arrivals on the fringe are subjects such as pharmacogenomics, pharmacoepidemiology and pharmacoconomics.

Pharmacogenomics. Pharmacogenetics, the study of genetic influences on responses to drugs, initially focused on familial idiosyncratic drug reactions, where affected individuals show an abnormal – usually adverse – response to a class of drug (see Nebert & Weber, 1990). Rebranded as pharmacogenomics, it now covers broader genetically based variations in drug response, where the genetic basis is more complex, the aim being to use genetic information to guide the choice of drug therapy on an individual basis – so-called personalised medicine (Ch. 12). The underlying principle is that differences between individuals in their response to therapeutic drugs can be predicted from their genetic make-up. Examples that confirm this are steadily accumulating (see Ch. 12). So far, they mainly involve genetic polymorphism of drug-metabolising enzymes or receptors. Ultimately, linking specific gene variations with variations in therapeutic or unwanted effects of a particular drug should enable the tailoring of therapeutic choices on the basis of an individual's genotype. Steady improvements in the cost and feasibility of individual genotyping will

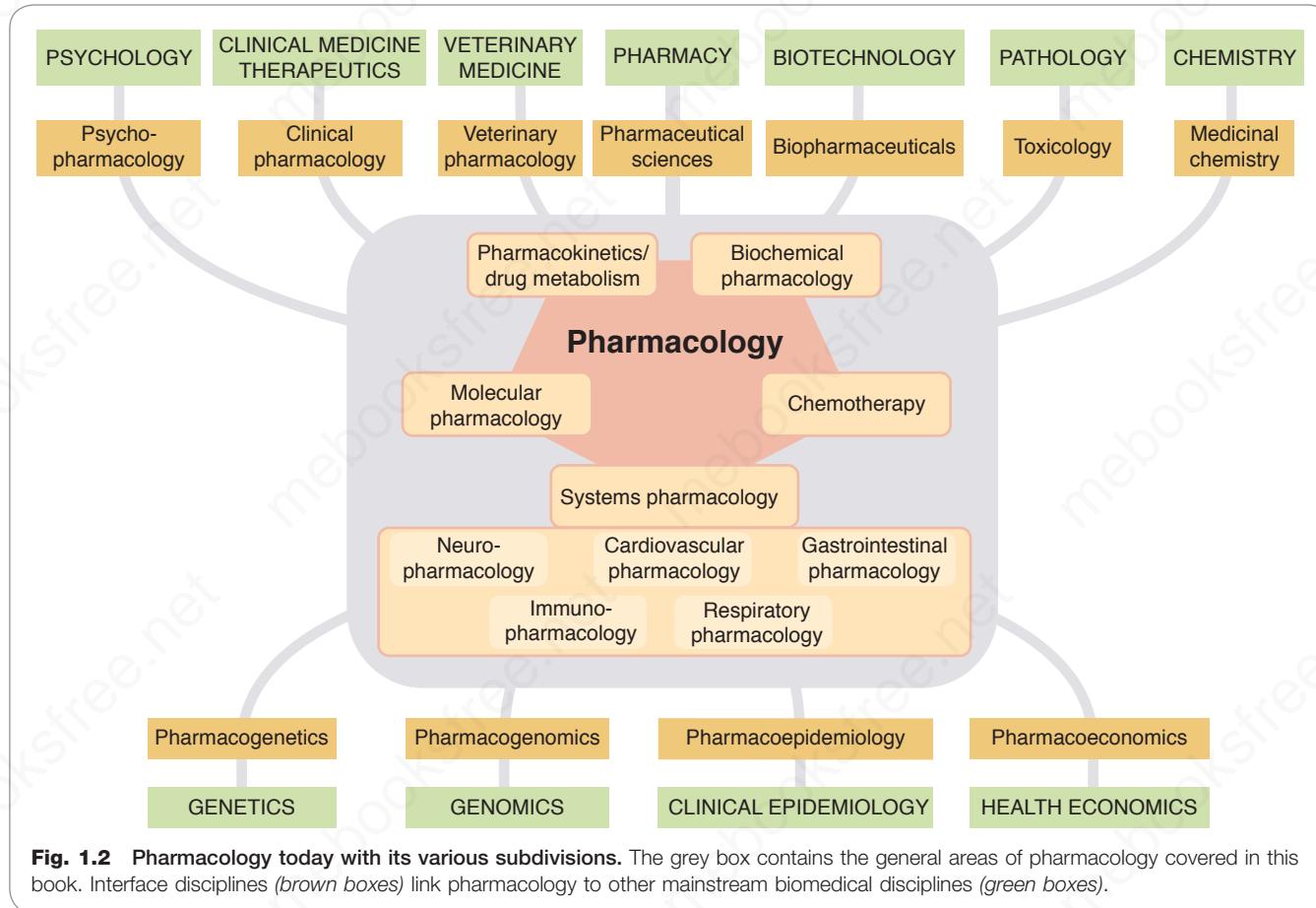


Fig. 1.2 Pharmacology today with its various subdivisions. The grey box contains the general areas of pharmacology covered in this book. Interface disciplines (brown boxes) link pharmacology to other mainstream biomedical disciplines (green boxes).

increase its applicability, potentially with far-reaching consequences for therapeutics (see Ch. 12).

Pharmacoepidemiology. This is the study of drug effects at the population level (see Strom et al., 2013). It is concerned with the variability of drug effects between individuals in a population, and between populations. It is an increasingly important topic in the eyes of the regulatory authorities who decide whether or not new drugs can be licensed for therapeutic use. Variability between individuals or populations detracts from the utility of a drug, even though its overall effect level may be satisfactory. Pharmacoepidemiological studies also take into account patient compliance and other factors that apply when the drug is used under real-life conditions.

Pharmacoeconomics. This branch of health economics aims to quantify in economic terms the cost and benefit of drugs used therapeutically. It arose from the concern of many governments to provide for healthcare from tax revenues, raising questions of what therapeutic procedures represent the best value for money. This, of course, raises fierce controversy, because it ultimately comes down to putting monetary value on health and longevity. As with pharmacoepidemiology, regulatory authorities are increasingly requiring economic analysis, as well as evidence of individual benefit, when making decisions on licensing. For more information on this complex subject, see [Rascati \(2013\)](#).

REFERENCES AND FURTHER READING

- Clark, D.P., Pazderink, N.J., 2015. Biotechnology. Elsevier, New York.
(*General account of biotechnology and its potential applications*)
- Nebert, D.W., Weber, W.W., 1990. Pharmacogenetics. In: Pratt, W.B., Taylor, P. (Eds.), Principles of Drug Action, third ed. Churchill Livingstone, New York. (*A detailed account of genetic factors that affect responses to drugs, with many examples from the pregenomic literature*)
- Porter, R., 1997. The Greatest Benefit to Mankind. Harper-Collins, London. (*An excellent and readable account of the history of medicine, with good coverage of the early development of pharmacology and the pharmaceutical industry*)

- Rascati, K.L., 2013. Essentials of Pharmacoeconomics, second ed. Lippincott Williams & Wilkins, Philadelphia. (*Introduction to a complex and fraught subject*)
- Strom, B.L., Kimmel, S.E., Hennessy, S., 2013. Textbook of Pharmacoepidemiology, second ed. Wiley, Chichester. (*A multiauthor book covering all aspects of a newly emerged discipline, including aspects of pharmacoeconomics*)

2

How drugs act: general principles

OVERVIEW

The emergence of pharmacology as a science came when the emphasis shifted from describing what drugs do to explaining how they work. In this chapter we set out some general principles underlying the interaction of drugs with living systems (Ch. 3 goes into the molecular aspects in more detail). The interaction between drugs and cells is described, followed by a more detailed examination of different types of drug-receptor interaction. The receptor concept has been described as the 'big idea' of pharmacology (Rang, 2006) and will be a recurring theme throughout this book.

INTRODUCTION

To begin with, we should gratefully acknowledge Paul Ehrlich for insisting that drug action must be explicable in terms of conventional chemical interactions between drugs and tissues, and for dispelling the idea that the remarkable potency and specificity of action of some drugs put them somehow out of reach of chemistry and physics and required the intervention of magical 'vital forces'. Although many drugs produce effects in extraordinarily low doses and concentrations, low concentrations still involve very large numbers of molecules. One drop of a solution of a drug at only 10^{-10} mol/L still contains about 3×10^9 drug molecules, so there is no mystery in the fact that it may produce an obvious pharmacological response. Some bacterial toxins (e.g. diphtheria toxin) act with such precision that a single molecule taken up by a target cell is sufficient to kill it.

One of the basic tenets of pharmacology is that drug molecules must exert some chemical influence on one or more cell constituents in order to produce a pharmacological response. In other words, drug molecules must get so close to these constituent cellular molecules that the two interact chemically in such a way that the function of the latter is altered. Of course, the molecules in the organism vastly outnumber the drug molecules, and if the drug molecules were merely distributed at random, the chance of interaction with any particular class of cellular molecule would be negligible. Therefore pharmacological effects require, in general, the non-uniform distribution of the drug molecule within the body or tissue, which is the same as saying that drug molecules must be 'bound' to particular constituents of cells and tissues in order to produce an effect. Ehrlich summed it up thus: '*Corpora non agunt nisi fixata*' (in this context, 'A drug will not work unless it is bound').¹

These critical binding sites are often referred to as 'drug targets' (an obvious allusion to Ehrlich's famous phrase 'magic bullets', describing the potential of antimicrobial drugs). The mechanisms by which the association of a drug molecule with its target leads to a physiological response constitute the major thrust of pharmacological research. Most drug targets are protein molecules. Even general anaesthetics (see Ch. 42), which were long thought to produce their effects by an interaction with membrane lipid, now appear to interact mainly with membrane proteins (see Franks, 2008).

All rules need exceptions, and many antimicrobial and antitumour drugs (Chs 52 and 57), as well as mutagenic and carcinogenic agents (Ch. 58), interact directly with DNA rather than protein; bisphosphonates, used to treat osteoporosis (Ch. 37), bind to calcium salts in the bone matrix, rendering them toxic to osteoclasts, much like rat poison. There are also exceptions among the new generation of *biopharmaceutical drugs* that include nucleic acids, proteins and antibodies (see Ch. 5).

PROTEIN TARGETS FOR DRUG BINDING

Four main kinds of regulatory protein are commonly involved as primary drug targets, namely:

- receptors
- enzymes
- carrier molecules (transporters)
- ion channels

Furthermore, many drugs bind (in addition to their primary targets) to plasma proteins (see Ch. 9) and other tissue proteins, without producing any obvious physiological effect. Nevertheless, the generalisation that most drugs act on one or other of the four types of protein listed above serves as a good starting point.

Further discussion of the mechanisms by which such binding leads to cellular responses is given in Chapters 3–4.

DRUG RECEPTORS

WHAT DO WE MEAN BY RECEPTORS?

▼ As emphasised in Chapter 1, the concept of receptors is central to pharmacology, and the term is most often used to describe the target molecules through which soluble physiological mediators – hormones, neurotransmitters, inflammatory mediators, etc. – produce their effects. Examples such as acetylcholine receptors, cytokine receptors, steroid receptors and growth hormone receptors abound in this book, and generally the term *receptor* indicates a recognition molecule for a chemical mediator through which a response is transduced.

'Receptor' is sometimes used to denote *any* target molecule with which a drug molecule (i.e. a foreign compound rather than an endogenous mediator) has to combine in order to elicit its specific

¹There are, if one looks hard enough, exceptions to Ehrlich's dictum – drugs that act without being bound to any tissue constituent (e.g. osmotic diuretics, osmotic purgatives, antacids and heavy metal chelating agents). Nonetheless, the principle remains true for the great majority.

Targets for drug action



- A drug is a chemical applied to a physiological system that affects its function in a specific way.
- With some exceptions, drugs act on target proteins, namely:
 - receptors
 - enzymes
 - carriers
 - ion channels.
- The term *receptor* is used in different ways. In pharmacology, it describes protein molecules whose function is to recognise and respond to endogenous chemical signals. Other macromolecules with which drugs interact to produce their effects are known as *drug targets*.
- Specificity is reciprocal: individual classes of drug bind only to certain targets, and individual targets recognise only certain classes of drug.
- No drugs are completely specific in their actions. In many cases, increasing the dose of a drug will cause it to affect targets other than the principal one, and this can lead to side effects.

effect. For example, the voltage-sensitive sodium channel is sometimes referred to as the ‘receptor’ for **local anaesthetics** (see Ch. 44), or the enzyme dihydrofolate reductase as the ‘receptor’ for **methotrexate** (Ch. 51). The term *drug target*, of which receptors are one type, is preferable in this context.

In the more general context of cell biology, the term receptor is used to describe various cell surface molecules (such as *T-cell receptors*, *integrins*, *Toll receptors*, etc; see Ch. 7) involved in the cell-to-cell interactions that are important in immunology, cell growth, migration and differentiation, some of which are also emerging as drug targets. These receptors differ from conventional pharmacological receptors in that they respond to proteins attached to cell surfaces or extracellular structures, rather than to soluble mediators.

Various carrier proteins are often referred to as receptors, such as the *low-density lipoprotein receptor* that plays a key role in lipid metabolism (Ch. 24) and the *transferrin receptor* involved in iron absorption (Ch. 26). These entities have little in common with pharmacological receptors. Though quite distinct from pharmacological receptors, these proteins play an important role in the action of drugs such as *statins* (Ch. 24).

RECEPTORS IN PHYSIOLOGICAL SYSTEMS

Receptors form a key part of the system of chemical communication that all multicellular organisms use to coordinate the activities of their cells and organs. Without them, we would be unable to function.

Some fundamental properties of receptors are illustrated by the action of **adrenaline** (epinephrine) on the heart. Adrenaline first binds to a receptor protein (the β_1 *adrenoceptor*, see Ch. 15) that serves as a recognition site for adrenaline and other catecholamines. When it binds to the receptor, a train of reactions is initiated (see Ch. 3), leading to an increase in force and rate of the heartbeat. In the absence of adrenaline, the receptor is normally functionally silent. This is true of most receptors for endogenous mediators (hormones, neurotransmitters, cytokines, etc.), although there are examples (see Ch. 3) of receptors that are ‘constitutively active’ – that is, they exert a controlling

influence even when no chemical mediator is present (see p. 14).

There is an important distinction between *agonists*, which ‘activate’ the receptors, and *antagonists*, which combine at the same site without causing activation, and block the effect of agonists on that receptor. The distinction between agonists and antagonists only exists for pharmacological receptors; we cannot usefully speak of ‘agonists’ for the other classes of drug target described above.

The characteristics and accepted nomenclature of pharmacological receptors are described by [Neubig et al. \(2003\)](#). The origins of the receptor concept and its pharmacological significance are discussed by [Rang \(2006\)](#).

DRUG SPECIFICITY

For a drug to be useful as either a therapeutic or a scientific tool, it must act selectively on particular cells and tissues. In other words, it must show a high degree of binding site specificity. Conversely, proteins that function as drug targets generally show a high degree of ligand specificity; they bind only molecules of a certain precise type.

These principles of binding site and ligand specificity can be clearly recognised in the actions of a mediator such as **angiotensin** (Ch. 23). This peptide acts strongly on vascular smooth muscle, and on the kidney tubule, but has very little effect on other kinds of smooth muscle or on the intestinal epithelium. Other mediators affect a quite different spectrum of cells and tissues, the pattern in each case reflecting the specific pattern of expression of the protein receptors for the various mediators. A small chemical change, such as conversion of one of the amino acids in angiotensin from L to D form, or removal of one amino acid from the chain, can inactivate the molecule altogether, because the receptor fails to bind the altered form. The complementary specificity of ligands and binding sites, which gives rise to the very exact molecular recognition properties of proteins, is central to explaining many of the phenomena of pharmacology. It is no exaggeration to say that the ability of proteins to interact in a highly selective way with other molecules – including other proteins – is the basis of living machines. Its relevance to the understanding of drug action will be a recurring theme in this book.

Finally, it must be emphasised that no drug acts with complete specificity. Thus tricyclic antidepressant drugs (Ch. 48) act by blocking monoamine transporters but are notorious for producing side effects (e.g. dry mouth) related to their ability to block various other receptors. In general, the lower the potency of a drug and the higher the dose needed, the more likely it is that sites of action other than the primary one will assume significance. In clinical terms, this is often associated with the appearance of unwanted ‘off-target’ side effects,² of which no drug is free.

Since the 1970s, pharmacological research has succeeded in identifying the protein targets of many different types of drug. Drugs such as opioid analgesics (Ch. 43), cannabinoids (Ch. 20) and benzodiazepine tranquillisers (Ch. 45), whose actions had been described in exhaustive detail for many years, are now known to target well-defined receptors, many of which have been fully characterised by

²‘On-target’ side effects are unwanted effects mediated through the same receptor as the clinically desired effect, for example constipation and respiratory depression by opioid analgesic drugs (see Ch. 43), whereas ‘off target’ side effects are mediated by a different mechanism.

gene-cloning and protein crystallography techniques (see Ch. 3).

RECEPTOR CLASSIFICATION

▼ Where the action of a drug can be associated with a particular receptor, this provides a valuable means for classification and refinement in drug design. For example, pharmacological analysis of the actions of histamine (see Ch. 18) showed that some of its effects (the H₁ effects, such as smooth muscle contraction) were strongly antagonised by the competitive histamine antagonists then known. Black and his colleagues suggested in 1970 that the remaining actions of histamine, which included its stimulant effect on gastric secretion, might represent a second class of histamine receptor (H₂). Testing a number of histamine analogues, they found that some were selective in producing H₂ effects, with little H₁ activity. By analysing which parts of the histamine molecule conferred this type of specificity, they were able to develop selective H₂ antagonists, which proved to be potent in blocking gastric acid secretion, a development of major therapeutic significance (Ch. 31).³ Two further types of histamine receptor (H₃ and H₄) were recognised later.

Receptor classification based on pharmacological responses continues to be a valuable and widely used approach. Subsequently, newer experimental approaches produced other criteria on which to base receptor classification. The direct measurement of ligand binding to receptors (see later) allowed many new receptor subtypes to be defined that could not easily be distinguished by studies of drug effects. Molecular sequencing of the amino acid structure (see Ch. 3) provided a completely new basis for classification at a much finer level of detail than can be reached through pharmacological analysis. Finally, analysis of the biochemical pathways that are linked to receptor activation (see Ch. 3) provides yet another basis for classification.

The result of this data explosion was that receptor classification suddenly became much more detailed, with a proliferation of receptor subtypes for all the main types of ligand. As alternative molecular and biochemical classifications began to spring up that were incompatible with the accepted pharmacologically defined receptor classes, the International Union of Basic and Clinical Pharmacology (IUPHAR) convened expert working groups to produce agreed receptor classifications for the major types, taking into account the pharmacological, molecular and biochemical information available. These wise people have a hard task; their conclusions will be neither perfect nor final but are essential to ensure a consistent terminology. To the student, this may seem an arcane exercise in taxonomy, generating much detail but little illumination. There is a danger that the tedious lists of drug names, actions and side effects that used to burden the subject will be replaced by exhaustive tables of receptors, ligands and transduction pathways. In this book, we have tried to avoid detail for its own sake and include only such information on receptor classification as seems interesting in its own right or is helpful in explaining the actions of important drugs. A comprehensive database of known receptor classes is available (see <www.guidetopharmacology.org/>), as well as a regularly updated summary (Alexander et al., 2015).

DRUG-RECEPTOR INTERACTIONS

Occupation of a receptor by a drug molecule may or may not result in *activation* of the receptor. By activation, we mean that the receptor is affected by the bound molecule in such a way as to alter the function of the cell and elicit a tissue response. The molecular mechanisms associated with receptor activation are discussed in Chapter 3. Binding and activation represent two distinct steps in the generation of the receptor-mediated response by an agonist (Fig. 2.1). If a drug binds to the receptor without causing activation and thereby prevents the agonist from binding, it is termed a *receptor antagonist*. The tendency of a drug to bind to the

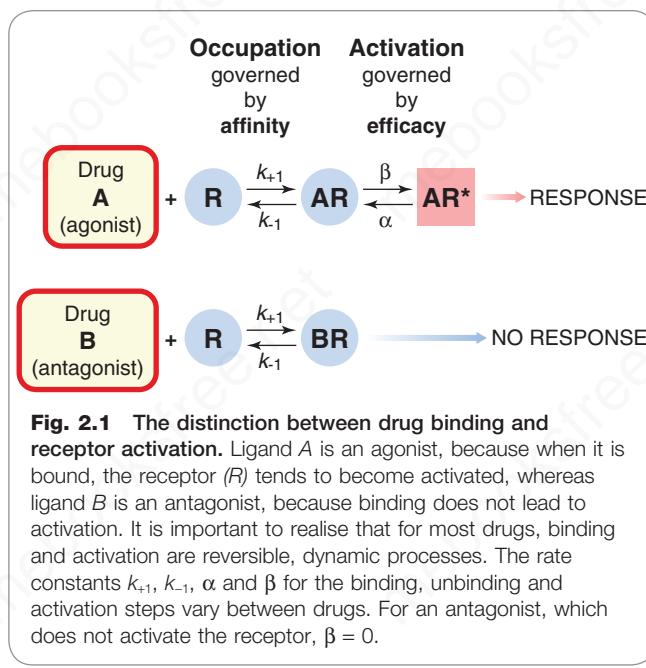


Fig. 2.1 The distinction between drug binding and receptor activation. Ligand A is an agonist, because when it is bound, the receptor (*R*) tends to become activated, whereas ligand B is an antagonist, because binding does not lead to activation. It is important to realise that for most drugs, binding and activation are reversible, dynamic processes. The rate constants k_{+1} , k_{-1} , α and β for the binding, unbinding and activation steps vary between drugs. For an antagonist, which does not activate the receptor, $\beta = 0$.

receptors is governed by its *affinity*, whereas the tendency for it, once bound, to activate the receptor is denoted by its *efficacy*. These terms are defined more precisely later (pp. 9 and 11). Drugs of high potency generally have a high affinity for the receptors and thus occupy a significant proportion of the receptors even at low concentrations. Agonists also possess significant efficacy, whereas antagonists, in the simplest case, have zero efficacy. Drugs with intermediate levels of efficacy, such that even when 100% of the receptors are occupied the tissue response is submaximal, are known as *partial agonists*, to distinguish them from *full agonists*, the efficacy of which is sufficient that they can elicit a maximal tissue response. These concepts, though clearly an oversimplified description of events at the molecular level (see Ch. 3), provide a useful basis for characterising drug effects.

We now discuss certain aspects in more detail, namely drug binding, agonist concentration-effect curves, competitive antagonism, partial agonists and the nature of efficacy. Understanding these concepts at a qualitative level is sufficient for many purposes, but for more detailed analysis a quantitative formulation is needed (see pp. 19–20).

THE BINDING OF DRUGS TO RECEPTORS

▼ The binding of drugs to receptors can often be measured directly by the use of drug molecules (agonists or antagonists) labelled with one or more radioactive atoms (usually ³H, ¹⁴C or ¹²⁵I). The usual procedure is to incubate samples of the tissue (or membrane fragments) with various concentrations of radioactive drug until equilibrium is reached (i.e. when the rate of association [binding] and dissociation [unbinding] of the radioactive drug are equal). The bound radioactivity is measured after removal of the supernatant.

In such experiments, the radiolabelled drug will exhibit both specific binding (i.e. binding to receptors, which is saturable as there are a finite number of receptors in the tissue) and a certain amount of 'non-specific binding' (i.e. drug taken up by structures other than receptors, which, at the concentrations used in such studies, is normally non-saturable), which obscures the specific component and needs to be kept to a minimum (Fig. 2.2A–B). The amount of non-specific

³For this work, and the development of β-adrenoceptor antagonists by a similar experimental approach, Sir James Black was awarded the 1984 Nobel Prize in Physiology or Medicine.

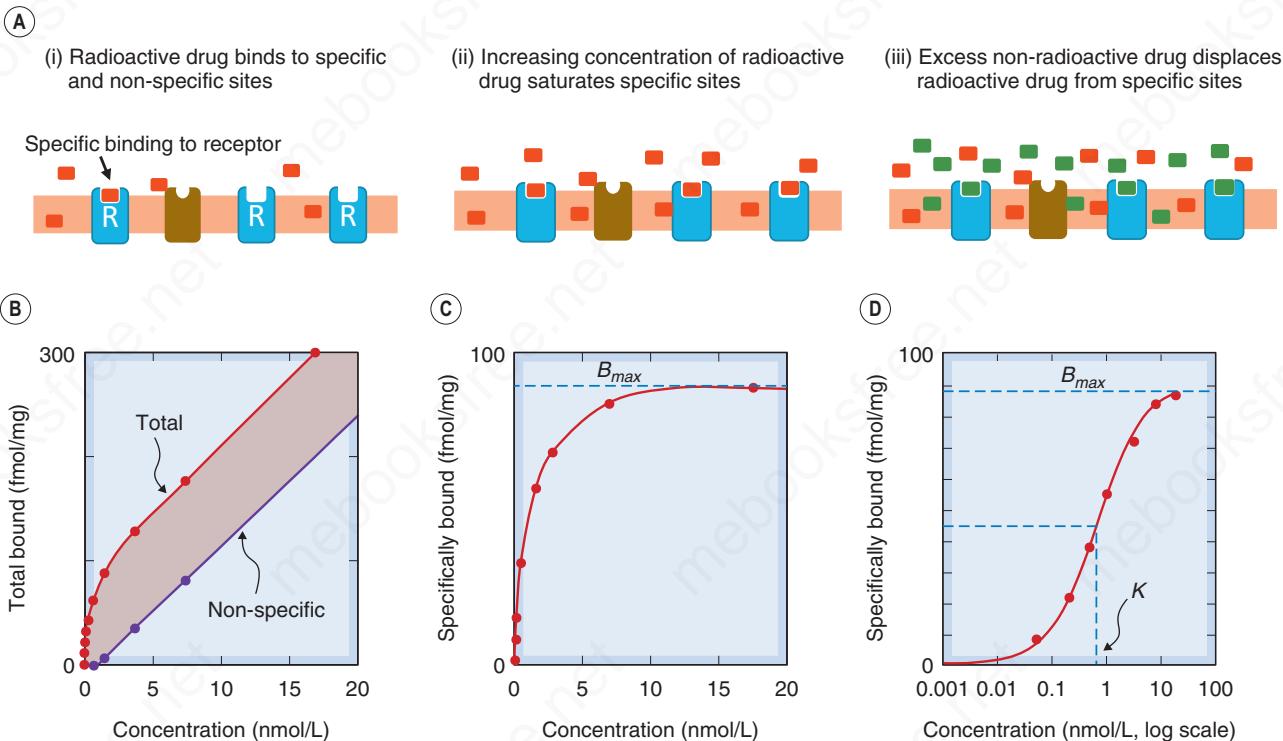


Fig. 2.2 Measurement of receptor binding. (A) (i) Cartoon depicting radioligand (shown in red) binding to its receptor (R) in the membrane as well as to non-specific sites on other proteins and lipid. In (ii) when the concentration of radioligand is increased all the specific sites become saturated but non-specific binding continues to increase. In (iii) addition of a high concentration of a non-radioactive drug (shown in green) that also binds to R displaces the radioactive drug from its receptors but not from the non-specific sites. (B–D) Illustrate actual experimental results for radioligand binding to β adrenoceptors in cardiac cell membranes. The ligand was [^3H]-cyanopindolol, a derivative of pindolol (see Ch. 15). (B) Measurements of total and non-specific binding at equilibrium. Non-specific binding is measured in the presence of a saturating concentration of a non-radioactive β -adrenoceptor agonist, which prevents the radioactive ligand from binding to β adrenoceptors. The difference between the two lines represents specific binding. (C) Specific binding plotted against concentration. The curve is a rectangular hyperbola (Eq. 2.5). (D) Specific binding as in (C) plotted against the concentration on a log scale. The sigmoid curve is a logistic curve representing the logarithmic scaling of the rectangular hyperbola plotted in panel (C) from which the binding parameters K (the equilibrium dissociation constant) and B_{max} (the binding capacity) can be determined.

binding is estimated by measuring the radioactivity taken up in the presence of a saturating concentration of a (non-radioactive) ligand that inhibits completely the binding of the radioactive drug to the receptors, leaving behind the non-specific component. This is then subtracted from the total binding to give an estimate of specific binding (Fig. 2.2C). The *binding curve* (Fig. 2.2C–D) defines the relationship between concentration and the amount of drug bound (B), and in most cases it fits well to the relationship predicted theoretically (see Fig. 2.14), allowing the affinity of the drug for the receptors to be estimated, as well as the *binding capacity* (B_{max}), representing the density of receptors in the tissue. When combined with functional studies, binding measurements have proved very valuable. It has, for example, been confirmed that the *spare receptor hypothesis* (p. 10) for muscarinic receptors in smooth muscle is correct; agonists are found to bind, in general, with rather low affinity, and a maximal biological effect occurs at low receptor occupancy. It has also been shown, in skeletal muscle and other tissues, that denervation leads to an increase in the number of receptors in the target cell, a finding that accounts, at least in part, for the phenomenon of *dennervation supersensitivity*. More generally, it appears that receptors tend to increase in number, usually over the course of a few days, if the relevant hormone or transmitter is absent or scarce, and to decrease in number if the receptors are activated for a prolonged period, a process of adaptation to continued administration of drugs or hormones (see p. 18).

Non-invasive imaging techniques, such as *positron emission tomography* (PET), using drugs labelled with an isotope of short half-life (such as ^{11}C or ^{18}F), can also be used to investigate the distribution of receptors in structures such as the living human brain. This technique has been used, for example, to measure the degree of dopamine-receptor blockade produced by antipsychotic drugs in the brains of schizophrenic patients (see Ch. 47).

Binding curves with agonists often reveal an apparent heterogeneity among receptors. For example, agonist binding to muscarinic receptors (Ch. 14) and also to β adrenoceptors (Ch. 15) suggests at least two populations of binding sites with different affinities. This may be because the receptors can exist either unattached or coupled within the membrane to another macromolecule, the G protein (see Ch. 3), which constitutes part of the transduction system through which the receptor exerts its regulatory effect. Antagonist binding does not show this complexity, probably because antagonists, by their nature, do not lead to the secondary event of G protein coupling. Because agonist binding results in activation, agonist affinity has proved to be a surprisingly elusive concept, about which aficionados love to argue.

THE RELATION BETWEEN DRUG CONCENTRATION AND EFFECT

Although binding can be measured directly, it is usually a biological response, such as a rise in blood pressure,

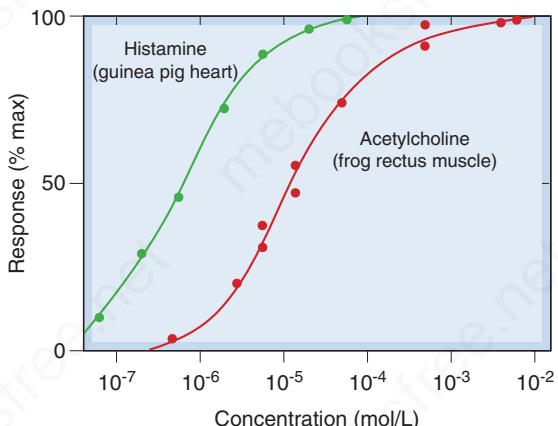


Fig. 2.3 Experimentally observed concentration–effect curves. Although the lines, drawn according to the binding Eq. 2.5, fit the points well, such curves do not give correct estimates of the affinity of drugs for receptors. This is because the relationship between receptor occupancy and response is usually non-linear.

contraction or relaxation of a strip of smooth muscle in an organ bath, the activation of an enzyme, or a behavioural response, that we are interested in, and this is often plotted as a *concentration–effect curve* (*in vitro*) or *dose–response curve* (*in vivo*), as in Fig. 2.3. This allows us to estimate the *maximal response* that the drug can produce (E_{\max}), and the concentration or dose needed to produce a 50% maximal response (EC_{50} or ED_{50}). A logarithmic concentration or dose scale is often used. This transforms the curve from a rectangular hyperbola to a sigmoidal curve in which the mid portion is essentially linear (the importance of the slope of the linear portion will become apparent later in this chapter when we consider antagonism and partial agonists). The E_{\max} , EC_{50} and slope parameters are useful for comparing different drugs that produce qualitatively similar effects (see Fig. 2.7 and Ch. 8). Although they look similar to the binding curve in Fig. 2.2D, concentration–effect curves cannot be used to measure the affinity of agonist drugs for their receptors, because the response produced is not, as a rule, directly proportional to receptor occupancy. This often arises because the maximum response of a tissue may be produced by agonists when they occupy less than 100% of the receptors. Under these circumstances the tissue is said to possess spare receptors (see later).

In interpreting concentration–effect curves, it must be remembered that the concentration of the drug at the receptors may differ from the known concentration in the bathing solution. Agonists may be subject to rapid enzymic degradation or uptake by cells as they diffuse from the surface towards their site of action, and a steady state can be reached in which the agonist concentration at the receptors is very much less than the concentration in the bath. In the case of acetylcholine, for example, which is hydrolysed by cholinesterase present in most tissues (see Ch. 14), the concentration reaching the receptors can be less than 1% of that in the bath, and an even bigger difference has been found with noradrenaline (norepinephrine), which is avidly taken up by sympathetic nerve terminals in many tissues (Ch. 15). The problem is reduced but not entirely eradicated

by the use of recombinant receptors expressed in cells in culture. Thus, even if the concentration–effect curve, as in Fig. 2.3, looks just like a facsimile of the binding curve (see Fig. 2.2D), it cannot be used directly to determine the affinity of the agonist for the receptors.

SPARE RECEPTORS

▼ Stephenson (1956), studying the actions of acetylcholine analogues in isolated tissues, found that many full agonists were capable of eliciting maximal responses at very low occupancies, often less than 1%. This means that the mechanism linking the response to receptor occupancy has a substantial reserve capacity. Such systems may be said to possess *spare receptors*, or a receptor reserve. The existence of spare receptors does not imply any functional subdivision of the receptor pool, but merely that the pool is larger than the number needed to evoke a full response. This surplus of receptors over the number actually needed might seem a wasteful biological arrangement. But in fact it is highly efficient in that a given number of agonist–receptor complexes, corresponding to a given level of biological response, can be reached with a lower concentration of hormone or neurotransmitter than would be the case if fewer receptors were provided. Economy of hormone or transmitter secretion is thus achieved at the expense of providing more receptors.

COMPETITIVE ANTAGONISM

Though one drug can inhibit the response to another in several ways (see p. 16), competition at the receptor level is particularly important, both in the laboratory and in the clinic, because of the high potency and specificity that can be achieved.

In the presence of a competitive antagonist, the agonist occupancy (i.e. proportion of receptors to which the agonist is bound) at a given agonist concentration is reduced, because the receptor can accommodate only one molecule at a time. However, because the two are in competition, raising the agonist concentration can restore the agonist occupancy (and hence the tissue response). The antagonism is therefore said to be *surmountable*, in contrast to other types of antagonism (see later) where increasing the agonist concentration fails to overcome the blocking effect. A simple theoretical analysis (see p. 20) predicts that in the presence of a fixed concentration of the antagonist, the log concentration–effect curve for the agonist will be shifted to the right, without any change in slope or maximum – the hallmark of competitive antagonism (Fig. 2.4A). The shift is expressed as a *dose ratio*, r (the ratio by which the agonist concentration has to be increased in the presence of the antagonist in order to restore a given level of response). Theory predicts that the dose ratio increases linearly with the concentration of the antagonist (see p. 20). These predictions are often borne out in practice (Fig. 2.5A), providing a relatively simple method for determining the equilibrium dissociation constant of the antagonist (K_B ; Fig. 2.5B). Examples of competitive antagonism are very common in pharmacology. The surmountability of the block by the antagonist may be important in practice, because it allows the functional effect of the agonist to be restored by an increase in concentration. With other types of antagonism (as detailed below), the block is usually insurmountable.

The salient features of competitive antagonism are:

- shift of the agonist log concentration–effect curve to the right, without change of slope or maximum (i.e. antagonism can be overcome by increasing the concentration of the agonist)

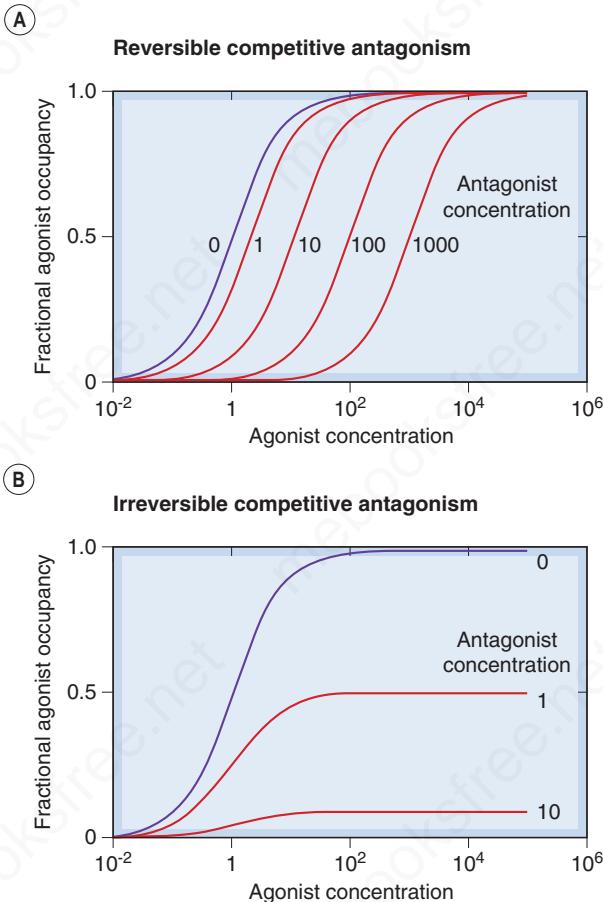


Fig. 2.4 Hypothetical agonist concentration-occupancy curves in the presence of reversible (A) and irreversible (B) competitive antagonists. The concentrations are normalised with respect to the equilibrium dissociation constants, K (i.e. 1.0 corresponds to a concentration equal to K and results in 50% occupancy). Note that in (A) increasing the agonist concentration overcomes the effect of a reversible antagonist (i.e. the block is surmountable), so that the maximal response is unchanged, whereas in (B) the effect of an irreversible antagonist is insurmountable and full agonist occupancy cannot be achieved.

- linear relationship between agonist dose ratio and antagonist concentration
- evidence of competition from binding studies.

Competitive antagonism is the most direct mechanism by which one drug can reduce the effect of another (or of an endogenous mediator).

▼ The characteristics of *reversible competitive antagonism* described above reflect the fact that agonist and competitive antagonist molecules do not stay bound to the receptor but dissociate and rebind continuously. The rate of dissociation of the antagonist molecules is sufficiently high that a new equilibrium is rapidly established on addition of the agonist. In effect, agonist molecules are able to replace the antagonist molecules on the receptors when the antagonist unbinds, although they cannot, of course, evict bound antagonist molecules. Displacement occurs because, by occupying a proportion of the vacant receptors, the agonist effectively reduces the rate of association of the antagonist molecules; consequently, the rate of dissociation temporarily exceeds that of association, and the overall antagonist occupancy falls.

Competitive antagonism

- Reversible competitive antagonism is the commonest and most important type of antagonism; it has two main characteristics.
 - In the presence of the antagonist, the agonist log concentration-effect curve is shifted to the right without change in slope or maximum, the extent of the shift being a measure of the *dose ratio*.
 - The dose ratio increases linearly with antagonist concentration.
- Antagonist affinity, measured in this way, is widely used as a basis for receptor classification.



IRREVERSIBLE COMPETITIVE ANTAGONISM

▼ *Irreversible competitive (or non-equilibrium) antagonism* occurs when the antagonist binds to the same site on the receptor as the agonist but dissociates very slowly, or not at all, from the receptors, with the result that no change in the antagonist occupancy takes place when the agonist is applied.⁴

The predicted effects of reversible and irreversible antagonists are compared in Fig. 2.4.

In some cases (Fig. 2.6A), the theoretical effect is accurately reproduced with the antagonist reducing the maximum response. However, the distinction between reversible and irreversible competitive antagonism (or even non-competitive antagonism) is not always so clear. This is because of the phenomenon of spare receptors (see p. 10); if the agonist occupancy required to produce a maximal biological response is very small (say 1% of the total receptor pool), then it is possible to block irreversibly nearly 99% of the receptors without reducing the maximal response. The effect of a lesser degree of antagonist occupancy will be to produce a parallel shift of the log concentration-effect curve that is indistinguishable from reversible competitive antagonism (Fig. 2.6B). Only when the antagonist occupancy exceeds 99% will the maximum response will be reduced.

Irreversible competitive antagonism occurs with drugs that possess reactive groups that form covalent bonds with the receptor. These are mainly used as experimental tools for investigating receptor function, and few are used clinically. Irreversible enzyme inhibitors that act similarly are clinically used, however, and include drugs such as **aspirin** (Ch. 27), **omeprazole** (Ch. 31) and monoamine oxidase inhibitors (Ch. 48).

PARTIAL AGONISTS AND THE CONCEPT OF EFFICACY

So far, we have considered drugs either as agonists, which in some way activate the receptor when they occupy it, or as antagonists, which cause no activation. However, the ability of a drug molecule to activate the receptor – namely its efficacy – is actually a graded, rather than an all-or-nothing, property. If a series of chemically related agonist drugs acting on the same receptors is tested on a given biological system, it is often found that the largest response that can be produced differs from one drug to another. Some compounds (known as *full agonists*) can produce a maximal response (the largest response that the tissue is capable of giving), whereas others (*partial agonists*) can produce only a submaximal response. Fig. 2.7A shows concentration-effect curves for several α -adrenoceptor agonists (see Ch. 15), which cause contraction of isolated

⁴This type of antagonism is sometimes called non-competitive, but that term is ambiguous and best avoided in this context.

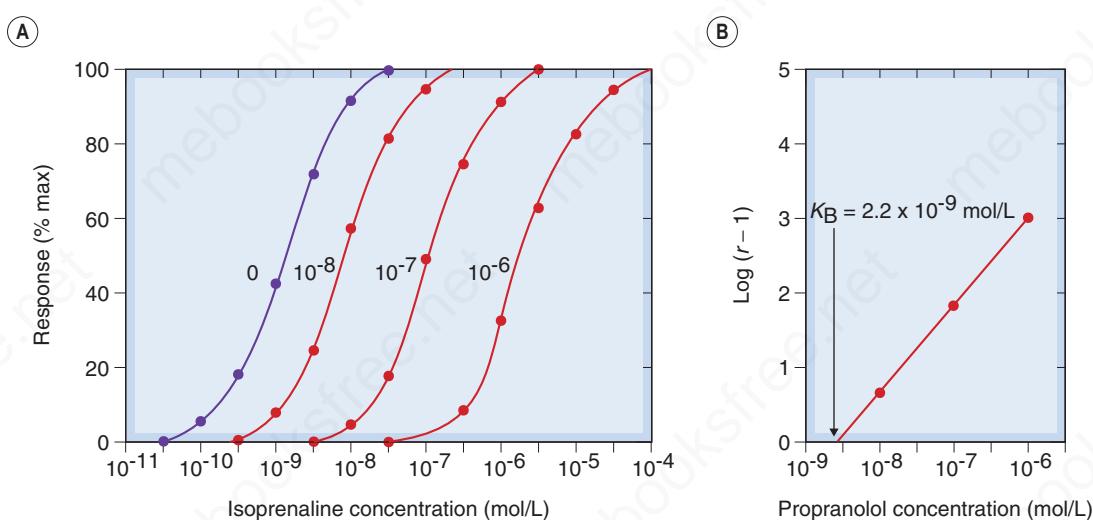


Fig. 2.5 Competitive antagonism of isoprenaline by propranolol measured on isolated guinea pig atria. (A) Concentration–effect curves at various propranolol concentrations (indicated on the curves). Note the progressive shift to the right without a change of slope or maximum. (B) Schild plot (Eq. 2.10). The equilibrium dissociation constant (K_B) for propranolol is given by the abscissal intercept, 2.2×10^{-9} mol/L. Note that the subscript ‘B’ is now used in ‘ K_B ’ to indicate that the equilibrium dissociation constant is that of the antagonist (designated drug B) measured in the presence of the agonist (designated drug A). (Results from Potter, L.T., 1967. Uptake of propranolol by isolated guinea pig atria. J. Pharmacol. Exp. Ther. 55, 91–100.)

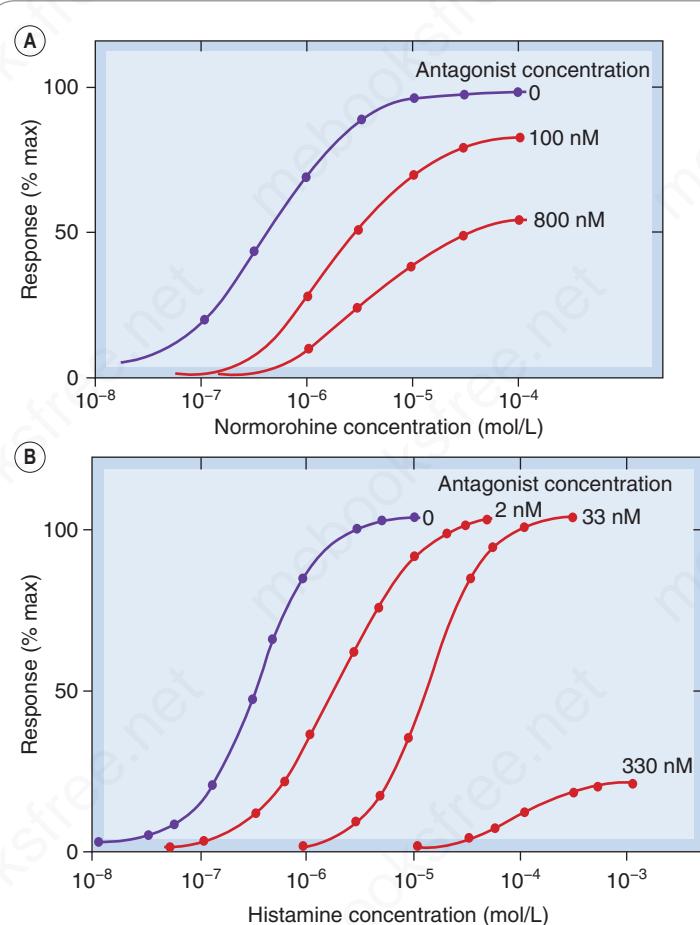


Fig. 2.6 Effects of irreversible competitive antagonists on agonist concentration–effect curves. (A) Rat brain neurones responding to the opioid agonist normorphine before and after being exposed to the irreversible competitive antagonist β -funaltrexamine for 30 minutes and then washed to remove the antagonist. Note the depression of the maximum response. (B) Responses of the guinea pig ileum to histamine before and after treatment with increasing concentrations of a receptor alkylating agent (GD121) for 5 minutes and then washed to remove the antagonist. Note the concentration–response curve is initially shifted to the right with no depression of the maximum response. (Panel [A] after Williams, J.T., North, R.A., 1984. Mol. Pharmacol. 26, 489–497; panel [B] after Nickerson, M., 1955. Nature 178, 696–697.)

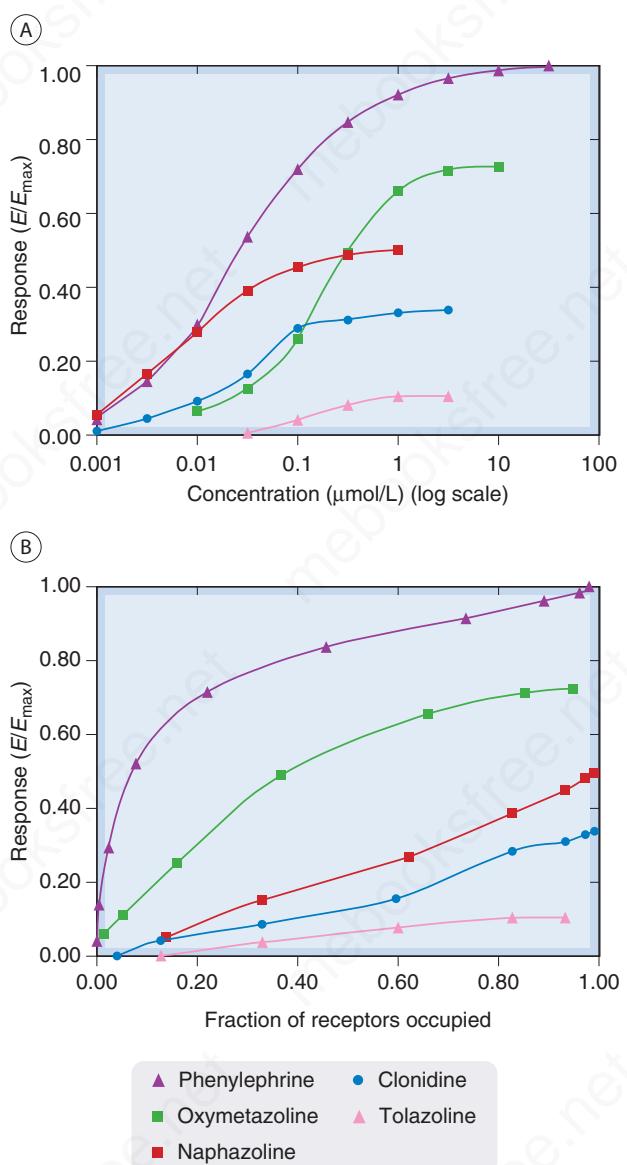


Fig. 2.7 Partial agonists. (A) Log concentration–effect curves for a series of α -adrenoceptor agonists causing contraction of an isolated strip of rabbit aorta. Phenylephrine is a full agonist. The others are partial agonists with different efficacies. The lower the efficacy of the drug the lower the maximum response and slope of the log concentration–response curve. (B) The relationship between response and receptor occupancy for the series. Note that the full agonist, phenylephrine, produces a near-maximal response when only about half the receptors are occupied, whereas partial agonists produce submaximal responses even when occupying all of the receptors. The efficacy of tolazoline is so low that it is classified as an α -adrenoceptor antagonist (see Ch. 15). In these experiments, receptor occupancy was not measured directly, but was calculated from pharmacological estimates of the equilibrium constants of the drugs. (Data from Ruffolo, R.R. Jr, et al., 1979. J. Pharmacol. Exp. Ther. 209, 429–436.)

strips of rabbit aorta. The full agonist **phenylephrine** produced the maximal response of which the tissue was capable; the other compounds could only produce submaximal responses and are partial agonists. The difference between full and partial agonists lies in the relationship between receptor occupancy and response. In the experiment shown in Fig. 2.7 it was possible to estimate the affinity of the various drugs for the receptor, and hence (based on the theoretical model described later; p. 19) to calculate the fraction of receptors occupied (known as *occupancy*) as a function of drug concentration. Plots of response as a function of occupancy for the different compounds are shown in Fig. 2.7B, showing that for partial agonists the response at a given level of occupancy is less than for full agonists. The weakest partial agonist, **tolazoline**, produces a barely detectable response even at 100% occupancy, and is usually classified as a *competitive antagonist* (see p. 10 and Ch. 15).

These differences can be expressed quantitatively in terms of *efficacy* (e), a parameter originally defined by **Stephenson (1956)** that describes the ‘strength’ of the agonist–receptor complex in evoking a response of the tissue. In the simple scheme shown in Fig. 2.1, efficacy describes the tendency of the drug–receptor complex to adopt the active (AR^*), rather than the resting (AR), state. A drug with zero efficacy ($e = 0$) has no tendency to cause receptor activation, and causes no tissue response. A full agonist is a drug whose efficacy⁵ is sufficient that it produces a maximal response when less than 100% of receptors are occupied. A partial agonist has lower efficacy, such that 100% occupancy elicits only a submaximal response.

▼ Subsequently it was appreciated that efficacy is composed of drug-dependent and tissue-dependent components. The drug-dependent component is referred to as the *intrinsic efficacy*, which is the ability of the agonist drug molecule, once bound, to activate the receptor protein (see Kelly, 2013). The tissue-dependent components of efficacy include the number of receptors that it expresses and the efficiency of coupling of receptor activation to the measured tissue response. The number of receptors expressed is especially relevant to the study of receptors in recombinant expression systems when receptors are often very highly expressed and intermediate efficacy agonists then appear as full agonists. Across different cell types expressing the same receptor but at different densities a given drug of intermediate efficacy may appear as a full agonist in one tissue (high level of receptor expression), a partial agonist in another (lower level of receptor expression), and even as an antagonist in another (very low level of receptor expression). The term ‘partial agonist’ is therefore only applicable when describing the action of a drug on a specific tissue or cell type.

For G protein-coupled receptors the elucidation of their X-ray crystal structures (described in Ch. 3) and the application of molecular dynamic simulations of drug binding are beginning to tease out the molecular basis of receptor activation and why some ligands are agonists and some are antagonists. For students starting to study pharmacology the simple theoretical two-state model described below provides a useful starting point.

PARTIAL AGONISTS AS ANTAGONISTS

In discussing the efficacy of partial agonists above we considered the situation in which the tissue was exposed

⁵In Stephenson’s formulation, efficacy is the reciprocal of the occupancy needed to produce a 50% maximal response, thus $e = 25$ implies that a 50% maximal response occurs at 4% occupancy. There is no theoretical upper limit to efficacy.

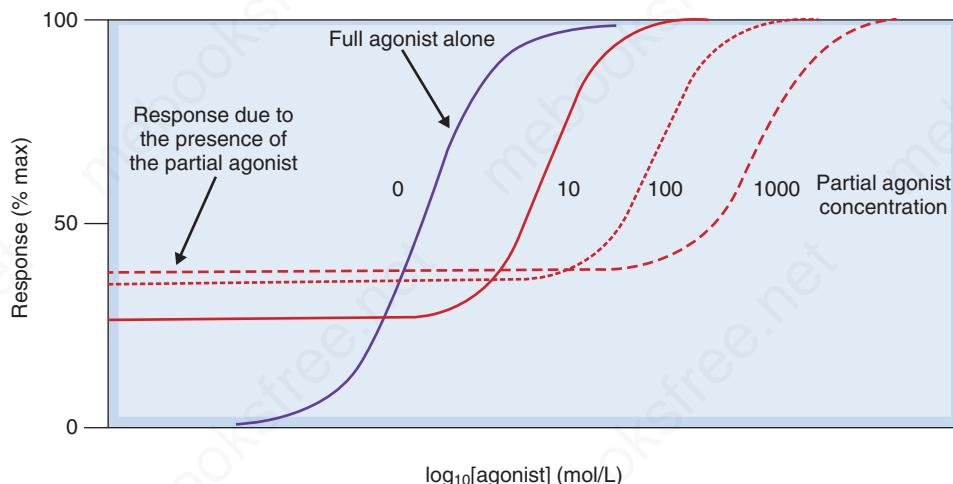


Fig. 2.8 Hypothetical concentration–response curves for a full agonist in the absence and presence of increasing concentrations of a partial agonist. The partial agonist will have agonist action and hence the initial response increases as the partial agonist concentration increases, reaching a maximum equal to the maximum response of the partial agonist. However, when the full agonist is added in the presence of the partial agonist its concentration–response curve is shifted to the right.

to only one drug, the partial agonist. What we should also consider is how the presence of a partial agonist would alter the response of a tissue to a higher efficacy agonist. This is depicted in Fig. 2.8 where it can be seen that the presence of the partial agonist induces some level of response dependent upon the concentration initially applied but in addition because the partial agonist is competing with the full agonist for the receptors it effectively acts as a competitive antagonist, shifting the concentration–response curve of the full agonist to the right. This is not just an obscure theoretical point but something which occurs in clinical practice. In the treatment of heroin users, buprenorphine, a weak partial agonist, not only acts as a weak opioid substitute but also acts as an antagonist and reduces the likelihood of overdose when users relapse and take heroin again (see Ch. 50).

CONSTITUTIVE RECEPTOR ACTIVATION AND INVERSE AGONISTS

▼ Although we are accustomed to thinking that receptors are activated only when an agonist molecule is bound, there are examples (see De Ligt et al., 2000) where an appreciable level of activation (*constitutive activation*) may exist even when no ligand is present. These include receptors for benzodiazepines (see Ch. 45), cannabinoids (Ch. 20), serotonin (Ch. 16) and several other mediators. Furthermore, receptor mutations occur – either spontaneously, in some disease states (see Bond & Ijzerman, 2006), or experimentally created (see Ch. 4) – that result in appreciable constitutive activation. If a ligand reduces the level of constitutive activation, such drugs are known as *inverse agonists* (Fig. 2.9; see De Ligt et al., 2000) to distinguish them from *neutral antagonists*, which do not by themselves affect the level of activation. Inverse agonists can be regarded as drugs with negative efficacy, to distinguish them from agonists (positive efficacy) and neutral antagonists (zero efficacy). Neutral antagonists, by binding to the agonist binding site, will antagonise both agonists and inverse agonists. Inverse agonism was first observed at the benzodiazepine receptor (Ch. 45) but such drugs are proconvulsive and thus not therapeutically useful! New examples of constitutively active receptors and inverse agonists are emerging with increasing frequency (mainly among G protein-coupled receptors). **Pimavanserin**, an inverse agonist at the 5-HT_{2A} receptor, has recently been developed for the treatment of

psychosis associated with Parkinson's disease (see Chs 41 and 47). It turns out that most of the receptor antagonists in clinical use are actually inverse agonists when tested in systems showing constitutive receptor activation. However, most receptors – like cats – show a preference for the inactive state, and for these there is no practical difference between a competitive antagonist and an inverse agonist. The following section describes a simple model that explains full, partial and inverse agonism in terms of the relative affinity of different ligands for the resting and activated states of the receptor.

The two-state receptor model

▼ As illustrated in Fig. 2.1, agonists and antagonists both bind to receptors, but only agonists activate them. How can we express this difference, and account for constitutive activity, in theoretical terms? The two-state model (Fig. 2.10) provides a simple but useful approach. As shown in Fig. 2.1, we envisage that the occupied receptor can switch from its 'resting' (R) state to an activated (R*) state, R* being favoured by binding of an agonist but not an antagonist molecule. As described above, receptors may show constitutive activation (i.e. the R* conformation can exist without any ligand being bound), so the added drug encounters an equilibrium mixture of R and R* (see Fig. 2.10). If it has a higher affinity for R* than for R, the drug will cause a shift of the equilibrium towards R* (i.e. it will promote activation and be classed as an agonist). If its preference for R* is very large, nearly all the occupied receptors will adopt the R* conformation and the drug will be a full agonist; if it shows only a modest degree of selectivity for R* (say 5- to 10-fold), a smaller proportion of occupied receptors will adopt the R* conformation and it will be a partial agonist; if it shows no preference, the prevailing R : R* equilibrium will not be disturbed and the drug will be a neutral antagonist (zero efficacy), whereas if it shows selectivity for R it will shift the equilibrium towards R and be an inverse agonist (negative efficacy). We can therefore think of efficacy as a property determined by the relative affinity of a ligand for R and R*, a formulation known as the *two-state model*, which is useful in that it puts a physical interpretation on the otherwise mysterious meaning of efficacy, as well as accounting for the existence of inverse agonists.

BIASED AGONISM

A major problem with the two-state model is that, as we now know, receptors are not actually restricted to two distinct states but have much greater conformational

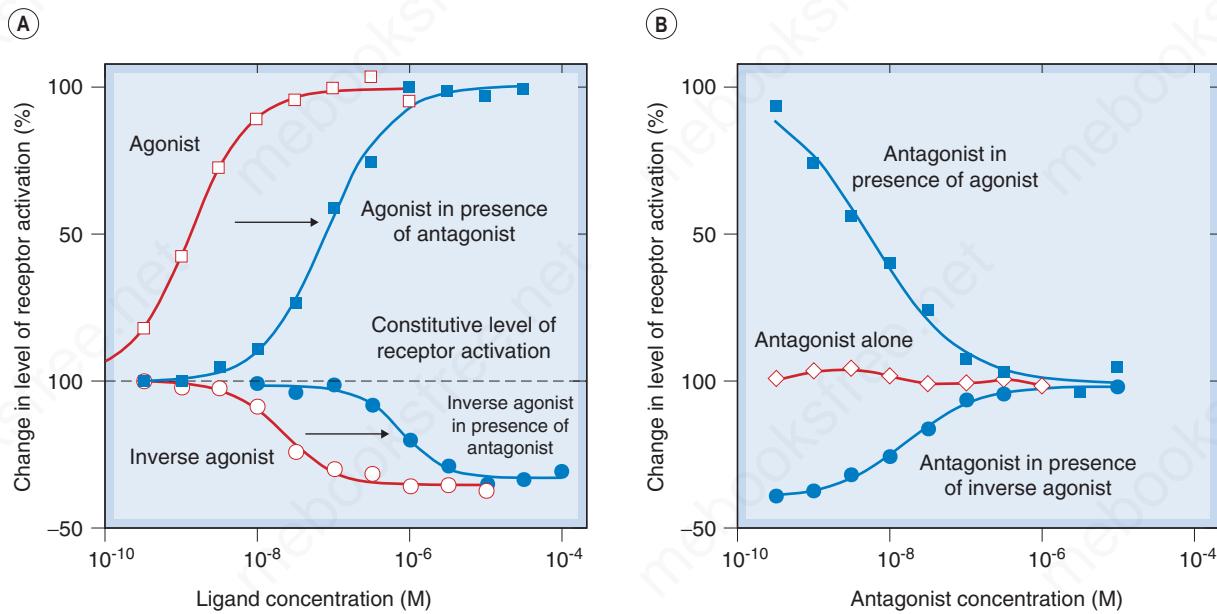


Fig. 2.9 Inverse agonism. The interaction of a competitive antagonist with normal and inverse agonists in a system that shows receptor activation in the absence of any added ligands (constitutive activation). (A) The degree of receptor activation (vertical scale) increases in the presence of an agonist (open squares) and decreases in the presence of an inverse agonist (open circles). Addition of a competitive antagonist shifts both curves to the right (closed symbols). (B) The antagonist on its own does not alter the level of constitutive activity (open symbols), because it has equal affinity for the active and inactive states of the receptor. In the presence of an agonist (closed squares) or an inverse agonist (closed circles), the antagonist restores the system towards the constitutive level of activity. These data (reproduced with permission from Newman-Tancredi, A., et al., 1997. Br. J. Pharmacol. 120, 737–739) were obtained with cloned human 5-hydroxytryptamine (5-HT) receptors expressed in a cell line. (Agonist, 5-carboxamidotryptamine; inverse agonist, spiperone; antagonist, WAY 100635; ligand concentration [M] = mol/L; see Ch. 16 for information on 5-HT receptor pharmacology.)

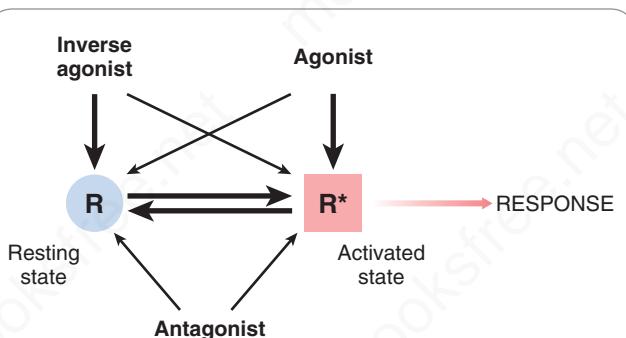


Fig. 2.10 The two-state model. The receptor is shown in two conformational states, resting (R) and activated (R^*), which exist in equilibrium. Normally, when no ligand is present, the equilibrium lies far to the left, and few receptors are found in the R^* state. For constitutively active receptors, an appreciable proportion of receptors adopt the R^* conformation in the absence of any ligand. Agonists have higher affinity for R^* than for R , so shift the equilibrium towards R^* . The greater the relative affinity for R^* with respect to R , the greater the efficacy of the agonist. An inverse agonist has higher affinity for R than for R^* and so shifts the equilibrium to the left. A neutral antagonist has equal affinity for R and R^* so does not by itself affect the conformational equilibrium but reduces by competition the binding of other ligands.

flexibility, so that there is more than one inactive and active conformation. The different conformations that they can adopt may be preferentially stabilised by different ligands, and may produce different functional effects by activating different signal transduction pathways (see Ch. 3).

Receptors that couple to second messenger systems (see Ch. 3) can couple to more than one intracellular effector pathway, giving rise to two or more simultaneous responses. One might expect that all agonists that activate the same receptor type would evoke the same array of responses (Fig. 2.11A). However, it has become apparent that different agonists can exhibit bias for the generation of one response over another even though they are acting through the same receptor (Fig. 2.11B), probably because they stabilise different activated states of the receptor (see Kelly, 2013). Agonist bias has become an important concept in pharmacology.

Redefining and attempting to measure agonist efficacy for such a multistate model is problematic, however, and requires a more complicated state transition model than the two-state model described above. The errors, pitfalls and a possible way forward have been outlined by Kenakin & Christopoulos (2013).

ALLOSTERIC MODULATION

- ▼ In addition to the agonist binding site (now referred to as the *orthosteric* binding site), to which competitive antagonists also bind, receptor proteins possess many other (*allosteric*) binding sites (see Ch. 3) through which drugs can influence receptor function in various ways, by increasing or decreasing the affinity of agonists

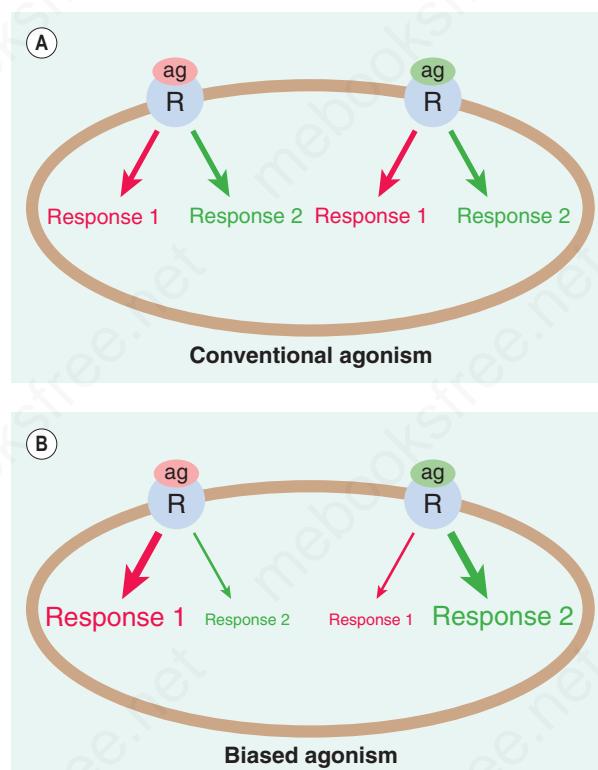


Fig. 2.11 Biased agonism. In (A), the receptor (*R*) is coupled to two intracellular responses – *response 1* and *response 2*. When different agonists indicated in red and green activate the receptor they evoke both responses in a similar manner. This is what we can consider as being conventional agonism. In (B), biased agonism is illustrated in which two agonists bind at the same site on the receptor yet the red agonist is better at evoking response 1 and the green agonist is better at evoking response 2.

for the agonist binding site, by modifying efficacy or by producing a response themselves (Fig. 2.12). Depending on the direction of the effect, the ligands may be allosteric antagonists or allosteric facilitators of the agonist effect, and the effect may be to alter the slope and maximum of the agonist log concentration–effect curve (see Fig. 2.12). This type of allosteric modulation of receptor function has attracted much attention recently and occurs at different types of receptors (see review by [Changeux & Christopoulos, 2016](#)). Well-known examples of allosteric facilitation include glycine at NMDA receptors (Ch. 39), benzodiazepines at GABA_A receptors (Ch. 45) and **cinacalcet** at the Ca²⁺ receptor (Ch. 37). One reason why allosteric modulation may be important to the pharmacologist and future drug development is that across families of receptors such as the muscarinic receptors (see Ch. 14) the orthosteric binding sites are very similar and it has proven difficult to develop selective agonists and antagonists for individual subtypes. The hope is that there will be greater variation in the allosteric sites and that receptor-selective allosteric ligands can be developed. Furthermore, positive allosteric modulators will exert their effects only on receptors that are being activated by endogenous ligands and have no effect on those that are not activated. This might provide a degree of selectivity (e.g. in potentiating spinal inhibition mediated by endogenous opioids, see Ch. 43) and a reduction in side effect profile.

OTHER FORMS OF DRUG ANTAGONISM

Other mechanisms can also account for inhibitory interactions between drugs.

Agonists, antagonists and efficacy



- Drugs acting on receptors may be *agonists* or *antagonists*.
- Agonists initiate changes in cell function, producing effects of various types; antagonists bind to receptors without initiating such changes.
- Agonist potency depends on two parameters: *affinity* (i.e. tendency to bind to receptors) and *efficacy* (i.e. ability, once bound, to initiate changes that lead to effects).
- For antagonists, efficacy is zero.
- *Full agonists* (which can produce maximal effects) have high efficacy; *partial agonists* (which can produce only submaximal effects) have intermediate efficacy.
- According to the two-state model, efficacy reflects the relative affinity of the compound for the resting and activated states of the receptor. Agonists show selectivity for the activated state; antagonists show no selectivity. This model, although helpful, fails to account for the complexity of agonist action.
- *Inverse agonists* show selectivity for the resting state of the receptor, this being of significance only in situations where the receptors show *constitutive activity*.
- *Allosteric modulators* bind to sites on the receptor other than the agonist binding site and can modify agonist activity.

The most important ones are:

- chemical antagonism
- pharmacokinetic antagonism
- block of receptor-response linkage
- physiological antagonism

CHEMICAL ANTAGONISM

Chemical antagonism refers to the uncommon situation where the two substances combine in solution; as a result, the effect of the active drug is lost. Examples include the use of chelating agents (e.g. **dimercaprol**) that bind to heavy metals and thus reduce their toxicity, and the use of the neutralising antibody **infliximab**, which has an anti-inflammatory action due to its ability to sequester the inflammatory cytokine tumour necrosis factor (TNF; see Ch. 19).

PHARMACOKINETIC ANTAGONISM

Pharmacokinetic antagonism describes the situation in which the 'antagonist' effectively reduces the concentration of the active drug at its site of action. This can happen in various ways. The rate of metabolic degradation of the active drug may be increased (e.g. the reduction of the anticoagulant effect of **warfarin** when an agent that accelerates its hepatic metabolism, such as **phenytoin**, is given; see Chs 10 and 58). Alternatively, the rate of absorption of the active drug from the gastrointestinal tract may be reduced, or the rate of renal excretion may be increased. Interactions of this sort, discussed in more detail in Chapter 58, are common and can be important in clinical practice.

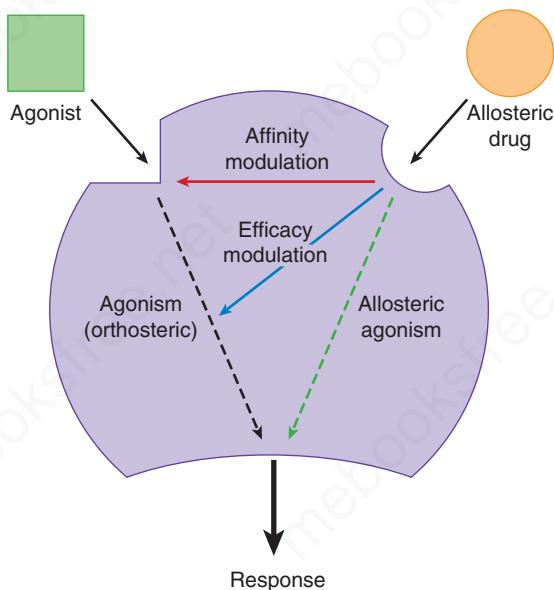
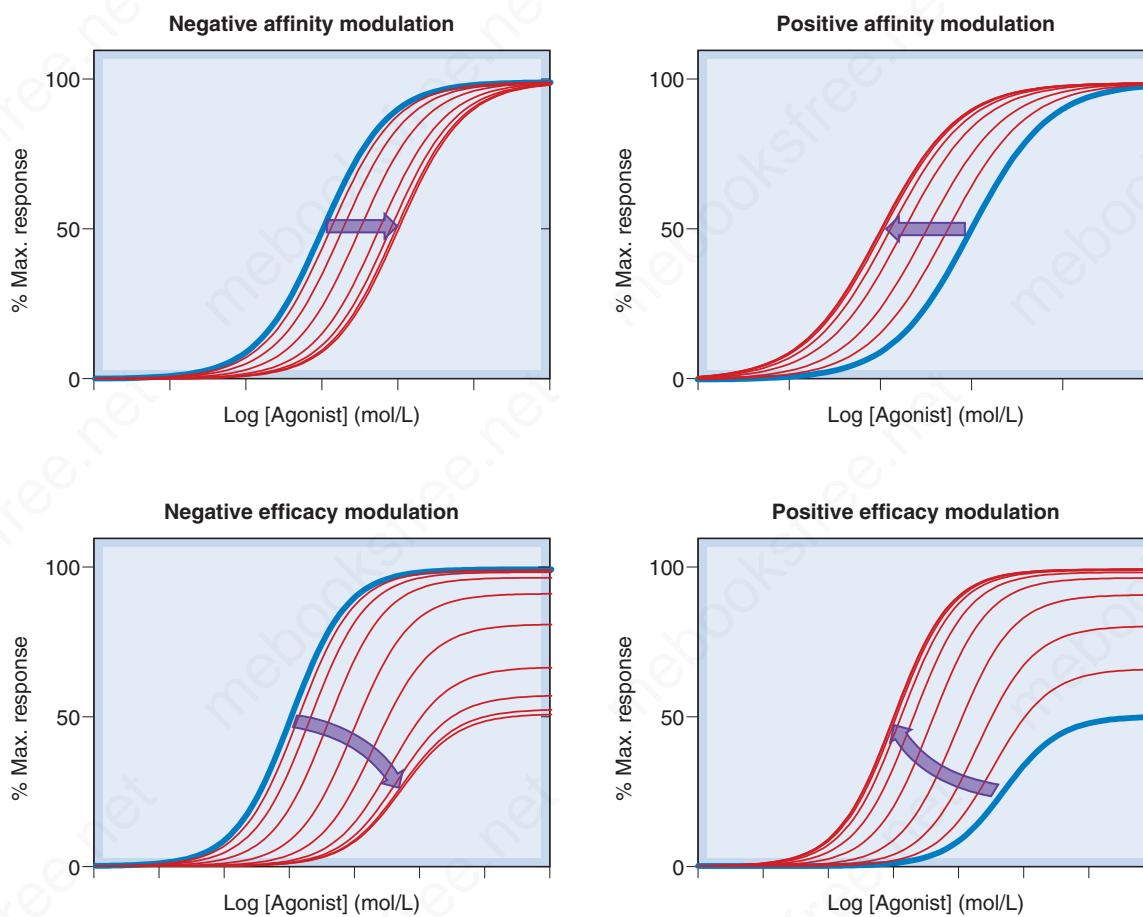
A**B**

Fig. 2.12 Allosteric modulation. (A) Allosteric drugs bind at a separate site on the receptor to ‘traditional’ agonists (now often referred to as ‘orthosteric’ agonists). They can modify the activity of the receptor by (i) altering agonist affinity, (ii) altering agonist efficacy or (iii) directly evoking a response themselves. (B) Effects of affinity- and efficacy-modifying allosteric modulators on the concentration–effect curve of an agonist (blue line). In the presence of the allosteric modulator the agonist concentration–effect curve (now illustrated in red) is shifted in a manner determined by the type of allosteric modulator until a maximum effect of the modulator is reached. (Panel [A] adapted with permission from Conn et al., 2009. *Nat. Rev. Drug Discov.* 8, 41–54; panel [B] courtesy of Christopoulos, A.)

BLOCK OF RECEPTOR-RESPONSE LINKAGE

Non-competitive antagonism describes the situation where the antagonist blocks at some point downstream from the agonist binding site on the receptor, and interrupts the chain of events that leads to the production of a response by the agonist. For example, **ketamine** enters the ion channel pore of the NMDA receptor (see Ch. 39) blocking it, thus preventing ion flux through the channels. Drugs such as **verapamil** and **nifedipine** prevent the influx of Ca^{2+} through the cell membrane (see Ch. 23) and thus non-selectively block the contraction of smooth muscle produced by drugs acting at any receptor that couples to these calcium channels. As a rule, the effect will be to reduce the slope and maximum of the agonist log concentration-response curve, although it is quite possible for some degree of rightward shift to occur as well.

PHYSIOLOGICAL ANTAGONISM

Physiological antagonism is a term used loosely to describe the interaction of two drugs whose opposing actions in the body tend to cancel each other. For example, **histamine** acts on receptors of the parietal cells of the gastric mucosa to stimulate acid secretion, while **omeprazole** blocks this effect by inhibiting the proton pump; the two drugs can be said to act as physiological antagonists.

Types of drug antagonism



Drug antagonism occurs by various mechanisms:

- chemical antagonism (interaction in solution)
- pharmacokinetic antagonism (one drug affecting the absorption, metabolism or excretion of the other)
- competitive antagonism (both drugs binding to the same receptors); the antagonism may be reversible or irreversible
- interruption of receptor-response linkage
- physiological antagonism (two agents producing opposing physiological effects)

DESENSITISATION AND TOLERANCE

Often, the effect of a drug gradually diminishes when it is given continuously or repeatedly. *Desensitisation* and *tachyphylaxis* are synonymous terms used to describe this phenomenon, which often develops in the course of a few minutes. The term *tolerance* is conventionally used to describe a more gradual decrease in responsiveness to a drug, taking hours, days or weeks to develop, but the distinction is not a sharp one. The term *refractoriness* is also sometimes used, mainly in relation to a loss of therapeutic efficacy. *Drug resistance* is a term used to describe the loss of effectiveness of antimicrobial or antitumour drugs (see Chs 51 and 57). Many different mechanisms can give rise to these phenomena. They include:

- change in receptors
- translocation of receptors
- exhaustion of mediators
- increased metabolic degradation of the drug
- physiological adaptation

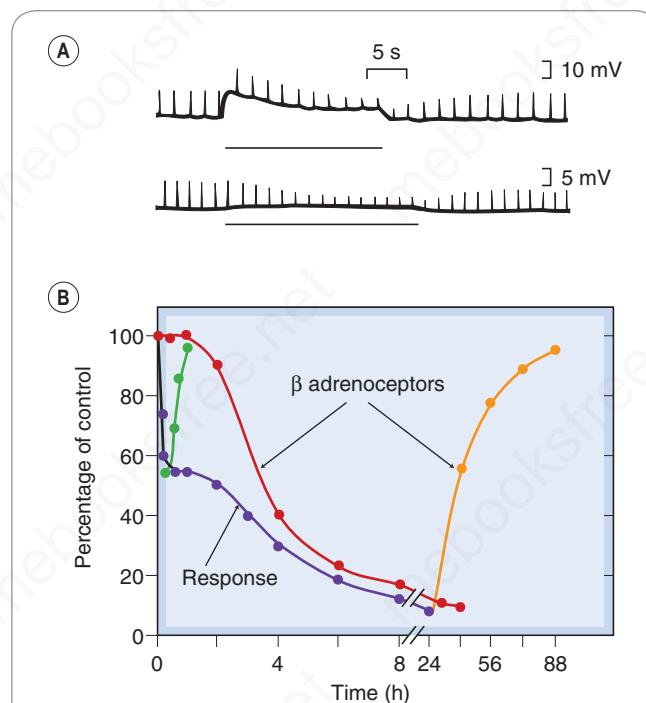


Fig. 2.13 Two kinds of receptor desensitisation.

(A) Acetylcholine (ACh) at the frog motor endplate. Brief depolarisations (*upward deflections*) are produced by short pulses of ACh delivered from a micropipette. A long pulse (*horizontal line*) causes the response to decline with a time course of about 20 seconds, owing to desensitisation, and it recovers with a similar time course. (B) β adrenoceptors of rat glioma cells in tissue culture. Isoproterenol (1 $\mu\text{mol/L}$) was added at time zero, and the adenylyl cyclase response and β -adrenoceptor density measured at intervals. During the early uncoupling phase, the response (blue line) declines with no change in receptor density (red line). Later, the response declines further concomitantly with disappearance of receptors from the membrane by internalisation. The green and orange lines show the recovery of the response and receptor density after the isoproterenol is washed out during the early or late phase. (Panel [A] from Katz B., Thesleff S., 1957. J. Physiol. 138, 63; panel [B] from Perkins, J.P., 1981. Trends Pharmacol. Sci. 2, 326.)

- active extrusion of drug from cells (mainly relevant in cancer chemotherapy; see Ch. 57)

CHANGE IN RECEPTORS

Among receptors directly coupled to ion channels (see Ch. 3), desensitisation is often rapid and pronounced. At the neuromuscular junction (Fig. 2.13A), the desensitised state is caused by a conformational change in the receptor, resulting in tight binding of the agonist molecule without the opening of the ionic channel. Phosphorylation of intracellular regions of the receptor protein is a second, slower mechanism by which ion channels become desensitised.

Most G protein-coupled receptors (see Ch. 3) also show desensitisation (Fig. 2.13B). Phosphorylation of the receptor interferes with its ability to activate second messenger cascades, although it can still bind the agonist molecule. The molecular mechanisms of this 'uncoupling' are considered further in Chapter 3. This type of desensitisation

usually takes seconds to minutes to develop, and recovers when the agonist is removed.

It will be realised that the two-state model in its simple form, discussed earlier, needs to be further elaborated to incorporate additional desensitised states of the receptor.

TRANSLOCATION OF RECEPTORS

Prolonged exposure to agonists often results in a gradual decrease in the number of receptors expressed on the cell surface, as a result of *internalisation* of the receptors. This is shown for β adrenoceptors in Fig. 2.13B and is a slower process than the uncoupling described above. Similar changes have been described for other types of receptor, including those for various peptides. The internalised receptors are taken into the cell by endocytosis of patches of the membrane, a process that normally depends on receptor phosphorylation and the subsequent binding of *arrestin* proteins to the phosphorylated receptor (see Ch. 3, Fig. 3.16). This type of adaptation is common for hormone receptors and has obvious relevance to the effects produced when drugs are given for extended periods. It is generally an unwanted complication when agonist drugs are used clinically.

EXHAUSTION OF MEDIATORS

In some cases, desensitisation is associated with depletion of an essential intermediate substance. Drugs such as **amphetamine**, which acts by releasing amines from nerve terminals (see Chs 15 and 49), show marked tachyphylaxis because the amine stores become depleted.

ALTERED DRUG METABOLISM

Tolerance to some drugs, for example **barbiturates** and **ethanol** (Ch. 49), occurs partly because repeated administration of the same dose produces a progressively lower plasma concentration, as a result of increased metabolic degradation. The degree of tolerance that results is generally modest, and in both of these examples other mechanisms contribute to the substantial tolerance that actually occurs. However, the pronounced tolerance to **nitrovasodilators** (see Chs 21 and 23) results mainly from decreased metabolism, which reduces the release of the active mediator, nitric oxide.

PHYSIOLOGICAL ADAPTATION

Diminution of a drug's effect may occur because it is nullified by a homeostatic response. For example, the blood pressure-lowering effect of **thiazide diuretics** is limited because of a gradual activation of the renin-angiotensin system (see Ch. 23). Such homeostatic mechanisms are very common, and if they occur slowly the result will be a gradually developing tolerance. It is a common experience that many side effects of drugs, such as nausea or sleepiness, tend to subside even though drug administration is continued. We may assume that some kind of physiological adaptation is occurring, presumably associated with altered gene expression resulting in changes in the levels of various regulatory molecules, but little is known about the mechanisms involved.

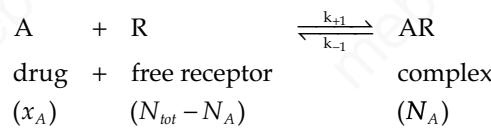
QUANTITATIVE ASPECTS OF DRUG-RECEPTOR INTERACTIONS

▼ Here we present some aspects of so-called receptor theory, which is based on applying the Law of Mass Action to the drug-receptor

interaction and which has served well as a framework for interpreting a large body of quantitative experimental data (see Colquhoun, 2006).

THE BINDING REACTION

▼ The first step in drug action on specific receptors is the formation of a reversible drug-receptor complex, the reactions being governed by the Law of Mass Action. Suppose that a piece of tissue, such as heart muscle or smooth muscle, contains a total number of receptors, N_{tot} , for an agonist such as adrenaline. When the tissue is exposed to adrenaline at concentration x_A and allowed to come to equilibrium, a certain number, N_A , of the receptors will become occupied, and the number of vacant receptors will be reduced to $N_{tot} - N_A$. Normally, the number of adrenaline molecules applied to the tissue in solution greatly exceeds N_{tot} , so that the binding reaction does not appreciably reduce x_A . The magnitude of the response produced by the adrenaline will be related (even if we do not know exactly how) to the number of receptors occupied, so it is useful to consider what quantitative relationship is predicted between N_A and x_A . The reaction can be represented by:



The Law of Mass Action (which states that the rate of a chemical reaction is proportional to the product of the concentrations of reactants) can be applied to this reaction.

$$\text{Rate of forward reaction} = k_{+1}x_A(N_{tot} - N_A) \quad (2.1)$$

$$\text{Rate of backward reaction} = k_{-1}N_A \quad (2.2)$$

At equilibrium, the two rates are equal:

$$k_{+1}x_A(N_{tot} - N_A) = k_{-1}N_A \quad (2.3)$$

The *affinity constant* of binding is given by k_{+1}/k_{-1} and from Eq. 2.3 equals $N_A/x_A(N_{tot} - N_A)$. Unfortunately, this has units of reciprocal concentration (L/mol) which for some of us is a little hard to get our heads around. Pharmacologists therefore tend to use the reciprocal of the affinity constant, the *equilibrium dissociation constant* (K), which has units of concentration (mol/L).

For drug A its equilibrium dissociation constant (K_A)⁶ can be represented as

$$K_A = k_{-1}/k_{+1} = x_A(N_{tot} - N_A)/N_A \quad (2.4)$$

The proportion of receptors occupied, or occupancy (P_A), is N_A/N_{tot} , which is independent of N_{tot} .

$$P_A = \frac{x_A}{x_A + k_{-1}/k_{+1}} = \frac{x_A}{x_A + K_A} \quad (2.5)$$

Thus if the equilibrium dissociation constant of a drug is known we can calculate the proportion of receptors it will occupy at any concentration.

Eq. 2.5 can be written:

$$P_A = \frac{x_A/K_A}{x_A/K_A + 1} \quad (2.6)$$

This important result is known as the Hill-Langmuir equation.⁷

⁶Here we now use ' K_A ' rather than just ' K ' because we will in the next section be going on to consider the situation when two drugs, A and B, are present and there we will use ' K_A' ' and ' K_B ' to denote the equilibrium dissociation constants of the two drugs.

⁷A.V. Hill first published it in 1909, when he was still a medical student. Langmuir, a physical chemist working on gas adsorption, derived it independently in 1916. Both subsequently won Nobel Prizes. Until recently, it was known to pharmacologists as the Langmuir equation, even though Hill deserves the credit.

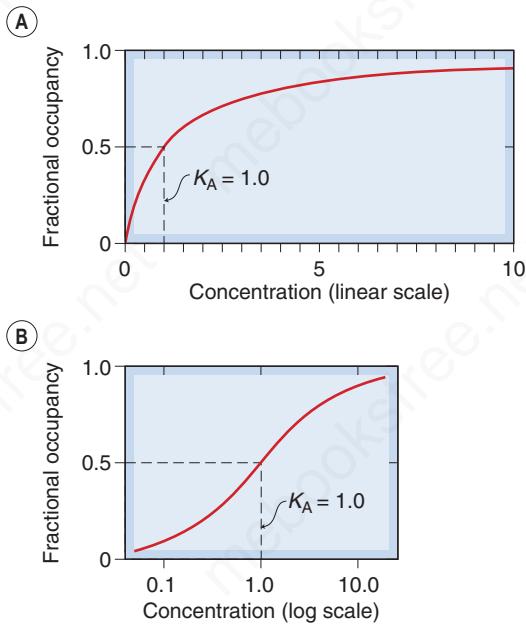


Fig. 2.14 Theoretical relationship between occupancy and ligand concentration. The relationship is plotted according to Eq. 2.5. (A) Plotted with a linear concentration scale, this curve is a rectangular hyperbola. (B) Plotted with a log concentration scale, it is a symmetrical sigmoid curve. K_A is defined in the text and footnote 6.

The *equilibrium dissociation constant*, K_A , is a characteristic of the drug and of the receptor; it has the dimensions of concentration and is numerically equal to the concentration of drug required to occupy 50% of the sites at equilibrium. (Verify from Eq. 2.5 that when $x_A = K_A$ then $P_A = 0.5$.) The higher the affinity of the drug for the receptors, the lower will be the value of K_A . Eq. 2.6 describes the relationship between occupancy and drug concentration, and it generates a characteristic curve known as a *rectangular hyperbola*, as shown in Fig. 2.14A. It is common in pharmacological work to use a logarithmic scale of concentration; this converts the hyperbola to a symmetrical sigmoid curve (Fig. 2.14B).

The same approach is used to analyse data from experiments in which drug binding is measured directly (see pp. 8–9, Fig. 2.2). In this case, the relationship between the amount bound (B) and ligand concentration (x_A) should be:

$$B = B_{\max} x_A / (x_A + K_A) \quad (2.7)$$

where B_{\max} is the total number of binding sites in the preparation (often expressed as pmol/mg of protein). To display the results in linear form, Eq. 2.6 may be rearranged to:

$$B/x_A = B_{\max}/K_A - B/K_A \quad (2.8)$$

A plot of B/x_A against B (known as a *Scatchard plot*) gives a straight line from which both B_{\max} and K_A can be estimated. Statistically, this procedure is not without problems, and it is now usual to estimate these parameters from the untransformed binding values by an iterative non-linear curve-fitting procedure.

To this point, our analysis has considered the binding of one ligand to a homogeneous population of receptors. To get closer to real-life pharmacology, we must consider (a) what happens when more than one ligand is present, and (b) how the tissue response is related to receptor occupancy.

BINDING WHEN MORE THAN ONE DRUG IS PRESENT

▼ Suppose that two drugs, A and B, which bind to the same receptor with equilibrium dissociation constants K_A and K_B , respectively, are present at concentrations x_A and x_B . If the two drugs compete (i.e. the receptor can accommodate only one at a time), then, by application of the same reasoning as for the one-drug situation described above, the occupancy by drug A is given by:

$$P_A = \frac{x_A/K_A}{x_A/K_A + x_B/K_B + 1} \quad (2.9)$$

Comparing this result with Eq. 2.5 shows that adding drug B, as expected, reduces the occupancy by drug A. Fig. 2.4A (p. 11) shows the predicted binding curves for A in the presence of increasing concentrations of B, demonstrating the shift without any change of slope or maximum that characterises the pharmacological effect of a competitive antagonist (see Fig. 2.5). The extent of the rightward shift, on a logarithmic scale, represents the ratio (r_A , given by x_A'/x_A where x_A' is the increased concentration of A) by which the concentration of A must be increased to overcome the competition by B. Rearranging Eq. 2.9 shows that

$$r_A = (x_B/K_B) + 1 \quad (2.10)$$

Thus r_A depends only on the concentration and equilibrium dissociation constant of the competing drug B, not on the concentration or equilibrium dissociation constant of A.

If A is an agonist, and B is a competitive antagonist, and we assume that the response of the tissue will be an unknown function of P_A , then the value of r_A determined from the shift of the agonist concentration–effect curve at different antagonist concentrations can be used to estimate the equilibrium dissociation constant K_B for the antagonist. Such pharmacological estimates of r_A are commonly termed *agonist dose ratios* (more properly concentration ratios, although most pharmacologists use the older term). This simple and very useful Eq. (2.10) is known as the *Schild equation*, after the pharmacologist who first used it to analyse drug antagonism.

Eq. 2.10 can be expressed logarithmically in the form:

$$\log(r_A - 1) = \log x_B - \log K_B \quad (2.11)$$

Thus a plot of $\log(r_A - 1)$ against $\log x_B$, usually called a Schild plot (as in Fig. 2.5, earlier), should give a straight line with unit slope (i.e. its gradient is equal to 1) and an abscissal intercept equal to $\log K_B$. Following the pH and pK notation, antagonist potency can be expressed as a pA_2 value; under conditions of competitive antagonism, $pA_2 = -\log K_B$. Numerically, pA_2 is defined as the negative logarithm of the molar concentration of antagonist required to produce an agonist dose ratio equal to 2. As with pH notation, its principal advantage is that it produces simple numbers, a pA_2 of 6.5 being equivalent to a K_B of 3.2×10^{-7} mol/L.

For competitive antagonism, r shows the following characteristics:

- It depends only on the concentration and equilibrium dissociation constant of the antagonist, and not on the size of response that is chosen as a reference point for the measurements (so long as it is submaximal).
- It does not depend on the equilibrium dissociation constant of the agonist.
- It increases linearly with x_B , and the slope of a plot of $(r_A - 1)$ against x_B is equal to $1/K_B$; this relationship, being independent of the characteristics of the agonist, should be the same for an antagonist against all agonists that act on the same population of receptors.

These predictions have been verified for many examples of competitive antagonism (see Fig. 2.5).

In this section, we have avoided going into great detail and have oversimplified the theory considerably. As we learn more about the actual molecular details of how receptors work to produce their biological effects (see Ch. 3), the shortcomings of this theoretical

treatment become more obvious. The two-state model can be incorporated without difficulty, but complications arise when we include the involvement of G proteins (see Ch. 3) in the reaction scheme (as they shift the equilibrium between R and R^{*}), and when we allow for the fact that receptor activation is not a simple on-off switch, as the two-state model assumes, but may take different forms. Despite strenuous efforts by theoreticians to allow for such possibilities, the molecules always seem to remain one step ahead. Nevertheless, this type of basic theory applied to the two-state model remains a useful basis for developing quantitative models of drug action. The book by Kenakin (1997) is recommended as an introduction, and the later review (Kenakin & Christopoulos, 2011) presents a detailed account of the value of quantification in the study of drug action.

Binding of drugs to receptors

- Binding of drugs to receptors necessarily obeys the *Law of Mass Action*.
- At equilibrium, receptor occupancy is related to drug concentration by the *Hill-Langmuir equation* (Eq. 2.6).
- The higher the affinity of the drug for the receptor, the lower the concentration at which it produces a given level of occupancy.
- The same principles apply when two or more drugs compete for the same receptors; each has the effect of reducing the apparent affinity for the other.



THE NATURE OF DRUG EFFECTS

In discussing how drugs act in this chapter, we have focused mainly on the rapid consequences of receptor activation. Details of the receptors and their linkage to effects at the cellular level are described in Chapter 3. We now have a fairly good understanding at this level. It is important, however, particularly when considering drugs in a therapeutic context, that their direct effects on cellular function generally lead to secondary, delayed effects, which are often highly relevant in a clinical situation in relation to both therapeutic efficacy and harmful effects (Fig. 2.15). For example, activation of cardiac β adrenoceptors (see Chs 3 and 22) causes rapid changes in the functioning of the heart muscle, but also slower (minutes to hours) changes in the functional state of the receptors (e.g. desensitisation), and even slower (hours to days) changes in gene expression that produce long-term changes (e.g. hypertrophy) in cardiac structure and function. Opioids (see Ch. 43) produce an immediate analgesic effect, but after a time, tolerance and dependence ensue, and in some cases long-term addiction. In these and many other examples, the nature of the intervening mechanism is unclear, although as a general rule any long-term phenotypic change necessarily involves alterations of gene expression. Drugs are often used to treat chronic conditions, and understanding long-term as well

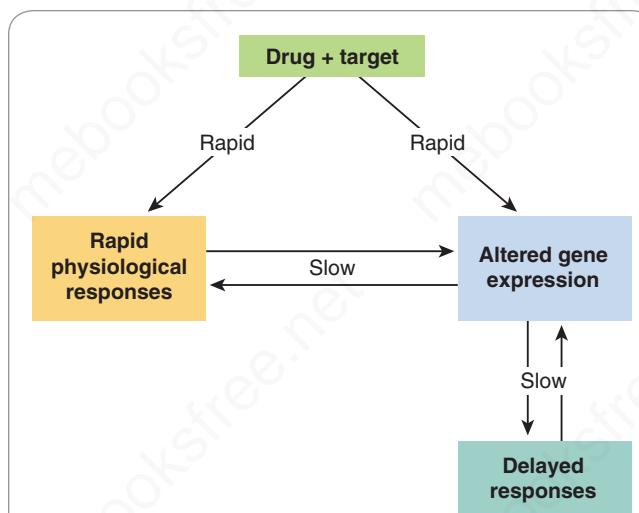


Fig. 2.15 Early and late responses to drugs. Many drugs act directly on their targets (left-hand arrow) to produce a rapid physiological response. If this is maintained, it is likely to cause changes in gene expression that give rise to delayed effects. Some drugs (right-hand arrow) have their primary action on gene expression, producing delayed physiological responses. Drugs can also work by both pathways. Note the bidirectional interaction between gene expression and response.

as acute drug effects is becoming increasingly important. Pharmacologists have traditionally tended to focus on short-term physiological responses, which are much easier to study, rather than on delayed effects. The focus is now clearly shifting.

Drug effects

- Drugs act mainly on cellular targets, producing effects at different functional levels (e.g. biochemical, cellular, physiological and structural).
- The direct effect of the drug on its target produces acute responses at the biochemical, cellular or physiological levels.
- Prolonged receptor activation generally leads to *delayed long-term effects*, such as desensitisation or down-regulation of receptors, hypertrophy, atrophy or remodelling of tissues, tolerance, dependence and addiction.
- Long-term delayed responses result from changes in gene expression, although the mechanisms by which the acute effects bring this about are often uncertain.
- Therapeutic effects may be based on acute responses (e.g. the use of bronchodilator drugs to treat asthma; Ch. 29) or delayed responses (e.g. antidepressants; Ch. 48).



REFERENCES AND FURTHER READING

General

- Alexander, S.P.H., Kelly, E., Marrion, N., et al., 2015. The Concise Guide to Pharmacology. Br. J. Pharmacol. 172, 5729–6202. (*Summary data on a vast array of receptors, ion channels, transporters and enzymes and of the drugs that interact with them – valuable for reference*)
- Colquhoun, D., 2006. The quantitative analysis of drug-receptor interactions: a short history. Trends Pharmacol. Sci. 27, 149–157. (*An illuminating account for those interested in the origins of one of the central ideas in pharmacology*)
- Franks, N.P., 2008. General anaesthesia: from molecular targets to neuronal pathways of sleep and arousal. Nat. Rev. Neurosci. 9, 370–386. (*Describes how we now understand that general anaesthetics interact with specific membrane proteins rather than by accumulating in membrane lipids*)
- Kenakin, T., 1997. Pharmacologic Analysis of Drug-Receptor Interactions, third ed. Lippincott-Raven, New York. (*Useful and detailed textbook covering most of the material in this chapter in greater depth*)
- Kenakin, T., Christopoulos, A., 2013. Signalling bias in new drug discovery: detection, quantification and therapeutic impact. Nat. Rev. Drug Discov. 12, 205–216. (*Detailed discussion of the difficulties in measuring agonist efficacy and bias*)
- Neubig, R., Spedding, M., Kenakin, T., Christopoulos, A., 2003. International Union of Pharmacology Committee on receptor nomenclature and drug classification: XXXVIII. Update on terms and symbols in quantitative pharmacology. Pharmacol. Rev. 55, 597–606. (*Summary of IUPHAR-approved terms and symbols relating to pharmacological receptors – useful for reference purposes*)
- Rang, H.P., 2006. The receptor concept: pharmacology's big idea. Br. J. Pharmacol. 147 (Suppl. 1), 9–16. (*Short review of the origin and status of the receptor concept*)
- Stephenson, R.P., 1956. A modification of receptor theory. Br. J. Pharmacol. 11, 379–393. (*Classic analysis of receptor action introducing the concept of efficacy*)
- Receptor mechanisms: agonists and efficacy**
- Bond, R.A., Ijzerman, A.P., 2006. Recent developments in constitutive receptor activity and inverse agonism, and their potential for GPCR drug discovery. Trends Pharmacol. Sci. 27, 92–96. (*Discussion of pathophysiological consequences of constitutive receptor activation and therapeutic potential of inverse agonists*)
- Changeux, J.P., Christopoulos, A., 2016. Allosteric modulation as a unifying mechanism for receptor function and regulation. Cell 166, 1084–1102. (*Extensive review describing allosteric modulation at different types of receptor*)
- De Ligt, R.A.F., Kourounakis, A.P., Ijzerman, A.P., 2000. Inverse agonism at G protein-coupled receptors: (patho)physiological relevance and implications for drug discovery. Br. J. Pharmacol. 130, 1–12. (*Useful review article giving many examples of constitutively active receptors and inverse agonists, and discussing the relevance of these concepts for disease mechanisms and drug discovery*)
- Kelly, E., 2013. Efficacy and ligand bias at the μ -opioid receptor. Br. J. Pharmacol. 169, 1430–1446. (*A readable account of the problems of measuring efficacy as well as a discussion of agonist bias at an important receptor*)
- Kenakin, T., Christopoulos, A., 2011. Analytical pharmacology: the impact of numbers on pharmacology. Trends Pharmacol. Sci. 32, 189–196. (*A theoretical treatment that attempts to take into account recent knowledge of receptor function at the molecular level*)
- May, L.T., Leach, K., Sexton, P.M., Christopoulos, A., 2007. Allosteric modulation of G protein-coupled receptors. Annu. Rev. Pharmacol. Toxicol. 47, 1–51. (*Comprehensive review describing the characteristics, mechanisms and pharmacological implications of allosteric interactions at GPCRs*)

How drugs act: molecular aspects

OVERVIEW

In this chapter, we move from the general principles of drug action outlined in Chapter 2 to the molecules that are involved in recognising chemical signals and translating them into cellular responses. Molecular pharmacology is advancing rapidly, and the new knowledge is changing our understanding of drug action and opening up many new therapeutic possibilities, further discussed in other chapters.

First, we consider the types of target proteins on which drugs act. Next, we describe the main families of receptors and ion channels. Finally, we discuss the various forms of receptor–effector linkage (signal transduction mechanisms) through which receptors are coupled to the regulation of cell function. The relationship between the molecular structure of a receptor and its functional linkage to a particular type of effector system is a principal theme. In the next two chapters, we see how these molecular events alter important aspects of cell function – a useful basis for understanding the effects of drugs on intact living organisms. We are confident that tomorrow's pharmacology will rest solidly on the advances in cellular and molecular biology that are discussed here.

PROTEIN TARGETS FOR DRUG ACTION

The protein targets for drug action on mammalian cells (Fig. 3.1) that are described in this chapter can be broadly divided into:

- receptors
- ion channels
- enzymes
- transporters (carrier molecules)

The great majority of important drugs act on one or other of these types of protein, but there are exceptions. For example, **colchicine** used to treat arthritic gout attacks (Ch. 27) interacts with the structural protein tubulin, while several immunosuppressive drugs (e.g. **ciclosporin**, Ch. 27) bind to cytosolic proteins known as immunophilins. Therapeutic antibodies that act by sequestering cytokines (protein mediators involved in inflammation; see Chs 5 and 27) are also used. Targets for chemotherapeutic drugs (Chs 51–57), where the aim is to suppress invading microorganisms or cancer cells, include DNA and cell wall constituents as well as other proteins.

RECEPTORS

Receptors (see Fig. 3.1A) are the sensing elements in the system of chemical communications that coordinates the

function and responses of all the different cells in the body, the chemical messengers being the various hormones, transmitters and other mediators discussed in Section 2 of this book. Many therapeutically useful drugs act, either as agonists or antagonists, on receptors for known endogenous mediators. In most cases, the endogenous mediator was discovered before – often many years before – the receptor was characterised pharmacologically and biochemically. In some cases, such as the cannabinoid and opioid receptors (see Chs 20 and 43), the endogenous mediators were identified later; in others, known as *orphan receptors* (see later) the mediator, if it exists, still remains unknown. The host defence system also utilises a set of receptors (e.g. the 'Toll' receptors) that are adept at recognising fragments of 'foreign' bacterial and other invading organisms. These are considered separately in Chapter 7.

ION CHANNELS

Ion channels¹ are essentially gateways in cell membranes that selectively allow the passage of particular ions, and that are induced to open or close by a variety of mechanisms. Two important types are *ligand-gated channels* and *voltage-gated channels*. The former open only when one or more agonist molecules are bound, and are properly classified as receptors, since agonist binding is needed to activate them. Voltage-gated channels are gated by changes in the transmembrane potential rather than by agonist binding.

In general, drugs can affect ion channel function in several ways:

1. By binding to the channel protein itself, either to the ligand-binding (*orthosteric*) site of ligand-gated channels, or to other (*allosteric*) sites, or, in the simplest case, exemplified by the action of local anaesthetics on the voltage-gated sodium channel (see Ch. 44), the drug molecule plugs the channel physically (see Fig. 3.1B), blocking ion permeation. Examples of drugs that bind to allosteric sites on the channel protein and thereby affect channel gating include:
 - **benzodiazepines** (see Ch. 45). These drugs bind to a region of the GABA_A receptor–chloride channel complex (a ligand-gated channel) that is distinct from the GABA binding site and facilitate the opening of the channel by the inhibitory neurotransmitter GABA (see Ch. 39)
 - vasodilator drugs of the **dihydropyridine** type (see Ch. 23), which inhibit the opening of L-type calcium channels (see Ch. 4).

¹'Ion channels and the electrical properties they confer on cells are involved in every human characteristic that distinguishes us from the stones in a field' (Armstrong, C.M., 2003. Voltage-gated K channels. Sci. STKE 188, re10).

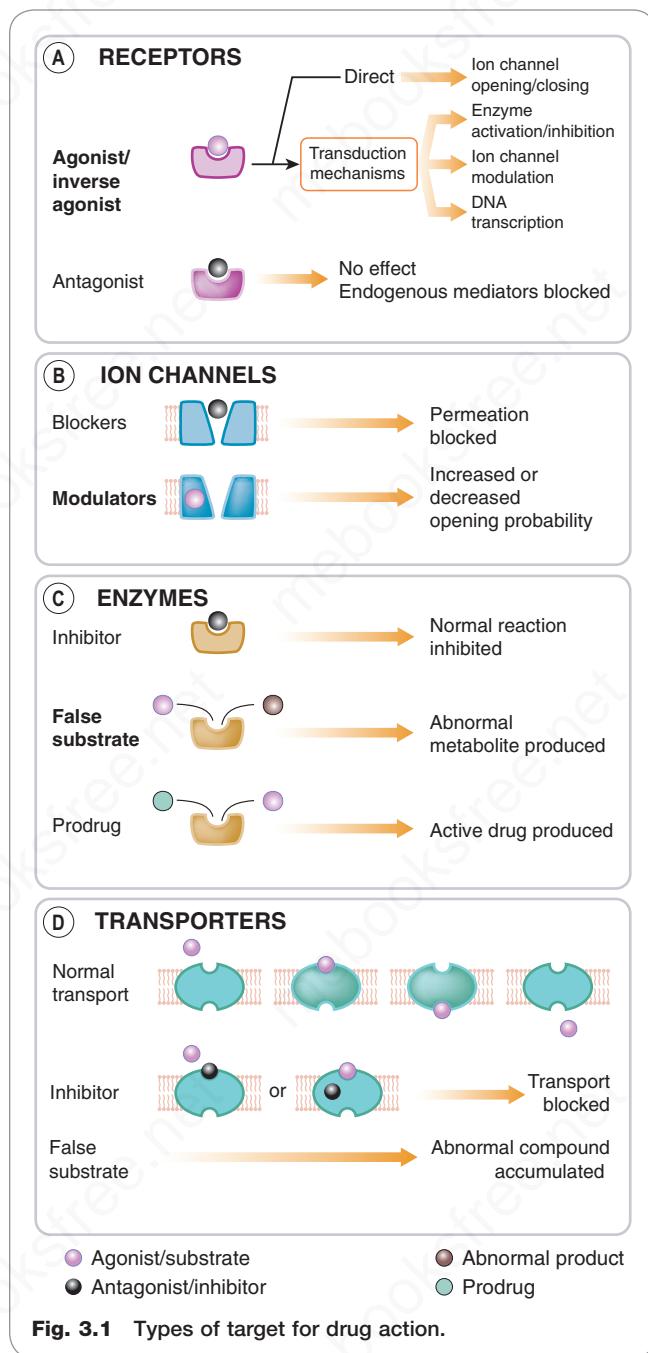


Fig. 3.1 Types of target for drug action.

- By an indirect interaction, involving an activated G protein subunit or other intermediary (see p. 34).
- By altering the level of expression of ion channels on the cell surface. For example, **gabapentin** reduces the insertion of neuronal calcium channels into the plasma membrane (Ch. 46).

A summary of the different ion channel families and their functions is given later.

ENZYMEs

Many drugs target enzymes (see Fig. 3.1C). Often, the drug molecule is a substrate analogue that acts as a competitive inhibitor of the enzyme (e.g. **captopril**, acting on

angiotensin-converting enzyme; Ch. 23); in other cases, the binding is irreversible and non-competitive (e.g. **aspirin**, acting on cyclo-oxygenase; Ch. 27). Drugs may also act as false substrates, where the drug molecule undergoes chemical transformation to form an abnormal product that subverts the normal metabolic pathway. An example is the anticancer drug **fluorouracil**, which replaces uracil as an intermediate in purine biosynthesis but cannot be converted into thymidylate, thus blocking DNA synthesis and preventing cell division (Ch. 57).

It should also be mentioned that drugs may require enzymic degradation to convert them from an inactive form, the prodrug (see Ch. 10), to an active form (e.g. **enalapril** is converted by esterases to enalaprilat, which inhibits angiotensin-converting enzyme). Furthermore, as discussed in Chapter 58, drug toxicity often results from the enzymic conversion of the drug molecule to a reactive metabolite. **Paracetamol** (see Ch. 27) causes liver damage in this way. As far as the primary action of the drug is concerned, this is an unwanted side reaction, but it is of major practical importance.

TRANSPORTERS

The movement of ions and small polar organic molecules across cell membranes generally occurs either through channels, or through the agency of a transport protein (see Fig. 3.1D), because the permeating molecules are often insufficiently lipid-soluble to penetrate lipid membranes on their own. Many such transporters are known; examples of particular pharmacological importance include those responsible for the transport of ions and many organic molecules across the renal tubule, the intestinal epithelium and the blood-brain barrier, the transport of Na^+ and Ca^{2+} out of cells, the uptake of neurotransmitter precursors (such as choline) or of neurotransmitters themselves (such as amines and amino acids) by nerve terminals, and the transport of drug molecules and their metabolites across cell membranes and epithelial barriers. We shall encounter transporters frequently in later chapters.

In many cases, hydrolysis of ATP provides the energy for transport of substances against their electrochemical gradient. Such transport proteins include a distinct ATP-binding site, and are termed ABC (ATP-Binding Cassette) transporters. Important examples include the sodium pump (Na^+-K^+ -ATPase; see Ch. 4) and multidrug resistance (MDR) transporters that eject cytotoxic drugs from cancer and microbial cells, conferring resistance to these therapeutic agents (see Ch. 57). In other cases, including the neurotransmitter transporters, the transport of organic molecules is coupled to the transport of ions (usually Na^+), either in the same direction (*symport*) or in the opposite direction (*antiport*), and therefore relies on the electrochemical gradient for Na^+ generated by the ATP-driven sodium pump. The carrier proteins embody a recognition site that makes them specific for a particular permeating species, and these recognition sites can also be targets for drugs whose effect is to block the transport system (e.g. cocaine blocks monoamine neurotransmitter uptake into nerve terminals; see Ch. 49).

The importance of transporters as a source of individual variation in the pharmacokinetic characteristics of various drugs is increasingly recognised (see Ch. 11).

RECEPTOR PROTEINS

CLONING OF RECEPTORS

In the 1970s, pharmacology entered a new phase when receptors, which had until then been theoretical entities, began to emerge as biochemical realities following the development of receptor-labelling techniques (see Ch. 2), which made it possible to extract and purify the receptor material.

Once receptor proteins were isolated and purified, it was possible to analyse the amino acid sequence of a short stretch, allowing the corresponding base sequence of the mRNA to be deduced and full-length DNA to be isolated by conventional cloning methods, starting from a cDNA library obtained from a tissue source rich in the receptor of interest. The first receptor clones were obtained in this way, but subsequently expression cloning and, with the sequencing of the entire genome of various species, including human, cloning strategies based on sequence homologies, which do not require prior isolation and purification of the receptor protein, were widely used, and now several hundred receptors of all four structural families (see Fig. 3.3) have been cloned. Sequence data so obtained has revealed many molecular variants (subtypes) of known receptors that had not been evident from pharmacological studies (see IUPHAR/BPS, *Guide to Pharmacology*). Much remains to be discovered about the pharmacological, functional and clinical significance of this abundant molecular polymorphism. It is expected, however, that such variations will account for part of the variability between individuals in response to therapeutic agents (see Ch. 12).

Endogenous ligands for many of the novel receptors identified by gene cloning are so far unknown, and they are described as 'orphan receptors'.² Identifying ligands for these presumed receptors is often difficult. Increasingly, there are examples (e.g. free fatty acid receptors) where important endogenous ligands have been linked to hitherto orphan receptors. There is optimism that novel therapeutic agents will emerge by targeting this pool of unclaimed receptors.

Much information has been gained by introducing the cloned DNA encoding individual receptors into cell lines, producing cells that express the foreign receptors in a functional form. Such engineered cells allow much more precise control of the expressed receptors than is possible with natural cells or intact tissues, and the technique is widely used to study the binding and pharmacological characteristics of cloned receptors. Expressed human receptors, which often differ in their sequence and pharmacological properties from their animal counterparts, can be studied in this way.

Obtaining crystals of a protein allows its structure to be analysed at very high resolution by X-ray diffraction techniques, but unfortunately, since many receptors are normally embedded in membrane lipid, they have, until relatively recently, proven difficult to crystallise. Much of the information obtained relates to how ligands bind to receptors, but we are now beginning to learn more about

²An oddly Dickensian term that seems inappropriately condescending. Because we can assume that these receptors play defined roles in physiological signalling, their 'orphanhood' reflects our ignorance, not their status. More information on orphan receptors can be found at <www.guidetopharmacology.org/GRAC/FamilyDisplayForward?familyId=115#16>.

agonist-induced receptor conformational changes and how signalling is initiated.

Now that the genes have been clearly identified, the emphasis has shifted to characterising the receptors pharmacologically and determining their molecular characteristics and physiological functions.

TYPES OF RECEPTOR

Receptors elicit many different types of cellular effect. Some of them are very rapid, such as those involved in fast synaptic transmission, operating within milliseconds, whereas other receptor-mediated effects, such as many of those produced by thyroid hormone or various steroid hormones, occur over hours or days. There are many examples of intermediate timescales – catecholamines, for example, usually act in a matter of seconds, whereas many peptides take rather longer to produce their effects. Not surprisingly, very different types of linkage between receptor occupation and the ensuing response are involved. Based on molecular structure and the nature of this linkage (the transduction mechanism), we can distinguish four receptor types, or superfamilies (Figs 3.2 and 3.3; Table 3.1).

- Type 1: **ligand-gated ion channels** (also known as **ionotropic receptors**)³. The chain of discoveries culminating in the molecular characterisation of these receptors is described by Halliwell (2007). Typically, these are the receptors on which fast neurotransmitters act (see Table 3.1).
- Type 2: **G protein-coupled receptors** (GPCRs). These are also known as **metabotropic receptors** or **7-transmembrane** (7-TM, serpentine or heptahelical) **receptors**. They are membrane receptors that are coupled to intracellular effector systems primarily via a G protein (see p. 32). They constitute the largest family,⁴ and include receptors for many hormones and slow transmitters (Table 3.1).
- Type 3: **kinase-linked and related receptors**. This is a large and heterogeneous group of membrane receptors responding mainly to protein mediators. They comprise an extracellular ligand-binding domain linked to an intracellular domain by a single transmembrane helix. In many cases, the intracellular domain is enzymic in nature (with protein kinase or guanylyl cyclase activity). Some lack enzymic activity themselves but link to intracellular effector enzymes through their binding of adaptor proteins. Examples of these latter receptor types include cytokine receptors (e.g. tumour necrosis factor [TNF] receptors) and pattern recognition receptors (PRRs) that recognise pathogen-associated molecular patterns (PAMPs) or danger-associated molecular patterns (DAMPs) found in pathogens, which stimulate the innate immune system host defence network (see Ch. 7). PRR receptors include the cell surface Toll-like receptors (TLRs), and the cytoplasmic receptors such

³Here, focusing on receptors, we include ligand-gated ion channels as an example of a receptor family. Other types of ion channels are described later (p. 46); many are also drug targets, although not receptors in the strict sense.

⁴There are 865 human GPCRs comprising 1.6% of the genome (Fredriksson & Schiöth, 2005). Nearly 500 of these are believed to be odorant receptors involved in smell and taste sensations, the remainder being receptors for known or unknown endogenous mediators – enough to keep pharmacologists busy for some time yet.

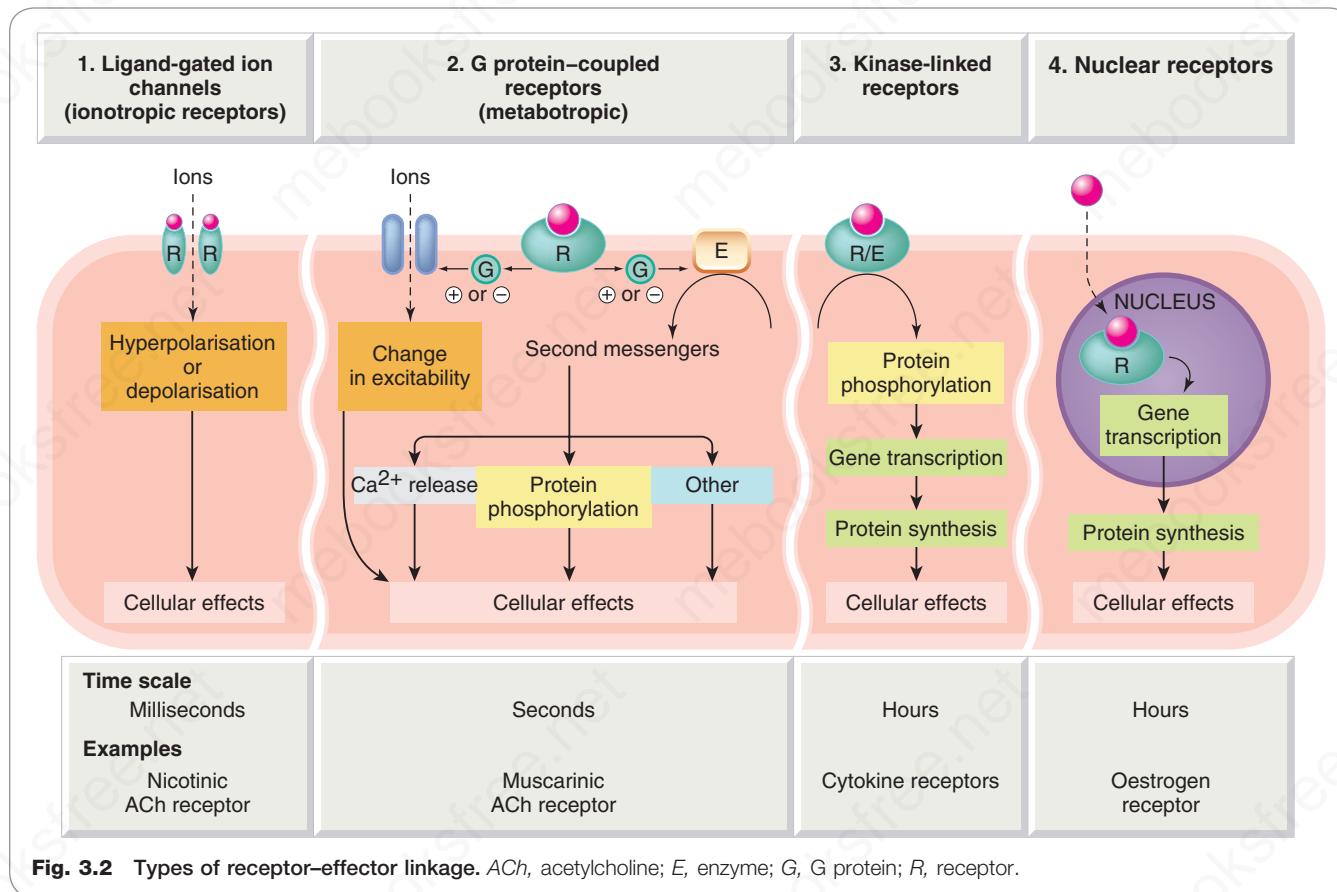


Table 3.1 The four main types of receptor

	Type 1: Ligand-gated ion channels	Type 2: G protein-coupled receptors	Type 3: Receptor kinases	Type 4: Nuclear receptors
Location	Membrane	Membrane	Membrane	Intracellular
Effector	Ion channel	Channel or enzyme	Protein kinases	Gene transcription
Coupling	Direct	G protein or arrestin	Direct	Via DNA
Examples	Nicotinic acetylcholine receptor, GABA _A receptor	Muscarinic acetylcholine receptor, adrenoceptors	Insulin, growth factors, cytokine receptors	Steroid receptors
Structure	Oligomeric assembly of subunits surrounding central pore	Monomeric or oligomeric assembly of subunits comprising seven transmembrane helices with intracellular G protein-coupling domain	Single transmembrane helix linking extracellular receptor domain to intracellular kinase domain	Monomeric structure with receptor- and DNA-binding domains

as RIG-I-like receptors (RLRs) and NOD-like receptors (NLRs). All these immune receptors signal their intracellular effects through adaptor proteins and kinases to alter the cell's transcription to elicit the correct immune response needed to fight against any pathogenic invaders.

- Type 4: **nuclear receptors.** These are receptors that regulate gene transcription.⁵ Receptors of this type

also recognise many foreign molecules, inducing the expression of enzymes that metabolise them.

MOLECULAR STRUCTURE OF RECEPTORS

The molecular organisation of typical members of each of these four receptor superfamilies is shown in Fig. 3.3. Although individual receptors show considerable sequence variation in particular regions, and the lengths of the main intracellular and extracellular domains also vary from one to another within the same family, the overall structural patterns and associated signal transduction pathways are

⁵The term *nuclear receptor* is something of a misnomer, because some are actually located in the cytosol and migrate to the nuclear compartment when a ligand is present.

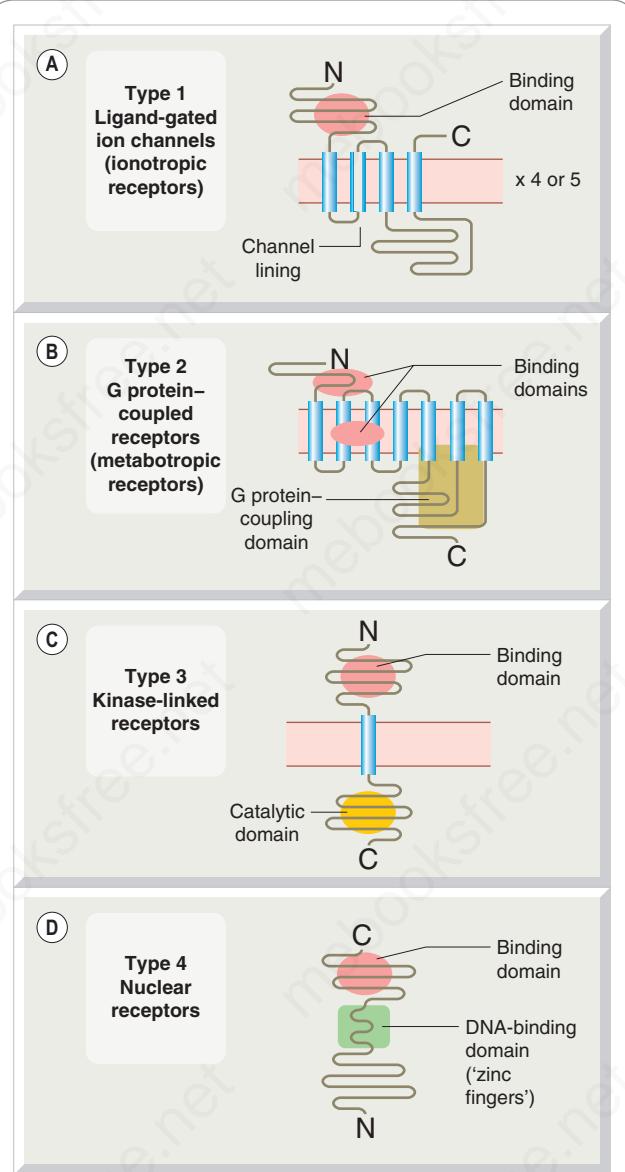


Fig. 3.3 General structure of four receptor families. The rectangular segments represent hydrophobic α-helical regions of the protein comprising approximately 20 amino acids, which form the membrane-spanning domains of the receptors. The pink shaded areas illustrate the region of the orthosteric ligand-binding domains. (A) Type 1: ligand-gated ion channels. The example illustrated here shows the subunit structure of the nicotinic acetylcholine receptor. The subunit structure of other ligand-gated ion channels is shown in Fig. 3.5. Many ligand-gated ion channels comprise four or five subunits of the type shown, the whole complex containing 16–20 membrane-spanning segments surrounding a central ion channel. (B) Type 2: G protein-coupled receptors (GPCRs). The two ligand-binding domains shown illustrate the position of the orthosteric ligand-binding domains on different types of GPCRs; there would be only one on each GPCR. (C) Type 3: kinase-linked receptors. Most growth factor receptors incorporate the ligand-binding and enzymatic (kinase) domains in the same molecule, as shown, whereas cytokine receptors lack an intracellular kinase domain but link to cytosolic kinase molecules. Other structural variants also exist. (D) Type 4: nuclear receptors that control gene transcription.

very consistent. The realisation that just four main receptor superfamilies provide a solid framework for interpreting the complex welter of information about the effects of a large proportion of the drugs that have been studied has been one of the most refreshing developments in modern pharmacology.

RECEPTOR HETEROGENEITY AND SUBTYPES

Receptors within a given family generally occur in several molecular varieties, or subtypes, with similar architecture but significant differences in their sequences, and often in their pharmacological properties.⁶ Nicotinic acetylcholine receptors are typical in this respect; distinct subtypes occur in different brain regions (see Table 40.2), and these differ from the muscle receptor. Some of the known pharmacological differences (e.g. sensitivity to blocking agents) between muscle and brain acetylcholine receptors correlate with specific sequence differences; however, as far as we know, all nicotinic acetylcholine receptors respond to the same physiological mediator and produce the same kind of synaptic response, so why many variants should have evolved is still a puzzle.

▼ Much of the sequence variation that accounts for receptor diversity arises at the genomic level, that is, different genes give rise to distinct receptor subtypes. Additional variation arises from alternative mRNA splicing, which means that a single gene can give rise to more than one receptor isoform. After translation from genomic DNA, the mRNA normally contains non-coding regions (introns) that are excised by mRNA splicing before the message is translated into protein. Depending on the location of the splice sites, splicing can result in inclusion or deletion of one or more of the mRNA coding regions, giving rise to long or short forms of the protein. This is an important source of variation, particularly for GPCRs, producing receptors with different binding characteristics and different signal transduction mechanisms, although its pharmacological relevance remains to be clarified. Another process that can produce different receptors from the same gene is mRNA editing, which involves the mischievous substitution of one base in the mRNA for another, and hence potentially a small variation in the amino acid sequence of the expressed receptor.

Molecular heterogeneity of this kind is a feature of all kinds of receptors – indeed of functional proteins in general. New receptor subtypes and isoforms continue to be discovered, and regular updates of the catalogue are available (www.guidetopharmacology.org/). The problems of classification, nomenclature and taxonomy resulting from this flood of data have been mentioned earlier.

We will now describe the characteristics of each of the four receptor superfamilies.

TYPE 1: LIGAND-GATED ION CHANNELS

The nicotinic acetylcholine receptor, which we find at the skeletal neuromuscular junction (Ch. 14), in autonomic ganglia (Ch. 14) and in the brain (Ch. 40), is a typical example of a ligand-gated ion channel, known as the cys-loop receptors (so called because they have in their structure a large intracellular domain between transmembrane domains 3 and 4 containing multiple cysteine residues [see Fig. 3.3A]). Others of this type include the GABA_A and glycine receptors (Ch. 39) as well as the 5-hydroxytryptamine type 3 (5-HT₃; Chs 16 and 40) receptor. Other types of ligand-gated ion

⁶Receptors for 5-hydroxytryptamine (see Ch. 16) are currently the champions with respect to diversity, with 13 subtypes of GPCR and 1 ligand-gated ion channel all responding to the same endogenous ligand.

channel exist – namely ionotropic glutamate receptors (Ch. 39) and purinergic P2X receptors (Chs 17 and 40) that differ in several respects from the nicotinic acetylcholine receptor (see Fig. 3.5). In addition to the ligand-gated ion channels found on the cell membrane that mediate fast synaptic transmission, there are also intracellular ligand-gated ion channels – namely the inositol trisphosphate (IP_3) and ryanodine receptors (see Ch. 4) that release Ca^{2+} from intracellular stores.

MOLECULAR STRUCTURE

Ligand-gated ion channels have structural features in common with other ion channels, described on p. 46. The nicotinic acetylcholine receptor cloned from the Torpedo electric ray (Fig. 3.4),⁷ consists of a pentameric assembly of different subunits, of which there are four types, termed α , β , γ and δ , each of molecular weight (M_r) 40–58 kDa. The subunits show marked sequence homology, and each contains four membrane-spanning α -helices, inserted into the membrane as shown in Fig. 3.4B. The pentameric structure ($\alpha_2, \beta, \gamma, \delta$) possesses two acetylcholine binding sites, each lying at the interface between one of the two α subunits and its neighbour. Both must bind acetylcholine molecules for the receptor to be activated. Fig. 3.4B shows the receptor structure. Each subunit spans the membrane four times, so the channel comprises no fewer than 20 membrane-spanning helices surrounding a central pore.

▼ One of the transmembrane helices (M_2) from each of the five subunits forms the lining of the ion channel (see Fig. 3.4). The five M_2 helices that form the pore are sharply kinked inwards halfway through the membrane, forming a constriction. When acetylcholine molecules bind, a conformation change occurs in the extracellular part of the receptor, which twists the α subunits, causing the kinked M_2 segments to swing out of the way, thus opening the channel. The channel lining contains a series of anionic residues, making the channel selectively permeable to cations (primarily Na^+ and K^+ , although some types of nicotinic receptor are permeable to Ca^{2+} as well).

The use of site-directed mutagenesis, which enables short regions, or single residues, of the amino acid sequence to be altered, has shown that a mutation of a critical residue in the M_2 helix changes the channel from being cation permeable (hence excitatory in the context of synaptic function) to being anion permeable (typical of receptors for inhibitory transmitters such as GABA and glycine). Other mutations affect properties such as gating and desensitisation of ligand-gated channels.

Other ligand-gated ion channels, such as glutamate receptors (see Ch. 39) and P2X receptors (see Chs 17 and 40), whose structures are shown in Fig. 3.5, have a different architecture. Ionotropic glutamate receptors are tetrameric and the pore is built from loops rather than transmembrane helices, in common with many other (non-ligand-gated) ion channels (see Fig. 3.20). P2X receptors are trimeric and each subunit has only two transmembrane domains (North, 2002). The nicotinic receptor and other cys-loop receptors are pentamers with two agonist binding sites on each receptor. Binding of one agonist molecule to one site increases the affinity of binding at the other site (positive cooperativity) and both sites need to be occupied for the receptor to be activated and the channel to open. Some ionotropic glutamate receptors have as many as four agonist binding sites and P2X receptors have three, but they appear to open when two agonist molecules are bound. Once again we realise that the simple model of receptor

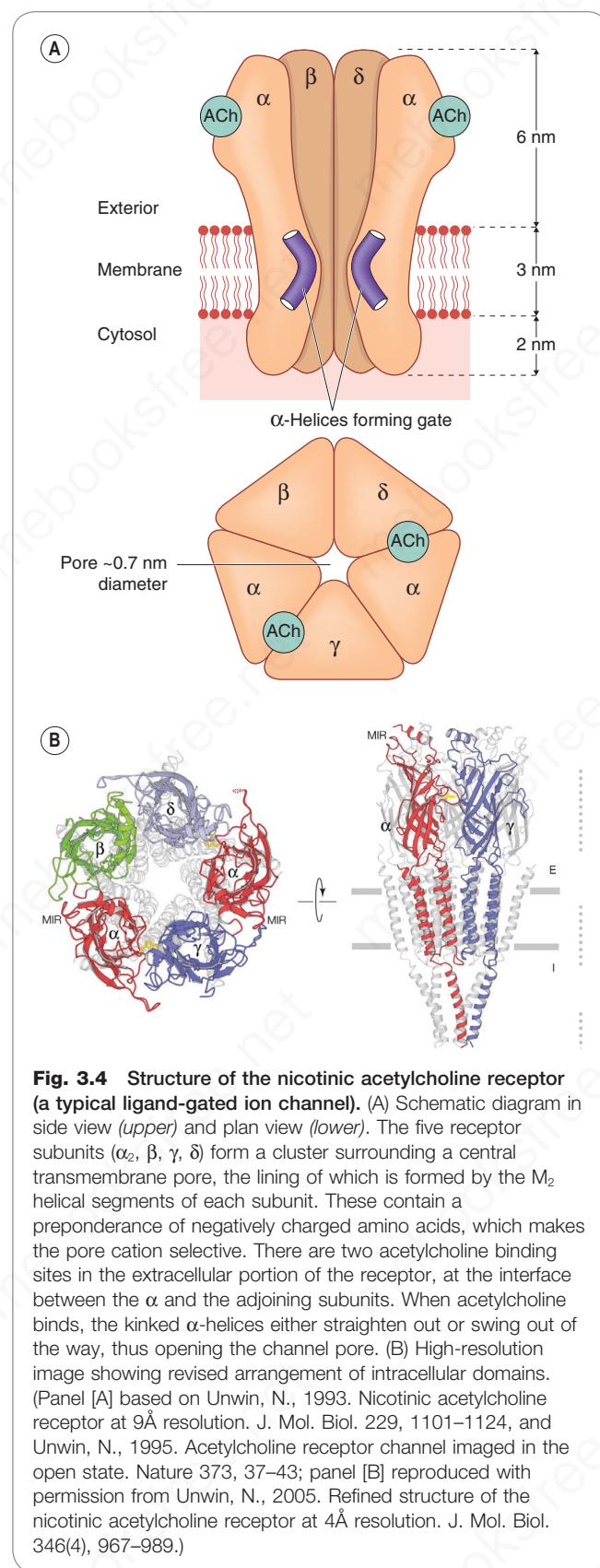


Fig. 3.4 Structure of the nicotinic acetylcholine receptor (a typical ligand-gated ion channel). (A) Schematic diagram in side view (upper) and plan view (lower). The five receptor subunits ($\alpha_2, \beta, \gamma, \delta$) form a cluster surrounding a central transmembrane pore, the lining of which is formed by the M_2 helical segments of each subunit. These contain a preponderance of negatively charged amino acids, which makes the pore cation selective. There are two acetylcholine binding sites in the extracellular portion of the receptor, at the interface between the α and the adjoining subunits. When acetylcholine binds, the kinked α -helices either straighten out or swing out of the way, thus opening the channel pore. (B) High-resolution image showing revised arrangement of intracellular domains. (Panel [A] based on Unwin, N., 1993. Nicotinic acetylcholine receptor at 9 Å resolution. *J. Mol. Biol.* 229, 1101–1124, and Unwin, N., 1995. Acetylcholine receptor channel imaged in the open state. *Nature* 373, 37–43; panel [B] reproduced with permission from Unwin, N., 2005. Refined structure of the nicotinic acetylcholine receptor at 4 Å resolution. *J. Mol. Biol.* 346(4), 967–989.)

⁷In early studies the Torpedo electric ray was used to isolate and purify the nicotinic receptor as it expresses a very high density of nicotinic receptors on its electroplaques. We now realise that the subunit compositions of the mammalian neuromuscular (Ch. 14) and neuronal (Chs 14 and 40) nicotinic receptors are different from that of the Torpedo but here we focus on the Torpedo receptor to keep it simple.

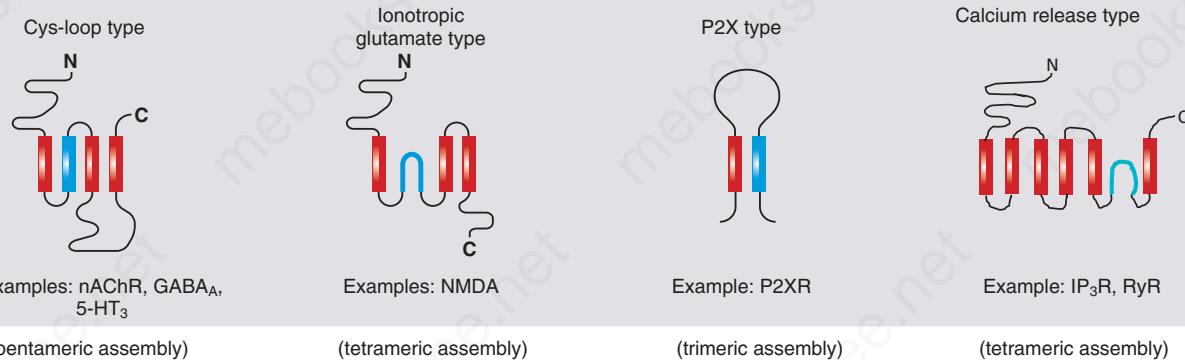


Fig. 3.5 Molecular architecture of ligand-gated ion channels. Red and blue rectangles represent membrane-spanning α -helices and blue hairpins represent the P loop pore-forming regions. 5-HT₃, 5-hydroxytryptamine type 3 receptor; GABA_A, GABA type A receptor; IP₃R, inositol trisphosphate receptor; nAChR, nicotinic acetylcholine receptor; NMDA, N-methyl-D-aspartic acid receptor; P2XR, purine P2X receptor; RyR, ryanodine receptor.

activation shown in Fig. 2.1 is an oversimplification as it only considered one agonist molecule binding to produce a response. For two or more agonist molecules binding, more complex mathematical models are needed (see Colquhoun, 2006).

THE GATING MECHANISM

Receptors of this type control the fastest synaptic events in the nervous system, in which a neurotransmitter acts on the postsynaptic membrane of a nerve or muscle cell and transiently increases its permeability to particular ions. Most excitatory neurotransmitters, such as acetylcholine at the neuromuscular junction (Ch. 14) or glutamate in the central nervous system (Ch. 39), cause an increase in Na⁺ and K⁺ permeability and in some instances Ca²⁺ permeability. At negative membrane potentials this results in a net inward current carried mainly by Na⁺, which depolarises the cell and increases the probability that it will generate an action potential. The action of the transmitter reaches a peak in a fraction of a millisecond, and usually decays within a few milliseconds. The sheer speed of this response implies that the coupling between the receptor and the ion channel is a direct one, and the molecular structure of the receptor-channel complex (see earlier) agrees with this. In contrast to other receptor families, no intermediate biochemical steps are involved in the transduction process.

▼ The patch clamp recording technique, devised by Neher and Sakmann, allows the very small current flowing through a single ion channel to be measured directly (Fig. 3.6). The patch clamp technique provides a view, rare in biology, of the physiological behaviour of individual protein molecules in real time, and has given many new insights into the gating reactions and permeability characteristics of both ligand-gated channels and voltage-gated channels. The magnitude of the single channel conductance confirms that permeation occurs through a physical pore through the membrane, because the ion flow is too large (about 10⁷ ions per second) to be compatible with a carrier mechanism. The channel conductance produced by different agonists is the same, whereas the mean channel lifetime varies. The ligand-receptor interaction scheme shown in Chapter 2 is a useful model for ion-channel gating. The conformation R*, representing the open state of the ion channel, is thought to be the same for all agonists, accounting for the finding that the channel conductance does not vary. Kinetically, the mean open time is determined mainly by the closing rate constant, α , and this varies from one drug to another. As explained in Chapter 2 (see Fig. 2.1), an agonist of high efficacy that activates a large proportion of the receptors that it occupies will

be characterised by $\beta/\alpha \gg 1$, whereas for a drug of low efficacy β/α has a lower value.

At some ligand-gated ion channels the situation is more complicated because different agonists may cause individual channels to open to one or more of several distinct conductance levels (see Fig. 3.6B). This implies that there is more than one R* conformation. Furthermore, desensitisation of ligand-gated ion channels (see Ch. 2) also involves one or more additional agonist-induced conformational states. These findings necessitate some elaboration of the simple scheme in which only a single open state, R*, is represented, and are an example of the way in which the actual behaviour of receptors makes our theoretical models look a little threadbare.

Ligand-gated ion channels



- These are sometimes called ionotropic receptors.
- They are involved mainly in fast synaptic transmission.
- There are several structural families, the commonest being heteromeric assemblies of four or five subunits, with transmembrane helices arranged around a central aqueous channel.
- Ligand binding and channel opening occur on a millisecond timescale.
- Examples include the nicotinic acetylcholine, GABA type A (GABA_A), glutamate (e.g. N-methyl-D-aspartic acid receptor [NMDA]) and ATP (P2X) receptors.

TYPE 2: G PROTEIN-COUPLED RECEPTORS

GPCRs constitute the commonest single class of targets for therapeutic drugs. The GPCR family comprises many of the receptors that are familiar to pharmacologists, such as muscarinic AChRs, adrenoceptors, dopamine receptors, 5-HT (serotonin) receptors, receptors for many peptides, purine receptors and many others, including the chemoreceptors involved in olfaction and pheromone detection, and also many ‘orphans’ (see Fredriksson & Schiöth, 2005). For most of these, pharmacological and molecular studies have revealed a variety of subtypes. All have the characteristic heptahelical structure (see Fig. 3.3B).

Many neurotransmitters, apart from peptides, can interact with both GPCRs and ligand-gated channels, allowing the

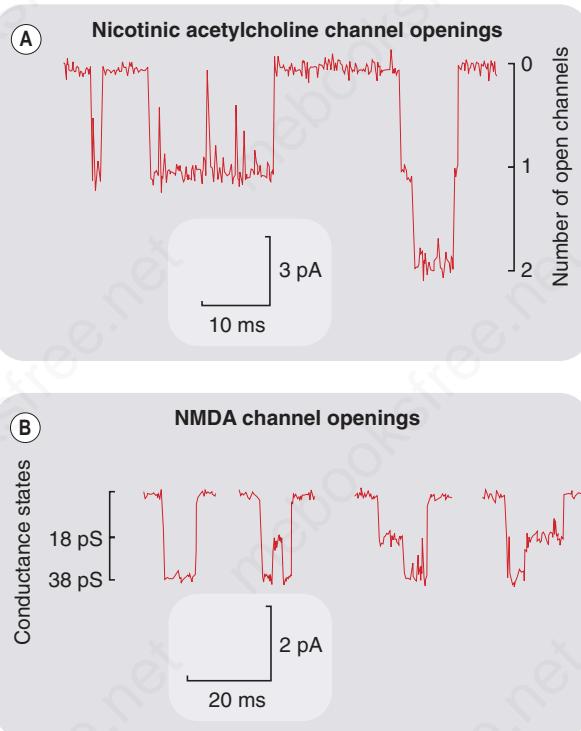


Fig. 3.6 Single channel openings recorded by the patch clamp technique. (A) Acetylcholine-operated ion channels at the frog motor endplate. The pipette, which was applied tightly to the surface of the membrane, contained 10 µmol/L ACh. The downward deflections show the currents flowing through single ion channels in the small patch of membrane under the pipette tip. Towards the end of the record, two channels can be seen to open with a discrete step from the first to the second. (B) Single-channel *N*-methyl-D-aspartic acid receptor (NMDA) receptor currents recorded from cerebellar neurons in the outside-out patch conformation. NMDA was added to the outside of the patch to activate the channel. The channel opens to multiple conductance levels. In (B) the openings to the higher conductance level and the subsequent closings are smooth, indicating that one channel is opening (two channels would not be expected to open and close simultaneously) whereas in (A) there are discrete steps indicating two channels. (Panel [A] courtesy D. Colquhoun and D.C. Ogden; panel [B] reproduced with permission from Cull-Candy, S.G. & Usowicz, M.M., 1987. Nature 325, 525–528.)

same molecule to produce fast (through ligand-gated ion channels) and relatively slow (through GPCRs) effects. Individual peptide hormones, however, generally act either on GPCRs or on kinase-linked receptors (see later), but rarely on both, and a similar choosiness applies to the many ligands that act on nuclear receptors.⁸

MOLECULAR STRUCTURE

In 1986 the first pharmacologically relevant GPCR, the β_2 adrenoceptor (Ch. 15), was cloned. Thereafter molecular

⁸Examples of promiscuity are increasing, however. Steroid hormones, normally faithful to nuclear receptors, make the occasional pass at ion channels and GPCRs, and some eicosanoids act on nuclear receptors as well as GPCRs. Nature is quite open-minded, although such examples are liable to make pharmacologists frown and students despair.

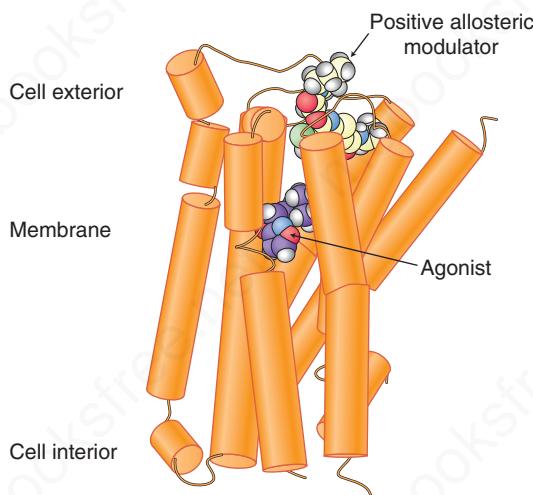


Fig. 3.7 Structure of the M₂ muscarinic receptor. High-resolution image showing the conformation of the M₂ muscarinic receptor bound with both an agonist (orthosteric) and a positive allosteric modulator. The brown cylinders represent the transmembrane domains. The full extent of the N- and C-terminal domains and the third intracellular loop are not shown. (Courtesy A. Christopoulos.)

biology caught up very rapidly with pharmacology, and with the sequencing of the human genome the amino acid sequence of all the GPCRs hitherto identified by their pharmacological properties was revealed, as was the structure of many novel GPCRs. More recently the difficulties of crystallising GPCRs have been overcome, allowing the use of the powerful technique of X-ray crystallography to study the three-dimensional molecular structure of these receptors in detail (Fig. 3.7) (Zhang et al., 2015). Also, computational molecular docking and nuclear magnetic resonance (NMR) methods have been developed to study ligand binding and subsequent conformational changes associated with activation (see Sounier et al., 2015). This is starting to provide important information on agonist- and antagonist-bound receptor conformations as well as receptor-G protein interactions. From such studies we are gaining a clearer picture of the mechanism of activation of GPCRs and the factors determining agonist efficacy, as well as having a better basis for designing new GPCR ligands.

GPCRs consist of a single polypeptide chain, usually of 350–400 amino acid residues, but in some cases up to 1100 residues. The general anatomy is shown in Fig. 3.3B. Their characteristic structure comprises seven transmembrane α -helices, similar to those of the ion channels discussed previously, with an extracellular N-terminal domain of varying length, and an intracellular C-terminal domain.

GPCRs are divided into three main classes – A, B and C (Table 3.2). There is considerable sequence homology between the members of one class, but little between different classes. They share the same seven transmembrane helix (heptahelical) structure, but differ in other respects, principally in the length of the extracellular N-terminus and the location of the agonist binding domain. Class A is by far the largest, comprising most monoamine, neuropeptide and chemokine receptors. Class B includes receptors for some other peptides, such as calcitonin and glucagon. Class C is the smallest, its main members being the

Table 3.2 Main G protein-coupled receptor classes^{a,b}

Class	Receptors ^b	Structural features
A: rhodopsin family	The largest group. Receptors for most amine neurotransmitters, many neuropeptides, purines, prostanooids, cannabinoids, etc.	Short extracellular (N-terminal) tail. Ligand binds to transmembrane helices (amines) or to extracellular loops (peptides)
B: secretin/glucagon receptor family	Receptors for peptide hormones, including secretin, glucagon, calcitonin	Intermediate extracellular tail incorporating ligand-binding domain
C: metabotropic glutamate receptor/calcium sensor family	Small group. Metabotropic glutamate receptors, GABA _B receptors, Ca ²⁺ -sensing receptors	Long extracellular tail incorporating ligand-binding domain

^aOther classes include frizzled G protein-coupled receptors (GPCRs), adhesion GPCRs and receptors for pheromones.

^bFor full lists, see <www.guidetopharmacology.org>.

metabotropic glutamate and GABA receptors, and the Ca²⁺-sensing receptors.⁹

▼ The understanding of the function of receptors of this type owes much to studies of a closely related protein, *rhodopsin*, which is responsible for transduction in retinal rods. This protein is abundant in the retina, and much easier to study than receptor proteins (which are anything but abundant); it is built on an identical plan to that shown in Fig. 3.3B and also produces a response in the rod (hyperpolarisation, associated with inhibition of Na⁺ conductance) through a mechanism involving a G protein (see p. 32, Fig. 3.9). The most obvious difference is that a photon, rather than an agonist molecule, produces the response. In effect, rhodopsin can be regarded as incorporating its own inbuilt agonist molecule, namely *retinal*, which isomerises from the *trans* (inactive) to the *cis* (active) form when it absorbs a photon.

For small molecules, such as noradrenaline (norepinephrine) and acetylcholine, the ligand-binding domain of class A receptors is buried in the cleft between the α -helical segments within the membrane (see Figs 3.3B and 3.7), similar to the slot occupied by retinal in the rhodopsin molecule.¹⁰ Peptide ligands, such as substance P (Ch. 19), bind more superficially to the extracellular loops, as shown in Fig. 3.3B. From crystal structures and single-site mutagenesis experiments, it is possible to map the ligand-binding domain of these receptors. Recent advances in computational molecular docking of ligands into the ligand-receptor-binding domain have made it possible to design novel synthetic ligands based primarily on knowledge of the receptor structure (see Manglik et al., 2016) – an important milestone in drug development, which has relied up to now mainly on the structure of endogenous mediators (such as histamine) or plant alkaloids (such as morphine) for its chemical inspiration.¹¹

⁹The Ca²⁺-sensing receptor (see Conigrave et al., 2000) is an unusual GPCR that is activated not by conventional mediators, but by extracellular Ca²⁺ in the range of 1–10 mmol/L – an extremely low affinity in comparison with other GPCR agonists. It is expressed by cells of the parathyroid gland, and serves to regulate the extracellular Ca²⁺ concentration by controlling parathyroid hormone secretion (Ch. 37). This homeostatic mechanism is quite distinct from the mechanisms for regulating intracellular Ca²⁺, discussed in Chapter 4.

¹⁰Hydrophilic small molecules access their ligand-binding domain from the extracellular space down the water-filled cleft, however for highly lipophilic molecules such as those activating the cannabinoid CB₁ and lysophospholipid S1P₁ receptors access appears to be through a membrane-embedded access channel in the side of the receptor.

¹¹In the past many lead compounds have come from screening huge chemical libraries (see Ch. 60). No inspiration was required, just robust assays, large computers and efficient robotics. Now with the generation of crystal structures we have moved to a more sophisticated age in drug discovery.

PROTEINASE-ACTIVATED RECEPTORS¹²

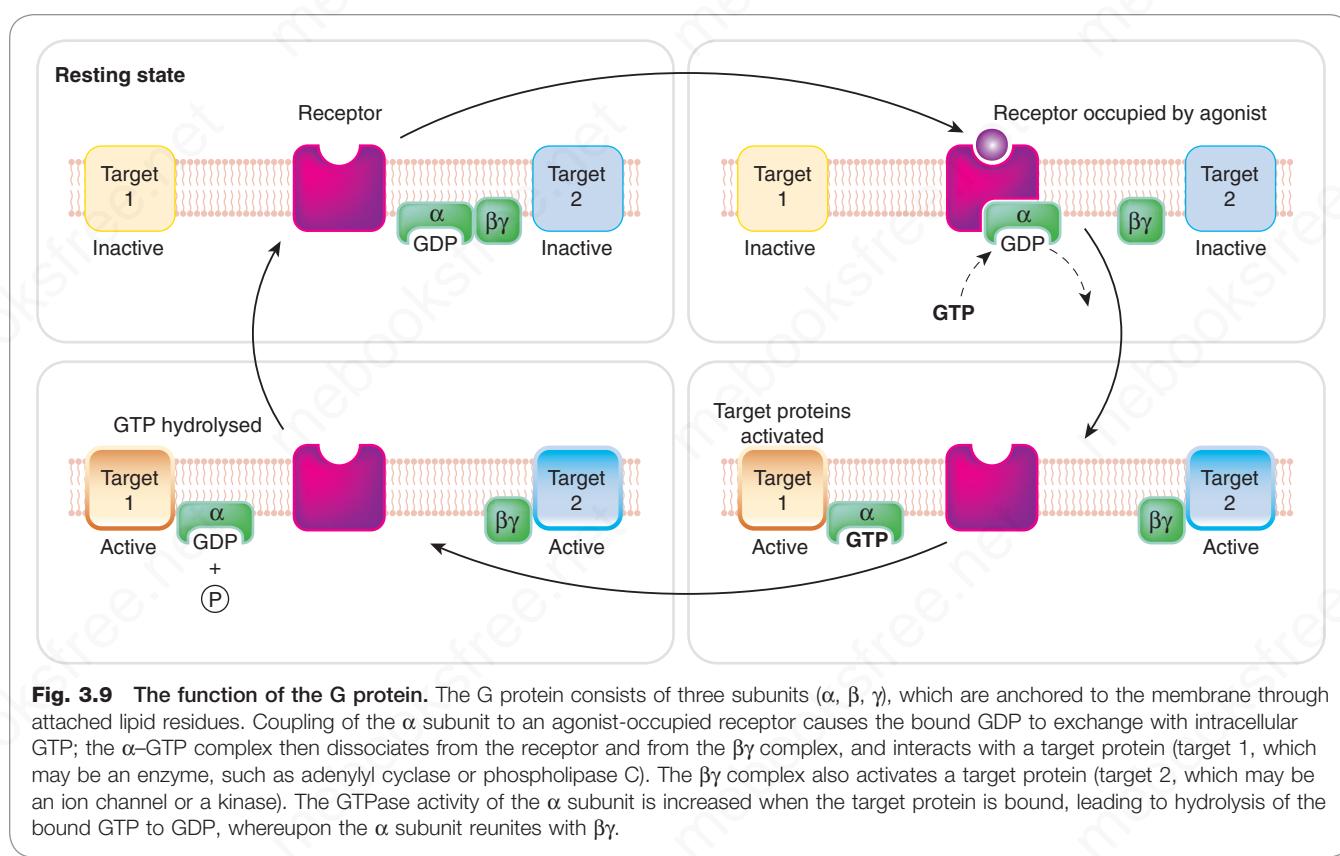
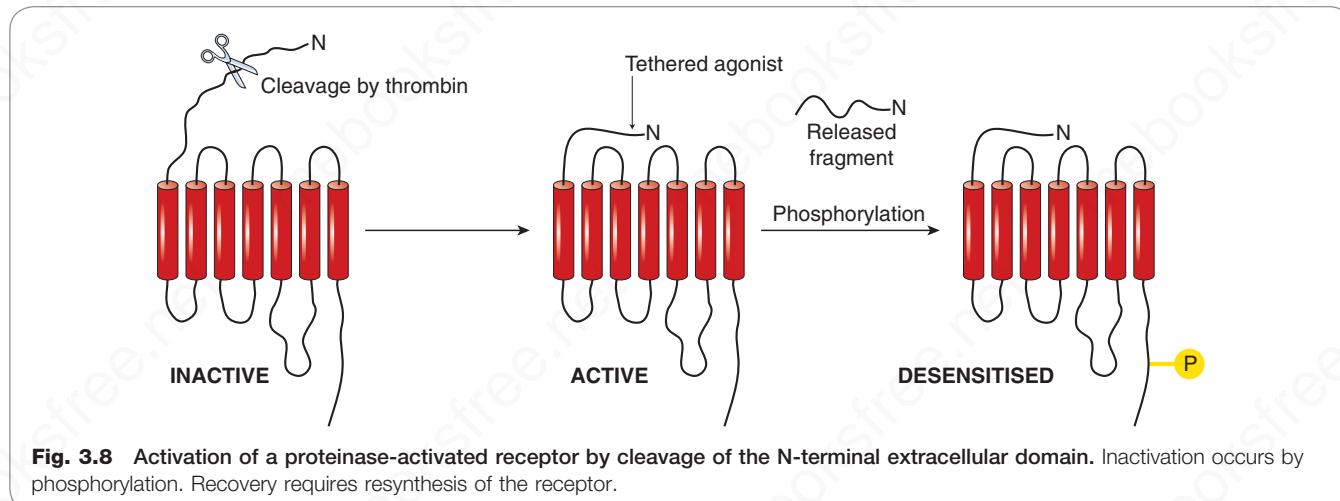
▼ Although activation of GPCRs is normally the consequence of a diffusible agonist, it can be the result of proteinase activation. Four types of protease-activated receptors (PARs) have been identified (see review by Ramachandran et al., 2012). Many proteinases, such as thrombin (a proteinase involved in the blood-clotting cascade; see Ch. 25), activate PARs by snipping off the end of the extracellular N-terminal tail of the receptor (Fig. 3.8) to expose five or six N-terminal residues that bind to receptor domains in the extracellular loops, functioning as a ‘tethered agonist’. Receptors of this type occur in many tissues and they appear to play a role in inflammation and other responses to tissue damage where tissue proteinases are released. A PAR molecule can be activated only once, because the cleavage cannot be reversed, and thus continuous resynthesis of the receptor protein is necessary. Inactivation occurs by a further proteolytic cleavage that frees the tethered ligand, or by desensitisation, involving phosphorylation (see Fig. 3.8), after which the receptor is internalised and degraded, to be replaced by newly synthesised protein.

G protein-coupled receptors



- These are sometimes called metabotropic or seven-transmembrane-domain (7-TMD) receptors.
- Structures comprise seven membrane-spanning α -helices.
- The G protein is a membrane protein comprising three subunits (α , β , γ), the α subunit possessing GTPase activity.
- The G protein interacts with a binding pocket on the intracellular surface of the receptor.
- When the G protein binds to an agonist-occupied receptor, the α subunit binds GTP, dissociates and is then free to activate an effector (e.g. a membrane enzyme). In some cases, the $\beta\gamma$ subunit is the activator species.
- Activation of the effector is terminated when the bound GTP molecule is hydrolysed, which allows the α subunit to recombine with $\beta\gamma$.
- There are several types of G protein, which interact with different receptors and control different effectors.
- Examples include muscarinic acetylcholine receptors, adrenoceptors, neuropeptide and chemokine receptors, and proteinase-activated receptors.

¹²These receptors were formerly called protease-activated receptors.



G PROTEINS AND THEIR ROLE

G proteins comprise a family of membrane-resident proteins whose function is to respond to GPCR activation and pass on the message inside the cell to the effector systems that generate a cellular response. They represent the level of middle management in the organisational hierarchy, intervening between the receptors – choosy mandarins, alert to the faintest whiff of their preferred chemical – and the effector enzymes or ion channels – the blue-collar brigade that gets the job done without needing to know which hormone authorised the process. They are the go-between

proteins, but were actually called G proteins because of their interaction with the guanine nucleotides, GTP and GDP. For more detailed information on the structure and functions of G proteins, see reviews by Milligan and Kostenis (2006), and Oldham and Hamm (2008). G proteins consist of three subunits: α , β and γ (Fig. 3.9). Guanine nucleotides bind to the α subunit, which has enzymic (GTPase) activity, catalysing the conversion of GTP to GDP. The β and γ subunits remain together as a $\beta\gamma$ complex. The ' γ ' subunit is anchored to the membrane through a fatty acid chain, coupled to the G protein through a reaction known as *prenylation*. In the 'resting' state (see Fig. 3.9), the G protein

Table 3.3 The main G protein subtypes and their functions^a

Subtypes	Main effectors	Notes
Gα subunits^b		
G α_s	Stimulates adenylyl cyclase, causing increased cAMP formation	Activated by cholera toxin, which blocks GTPase activity, thus preventing inactivation
G α_i	Inhibits adenylyl cyclase, decreasing cAMP formation	Blocked by pertussis toxin, which prevents dissociation of $\alpha\beta\gamma$ complex
G α_o	? Limited effects of α subunit (effects mainly due to $\beta\gamma$ subunits)	Blocked by pertussis toxin. Occurs mainly in nervous system
G α_q	Activates phospholipase C, increasing production of second messengers inositol trisphosphate and diacylglycerol (see pp. 36–38) thus releasing Ca ²⁺ from intracellular stores and activating protein kinase C (PKC)	
G $\alpha_{12/13}$	Activates Rho and thus Rho kinase	
G$\beta\gamma$ subunits		
	Activate potassium channels Inhibit voltage-gated calcium channels Activate GPCR kinases (GRKs, pp. 38–39) Activate mitogen-activated protein kinase cascade Interact with some forms of adenylyl cyclase and with phospholipase C β	Many $\beta\gamma$ isoforms identified, but specific functions are not yet known

^aThis table lists only those isoforms of major pharmacological significance. Many more have been identified, some of which play roles in olfaction, taste, visual transduction and other physiological functions (see Offermanns, 2003).

^bInitially the subscripts 's' and 'i' were used to denote stimulatory and inhibitory actions on adenylyl cyclase but, subsequently, the terms used, 'q' and '12/13', have little logic behind their use.

GPCR, G protein-coupled receptor.

exists as an $\alpha\beta\gamma$ trimer, which may or may not be precoupled to the receptor, with GDP occupying the site on the α subunit. When a GPCR is activated by an agonist this induces small changes in residues around the ligand-binding pocket that translate to larger rearrangements of the intracellular regions of the receptor that open a cavity on the intracellular side of the receptor into which the G protein can bind, resulting in a high-affinity interaction of $\alpha\beta\gamma$ and the receptor. This agonist-induced interaction of $\alpha\beta\gamma$ with the receptor occurs within about 50 ms, causing the bound GDP to dissociate and to be replaced with GTP (GDP–GTP exchange), which in turn causes dissociation of the G protein trimer, releasing α -GTP from the $\beta\gamma$ subunits; these are the 'active' forms of the G protein, which diffuse in the membrane and can associate with various enzymes and ion channels, causing activation of the target (see Fig. 3.9). It was originally thought that only the α subunit had a signalling function, the $\beta\gamma$ complex serving merely as a chaperone to keep the flighty α subunits out of range of the various effector proteins that they might otherwise excite. However, the $\beta\gamma$ complexes actually make assignments of their own, and control effectors in much the same way as the α subunits. Association of α or $\beta\gamma$ subunits with target enzymes or channels can cause either activation or inhibition, depending on which G protein is involved (see Table 3.3). G protein activation results in amplification, because a single agonist–receptor complex can activate several G protein molecules in turn, and each of these can remain associated with their effector enzyme for long enough to produce many molecules of product. The product (see later) is often a 'second messenger', and further amplification occurs before the final cellular response is produced.

Signalling is terminated when the hydrolysis of GTP to GDP occurs through the inherent GTPase activity of the α subunit. The resulting α -GDP then dissociates from the effector, and reunites with $\beta\gamma$, completing the cycle.

▼ Attachment of the α subunit to an effector molecule actually increases its GTPase activity, the magnitude of this increase being different for different types of effector. Because GTP hydrolysis is the step that terminates the ability of the α subunit to produce its effect, regulation of its GTPase activity by the effector protein means that the activation of the effector tends to be self-limiting. In addition, there is a family of about 20 cellular proteins, regulators of G protein signalling (RGS) proteins (see review by Sjögren, 2017), that possess a conserved sequence that binds specifically to α subunits to increase greatly their GTPase activity, so hastening the hydrolysis of GTP and inactivating the complex. RGS proteins thus exert an inhibitory effect on G protein signalling, a mechanism that is thought to have a regulatory function in many situations.

Different GPCRs couple to different G proteins and thus produce distinct cellular responses. For example, M₂ muscarinic acetylcholine receptors (mAChRs) and β_1 adrenoceptors, both of which occur in cardiac muscle cells, produce opposite functional effects (Chs 14 and 15). Four main classes of G protein (G_s, G_i, G_o and G_q) are of pharmacological importance (Table 3.3). These differ primarily in the α subunit they contain.¹³ G proteins show selectivity with

¹³In humans there are 21 known subtypes of G α , 6 of G β and 12 of G γ , providing, in theory, about 1500 variants of the trimer. We know little about the role of different α , β and γ subtypes, but it would be rash to assume that the variations are functionally irrelevant. By now, you will be unsurprised (even if somewhat bemused) by such a display of molecular heterogeneity, for it is the way of evolution.

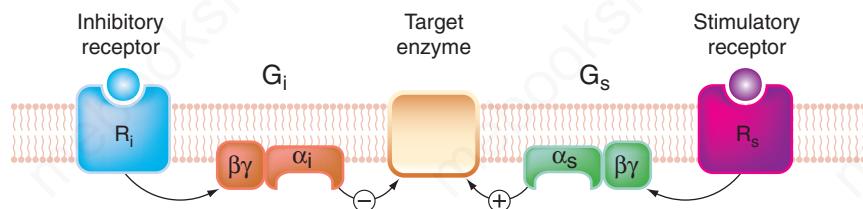


Fig. 3.10 Bidirectional control of a target enzyme, such as adenylyl cyclase by G_s and G_i. Heterogeneity of G proteins allows different receptors to exert opposite effects on a target enzyme.

respect to both the receptors and the effectors with which they couple, having specific recognition domains in their structure complementary to specific G protein-binding domains in the receptor and effector molecules. For example, G_s and G_i produce, respectively, stimulation and inhibition of the enzyme *adenylyl cyclase* (Fig. 3.10).

One functional difference that has been useful as an experimental tool to distinguish which type of G protein is involved in different situations concerns the action of two bacterial toxins, *cholera toxin* and *pertussis toxin* (see Table 3.3). These toxins, which are enzymes, catalyse a conjugation reaction (ADP ribosylation) on the α subunit of G proteins. Cholera toxin acts only on G_s, and it causes persistent activation. Many of the symptoms of cholera, such as the excessive secretion of fluid from the gastrointestinal epithelium (leading to 'rice-water stools'), are due to the uncontrolled activation of adenylyl cyclase that occurs. Pertussis toxin specifically blocks G_i and G_o by preventing dissociation of the G protein trimer. Pertussis toxin is released from *Bordetella pertussis* bacteria, which cause whooping cough. As with cholera toxin, the symptoms caused by pertussis toxin are related to its effects on G proteins, but in this case by inhibiting G_i and G_o rather than activating G_s and leading to changes in respiratory tract secretion and a distinctive cough rather than the copious diarrhoea of cholera.

TARGETS FOR G PROTEINS

The main targets for G proteins, through which GPCRs control different aspects of cell function (see Table 3.3), are:

- *adenylyl cyclase*, the enzyme responsible for cAMP formation;
- *phospholipase C*, the enzyme responsible for inositol phosphate and diacylglycerol (DAG) formation;
- *ion channels*, particularly calcium and potassium channels;
- *Rho A/Rho kinase*, a system that regulates the activity of many signalling pathways controlling cell growth, proliferation and motility, smooth muscle contraction, etc.;
- *mitogen-activated protein kinase* (MAP kinase), a system that controls many cell functions, including cell division and is also a target of several kinase-linked receptors.

The adenylyl cyclase/cAMP system

The discovery by Sutherland and his colleagues of the role of cAMP (cyclic 3',5'-adenosine monophosphate) as an intracellular mediator demolished at a stroke the barriers that existed between biochemistry and pharmacology, and

introduced the concept of second messengers in signal transduction. cAMP is a nucleotide synthesised within the cell from ATP by the action of a membrane-bound enzyme, adenylyl cyclase. It is produced continuously and inactivated by hydrolysis to 5'-AMP by the action of a family of enzymes known as phosphodiesterases (PDEs). Many different drugs, hormones and neurotransmitters act on GPCRs and increase or decrease the catalytic activity of adenylyl cyclase (see Fig. 3.10), thus raising or lowering the concentration of cAMP within the cell. In mammalian cells there are 10 different molecular isoforms of the enzyme, some of which respond selectively to Gα_s or Gα_i.

Cyclic AMP regulates many aspects of cellular function including, for example, enzymes involved in energy metabolism, cell division and cell differentiation, ion transport, ion channels and the contractile proteins in smooth muscle. These varied effects are, however, all brought about by a common mechanism, namely the activation of protein kinases by cAMP (known as cyclic AMP-dependent protein kinases) in eukaryotic cells. One important cyclic AMP-dependent protein kinase is *protein kinase A* (PKA). Protein kinases regulate the function of many different cellular proteins by controlling protein phosphorylation. Fig. 3.11 shows how increased cAMP production in response to β-adrenoceptor activation affects enzymes involved in glycogen and fat metabolism in liver, fat and muscle cells. The result is a coordinated response in which stored energy in the form of glycogen and fat is made available as glucose to fuel muscle contraction.

Other examples of regulation by PKA include the increased activity of voltage-gated calcium channels in heart muscle cells (see Ch. 22). Phosphorylation of these channels increases the amount of Ca²⁺ entering the cell during the action potential, and thus increases the force of contraction of the heart.

In smooth muscle, PKA phosphorylates (thereby inactivating) another enzyme, *myosin light-chain kinase*, which is required for contraction. This accounts for the smooth muscle relaxation produced by many drugs that increase cAMP production in smooth muscle (see Ch. 4).

As mentioned earlier, receptors linked to G_i rather than G_s inhibit adenylyl cyclase, and thus reduce cAMP formation to elicit opposing responses to those receptors which activate G_s. Examples include certain types of mAChR (e.g. the M₂ receptor of cardiac muscle; see Ch. 14), α₂ adrenoceptors in smooth muscle (Ch. 15) and opioid receptors (see Ch. 43). Adenylyl cyclase can be activated directly by drugs such as **forskolin**, which is used experimentally to study the role of the cAMP system.

Cyclic AMP is hydrolysed within cells by PDEs, an important and ubiquitous family of enzymes. Twenty-four

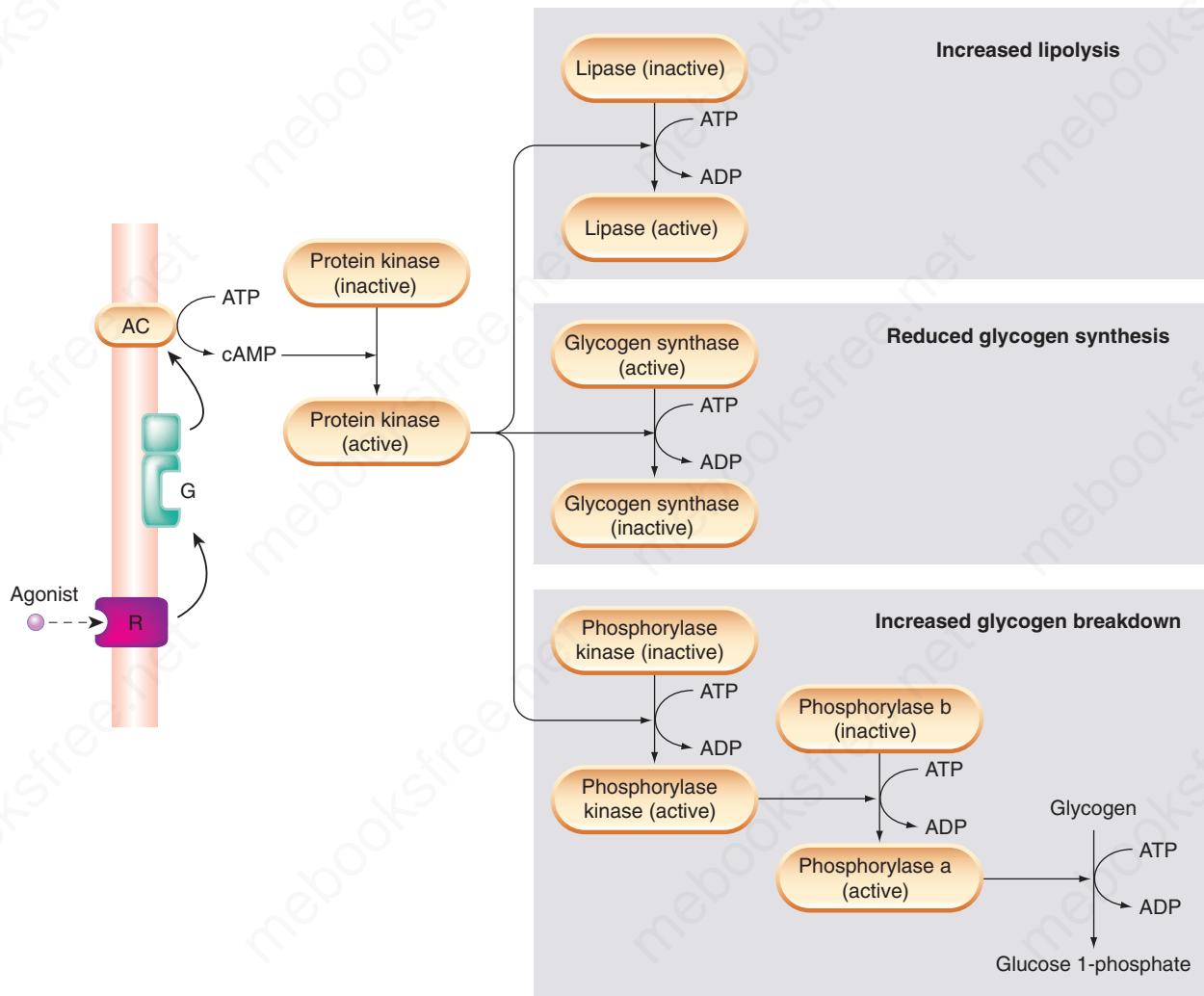


Fig. 3.11 Regulation of energy metabolism by cAMP. AC, adenyl cyclase.

PDE subtypes exist, of which some are more selective for cAMP, while others are more selective for cGMP. Most are weakly inhibited by drugs such as methylxanthines (e.g. **theophylline** and **caffeine**; see Chs 29 and 49). **Roflumilast** (used to treat chronic obstructive pulmonary disease [COPD]; Ch. 29) is selective for PDE_{4B}, expressed in inflammatory cells; **milrinone** (a positive inotrope that makes the heart beat harder and is sometimes used for symptoms in patients awaiting heart transplantation; Ch. 22) is selective for PDE_{2A}, which is expressed in heart muscle; **sildenafil** (better known as Viagra; Ch. 36) is selective for PDE_{5A}, and consequently enhances the vasodilator effects of nitric oxide (NO) and drugs that release NO, whose effects are mediated by cGMP (see Ch. 21). The similarity of some of the actions of these drugs to those of sympathomimetic amines (Ch. 15) probably reflects their common property of increasing the intracellular concentration of cAMP.

The phospholipase C/inositol phosphate system

The *phosphoinositide* system, an important intracellular second messenger system, was first discovered in the 1950s

by Hokin and Hokin, whose recondite interests centred on the mechanism of salt secretion by the nasal glands of seabirds. They found that secretion was accompanied by increased turnover of a minor class of membrane phospholipids known as phosphoinositides (collectively known as PIs; Fig. 3.12). Subsequently, Michell and Berridge found that many hormones that produce an increase in free intracellular Ca²⁺ concentration (which include, for example, muscarinic agonists and α-adrenoceptor agonists acting on smooth muscle and salivary glands) also increase PI turnover. It was later found that one particular member of the PI family, namely phosphatidylinositol (4,5) bis-phosphate (PIP₂), which has additional phosphate groups attached to the inositol ring, plays a key role. PIP₂ is the substrate for a membrane-bound enzyme, phospholipase Cβ (PLCβ), which splits it into DAG and *inositol* (1,4,5) *trisphosphate* (IP₃; Fig. 3.13), both of which function as second messengers as discussed later (p. 36). The activation of PLCβ by various agonists is mediated through a G protein (G_q, see Table 3.3). After cleavage of PIP₂, the status quo is restored, as shown in Fig. 3.13, DAG being phosphorylated to form phosphatidic acid (PA), while the IP₃ is

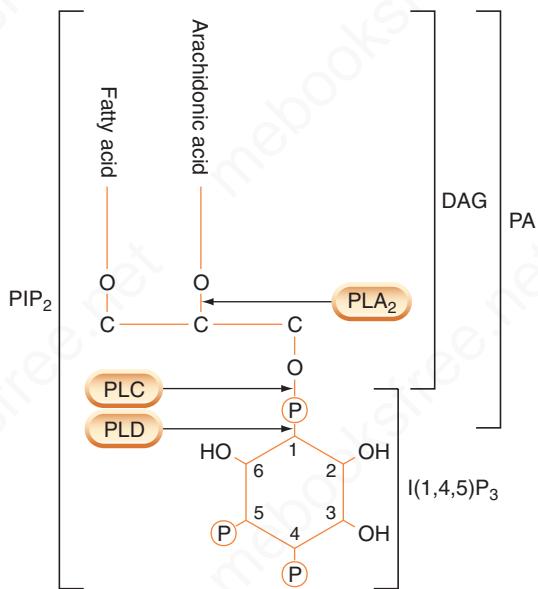


Fig. 3.12 Structure of phosphatidylinositol bisphosphate (PIP_2), showing sites of cleavage by different phospholipases to produce active mediators. Cleavage by phospholipase A2 (PLA_2) yields arachidonic acid. Cleavage by phospholipase C (PLC) yields inositol triphosphate ($\text{I}(1,4,5)\text{P}_3$) and diacylglycerol (DAG). PA, phosphatidic acid; PLD , phospholipase D.

dephosphorylated and then recoupled with PA to form PIP_2 once again.¹⁴ Lithium, an agent used in psychiatry (see Ch. 48), blocks this recycling pathway (see Fig. 3.13).

Inositol phosphates and intracellular calcium

Inositol (1,4,5) trisphosphate (IP_3) is a water-soluble mediator that is released into the cytosol and acts on a specific receptor – the IP_3 receptor – which is a ligand-gated calcium channel present on the membrane of the endoplasmic reticulum (see Fig. 3.5). The main role of IP_3 , described in more detail in Chapter 4, is to control the release of Ca^{2+} from intracellular stores. Because many drug and hormone effects involve intracellular Ca^{2+} , this pathway is particularly important.

Diacylglycerol and protein kinase C

DAG is produced, as well as IP_3 , whenever receptor-induced PI hydrolysis occurs. The main effect of DAG is to activate a protein kinase, *protein kinase C* (PKC), which catalyses the phosphorylation of several intracellular proteins. DAG, unlike the inositol phosphates, is highly lipophilic and remains within the membrane. It binds to a specific site on the PKC molecule, causing the enzyme to migrate from the cytosol to the cell membrane, thereby becoming activated. There are at least 10 different mammalian PKC subtypes, which have distinct cellular distributions and phosphorylate different proteins. Several are activated by DAG and raised intracellular Ca^{2+} , both of which are produced by activation of GPCRs.¹⁵ PKCs are also activated

by *phorbol esters* (highly irritant, tumour-promoting compounds produced by certain plants), which have been extremely useful in studying the functions of PKC. One of the subtypes is activated by the lipid mediator *arachidonic acid* (see Ch. 18) generated by the action of phospholipase A₂ on membrane phospholipids, so PKC activation can also occur with agonists that activate this enzyme. The various PKC isoforms, like the tyrosine kinases discussed later (p. 40), act on many different functional proteins, such as ion channels, receptors, enzymes (including other kinases), transcription factors and cytoskeletal proteins. Protein phosphorylation by kinases plays a central role in signal transduction, and controls many different aspects of cell function. The DAG–PKC link provides a mechanism whereby GPCRs can mobilise this army of control freaks.

Ion channels as targets for G proteins

Another major function of GPCRs is to control ion channel function directly by mechanisms that do not involve second messengers such as cAMP or inositol phosphates. Direct G protein–channel interaction, through the $\beta\gamma$ subunits of G_i and G_o proteins, appears to be a general mechanism for controlling K^+ and Ca^{2+} channels. In cardiac muscle, for example, mACHRs enhance K^+ permeability in this way (thus hyperpolarising the cells and inhibiting electrical activity; see Ch. 22). Similar mechanisms operate in neurons, where many inhibitory drugs, such as opioid analgesics, reduce excitability by opening certain K^+ channels – known as G protein-activated inwardly rectifying K^+ channels (GIRK) – or by inhibiting voltage-activated N and P/Q type Ca^{2+} channels, thus reducing neurotransmitter release (see Chs 4 and 43).

The Rho/Rho kinase system

▼ This signal transduction pathway (see Bishop & Hall, 2000) is activated by certain GPCRs (and also by non-GPCR mechanisms), which couple to G proteins of the $\text{G}_{12/13}$ type. The free G protein α subunit interacts with a *guanosine nucleotide exchange factor*, which facilitates GDP–GTP exchange at another GTPase, Rho. Rho–GDP, the resting form, is inactive, but when GDP–GTP exchange occurs, Rho is activated, and in turn activates Rho kinase. Rho kinase phosphorylates many substrate proteins and controls a wide variety of cellular functions, including smooth muscle contraction and proliferation, cell movement and migration, angiogenesis and synaptic remodelling. By enhancing hypoxia-induced pulmonary artery vasoconstriction, activation of Rho kinase is thought to be important in the pathogenesis of pulmonary hypertension (see Ch. 23). Specific Rho kinase inhibitors are in development for several clinical indications including glaucoma – an area to watch.

The MAP kinase system

▼ The MAP kinase system involves several signal transduction pathways (Fig. 3.15) that are activated not only by various cytokines and growth factors acting on kinase-linked receptors (see p. 42, Fig. 3.17), but also by ligands activating GPCRs. The coupling of GPCRs to different families of MAP kinases can involve G protein α and $\beta\gamma$ subunits as well as *Src* and *arrestins* – proteins also involved in GPCR desensitisation (see p. 38). The MAP kinase system controls many processes involved in gene expression, cell division, apoptosis and tissue regeneration.

¹⁴Alternative abbreviations for these mediators are PtdIns (PI), PtdIns(4,5)-P₂ (PIP_2), Ins (1,4,5)-P₃ (IP_3).

¹⁵PKCs were originally named as Ca^{2+} -dependent protein kinases (PKC), as opposed to cAMP-dependent PKA. Although later subtypes were found not to be Ca^{2+} -dependent, the PKC name has stuck.

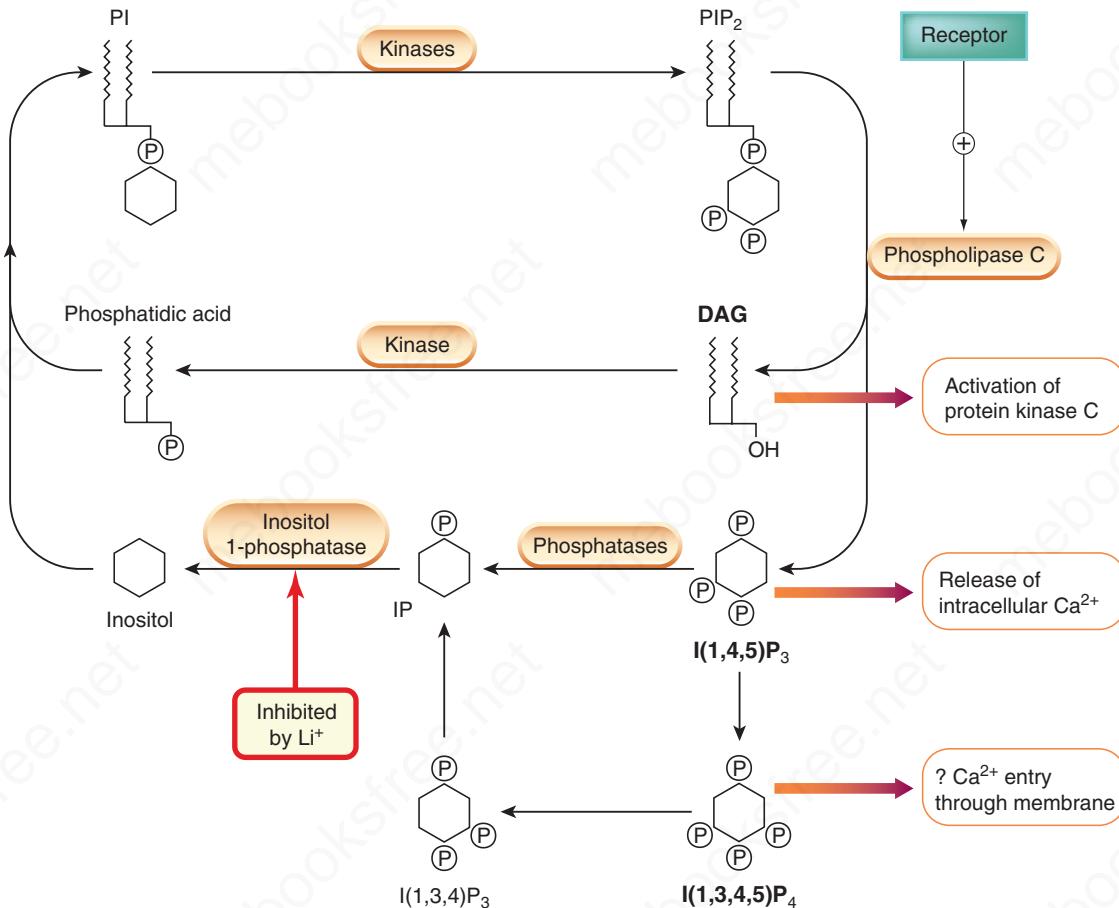


Fig. 3.13 The phosphatidylinositol (PI) cycle. Receptor-mediated activation of phospholipase C results in the cleavage of phosphatidylinositol bisphosphate (PIP_2), forming diacylglycerol (DAG) (which activates protein kinase C) and inositol trisphosphate (IP_3) (which releases intracellular Ca^{2+}). The role of inositol tetraphosphate (IP_4), which is formed from IP_3 and other inositol phosphates, is unclear, but it may facilitate Ca^{2+} entry through the plasma membrane. IP_3 is inactivated by dephosphorylation to inositol. DAG is converted to phosphatidic acid, and these two products are used to regenerate PI and PIP_2 .

Effectors controlled by G proteins

Two key second messenger pathways are controlled by receptors via G proteins:

- Adenyl cyclase/cAMP:
 - can be activated or inhibited by pharmacological ligands, depending on the nature of the receptor and G protein;
 - adenyl cyclase catalyses formation of the intracellular messenger cAMP;
 - cAMP activates protein kinases such as protein kinase A (PKA) that control cell function in many different ways by causing phosphorylation of various enzymes, carriers and other proteins.
- Phospholipase C/inositol trisphosphate (IP_3)/diacylglycerol (DAG):
 - catalyses the formation of two intracellular messengers, IP_3 and DAG, from membrane phospholipid;
 - IP_3 acts to increase free cytosolic Ca^{2+} by releasing Ca^{2+} from intracellular compartments

– increased free Ca^{2+} initiates many events, including contraction, secretion, enzyme activation and membrane hyperpolarisation;

– DAG activates various protein kinase C (PKC) isoforms, which control many cellular functions by phosphorylating a variety of proteins.

Receptor-linked G proteins also control:

- Ion channels:
 - opening potassium channels, resulting in membrane hyperpolarisation;
 - inhibiting calcium channels, thus reducing neurotransmitter release.
- Phospholipase A₂ (and thus the formation of arachidonic acid and eicosanoids).

The main postulated roles of GPCRs in controlling enzymes and ion channels are summarised in Fig. 3.14.



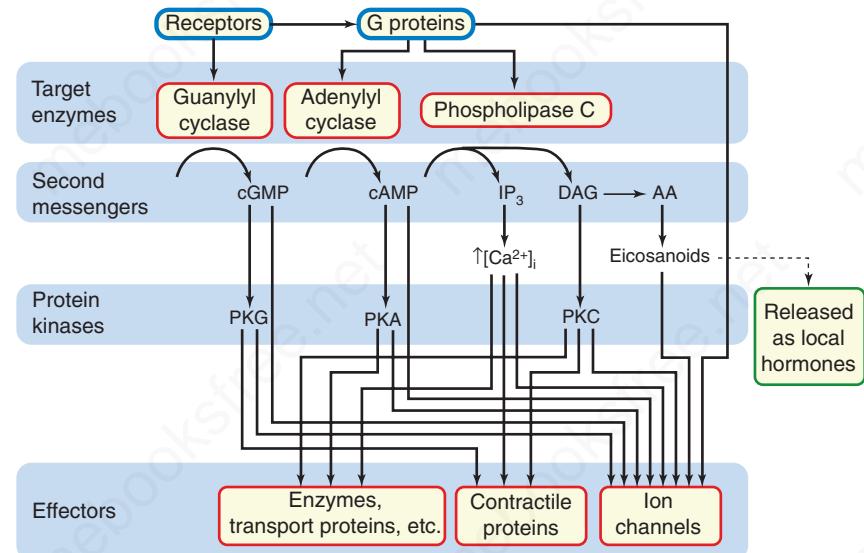


Fig. 3.14 G protein and second messenger control of cellular effector systems. Not shown in this diagram are signalling pathways where arrestins, rather than G proteins, link G protein-coupled receptors to downstream events (see text and *Fig. 3.15*). AA, arachidonic acid; DAG, diacylglycerol; IP₃, inositol trisphosphate.

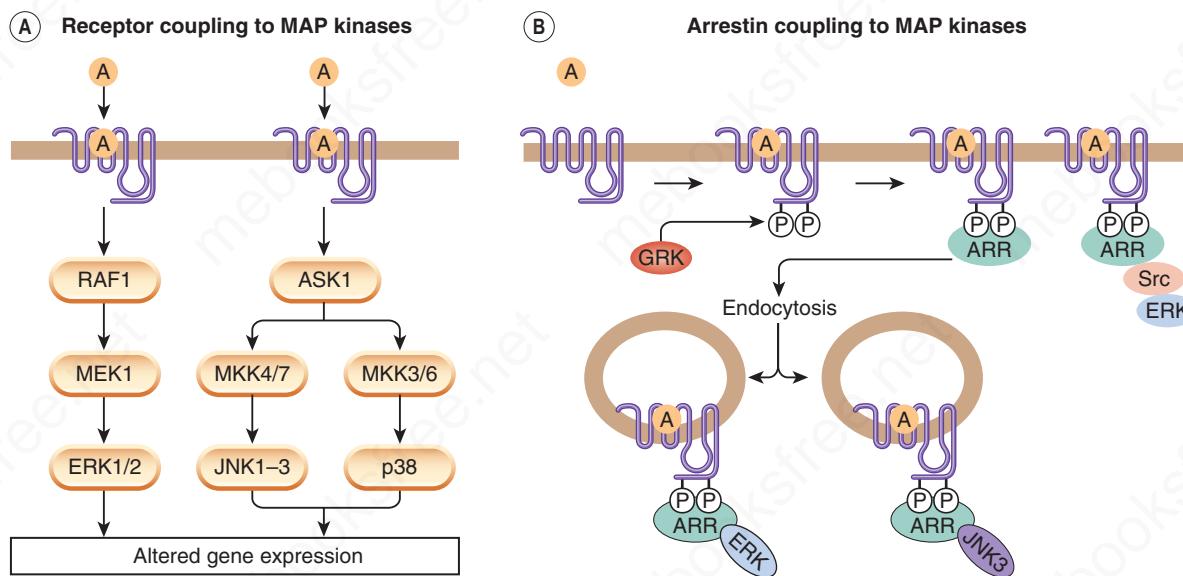


Fig. 3.15 G protein-coupled receptor (GPCR) activation of mitogen-activated protein (MAP) kinase cascade. (A) Sequential activation of the multiple components of the MAP kinase cascade. GPCR activation of MAP kinases can involve G α and G $\beta\gamma$ subunits (not shown). (B) Activation of ERK and JNK3 through interaction with arrestins (βARR). Activation of ERK can occur either at the plasma membrane involving Src, or by direct activation after internalisation of the receptor/arrestin complex. ARR, arrestin; GRK, G protein-coupled receptor kinase.

FURTHER DEVELOPMENTS IN GPCR BIOLOGY

▼ By the early 1990s, we thought we had more or less got the measure of GPCR function, as described previously. Since then, the plot has thickened, and further developments have necessitated a substantial overhaul of the basic model.

GPCR desensitisation

▼ As described in Chapter 2, desensitisation is a feature of most GPCRs, and the mechanisms underlying it have been extensively studied. *Homologous desensitisation* is restricted to the receptors activated

by the desensitising agonist, while *heterologous desensitisation* affects other GPCRs in addition. Two main processes are involved (see Kelly et al., 2008):

- receptor phosphorylation
- receptor internalisation (endocytosis)

The sequence of GPCRs includes certain residues (serine and threonine), mainly in the C-terminal cytoplasmic tail, which can be phosphorylated by specific GPCR kinases (GRKs) and by kinases such as PKA and PKC.

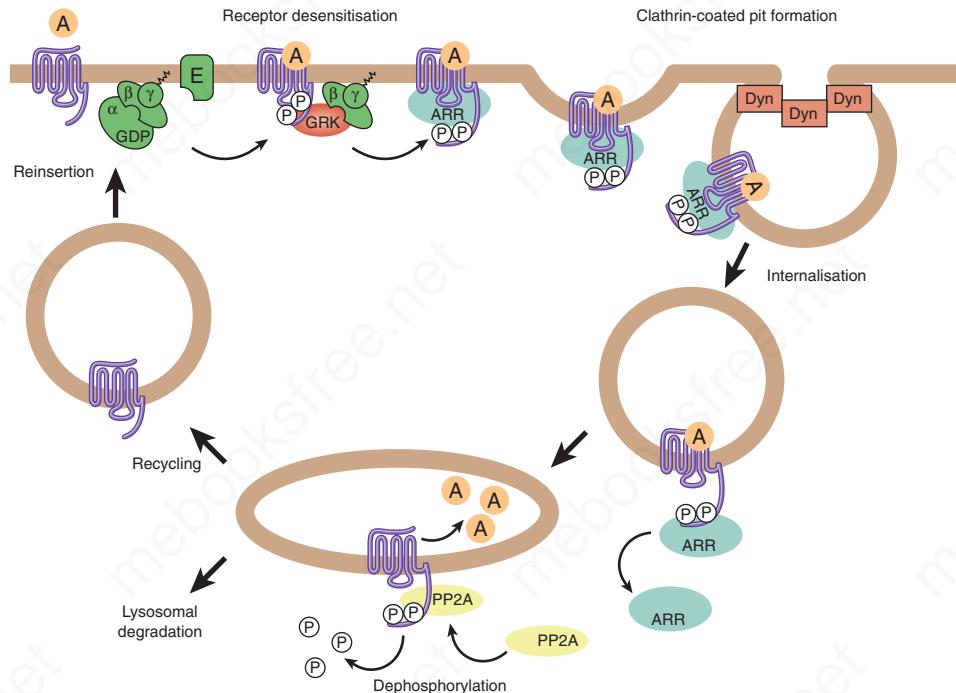


Fig. 3.16 Desensitisation and trafficking of G protein-coupled receptors (GPCRs). On prolonged agonist activation of the GPCR, selective GPCR kinases (GRKs) are recruited to the plasma membrane and phosphorylate the receptor. Arrestin (ARR) then binds and traffics the GPCR to clathrin-coated pits for subsequent internalisation into endosomes in a dynamin-dependent process. The GPCR is then dephosphorylated by a phosphatase (PP2A) and either recycled back to the plasma membrane or trafficked to lysosomes for degradation. *Dyn*, dynamin; *GRK*, G protein-coupled receptor kinase; *PP2A*, phosphatase 2A.

On receptor activation GRK2 and GRK3 are recruited to the plasma membrane by binding to free G protein $\beta\gamma$ subunits. GRKs then phosphorylate the receptors in their activated (i.e. agonist-bound) state. The phosphorylated receptor serves as a binding site for arrestins, intracellular proteins that block the interaction between the receptor and the G proteins producing a selective *homologous desensitisation*. Arrestin binding also targets the receptor for endocytosis in clathrin-coated pits (Fig. 3.16). The internalised receptor can then either be dephosphorylated and reinserted into the plasma membrane (*resensitisation*) or trafficked to lysosomes for degradation (*inactivation*). This type of desensitisation seems to occur with most GPCRs but with subtle differences that fascinate the aficionados.

Phosphorylation by PKA and PKC at residues different from those targeted by GRKs generally leads to impaired coupling between the activated receptor and the G protein, so the agonist effect is reduced. This can give rise to either homologous or heterologous desensitisation, depending on whether or not receptors other than that for the desensitising agonist are simultaneously phosphorylated by the kinases, some of which are not very selective. Receptors phosphorylated by second messenger kinases are probably not internalised and are reactivated by dephosphorylation by phosphatases when the agonist is removed.

GPCR oligomerisation

▼ The earlier view that GPCRs exist and function as monomeric proteins (in contrast to ion channels, which generally comprise multimeric complexes; see p. 28) was first overturned by work on

the GABA_B receptor. Two subtypes of this GPCR exist, encoded by different genes, and the functional receptor consists of a heterodimer of the two (see Ch. 39). A similar situation arises with G protein-coupled glutamate receptors. Oddly, although the GABA_B dimer has two potential agonist binding sites, one on each subunit, only one is functional and signalling is transmitted through the dimer to the other receptor in the dimer which couples to the G protein (see Fig. 39.9).

Other GPCRs are functional as monomers but it now seems likely that most, if not all, GPCRs can exist as either homomeric or heteromeric oligomers (i.e. dimers or larger oligomers) (Ferré et al., 2015). Within the opioid receptor family (see Ch. 43), the μ receptor was crystallised as a dimer and stable and functional heterodimers of κ and δ receptors, whose pharmacological properties differ from those of either parent, have been created in cell lines. More diverse GPCR combinations have also been found, such as that between dopamine (D₂) and somatostatin receptors, on which both ligands act with increased potency. Roaming even further afield in search of functional assignations, the dopamine receptor D₅ can couple directly with a ligand-gated ion channel, the GABA_A receptor, inhibiting the function of the latter without the intervention of any G protein (Liu et al., 2000). These interactions have so far been studied mainly in engineered cell lines, but they also occur in native cells. Functional dimeric complexes between angiotensin (AT₁) and bradykinin (B₂) receptors occur in human platelets and show greater sensitivity to angiotensin than 'pure' AT₁ receptors (AbdAlla et al., 2001). In women suffering from pregnancy-related hypertension (pre-eclamptic toxæmia), the number of these dimers increases due to increased expression of B₂ receptors, resulting – paradoxically – in increased sensitivity to the vasoconstrictor action of angiotensin.

It is too early to say what impact this newly discovered versatility of GPCRs in linking up with other receptors to form functional combinations will have on conventional pharmacology and therapeutics, but it could be considerable.

Constitutively active receptors

▼ GPCRs may be constitutively (i.e. spontaneously) active in the absence of any agonist (see Ch. 2 and review by Costa & Cotecchia, 2005). This was first shown for δ opioid receptors (see Ch. 43). There are now many other examples of native GPCRs that show constitutive activity when studied in vitro. The histamine H₃ receptor also shows constitutive activity in vivo, and this may prove to be a quite general phenomenon. It means that inverse agonists (see Ch. 2), which suppress this basal activity, may exert effects distinct from those of neutral antagonists, which block agonist effects without affecting basal activity.

Agonist specificity

▼ It was thought that the linkage of a particular GPCR to a particular signal transduction pathway depends mainly on the structure of the receptor, which confers specificity for a particular G protein, from which the rest of the signal transduction pathway follows. This would imply, in line with the two-state model discussed in Chapter 2, that all agonists acting on a particular receptor stabilise the same activated (R*) state and should activate the same signal transduction pathway, and produce the same type of cellular response. It is now clear that this is an oversimplification. In many cases, for example, with agonists acting on angiotensin receptors, or with inverse agonists on β adrenoceptors, the cellular effects are qualitatively different with different ligands, implying the existence of more than one – probably many – R* states (sometimes referred to as *biased agonism*; see Ch. 2). Binding of arrestins to GPCRs initiates MAP kinase signalling, such that agonists that induce GRK/arrestin ‘desensitisation’ will terminate some GPCR signalling but may also activate signalling through arrestins that may continue even after the receptor/arrestin complex has been internalised (see Fig. 3.15).

Biased agonism has profound implications – indeed heretical to many pharmacologists, who are accustomed to thinking of agonists in terms of their affinity and efficacy, and nothing else; it has added a new dimension to the way in which we think about drug efficacy and specificity (see Kenakin and Christopoulos, 2013).

Receptor activity-modifying proteins

▼ Receptor activity-modifying proteins (RAMPs) are a family of membrane proteins that associate with some GPCRs and alter their functional characteristics. They were discovered in 1998 when it was found that the functionally active receptor for the neuropeptide calcitonin gene-related peptide (CGRP) (see Chs 16 and 19) consisted of a complex of a GPCR – called calcitonin receptor-like receptor (CRLR) – that by itself lacked activity, with another membrane protein (RAMP1). More surprisingly, CRLR when coupled with another RAMP (RAMP2) showed a quite different pharmacology, being activated by an unrelated peptide, *adrenomedulin*. In other words, the agonist specificity is conferred by the associated RAMP as well as by the GPCR itself. More RAMPs have emerged, and so far nearly all the examples involve Class B peptide receptors (see Table 3.2), the calcium-sensing receptor being an exception. RAMPs are an example of how protein–protein interactions influence the pharmacological behaviour of the receptors in a highly selective way and may be novel targets for drug development (Sexton et al., 2012).

G protein-independent signalling

▼ In using the term *G protein-coupled receptor* to describe the class of receptors characterised by their heptahelical structure, we are following conventional textbook dogma but neglecting the fact that G proteins are not the only link between GPCRs and the various effector systems that they regulate. In this context, signalling mediated through arrestins bound to the receptor (see p. 36), rather than through G proteins, is important (see reviews by Pierce & Lefkowitz, 2001; Delcourt et al., 2007). Arrestins can act as an intermediary for GPCR activation of the MAP kinase cascade (see Fig. 3.15B).

There are many examples where the various ‘adapter proteins’ that link receptors of the tyrosine kinase type to their effectors (see p. 42) can also interact with GPCRs (see Brzostowski & Kimmel, 2001), allowing the same effector systems to be regulated by receptors of either type.

In summary, the simple dogma that has underpinned much of our understanding of GPCRs, namely, one GPCR gene – one GPCR protein – one functional GPCR – one G protein – one response, is showing distinct signs of wear. In particular:

- one gene, through alternative splicing, RNA editing, etc., can give rise to more than one receptor protein;
- one GPCR protein can associate with others, or with other proteins such as RAMPs, to produce more than one type of functional receptor;
- different agonists may affect a receptor in different ways and elicit qualitatively different responses;
- the signal transduction pathway from ‘GPCR’ does not invariably require G proteins, and there can be cross-talk with tyrosine kinase-linked receptors.

GPCRs are evidently versatile and adventurous molecules around which much modern pharmacology revolves, and nobody imagines that we have reached the end of the story.

TYPE 3: KINASE-LINKED AND RELATED RECEPTORS

These membrane receptors are quite different in structure and function from ligand-gated channels and GPCRs. They are activated by a wide variety of protein mediators, including growth factors and cytokines (see Ch. 19), and hormones such as insulin (see Ch. 32) and leptin (Ch. 33), whose effects are exerted mainly at the level of gene transcription. Most of these receptors are large proteins consisting of a single chain of up to 1000 residues, with a single membrane-spanning helical region, linking a large extracellular ligand-binding domain to an intracellular domain of variable size and function. The basic structure is shown in Fig. 3.1C, but many variants exist (see later). Over 100 such receptors have been cloned, and many structural variations exist. For more detail, see the review by Hubbard & Miller (2007). These receptors play a major role in controlling cell division, intermediary metabolism, growth, differentiation, inflammation, tissue repair, apoptosis and immune responses, discussed further in Chapters 6 and 19.

The main types are as follows:

Receptor tyrosine kinases (RTKs). These receptors have the basic structure shown in Fig. 3.17A, incorporating a tyrosine kinase moiety in the intracellular region. They include receptors for many growth factors, such as **epidermal growth factor** and **nerve growth factor**, and also the group of TLRs that recognise bacterial lipopolysaccharides and play an important role in the body’s reaction to infection (see Ch. 7). The insulin receptor (see Ch. 32) also belongs to the RTK class, although it has a more complex dimeric structure, and links indirectly to intracellular tyrosine kinases.

Receptor serine/threonine kinases. This smaller class is similar in structure to RTKs but they phosphorylate serine and/or threonine residues rather than tyrosine. The main example is the receptor for **transforming growth factor** (TGF).

Cytokine receptors. These receptors (Fig. 3.17B) lack intrinsic enzyme activity. When occupied, they activate various tyrosine kinases, such as Jak (the Janus kinase). Ligands for these receptors include cytokines such as **interferons** and **colony-stimulating factors** involved in immunological responses as well as cell growth and differentiation.

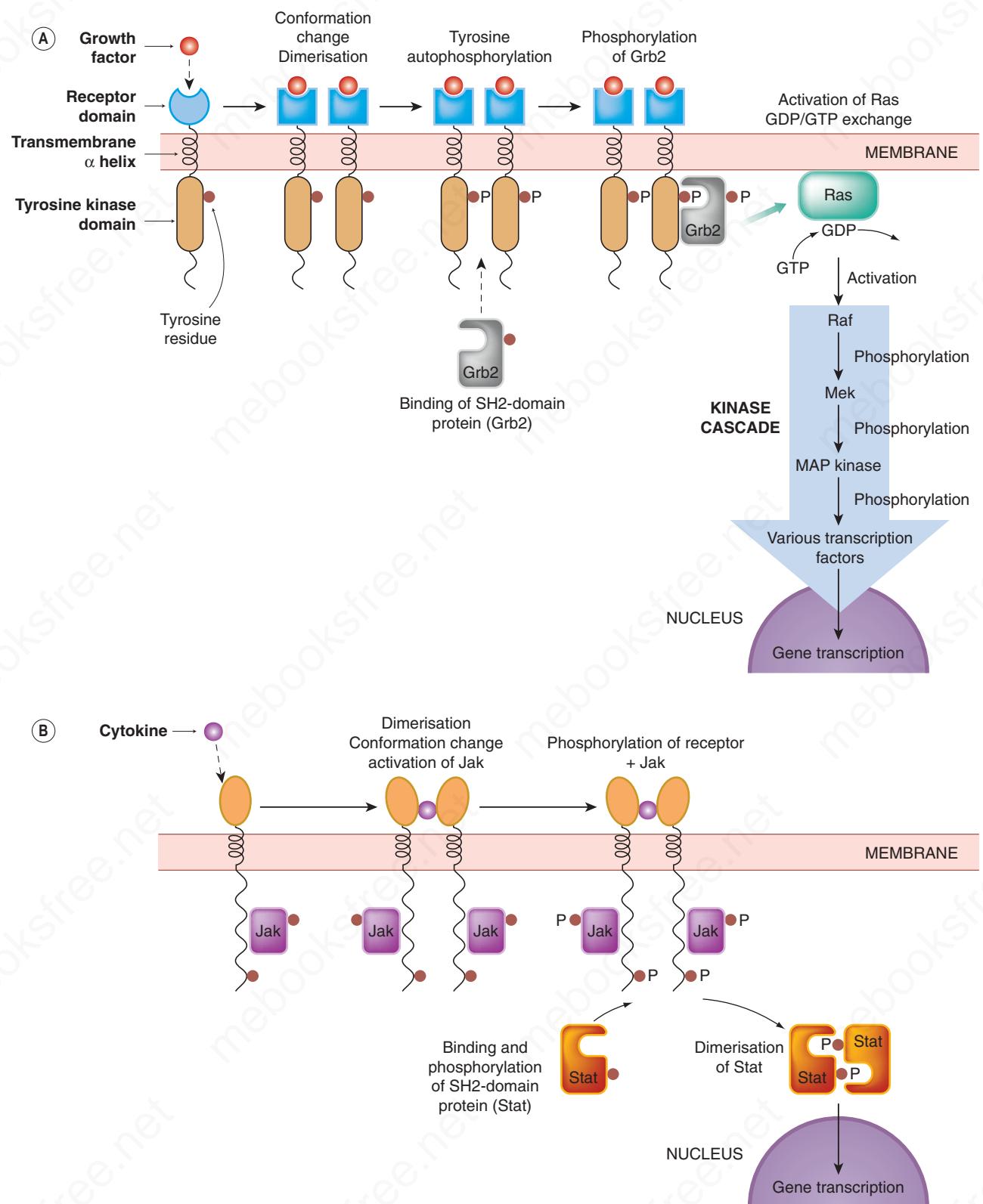


Fig. 3.17 Transduction mechanisms of kinase-linked receptors. The first step following agonist binding is dimerisation, which leads to autophosphorylation of the intracellular domain of each receptor. SH2-domain proteins then bind to the phosphorylated receptor and are themselves phosphorylated. Two well-characterised pathways are shown: (A) the growth factor (Ras/Raf/mitogen-activated protein [MAP] kinase) pathway (see also Ch. 6). Grb2 can also be phosphorylated but this negatively regulates its signalling. (B) Simplified scheme of the cytokine (Jak/Stat) pathway (see also Ch. 19). Some cytokine receptors may pre-exist as dimers rather than dimerise on cytokine binding. Several other pathways exist, and these phosphorylation cascades interact with components of G protein systems.

Kinase-linked receptors



- Receptors for various growth factors incorporate tyrosine kinase in their intracellular domain.
- Cytokine receptors have an intracellular domain that binds and activates cytosolic kinases when the receptor is occupied.
- The receptors all share a common architecture, with a large extracellular ligand-binding domain connected via a single membrane-spanning helix to the intracellular domain.
- Signal transduction generally involves dimerisation of receptors, followed by autophosphorylation of tyrosine residues. The phosphotyrosine residues act as acceptors for the SH2 domains of a variety of intracellular proteins, thereby allowing control of many cell functions.
- They are involved mainly in events controlling cell growth and differentiation, and act indirectly by regulating gene transcription.
- Two important pathways are:
 - the Ras/Raf/mitogen-activated protein (MAP) kinase pathway, which is important in cell division, growth and differentiation
 - the Jak/Stat pathway activated by many cytokines, which controls the synthesis and release of many inflammatory mediators.

PROTEIN PHOSPHORYLATION AND KINASE CASCADE MECHANISMS

Protein phosphorylation (see Cohen, 2002) is a key mechanism for controlling the function of proteins (e.g. enzymes, ion channels, receptors, transport proteins) involved in regulating cellular processes. Phosphorylation and dephosphorylation are accomplished by *kinases* and *phosphatases*, respectively – enzymes, of which several hundred subtypes are represented in the human genome – which are themselves subject to regulation dependent on their phosphorylation status. Much effort is currently being invested in mapping the complex interactions between signalling molecules that are involved in drug effects and pathophysiological processes such as oncogenesis, neurodegeneration, inflammation and much else. Here we can present only a few pharmacologically relevant aspects of what has become an enormous subject.

In many cases, ligand binding to the receptor leads to dimerisation. The association of the two intracellular kinase domains allows a mutual autophosphorylation of intracellular tyrosine residues to occur. The phosphorylated tyrosine residues then serve as high-affinity docking sites for other intracellular proteins that form the next stage in the signal transduction cascade. One important group of such proteins is known as the *SH2 domain proteins* (standing for *Src* homology, because they were first identified in the *Src* oncogene product).¹⁶ These possess a highly conserved

¹⁶*v-Src* is a gene found in Rous sarcoma virus that encodes a tyrosine kinase which causes sarcoma (a malignant tumour) in chickens – it was found to have a closely related sequence to the chicken's own gene termed *c-Src* (for cellular rather than viral *Src*). This was the first oncogene to be discovered, in 1979.

sequence of about 100 amino acids, forming a recognition site for the phosphotyrosine residues of the receptor. Individual SH2 domain proteins, of which many are now known, bind selectively to particular receptors, so the pattern of events triggered by particular growth factors is highly specific. The mechanism is summarised in Fig. 3.17.

What happens when the SH2 domain protein binds to the phosphorylated receptor varies greatly according to the receptor that is involved; many SH2 domain proteins are enzymes, such as protein kinases or phospholipases. Some growth factors activate a specific subtype of phospholipase C (PLC γ), thereby causing phospholipid breakdown, IP₃ formation and Ca²⁺ release (see p. 35). Other SH2-containing proteins couple phosphotyrosine-containing proteins with a variety of other functional proteins, including many that are involved in the control of cell division and differentiation. The end result is to activate or inhibit, by phosphorylation, a variety of transcription factors that migrate to the nucleus and suppress or induce the expression of particular genes. For more detail, see Jin and Pawson (2012). Nuclear factor kappa B (NF κ B) is a transcription factor that plays a key role in multiple disorders including inflammation and cancer (see Chs 18 and 57; Karin et al., 2004). It is normally present in the cytosol, complexed with an inhibitor (I κ B). Phosphorylation of I κ B occurs when a specific kinase (IKK) is activated in response to various inflammatory cytokines and GPCR agonists. This results in dissociation of I κ B from NF κ B and migration of NF κ B to the nucleus, where it switches on various proinflammatory and anti-apoptotic genes.

▼ Two well-defined signal transduction pathways are summarised in Fig. 3.17. The Ras/Raf pathway mediates the effect of many growth factors and mitogens. Ras, which is a proto-oncogene product, functions like a G protein, and conveys the signal (by GDP/GTP exchange) from the SH2-domain protein, Grb. Activation of Ras in turn activates Raf, which is the first of a sequence of three serine/threonine kinases, each of which phosphorylates, and activates, the next in line. The last of these, MAP kinase (which is also activated by GPCRs, see earlier), phosphorylates one or more transcription factors that initiate gene expression, resulting in a variety of cellular responses, including cell division. This three-tiered MAP kinase cascade forms part of many intracellular signalling pathways involved in a wide variety of disease processes, including malignancy, inflammation, neurodegeneration, atherosclerosis and much else. The kinases form a large family, with different subtypes serving specific roles. They are thought to represent an important target for future therapeutic drugs. Many cancers are associated with mutations in the genes coding for proteins involved in this cascade, leading to activation of the cascade in the absence of the growth factor signal (see Chs 6 and 57). For more details, see the review by Avruch (2007).

A second pathway, the Jak/Stat pathway (see Fig. 3.17B), is involved in responses to many cytokines. Dimerisation of these receptors occurs when the cytokine binds, and this attracts a cytosolic tyrosine kinase unit (Jak) to associate with, and phosphorylate, the receptor dimer. Jaks belong to a family of proteins, different members having specificity for different cytokine receptors. Among the targets for phosphorylation by Jak are a family of transcription factors (Stats). These are SH2-domain proteins that bind to the phosphotyrosine groups on the receptor-Jak complex, and are themselves phosphorylated. Thus activated, Stat migrates to the nucleus and activates gene expression.

Other important mechanisms centre on *phosphatidylinositol-3 kinase* (PI₃ kinases, see Vanhaesebroeck et al., 1997), a ubiquitous enzyme family that is activated both by GPCRs and RTKs and attaches a phosphate group to position 3 of PIP₂ to form PIP₃. Other protein

¹⁷Protein kinase B was named to fill in the gap between protein kinase A (cAMP-dependent) and protein kinase C (Ca²⁺-dependent). As you can see, nomenclature is highly imaginative!

kinases, particularly protein kinase B (PKB,¹⁷ also known as Akt), have recognition sites for PIP₃ and are thus activated, controlling a wide variety of cellular functions, including apoptosis, differentiation, proliferation and trafficking. Akt also causes NO synthase activation in the vascular endothelium (see Ch. 21).

Recent work on signal transduction pathways has produced a bewildering profusion of molecular detail, often couched in a jargon that is apt to deter the faint-hearted. Perseverance will be rewarded, however, for there is no doubt that important new drugs, particularly in the areas of inflammation, immunology and cancer, will come from the targeting of these proteins. A breakthrough in the treatment of chronic myeloid leukaemia was achieved with the introduction of the first explicitly-designed kinase inhibitor, **imatinib**, a drug that inhibits a specific tyrosine kinase involved in the pathogenesis of the disease (see Ch. 57).

Fig. 3.18 illustrates the central role of protein kinases in signal transduction pathways in a highly simplified and schematic way. Many, if not all, of the proteins involved, including the receptors and the kinases themselves, are substrates for kinases, so there are many mechanisms for feedback and cross-talk between the various signalling pathways. Given that there are over 500 protein kinases, and similarly large numbers of receptors and other signalling molecules, the network of interactions can look bewilderingly complex. Dissecting out the details has become a major theme in cell biology. For pharmacologists, the idea of a simple connection between receptor and response, which guided thinking throughout the 20th century, is undoubtedly crumbling, although it will take some time before the complexities of signalling pathways are assimilated into a new way of thinking about drug action.

Protein phosphorylation in signal transduction

- Many receptor-mediated events involve protein phosphorylation, which controls the functional and binding properties of intracellular proteins.
- Receptor-linked tyrosine kinases, cyclic nucleotide-activated tyrosine kinases and intracellular serine/threonine kinases comprise a 'kinase cascade' mechanism that leads to amplification of receptor-mediated events.
- There are many kinases, with differing substrate specificities, allowing specificity in the pathways activated by different hormones.
- Desensitisation of G protein-coupled receptors occurs as a result of phosphorylation by specific receptor kinases, causing the receptor to become non-functional and to be internalised.
- There is a large family of phosphatases that act to dephosphorylate proteins and thus reverse the effects of kinases.

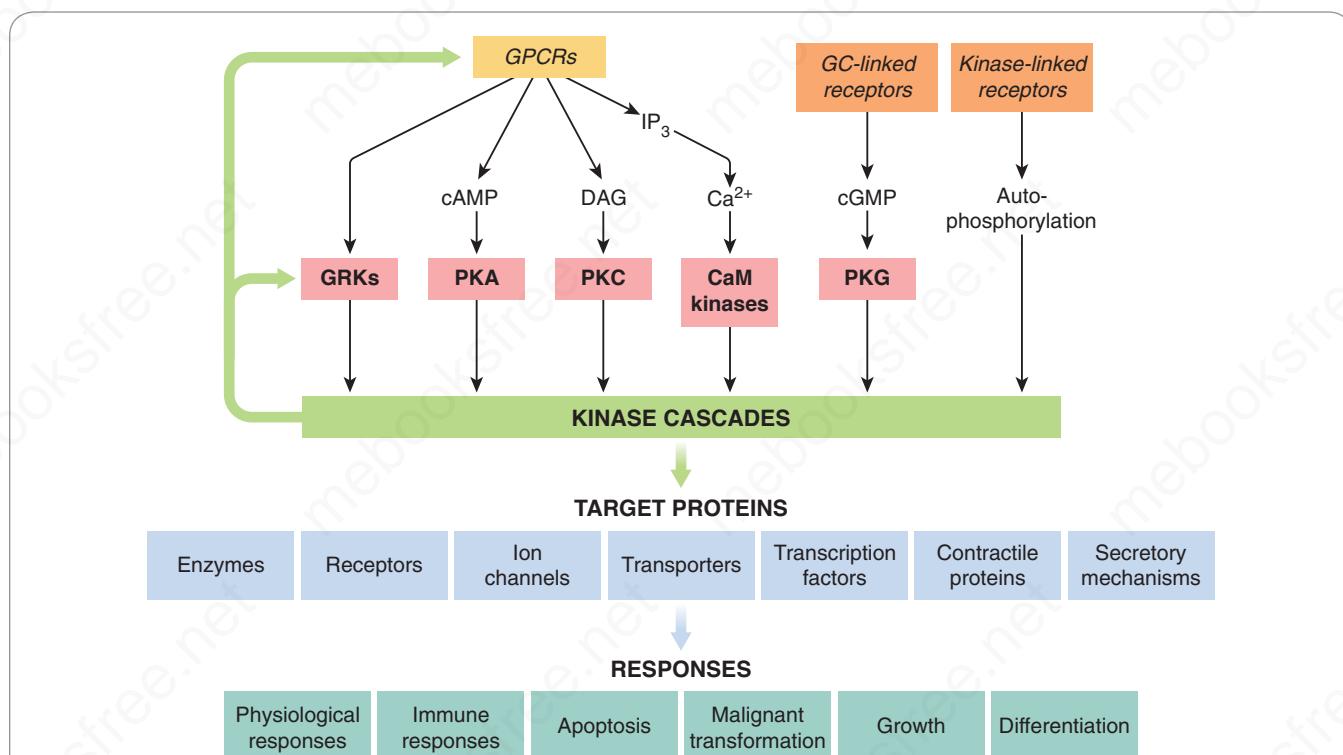


Fig. 3.18 Central role of kinase cascades in signal transduction. Kinase cascades (e.g. those shown in **Fig. 3.15**) are activated by G protein-coupled receptors (GPCRs), either directly or via different second messengers, by receptors that generate cGMP, or by kinase-linked receptors. The kinase cascades regulate various target proteins, which in turn produce a wide variety of short- and long-term effects. *CaM kinase*, Ca^{2+} /calmodulin-dependent kinase; *DAG*, diacylglycerol; *GC*, guanylyl cyclase; *GRK*, GPCR kinase; *IP₃*, inositol trisphosphate; *PKA*, cAMP-dependent protein kinase; *PKC*, protein kinase C; *PKG*, cGMP-dependent protein kinase.

TYPE 4: NUCLEAR RECEPTORS

By the 1970s, it was clear that receptors for steroid hormones such as oestrogen and the glucocorticoids (Chs. 36 and 34) were present in the cytoplasm of cells and translocated into the nucleus after binding with their steroid partner. Other hormones, such as the thyroid hormone T₃ (Ch. 35) and the fat-soluble vitamins D and A (retinoic acid), were found to act in a similar fashion. Comparisons of gene and protein sequence data led to the recognition that these receptors were members of a much larger family of related proteins. We now know these as the *nuclear receptor (NR) family*.

As well as NRs such as the glucocorticoid and retinoic acid receptor, whose ligands are well characterised, this family includes a great many (~40%) orphan receptors – receptors with no known well-defined ligands (see earlier). The first of these to be described, in the 1990s, was the *retinoid X receptor (RXR)*, a receptor cloned on the basis of its similarity with the vitamin A receptor, and which was subsequently found to bind the vitamin A derivative 9-cis-retinoic acid. This event triggered intense interest in the NR field and, during the intervening years, specific binding partners have been characterised for many NRs ('adopted orphans', e.g. RXR) although in the case of many others ('true orphans') these have yet to be identified – or perhaps do not exist as such, as one possible function of these receptors is their 'promiscuous' ability to bind to many related compounds (such as dietary factors) with low affinity.

Unlike the other receptors described in this chapter, NRs can interact with DNA directly, and can be regarded as *ligand-activated transcription factors* that produce their effects by modifying gene transcription. Through this mechanism they can control the transcription and expression of many genes and proteins so, as it might be imagined, they are key players in regulating metabolic, developmental and other critical physiological processes. Another unique property is that NRs are not generally embedded in membranes like GPCRs or ion channels, but are present in other compartments of the cell. Some, such as the steroid receptors, which are predominately located in the cytoplasm, are activated by their ligand and translocate from the cytoplasm to the nucleus, while others, such as the RXR, probably dwell mainly within the nuclear compartment. Having said this, there is increasing evidence for the existence of small pools of some NRs, such as oestrogen and glucocorticoid receptors (ER and GR) at the plasma membrane and in organelles such as the mitochondria (Levin and Hammes, 2016), where they can act directly on other targets such as protein kinases to bring about immediate biological actions.

The NR superfamily probably evolved from a single distant evolutionary ancestral gene by duplication and other events. In man, there are at least 48 members, but more proteins may arise through alternative splicing events. While this represents a rather small proportion of all receptors (less than 10% of the total number of GPCRs), the NRs are very important drug targets (Burris et al., 2013), being responsible for the biological effects of approximately 10%–15% of all prescription drugs. They can recognise an extraordinarily diverse group of substances (mostly small hydrophobic molecules), which may exhibit full or partial agonist, antagonist or inverse agonist activity. Some NRs which bind their ligands with high affinity (e.g. ER and GR) are involved predominantly in endocrine

signalling, but many bind their ligands with low affinity and probably act as metabolic (e.g. lipid) sensors. They are thus crucial links between our dietary and metabolic status and the expression of genes that regulate the metabolism and disposition of lipids. NRs also regulate expression of many drug-metabolising enzymes and transporters.

STRUCTURE OF NUCLEAR RECEPTORS

▼ All NRs are monomeric proteins of 50–100 kDa, which share a broadly similar structural design (see Fig. 3.19 and Bourguet et al., 2000, for further details). The *N-terminal domain* displays the most heterogeneity. It harbours the *activation function 1 (AF1)* site that binds to other cell-specific transcription factors in a ligand-independent way and modifies the binding or regulatory capacity of the receptor itself. In the presence of the ligand, it synergises with AF2 to produce the fully active complex. Alternative splicing of genes may yield several receptor isoforms, each with slightly different N-terminal regions. The *core domain* of the receptor is highly conserved and consists of the structure responsible for DNA recognition and binding. At the molecular level, this comprises two *zinc fingers* – cysteine- (or cystine-/histidine-) rich loops in the amino acid chain that are held in a particular conformation by zinc ions. The main function of this portion of the molecule is to recognise and bind to the *hormone response elements (HREs)* located in the genes that are regulated by this family of receptors, but it also plays a part in regulating receptor dimerisation which is crucial to the function of most NRs.

It is the highly flexible *hinge region* in the molecule that allows it to dimerise with other NRs and regulates the intracellular trafficking of the receptor. This can produce molecular complexes with diverse configurations, able to interact differently with DNA. Finally, the *C-terminal domain* contains the *ligand-binding module* and is specific to each class of receptor, although structurally highly conserved. It is also important in dimerisation and binding co-activator and co-repressor proteins (see later). The AF2 region is important in ligand-dependent activation and is generally highly conserved, although it is absent in *Rev-erbAα* and *Rev-erbAβ*, NRs that regulate metabolism (and also function as part of a circadian molecular clock mechanism). Also located near the C-terminal are motifs that contain *nuclear localisation signals* and others that may, in the case of some receptors, bind *accessory heat shock* and other proteins.

CONTROL OF GENE TRANSCRIPTION

▼ HREs are short (usually 4–6 base pairs) sequences of DNA to which the NRs bind to modify gene transcription. They are generally present symmetrically in pairs or *half-sites*, although these may be arranged together in different ways (e.g. *simple repeats* or *inverted repeats*). Each NR exhibits a preference for a particular *consensus sequence* and the nucleotide spacing between them, but because of the family homology, they share a close similarity. In the nucleus, the AF1 and AF2 domains of the ligand-bound receptor recruit large complexes of other proteins including co-activators or co-repressors to modify gene expression. Some of these co-activators are enzymes involved in chromatin remodelling, such as histone acetylase/deacetylase which, together with other enzymes, regulate the unravelling of the DNA to facilitate access by polymerase enzymes and hence gene transcription. Co-repressor complexes are recruited by some receptors and comprise histone deacetylase and other factors that cause the chromatin to become tightly packed, preventing further transcriptional activation. The case of the *constitutive androstane receptor (CAR)* (see later) is particularly interesting: like some G proteins described earlier in this chapter, CAR can form a constitutively active complex that is terminated when it binds its ligand. The mechanisms of negative gene regulation by NRs are particularly complex (see Santos et al., 2011 for a good account of this phenomenon). In addition to agonists, NRs can also be targeted by competitive antagonists, which prevent occupation of the binding site by the endogenous ligand or by inverse agonists (or antagonists), which sterically prevent the binding of co-activator factors, thus reducing the constitutive activity of these

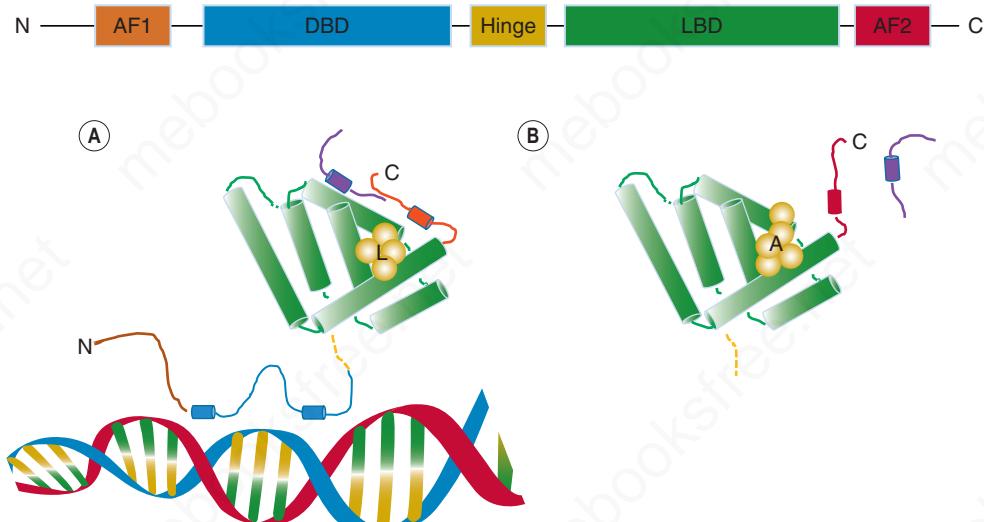


Fig. 3.19 Schematic diagram of a nuclear receptor. A greatly simplified diagram of the functional topology of a nuclear receptor (the oestrogen receptor is picked as an example). A schematic diagram shows the various regions of the receptor including the DNA-binding domain (DBD). Below is a diagram illustrating, in the corresponding colours, the configuration of the liganded receptor showing its binding to hormone response elements (HREs) on DNA. In panel A, the ligand (L) is bound in the ligand-binding domain (LBD) and this enables the C-terminal AF2 region to bind to the LBD. In turn, this allows the binding of a co-activator protein at the LBD (only a partial structure shown), which allows gene transcription to proceed. In panel B, an antagonist (A) is bound to the LBD. This sterically inhibits the binding of AF2 and thus the attachment of the co-activator protein. Most nuclear receptors operate as dimers but only a monomer is shown here for clarity. (Based largely upon Shiu et al., 1998.) Cylindrical structures represent regions of α -helical protein structure.

receptors. A very interesting development is the identification of selective receptor modulators (e.g. selective oestrogen receptor modulators - SERMs) which, by altering the binding of co-activator and co-repressor proteins, have agonist activity in some tissues and antagonist activities in others.

CLASSIFICATION OF NUCLEAR RECEPTORS

NRs are usually classified into subfamilies according to their phylogeny. For our purposes, however, it is more useful to classify them on the basis of their molecular action into two main classes (I and II), and two other minor groups of receptors (III, IV).

Class I consists largely of endocrine steroid receptors, including the GRs and mineralocorticoid receptors (MRs), as well as the oestrogen, progesterone and androgen receptors (ER, PR and AR, respectively). The hormones (e.g. glucocorticoids) recognised by these receptors generally act in a negative feedback fashion to control biological events (see Ch. 34 for more details). In the absence of their ligand, these NRs are predominantly located in the cytoplasm, complexed with heat shock and other proteins, and possibly reversibly attached to the cytoskeleton or other intracellular structures. Following diffusion (or possibly transportation) into the cell from the blood, ligands bind their NR partner with high affinity. These liganded receptors generally form homodimers and translocate to the nucleus, where they can *transactivate* or *transrepress* genes by binding to 'positive' or 'negative' HREs. Once bound, the NR recruits other proteins to form complexes that promote transcription of multiple genes. For example, it is estimated that the activated GR itself can regulate

transcription of ~1% of the genome either directly or indirectly.

Class II NRs function in a slightly different way. Their ligands are generally lipids or other metabolites already present to some extent within the cell. This group includes the *peroxisome proliferator-activated receptor* (PPAR) that recognises fatty acids; the *liver oxysterol receptor* (LXR) that recognises and acts as a cholesterol sensor, the *farnesoid (bile acid) receptor* (FXR), a *xenobiotic receptor* (SXR; in rodents the PXR) that recognises a great many foreign substances, including therapeutic drugs, and the CAR, which not only recognises the steroid androstane but also some drugs such as **phenobarbital** (see Ch. 46). Indeed, PXR and CAR are akin to airport security guards who alert the bomb disposal squad when suspicious luggage is found. When they sense foreign molecules (xenobiotics), they induce drug-metabolising enzymes such as CYP3A (which is responsible for metabolising about 60% of all prescription drugs; see Ch. 10 and [di Masi et al., 2009](#)). They also bind some prostaglandins and non-steroidal drugs, as well as the antidiabetic **thiazolidinediones** (see Ch. 32) and **fibrates** (see Ch. 24).

Unlike the receptors in class I, these NRs almost always operate as heterodimers together with RXR, the retinoid X receptor. Two types of heterodimer may then be formed: a *non-permissive heterodimer*, which can be activated only by the RXR ligand itself, and the *permissive heterodimer*, which can be activated either by retinoic acid itself or by its partner's ligand. Class II NRs are generally bound to co-repressor proteins. These dissociate when the ligand binds and allows recruitment of co-activator proteins

Table 3.4 Some common pharmacologically significant nuclear receptors

Receptor name	Abbreviation	Ligand	Drugs	Location	Ligand binding	Mechanism of action
Type I						
Androgen	AR	Testosterone	All natural and synthetic glucocorticoids (Ch. 34), mineralocorticoids (Ch. 30) and sex steroids (Ch. 36) together with their antagonists (e.g. raloxifene, 4-hydroxy-tamoxifen and mifepristone).			Translocation to nucleus. Binding to HREs with two half-sites with an inverted sequence. Recruitment of co-activators, transcription factors and other proteins.
Oestrogen	ER α , β	17 β -oestradiol				
Glucocorticoid	GR α	Cortisol, corticosterone		Cytosolic	Homodimers	
Progesterone	PR	Progesterone				
Mineralocorticoid	MR	Aldosterone				
Type II						
Retinoid X	RXR α , β , γ	9-cis-retinoic acid	Retinoid drugs (Ch. 28)			Binding to HREs with two half-sites with an inverted or simple repeat sequence.
Retinoic acid	RAR α , β , γ	Vitamin A				
Thyroid hormone	TR α , β	T3, T4	Thyroid hormone drugs (Ch. 35)	Nuclear	Heterodimers often with RXR	Complexed with co-repressors, which are displaced following ligand binding, allowing the binding of co-activators
Peroxisome proliferator	PPAR α , β , γ , δ	Fatty acids, prostaglandins	Rosiglitazone, pioglitazone (Ch. 32)			
Constitutive androstane	CAR	Androstane	Stimulation of CYP synthesis and alteration of drug metabolism (Ch. 10)			
Pregnane X	PXR	Xenobiotics				

Only examples from Classes I and II are included.

and hence changes in gene transcription. They tend to mediate positive feedback effects (e.g. occupation of the receptor amplifies rather than inhibits a particular biological event).

Class III NRs are very similar to Class I in the sense that they form homodimers, but they can bind to HREs, which do not have an inverted repeat sequence. Class IV NRs may function as monomers or dimers but only bind to one HRE half site. Many of the remaining orphan receptors belong to these latter classes.

The discussion here must be taken only as a broad guide to the action of NRs, as many other types of interaction have also been discovered. For example, some of these receptors may bring about non-genomic – or even genomic – actions by directly interacting with factors in the cytosol, or they may be covalently modified by phosphorylation or by protein–protein interactions with other transcription factors such that their function is altered (see Falkenstein et al., 2000).

Table 3.4 summarises the properties of some common NRs of importance to pharmacologists.

ION CHANNELS AS DRUG TARGETS

We have discussed ligand-gated ion channels as one of the four main types of drug receptor. There are many other types of ion channel that represent important drug targets, even though they are not generally classified as ‘receptors’

Nuclear receptors

- A family of 48 soluble receptors that sense lipid and hormonal signals and modulate gene transcription.
- Their ligands are many and varied, including steroid drugs and hormone, thyroid hormones, vitamins A and D, various lipids and xenobiotics
- There are two main categories:
 - Class I nuclear receptors (NRs) are present in the cytoplasm, form homodimers in the presence of their ligand, and migrate to the nucleus. Their ligands are mainly endocrine in nature (e.g. steroid hormones);
 - Class II NRs are generally constitutively present in the nucleus and form heterodimers with the retinoid X receptor. Their ligands are usually lipids (e.g. the fatty acids).
- The liganded receptor complexes initiate changes in gene transcription by binding to hormone response elements in gene promoters and recruiting co-activator or co-repressor factors.
- The receptor family is the target of approximately 10% of prescription drugs, and the enzymes that it regulates affect the pharmacokinetics of some 60% of all prescription drugs.

because they are not the immediate targets of fast neurotransmitters, but drugs can act upon them to alter their ability to open and close.¹⁸

Here we discuss the structure and function of ion channels at the molecular level; their role as regulators of cell function is described in Chapter 4.

Ions are unable to penetrate the lipid bilayer of the cell membrane, and can get across only with the help of membrane-spanning proteins in the form of channels or transporters. The concept of ion channels was developed in the 1950s on the basis of electrophysiological studies on the mechanism of membrane excitation (see Ch. 4). Electrophysiology, particularly the *voltage clamp technique*, remains an essential tool for studying the physiological and pharmacological properties of ion channels. Since the mid-1980s, when the first ion channels were cloned by Numa in Japan, much has been learned about the structure and function of these complex molecules. The use of patch clamp recording, which allows the behaviour of individual channels to be studied in real time, has been particularly valuable in distinguishing channels on the basis of their conductance and gating characteristics. Accounts by Hille (2001), Ashcroft (2000) and Catterall (2000) give background information.

Ion channels consist of protein molecules designed to form water-filled pores that span the membrane, and can switch between open and closed states. The rate and direction of ion movement through the pore is governed by the electrochemical gradient for the ion in question, which is a function of its concentration on either side of the membrane, and of the membrane potential. Ion channels are characterised by:

- their selectivity for particular ion species, determined by the size of the pore and the nature of its lining;
- their gating properties (i.e. the nature of the stimulus that controls the transition between open and closed states of the channel);
- their molecular architecture.

ION SELECTIVITY

Channels are generally either cation selective or anion selective. The main cation-selective channels are selective for Na^+ , Ca^{2+} or K^+ , or non-selective and permeable to all three. Anion channels are mainly permeable to Cl^- , although other types also occur. The effect of modulation of ion channels on cell function is discussed in Chapter 4.

GATING

VOLTAGE-GATED CHANNELS

In the main these channels open when the cell membrane is depolarised.¹⁹ They form a very important group because

¹⁸In truth, the distinction between ligand-gated channels and other ion channels is an arbitrary one. In grouping ligand-gated channels with other types of receptor in this book, we are respecting the historical tradition established by Langley and others, who first defined receptors in the context of the action of acetylcholine at the neuromuscular junction. The advance of molecular biology may force us to reconsider this semantic issue in the future, but for now we make no apology for upholding the pharmacological tradition.

¹⁹There is always an exception to the rule! The members of the HCN family of potassium channels found in neurons and cardiac muscle cells are activated by hyperpolarisation.

they underlie the mechanism of membrane excitability (see Ch. 4). The most important channels in this group are selective sodium, potassium or calcium channels.

Commonly, the channel opening (activation) induced by membrane depolarisation is short lasting, even if the depolarisation is maintained. This is because, with some channels, the initial activation of the channels is followed by a slower process of inactivation.

The role of voltage-gated channels in the generation of action potentials and in controlling other cell functions is described in Chapter 4.

LIGAND-GATED CHANNELS

These (see Fig. 3.5) are activated by binding of a chemical ligand to a site on the channel molecule. Fast neurotransmitters, such as glutamate, acetylcholine, GABA, 5-HT and ATP (see Chs 14, 16, 17 and 39) act in this way, binding to sites on the outside of the membrane. In addition, there are also ligand-gated ion channels that do not respond to neurotransmitters but to changes in their local environment. For example, the TRPV1 channel on sensory nerves that mediates the pain-producing effect of the chilli pepper ingredient capsaicin responds to extracellular protons when tissue pH falls, as occurs in inflamed tissue, as well as to the physical stimulus, heat (see Ch. 43).

Some ligand-gated channels in the plasma membrane respond to intracellular rather than extracellular signals, the most important being the following:

- Calcium-activated potassium channels, which occur in most cells and open, thus hyperpolarising the cell, when $[\text{Ca}^{2+}]_i$ increases.
- Calcium-activated chloride channels, widely expressed in excitable and non-excitable cells where they are involved in diverse functions such as epithelial secretion of electrolytes and water, sensory transduction, regulation of neuronal and cardiac excitability and regulation of vascular tone.
- ATP-sensitive potassium channels, which open when the intracellular ATP concentration falls because the cell is short of energy. These channels, which are quite distinct from those mediating the excitatory effects of extracellular ATP, occur in many nerve and muscle cells, and also in insulin-secreting cells (see Ch. 32), where they are part of the mechanism linking insulin secretion to blood glucose concentration.

Other examples of cell membrane channels that respond to intracellular ligands include arachidonic acid-sensitive potassium channels and DAG-sensitive calcium channels, whose functions are not well understood.

CALCIUM RELEASE CHANNELS

The main ones, IP_3 and ryanodine receptors (see Ch. 4), are a special class of ligand-gated calcium channels that are present on the endoplasmic or sarcoplasmic reticulum rather than the plasma membrane and control the release of Ca^{2+} from intracellular stores. Ca^{2+} can also be released from lysosomal stores by nicotinic acid adenine dinucleotide phosphate, which activates two-pore domain calcium channels.

STORE-OPERATED CALCIUM CHANNELS

When the intracellular Ca^{2+} stores are depleted, 'store-operated' channels (SOCs) in the plasma membrane open to allow Ca^{2+} entry. The mechanism by which this linkage

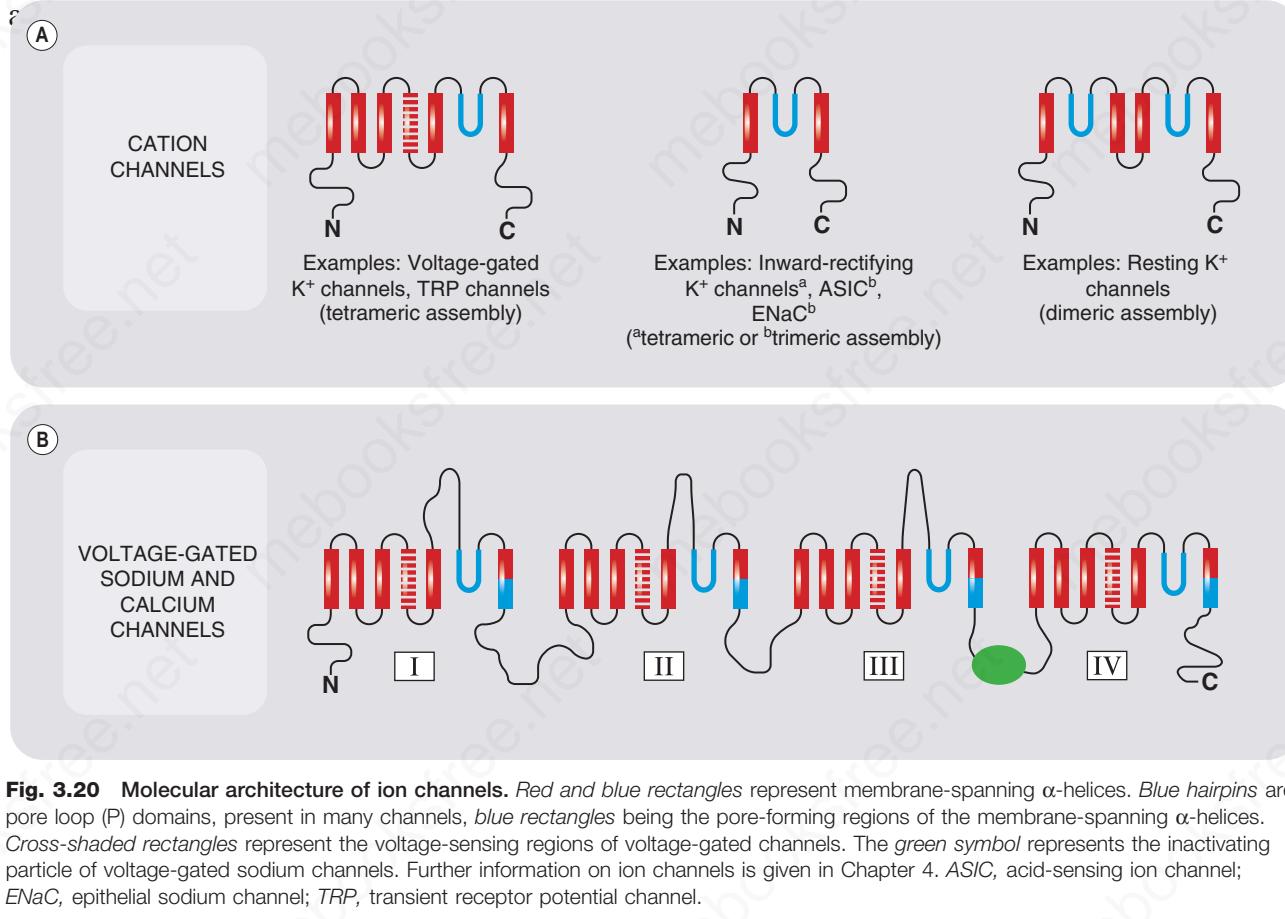


Fig. 3.20 Molecular architecture of ion channels. Red and blue rectangles represent membrane-spanning α -helices. Blue hairpins are pore loop (P) domains, present in many channels, blue rectangles being the pore-forming regions of the membrane-spanning α -helices. Cross-shaded rectangles represent the voltage-sensing regions of voltage-gated channels. The green symbol represents the inactivating particle of voltage-gated sodium channels. Further information on ion channels is given in Chapter 4. ASIC, acid-sensing ion channel; ENaC, epithelial sodium channel; TRP, transient receptor potential channel.

occurs involves interaction of a Ca²⁺-sensor protein in the endoplasmic reticulum membrane with a dedicated Ca²⁺ channel in the plasma membrane (see Stathopoulos & Ikura, 2017). In response to GPCRs that elicit Ca²⁺ release, the opening of these channels allows the cytosolic free Ca²⁺ concentration, [Ca²⁺]_i, to remain elevated even when the intracellular stores are running low, and also provides a route through which the stores can be replenished (see Ch. 4).

MOLECULAR ARCHITECTURE OF ION CHANNELS

▼ Ion channels are large and elaborate molecules. Their characteristic structural motifs have been revealed as knowledge of their sequence and structure has accumulated since the mid-1980s, when the first voltage-gated sodium channel was cloned. The main structural subtypes are shown in Fig. 3.20. All consist of several (often four) domains, which are similar or identical to each other, organised either as an oligomeric array of separate subunits, or as one large protein. Each subunit or domain contains a bundle of two to six membrane-spanning helices.

Voltage-gated channels generally include one transmembrane helix that contains an abundance of basic (i.e. positively charged) amino acids. When the membrane is depolarised, so that the interior of the cell becomes less negative, this region – the voltage sensor – moves slightly towards the outer surface of the membrane, which has the effect of opening the channel (see Bezanilla, 2008). Many voltage-activated channels also show *inactivation*, which happens when an intracellular appendage of the channel protein moves to plug the channel from the inside. Voltage-gated sodium and calcium channels

are remarkable in that the whole structure with four six-helix domains consists of a single huge protein molecule, the domains being linked together by intracellular loops of varying length (see Fig. 3.20B). Potassium channels comprise the most numerous and heterogeneous class.²⁰ Voltage-gated potassium channels resemble sodium channels, except that they are made up of four subunits rather than a single long chain. The class of potassium channels known as ‘inward rectifier channels’ because of their biophysical properties has the two-helix structure shown in Fig. 3.20A, whereas others are classed as ‘two-pore domain’ channels, because each subunit contains two P loops.

The various architectural motifs shown in Fig. 3.20 only scrape the surface of the molecular diversity of ion channels. In all cases, the individual subunits come in several molecular varieties, and these can unite in different combinations to form functional channels as *hetero-oligomers* (as distinct from *homo-oligomers* built from identical subunits). Furthermore, the channel-forming structures described are usually associated with other membrane proteins, which significantly affect their functional properties. For example, the ATP-gated potassium channel exists in association with the *sulfonylurea receptor* (SUR), and it is through this linkage that various drugs (including antidiabetic drugs of the sulfonylurea class; see Ch. 32) regulate the channel. Good progress is being made in understanding the relation between molecular structure and ion channel function, but we still have only a fragmentary understanding of the physiological role of many of these channels. Many important drugs exert their effects by influencing channel function, either directly or indirectly.

²⁰The human genome encodes more than 70 distinct potassium channel subtypes – either a nightmare or a golden opportunity for the pharmacologist, depending on one’s perspective.

PHARMACOLOGY OF ION CHANNELS

▼ Many drugs and physiological mediators described in this book exert their effects by altering the behaviour of ion channels.

The gating and permeation of both voltage-gated and ligand-gated ion channels is modulated by many factors, including the following.

- *Ligands that bind directly to various sites on the channel protein.* These include a variety of drugs and toxins that act in different ways, for example by blocking the channel or by affecting the gating process, thereby either facilitating or inhibiting the opening of the channel.
- *Mediators and drugs that act indirectly, mainly by activation of GPCRs.* The latter produce their effects mainly by affecting the state of phosphorylation of individual amino acids located on the intracellular region of the channel protein. As described above, this modulation involves the production of second messengers that activate protein kinases. The opening of the channel may be facilitated or inhibited, depending on which residues are phosphorylated. Drugs such as β -adrenoceptor agonists (Ch. 15) affect calcium and potassium channel function in this way, producing a wide variety of cellular effects.
- *Intracellular signals, particularly Ca^{2+} and nucleotides such as ATP and GTP* (see Ch. 4). Many ion channels possess binding sites for these intracellular mediators. Increased $[Ca^{2+}]_i$ opens certain types of potassium and chloride channels, and inactivates voltage-gated calcium channels. As described in Chapter 4, $[Ca^{2+}]_i$ is itself affected by the function of ion channels and GPCRs. Intracellular ATP binds to and closes a family of potassium channels known as the ATP-gated potassium channels (see Ch. 32) that are also sensitive to sulfonylurea drugs. Intracellular cyclic nucleotides, cAMP and cGMP, activate channels permeable to either calcium and sodium ions or to potassium ions.

Fig. 3.21 summarises the main sites and mechanisms by which drugs affect voltage-gated sodium channels, a typical example of this type of drug target.

CONTROL OF RECEPTOR EXPRESSION

Receptor proteins are synthesised by the cells that express them, and the level of expression is itself controlled, via the pathways discussed previously, by receptor-mediated events. We can no longer think of the receptors as the fixed elements in cellular control systems, responding to changes in the concentration of ligands, and initiating effects through the signal transduction pathway – they are themselves subject to regulation. Short-term regulation of receptor function generally occurs through *desensitisation*, as discussed earlier. Long-term regulation occurs through *an increase or decrease of receptor expression*. Examples of this type of control include the proliferation of various postsynaptic receptors after denervation (see Ch. 13), the up-regulation of various G protein-coupled and cytokine receptors in response to inflammation (see Ch. 18), and the induction of growth factor receptors by certain tumour viruses (see Ch. 6). Long-term drug treatment invariably induces adaptive responses, which, particularly with drugs that act on the central nervous system, can limit their effectiveness as in opioid tolerance (see Ch. 43) or can be the basis for therapeutic efficacy. In the latter instance this may take the form of a very slow onset of the therapeutic effect (e.g. with antidepressant drugs; see Ch. 48). It is likely that changes in receptor expression, secondary to the immediate action of the drug, are involved in delayed effects of this sort – a kind of ‘secondary pharmacology’, the importance of which is only now becoming clearer. The same principles apply

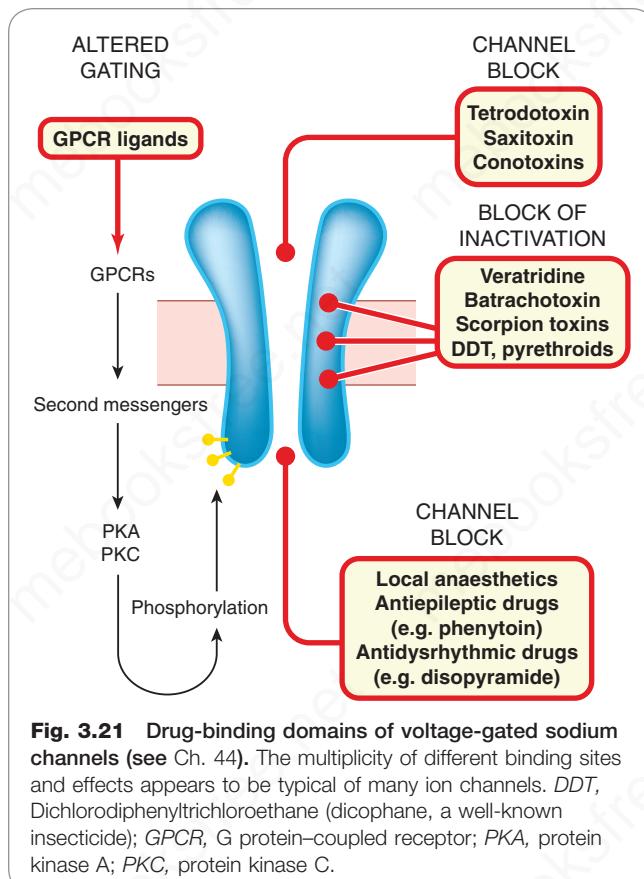


Fig. 3.21 Drug-binding domains of voltage-gated sodium channels (see Ch. 44). The multiplicity of different binding sites and effects appears to be typical of many ion channels. *DDT*, Dichlorodiphenyltrichloroethane (dicophane, a well-known insecticide); *GPCR*, G protein-coupled receptor; *PKA*, protein kinase A; *PKC*, protein kinase C.

to drug targets other than receptors (ion channels, enzymes, transporters, etc.) where adaptive changes in expression and function follow long-term drug administration, resulting, for example, in resistance to certain anticancer drugs (Ch. 57).

RECEPTORS AND DISEASE

Increasing understanding of receptor function in molecular terms has revealed a number of disease states directly linked to receptor malfunction. The principal mechanisms involved are:

- autoantibodies directed against receptor proteins;
- mutations in genes encoding receptors, ion channels and proteins involved in signal transduction.

An example of the former is *myasthenia gravis* (see Ch. 14), a disease of the neuromuscular junction due to autoantibodies that inactivate nicotinic acetylcholine receptors. Autoantibodies can also mimic the effects of agonists, as in many cases of thyroid hypersecretion, caused by activation of **thyrotropin** receptors (Ch. 35).

Inherited mutations of genes encoding GPCRs account for various disease states (see Stoy & Gurevich, 2015). Mutated **vasopressin** and **adrenocorticotropic hormone** receptors (see Chs 30 and 34) can result in resistance to these hormones. Receptor mutations can result in activation of effector mechanisms in the absence of agonists. One of these involves the receptor for thyrotropin, producing continuous oversecretion of thyroid hormone; another

involves the receptor for luteinising hormone and results in precocious puberty. Adrenoceptor polymorphisms are common in humans, and recent studies suggest that certain mutations of the β_2 adrenoceptor, although they do not directly cause disease, are associated with a reduced efficacy of β -adrenoceptor agonists in treating asthma (Ch. 29) and a poor prognosis in patients with cardiac failure, potentially through *constitutively active mutations* that render receptors active in the absence of any agonists (Ch. 22). Mutations in G proteins can also cause disease (see Spiegel & Weinstein, 2004). For example, mutations of a particular $G\alpha$ subunit cause one form of *hypoparathyroidism*, while mutations of a $G\beta$ subunit result in hypertension. Many cancers are associated with mutations of the genes encoding growth factor receptors, kinases and other proteins involved in signal transduction (see Ch. 6).

Mutations in ligand-gated ion channels ($GABA_A$ and nicotinic) and other ion channels (Na^+ and K^+) that alter

their function give rise to some forms of idiopathic epilepsy (see Ch. 46 and Poduri & Lowenstein, 2011).

Given the fact that the NR family of receptors plays a key part in the regulation and coordination of growth, development and organogenesis, reproduction, the immune system and many other fundamental biological processes, it is not surprising that many illnesses are associated with malfunctioning of the NR system. Such conditions include inflammation, cancer, diabetes, cardiovascular disease, obesity and reproductive disorders (see Kersten et al., 2000; Murphy & Holder, 2000).

Research on genetic polymorphisms affecting receptors, signalling molecules, ion channels and effector enzymes is continuing apace, and it is expected that a clearer understanding of the variability between individuals in their disease susceptibility and response to therapeutic drugs (see Ch. 58) will result, in the foreseeable future.

REFERENCES AND FURTHER READING

General

- IUPHAR/BPS. Guide to Pharmacology. www.guidetopharmacology.org/. (Comprehensive catalogue of molecular and pharmacological properties of known receptor and ion channels – also transporters and some enzymes involved in signal transduction)
- Nelson, N., 1998. The family of Na^+ / Cl^- neurotransmitter transporters. J. Neurochem. 71, 1785–1803. (Review article describing the molecular characteristics of the different families of neurotransmitters)

Ion channels

- Ashcroft, F.M., 2000. Ion Channels and Disease. Academic Press, London. (A useful textbook covering all aspects of ion channel physiology and its relevance to disease, with a lot of pharmacological information for good measure)
- Bezanilla, F., 2008. How membrane proteins sense voltage. Nat. Rev. Mol. Cell Biol. 9, 323–332. (Review of recent studies of how membrane proteins respond to changes in transmembrane potential)
- Catterall, W.A., 2000. From ionic currents to molecular mechanisms: the structure and function of voltage-gated sodium channels. Neuron 26, 13–25. (General review of sodium channel structure, function and pharmacology)
- Colquhoun, D., 2006. Agonist-activated ion channels. Br. J. Pharmacol. 147, S17–S26. (Review article discussing the relationship between agonist binding and channel opening)
- Halliwell, R.F., 2007. A short history of the rise of the molecular pharmacology of ionotropic drug receptors. Trends Pharmacol. Sci. 28, 214–219. (Good account of the main discoveries in this active field)
- Hille, B., 2001. Ionic Channels of Excitable Membranes. Sinauer Associates, Sunderland. (A clear and detailed account of the basic principles of ion channels, with emphasis on their biophysical properties)
- North, R.A., 2002. Molecular physiology of P2X receptors. Physiol. Rev. 82, 1013–1067. (Encyclopedic review of P2X receptor structure and function)
- Poduri, A., Lowenstein, D., 2011. Epilepsy genetics – past, present, and future. Curr. Opin. Genet. Dev. 21, 325–332.
- Stathopoulos, P.B., Ikura, M., 2017. Store operated calcium entry: From concept to structural mechanisms. Cell Calcium 63, 3–7.
- G protein-coupled receptors**
- AbdAlla, S., Loether, H., El Massiery, A., Quitterer, U., 2001. Increased AT_1 receptor heterodimers in preeclampsia mediate enhanced angiotensin II responsiveness. Nat. Med. 7, 1003–1009. (The first instance of disturbed GPCR heterodimerisation in relation to human disease)
- Adams, M.N., Ramachandran, R., Yau, M.K., et al., 2011. Structure, function and pathophysiology of protease activated receptors. Pharmacol. Ther. 130, 248–282. (Extensive review of the topic)
- Audet, M., Bouvier, M., 2012. Restructuring G protein-coupled receptor activation. Cell 151, 14–23. (Review of recent developments related to G protein-coupled receptor crystallisation)
- Conigrave, A.D., Quinn, S.J., Brown, E.M., 2000. Cooperative multi-modal sensing and therapeutic implications of the extracellular Ca^{2+} -sensing receptor. Trends Pharmacol. Sci. 21, 401–407. (Short account of the Ca^{2+} -sensing receptor, an anomalous type of GPCR)
- Costa, T., Cotecchia, S., 2005. Historical review: negative efficacy and the constitutive activity of G protein-coupled receptors. Trends Pharmacol. Sci. 26, 618–624. (A clear and thoughtful review of ideas relating to constitutive receptor activation and inverse agonists)
- Ferré, S., Casadó, V., Devi, L.A., et al., 2015. G protein-coupled receptor oligomerization revisited: functional and pharmacological perspectives. Pharmacol. Rev. 66, 413–434. (Comprehensive review of GPCR oligomerisation and the implications for pharmacology and cellular signalling)
- Fredriksson, R., Schiöth, H.B., 2005. The repertoire of G protein-coupled receptors in fully sequenced genomes. Mol. Pharmacol. 67, 1414–1425. (Estimation of the number of GPCR genes in different species – nearly 500 more in mouse than in human!)
- Kelly, E., Bailey, C.P., Henderson, G., 2008. Agonist-selective mechanisms of GPCR desensitization. Br. J. Pharmacol. 153 (Suppl. 1), S379–S388. (Short review of main mechanisms of GPCR desensitisation)
- Kenakin, T., Christopoulos, A., 2013. Signalling bias in new drug discovery: detection, quantification and therapeutic impact. Nat. Rev. Drug Discov. 12, 205–216. (Detailed discussion of the difficulties in measuring agonist efficacy and bias)
- Liu, F., Wan, Q., Pristupa, Z., et al., 2000. Direct protein–protein coupling enables cross-talk between dopamine D_5 and γ -aminobutyric acid A receptors. Nature 403, 274–280. (The first demonstration of direct coupling of a GPCR with an ion channel. Look, no G protein!)
- Manglik, A., Lin, H., Aryal, D.K., et al., 2016. Structure-based discovery of opioid analgesics with reduced side effects. Nature 537, 185–190. (Exciting report of the use of computational molecular modelling in the design of a new opioid receptor agonist)
- Milligan, G., Kostenis, E., 2006. Heterotrimeric G proteins: a short history. Br. J. Pharmacol. 147 (Suppl. 1), 46–55.
- Offermanns, S., 2003. G proteins as transducers in transmembrane signalling. Prog. Biophys. Mol. Biol. 83, 101–130. (Detailed review of G protein subtypes and their function in signal transduction)
- Oldham, W.M., Hamm, H.E., 2008. Heterotrimeric G protein activation by G protein-coupled receptors. Nat. Rev. Mol. Cell Biol. 9, 60–71. (Useful review of the structure and function of G proteins)
- Pierce, K.L., Lefkowitz, R.J., 2001. Classical and new roles of β -arrestins in the regulation of G protein-coupled receptors. Nat. Rev. Neurosci. 2, 727–733. (Good review of non-G protein-mediated signalling of GPCRs through arrestin)
- Ramachandran, R., Noorbakhsh, F., defea, K., Hollenberg, M.D., 2012. Targeting proteinase-activated receptors: therapeutic potential and challenges. Nat. Rev. Drug Discov. 11, 69–86.
- Sexton, P.M., Poyner, D.R., Simms, J., Christopoulos, A., Hay, D.L., 2012. RAMPs as drug targets. Adv. Exp. Med. Biol. 744, 61–74.

- Simonds, W.F., 1999. G protein regulation of adenylate cyclase. *Trends Pharmacol. Sci.* 20, 66–72. (*Review of mechanisms by which G proteins affect adenyl cyclase at the level of molecular structure*)
- Sjögren, B., 2017. The evolution of regulators of G protein signalling proteins as drug targets – 20 years in the making: IUPHAR Review 21. *Br. J. Pharmacol.* 174, 427–437. (*Useful description of how RGS proteins function and their potential as drug targets*)
- Sounier, R., Mas, C., Steyaert, J., et al., 2015. Propagation of conformational changes during μ -opioid receptor activation. *Nature* 524, 375–378. (*Describes the use of NMR to study conformational changes in a GPCR upon agonist activation*)
- Spiegel, A.M., Weinstein, L.S., 2004. Inherited diseases involving G proteins and G protein-coupled receptors. *Annu. Rev. Med.* 55, 27–39. (*Short review article*)
- Stoy, H., Gurevich, V.V., 2015. How genetic errors in GPCRs affect their function: Possible therapeutic strategies. *Genes Dis.* 2, 108–132. (*Extensive review with many examples of GPCR mutations associated with disease*)
- Xie, G.X., Palmer, P.P., 2007. How regulators of G protein signalling achieve selective regulation. *J. Mol. Biol.* 366, 349–365. (*General review about RGS proteins and how they work*)
- Zhang, D., Zhao, Q., Wu, B., 2015. Structural studies of G protein-coupled receptors. *Mol. Cells* 38, 836–842. (*Informative review of GPCR crystal structures*)
- Signal transduction**
- Avruch, J., 2007. MAP kinase pathways: the first twenty years. *Biochim. Biophys. Acta* 1773, 1150–1160. (*Short general review. One of a series of articles on MAP kinases in this issue*)
- Bishop, A.L., Hall, R.A., 2000. Rho-GTPases and their effector proteins. *Biochem. J.* 348, 241–255. (*General review article on the Rho/Rho kinase system and the various pathways and functions that it controls*)
- Brzostowski, J.A., Kimmel, A.R., 2001. Signaling at zero: G protein-independent functions for 7TM receptors. *Trends Biochem. Sci.* 26, 291–297. (*Review of evidence for GPCR signalling that does not involve G proteins, thus conflicting with the orthodox dogma*)
- Vanhaesebroeck, B., Leevers, S.J., Panayotou, G., Waterfield, M.D., 1997. Phosphoinositide 3-kinases: a conserved family of signal transducers. *Trends Biochem. Sci.* 22, 267–272. (*Review by the group that discovered PI-3 kinases, summarising the multiple roles of this signal transduction mechanism – much expanded since 1997*)
- Kinase-linked receptors**
- Cohen, P., 2002. Protein kinases – the major drug targets of the twenty-first century. *Nat. Rev. Drug Discov.* 1, 309–315. (*General review on pharmacological aspects of protein kinases*)
- Cook, D.N., Pisetsky, D.S., Schwartz, D.A., 2004. Toll-like receptors in the pathogenesis of human disease. *Nat. Immunol.* 5, 975–979. (*Review emphasising the role of this class of receptor tyrosine kinases in many human disease states*)
- Delcourt, N., Bockaert, J., Marin, P., 2007. GPCR-jacking: from a new route in RTK signalling to a new concept in GPCR activation. *Trends Pharmacol. Sci.* 28, 602–607. (*Gives examples of ‘cross-talk’ between GPCR and RTK signalling pathways*)
- Hubbard, S.R., Miller, W.T., 2007. Receptor tyrosine kinases: mechanisms of activation and signaling. *Curr. Opin. Cell Biol.* 19, 117–123. (*Reviews recent structural results showing mechanism of RTK dimerisation and signalling*)
- Ihle, J.N., 1995. Cytokine receptor signalling. *Nature* 377, 591–594.
- Jin, J., Pawson, T., 2012. Modular evolution of phosphorylation-based signalling systems. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 367, 2540–2555. (*Informative review on receptor kinase signalling*)
- Karin, M., Yamamoto, Y., Wang, M., 2004. The IKK-NF κ B system: a treasure trove for drug development. *Nat. Rev. Drug Discov.* 3, 17–26. (*Describes the transcription factor NF κ B, which plays a key role in inflammation, and its control by kinase cascades*)
- Nuclear receptors**
- Bourguel, W., Germain, P., Gronemeyer, H., 2000. Nuclear receptor ligand-binding domains: three-dimensional structures, molecular interactions and pharmacological implications. *Trends Pharmacol. Sci.* 21, 381–388. (*Accessible review concentrating on distinction between agonist and antagonist effects at the molecular level*)
- Burris, T.P., Solt, L.A., Wang, Y., et al., 2013. Nuclear receptors and their selective pharmacologic modulators. *Pharmacol. Rev.* 65, 710–778. (*A very comprehensive account of the action of drugs at nuclear receptors. Not light reading, but worthwhile if you really want to dig deeply into this fascinating subject*)
- Evans, R.M., Mangelsdorf, D.J., 2014. Nuclear receptors, RXR, and the big bang. *Cell* 157, 255–266. (*Easy-to-read review of nuclear receptors focusing on the key role that RXR plays in many crucial physiological signalling systems. Contains a lot of useful background material on nuclear receptors and their history*)
- Falkenstein, E., Tillmann, H.C., Christ, M., Feuring, M., Wehling, M., 2000. Multiple actions of steroid hormones – a focus on rapid non-genomic effects. *Pharm. Rev.* 52, 513–553. (*Comprehensive review article describing the ‘non-classical’ effects of steroids*)
- Germain, P., Staels, B., Dacquet, C., Spedding, M., Laudet, V., 2006. Overview of nomenclature of nuclear receptors. *Pharmacol. Rev.* 58, 685–704. (*Comprehensive and authoritative review that deals with receptor biology as well as nomenclature. Recommended*)
- Kersten, S., Desvergne, B., Wahli, W., 2000. Roles of PPARs in health and disease. *Nature* 405, 421–424. (*General review of an important class of nuclear receptors*)
- di Masi, A., De Marinis, E., Ascenzi, P., Marino, M., 2009. Nuclear receptors CAR and PXR: molecular, functional, and biomedical aspects. *Mol. Aspects Med.* 30, 297–343. (*A very comprehensive account of the role played by these nuclear receptors in xenobiotic metabolism but also contains some useful general background information on the receptor family*)
- Levin, E.R., Hammes, S.R., 2016. Nuclear receptors outside the nucleus: extranuclear signalling by steroid receptors. *Nat. Rev. Mol. Cell Biol.* 17, 783–797. (*This paper reviews the increasing evidence that some of the key effects of steroids are mediated by populations of nuclear receptors outside the nucleus and describes the often rather complex networks through which they bring about these effects*)
- Murphy, G.J., Holder, J.C., 2000. PPAR- γ agonists: therapeutic role in diabetes, inflammation and cancer. *Trends Pharmacol. Sci.* 21, 469–474. (*Account of the emerging importance of nuclear receptors of the PPAR family as therapeutic targets*)
- Santos, G.M., Fairall, L., Schwabe, J.W.R., 2011. Negative regulation by nuclear receptors: a plethora of mechanisms. *Trends Endocrinol. Metab.* 22, 87–93. (*A very accessible and well written introduction to a very complex subject. Highly recommended*)
- Shiau, A.K., Barstad, D., Loria, P.M., et al., 1998. The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. *Cell* 95, 927–937.

4

How drugs act: cellular aspects – excitation, contraction and secretion

OVERVIEW

The link between a drug interacting with a molecular target and its effect at the pathophysiological level, such as a change in blood glucose concentration or the shrinkage of a tumour, involves events at the cellular level. Whatever their specialised physiological function, cells generally share much the same repertoire of signalling mechanisms. In the next four chapters, we describe the parts of this repertoire that are of particular significance in understanding drug action at the cellular level. In this chapter, we describe mechanisms that operate mainly over a short timescale (milliseconds to hours), particularly excitation, contraction and secretion, which account for many physiological responses; Chapter 5 looks at how biopharmaceuticals and gene therapy may alter the cell's chemical behaviours to have their desired effects; Chapter 6 deals with the slower processes (generally days to months), including cell division, growth, differentiation and cell death, that determine the body's structure and constitution; Chapter 7 describes host defence mechanisms.

The short-term regulation of cell function depends mainly on the following components and mechanisms, which regulate, or are regulated by, the free concentration of Ca^{2+} in the cytosol, $[\text{Ca}^{2+}]_i$:

- ion channels and transporters in the plasma membrane
- the storage and release of Ca^{2+} by intracellular organelles
- Ca^{2+} -dependent regulation of a variety of functional proteins, including enzymes, contractile proteins and vesicle proteins

More detailed coverage of the topics presented in this chapter can be found in [Berridge \(2014\)](#) and [Kandel et al. \(2012\)](#).

Because $[\text{Ca}^{2+}]_i$ plays such a key role in cell function, a wide variety of drug effects result from interference with one or more of these mechanisms. Knowledge of the molecular and cellular details is extensive, and here we focus on the aspects that help to explain drug effects.

REGULATION OF INTRACELLULAR CALCIUM

Ever since the famous accident in 1882 by Sidney Ringer's technician, which showed that using tap water rather than distilled water to make up the bathing solution for isolated

frog hearts would allow them to carry on contracting, the role of Ca^{2+} as a major regulator of cell function has never been in question. Many drugs and physiological mechanisms operate, directly or indirectly, by influencing $[\text{Ca}^{2+}]_i$. Here we consider the main ways in which it is regulated, and later we describe some of the ways in which $[\text{Ca}^{2+}]_i$ controls cell function. Details of the molecular components and drug targets are presented in Chapter 3, and descriptions of drug effects on integrated physiological function are given in later chapters.

The study of Ca^{2+} regulation took a big step forward in the 1970s with the development of optical techniques based on the Ca^{2+} -sensitive photoprotein *aequorin*, and fluorescent dyes such as *Fura-2*, which, for the first time, allowed free $[\text{Ca}^{2+}]_i$ to be continuously monitored in living cells with a high level of temporal and spatial resolution.

Most of the Ca^{2+} in a resting cell is sequestered in organelles, particularly the *endoplasmic* or *sarcoplasmic reticulum* (ER or SR) and the mitochondria, and the free $[\text{Ca}^{2+}]_i$ is kept to a low level, about 100 nmol/L. The Ca^{2+} concentration in extracellular fluid, $[\text{Ca}^{2+}]_o$, is about 2.4 mmol/L, so there is a large concentration gradient favouring Ca^{2+} entry. $[\text{Ca}^{2+}]_i$ is kept low (a) by the operation of active transport mechanisms that eject cytosolic Ca^{2+} through the plasma membrane and pump it into the ER, and (b) by the normally low Ca^{2+} permeability of the plasma and ER membranes. Regulation of $[\text{Ca}^{2+}]_i$ involves three main mechanisms:

- control of Ca^{2+} entry
- control of Ca^{2+} extrusion
- exchange of Ca^{2+} between the cytosol and the intracellular stores

These mechanisms are described in more detail later and are summarised in [Fig. 4.1](#).

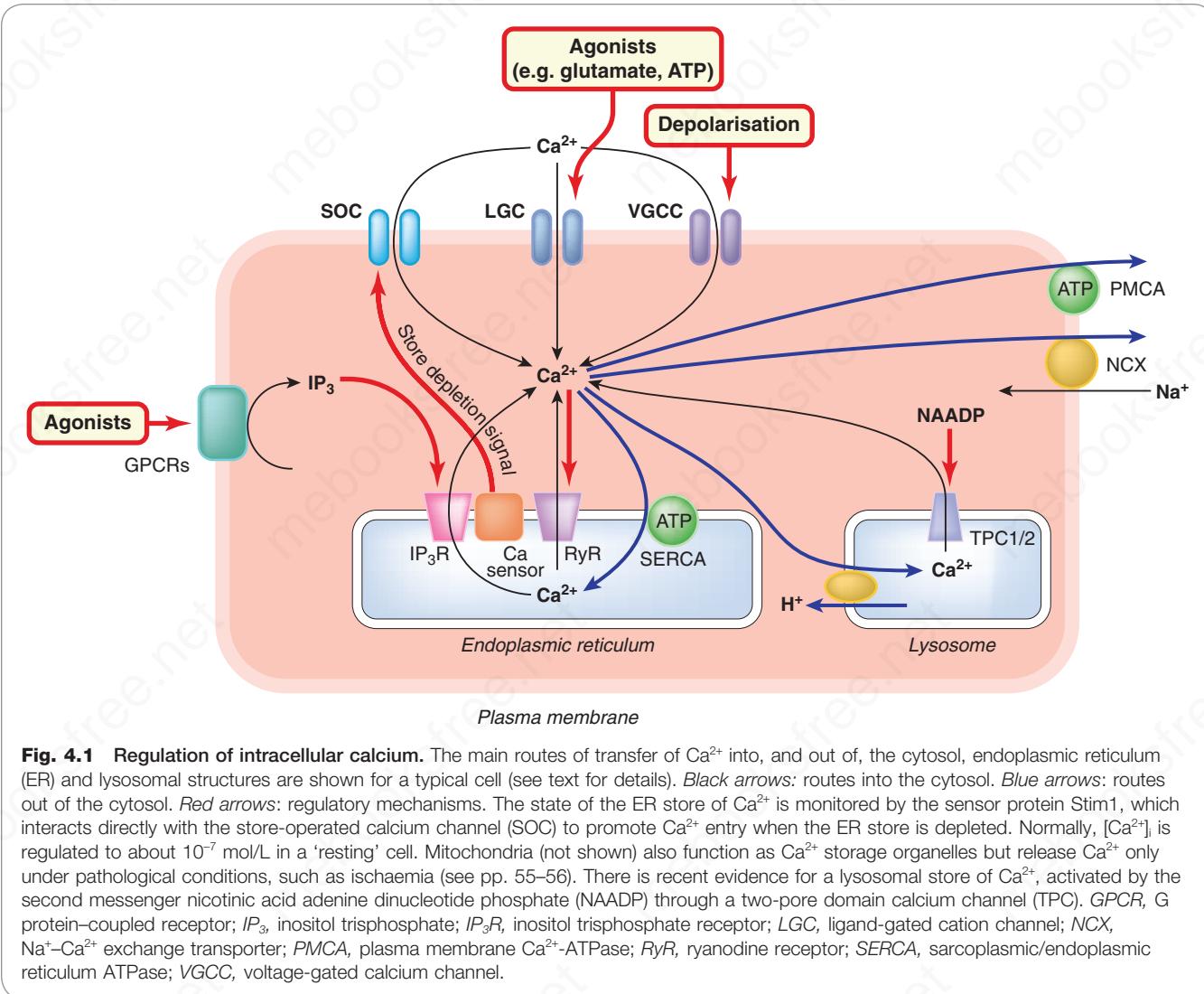
CALCIUM ENTRY MECHANISMS

There are four main routes by which Ca^{2+} enters cells across the plasma membrane:

- voltage-gated calcium channels
- ligand-gated calcium channels
- store-operated calcium channels (SOCs)
- $\text{Na}^+-\text{Ca}^{2+}$ exchange (can operate in either direction; see [Calcium extrusion mechanisms](#), p. 55)

VOLTAGE-GATED CALCIUM CHANNELS

The pioneering work of Hodgkin and Huxley on the ionic basis of the nerve action potential (see pp. 56–58) identified voltage-dependent Na^+ and K^+ conductances as the main participants. It was later found that some invertebrate nerve and muscle cells could produce action potentials that depended on Ca^{2+} rather than Na^+ , and it was then found that vertebrate cells also possess voltage-activated calcium channels capable of allowing substantial amounts



of Ca^{2+} to enter the cell when the membrane is depolarised. These voltage-gated channels are highly selective for Ca^{2+} (although they also conduct Ba^{2+} ions, which are often used as a substitute in electrophysiological experiments), and do not conduct Na^+ or K^+ ; they are ubiquitous in excitable cells and cause Ca^{2+} to enter the cell whenever the membrane is depolarised, for example by a conducted action potential.

A combination of electrophysiological and pharmacological criteria have revealed five distinct subtypes of voltage-gated calcium channels: L, T, N, P/Q and R¹. The subtypes vary with respect to their activation and inactivation kinetics, their voltage threshold for activation, their conductance and their sensitivity to blocking agents, as summarised in Table 4.1. The molecular basis for this heterogeneity has been worked out in some detail. The main pore-forming subunit (termed

¹P and Q are so similar that they usually get lumped together. The terminology is less than poetic: L stands for *long-lasting*; T stands for *transient*; N stands for *neither long-lasting nor transient*. Although P stands for *Purkinje* – this type of channel was first observed in cerebellar Purkinje cells – it continued the alphabetical sequence (missing out O of course) and so the next discovered were termed Q and R.

α_1 , see Fig. 3.20) occurs in at least 10 molecular subtypes, and associates with other subunits (β , γ and two subunits from the same gene, $\alpha_2\delta$, linked by a disulfide bond) that also exist in different subtypes to form the functional channel. Different combinations of these subunits give rise to the different physiological subtypes². In general, L channels are particularly important in regulating contraction of cardiac and smooth muscle (see p. 64), and N channels (and also P/Q) are involved in neurotransmitter and hormone release, while T channels mediate Ca^{2+} entry into neurons around the resting membrane potential and can control the rate of repolarisation of neurons and cardiac cells as well as various Ca^{2+} -dependent functions such as regulation of other channels, enzymes, etc. Clinically used drugs that act directly on some forms of calcium channel include the group of ' Ca^{2+} antagonists' consisting of *dihydropyridines* (e.g. **nifedipine**), **verapamil** and **diltiazem** (used for their

²Readers interested in knowing more about the subunit composition of different voltage-gated calcium channels should consult the Guide to Pharmacology at <http://www.guidetopharmacology.org/GRAC/FamilyDisplayForward?familyId=80>

Table 4.1 Types and functions of Ca^{2+} channels

Gated by	Main types	Characteristics	Location and function	Drug effects
Voltage	L	High activation threshold Slow inactivation	Plasma membrane of many cells Main Ca^{2+} source for contraction in smooth and cardiac muscle	Blocked by dihydropyridines, verapamil, diltiazem; and calciseptine (peptide from snake venom) Activated by BayK 8644 Phosphorylation by PKA (e.g. following β_1 adrenoceptor activation) increases channel opening
	N	Low activation threshold Slow inactivation	Main Ca^{2+} source for transmitter release by nerve terminals	Blocked by ω -conotoxin GV1A (component of <i>Conus</i> snail venom) and ziconotide (marketed preparation of ω -conotoxin used to control pain) (Ch. 43)
	T	Low activation threshold Fast inactivation	Widely distributed Important in cardiac pacemaker and atria (role in dysrhythmias), also neuronal firing patterns	Blocked by mibepradil
	P/Q	Low activation threshold Slow inactivation	Nerve terminals Transmitter release	Blocked by ω -agatoxin-4A (component of funnel-web spider venom)
	R	Low threshold Fast inactivation	Neurons and dendrites Control of firing patterns	Blocked by low concentrations of SNX-482 (a toxin from a member of the tarantula family)
IP ₃	IP ₃ receptor	Activated by binding of IP ₃ and Ca^{2+}	Located in endoplasmic/sarcoplasmic reticulum Mediates Ca^{2+} release produced by GPCR activation	Not directly targeted by drugs Some experimental blocking agents known Responds to GPCR agonists and antagonists in many cells
Ca^{2+}	Ryanodine receptor	Directly activated in skeletal muscle via dihydropyridine receptor of T-tubules. Activated by Ca^{2+} in cardiac muscle	Located in endoplasmic/sarcoplasmic reticulum. Pathway for Ca^{2+} release in striated muscle	Activated by caffeine and ATP in the presence of Ca^{2+} Ryanodine both activates (low concentrations) and closes (high concentrations) the channel. Also closed by Mg^{2+} , K^+ channel blockers and dantrolene Mutations may lead to drug-induced malignant hypothermia, sudden cardiac death and central core disease
Store depletion	Store-operated channels	Activated by sensor protein that monitors level of ER Ca^{2+} stores	Located in plasma membrane	Activated indirectly by agents that deplete intracellular stores (e.g. GPCR agonists, thapsigargin) Not directly targeted by drugs

ER, endoplasmic reticulum; GPCR, G protein-coupled receptor; IP₃, inositol trisphosphate; PKA, protein kinase A.

cardiovascular effects; see Chs 22 and 23), and **gabapentin** and **pregabalin** (used to treat epilepsy, pain and anxiety; see Chs 43, 45 and 46). Many drugs affect calcium channels indirectly by acting on G protein-coupled receptors (see Ch. 3). A number of toxins act selectively on one or other type of calcium channel (see Table 4.1), and these are used as experimental tools.

LIGAND-GATED CHANNELS

Most ligand-gated cation channels (see Ch. 3) that are activated by excitatory neurotransmitters are relatively non-selective, and conduct Ca^{2+} ions as well as other cations. Most important in this respect is the glutamate receptor of the NMDA type (Ch. 39), which has a particularly high

permeability to Ca^{2+} and is a major contributor to Ca^{2+} uptake by postsynaptic neurons (and also glial cells) in the central nervous system. Activation of this receptor can readily cause so much Ca^{2+} entry that the cell dies, mainly through activation of Ca^{2+} -dependent proteases but also by triggering *apoptosis* (see Ch. 6). This mechanism, termed *excitotoxicity*, probably plays a part in various neurodegenerative disorders (see Ch. 41).

For many years, there was dispute about the existence of 'receptor-operated channels' in smooth muscle, responding directly to mediators such as adrenaline (epinephrine), acetylcholine and histamine. Now it seems that the P2X receptor (see Ch. 3), activated by ATP, is the only example of a true ligand-gated channel in smooth muscle, and this

constitutes an important route of entry for Ca^{2+} . As mentioned above, many mediators acting on G protein-coupled receptors affect Ca^{2+} entry indirectly, mainly by regulating voltage-gated calcium channels or potassium channels.

STORE-OPERATED CALCIUM CHANNELS (SOCs)

SOCs are very low-conductance channels that occur in the plasma membrane and open to allow entry when the ER stores are depleted, but are not sensitive to cytosolic $[\text{Ca}^{2+}]_i$. The linkage between the ER and the plasma membrane involves a Ca^{2+} -sensor protein (*Stim1*) in the ER membrane, which connects directly to the channel protein (*Orai1*) in the plasma membrane. Depletion of ER Ca^{2+} causes *Stim1* to accumulate at junctions between the ER and the plasma membrane where it traps and activates *Orai1* resulting in Ca^{2+} entry (see [Prakriya & Lewis, 2015](#)).

Like the ER and SR channels, these channels can serve to amplify the rise in $[\text{Ca}^{2+}]_i$ resulting from Ca^{2+} release from the stores. So far, only experimental compounds are known to block these channels, but efforts are being made to develop specific blocking agents for therapeutic use as relaxants of smooth muscle.

CALCIUM EXTRUSION MECHANISMS

Active transport of Ca^{2+} outwards across the plasma membrane, and inwards across the membranes of the ER or SR, depends on the activity of distinct Ca^{2+} -dependent ATPases,³ similar to the Na^+/K^+ -dependent ATPase that pumps Na^+ out of the cell in exchange for K^+ . **Thapsigargin** (derived from a Mediterranean plant, *Thapsia gargarica*) specifically blocks the ER pump, causing loss of Ca^{2+} from the ER. It is a useful experimental tool but has no therapeutic significance.

Calcium is also extruded from cells in exchange for Na^+ , by $\text{Na}^+-\text{Ca}^{2+}$ exchange. The exchanger transfers three Na^+ ions for one Ca^{2+} , and therefore produces a net depolarising current when it is extruding Ca^{2+} . The energy for Ca^{2+} extrusion comes from the electrochemical gradient for Na^+ , not directly from ATP hydrolysis. This means that a reduction in the Na^+ concentration gradient resulting from Na^+ entry will reduce Ca^{2+} extrusion by the exchanger, causing a secondary rise in $[\text{Ca}^{2+}]_i$, a mechanism that is particularly important in cardiac muscle (see Ch. 22). **Digoxin** (derived from the *Digitalis* or ‘Foxglove’ plant), which inhibits Na^+ extrusion, acts on cardiac muscle in this way (Ch. 22), causing $[\text{Ca}^{2+}]_i$ to increase.

CALCIUM RELEASE MECHANISMS

There are two main types of calcium channel in the ER and SR membrane, which play an important part in controlling the release of Ca^{2+} from these stores.

- The *inositol trisphosphate receptor* (IP₃R) is activated by inositol trisphosphate (IP₃), a second messenger produced by the action of many ligands on G protein-coupled receptors (see Ch. 3). IP₃R is a ligand-gated ion channel, although its molecular structure differs from that of ligand-gated channels in the plasma membrane (see [Berridge, 2016](#)). This is the main mechanism by which activation of Gq-coupled receptors causes an increase in $[\text{Ca}^{2+}]_i$.

³These pumps have been likened to Sisyphus, condemned endlessly to push a stone up a hill (also consuming ATP, no doubt), only for it to roll down again.

- Ryanodine receptors (RyR) are so called because they were first identified through the specific blocking action of the plant alkaloid **ryanodine**. There are three isoforms – RyR1-3 ([Van Petegem, 2012](#)), which are expressed in many different cell types. RyR1 is highly expressed in skeletal muscle, RyR2 in the heart and RyR3 in brain neurons. In skeletal muscle, RyRs on the SR are physically coupled to *dihydropyridine receptors* on the T-tubules (see [Fig. 4.9](#)); this coupling results in rapid Ca^{2+} release following the action potential in the muscle fibre. In other muscle types, RyRs respond to Ca^{2+} that enters the cell through membrane calcium channels by a mechanism known as *calcium-induced calcium release* (CICR).

The functions of IP₃Rs and RyRs are modulated by a variety of other intracellular signals (see [Berridge, 2016](#); [Van Petegem, 2012](#)), which affect the magnitude and spatiotemporal patterning of Ca^{2+} signals. Fluorescence imaging techniques have revealed a remarkable level of complexity of Ca^{2+} signals, and much remains to be discovered about the importance of this patterning in relation to physiological and pharmacological mechanisms. The Ca^{2+} sensitivity of RyRs is increased by **caffeine**, causing Ca^{2+} release from the SR even at resting levels of $[\text{Ca}^{2+}]_i$. This is used experimentally but rarely happens in humans, because the other pharmacological effects of caffeine (see Ch. 49) occur at much lower doses. The blocking effect of **dantrolene**, a compound related to ryanodine, is used therapeutically to relieve muscle spasm in the rare condition of *malignant hyperthermia* (see Ch. 42), which is associated with inherited abnormalities in the RyR protein.

A typical $[\text{Ca}^{2+}]_i$ signal resulting from activation of a Gq-coupled receptor is shown in [Fig. 4.2A](#). The response produced in the absence of extracellular Ca^{2+} represents release from intracellular stores. The larger and more prolonged response when extracellular Ca^{2+} is present shows the contribution of SOC-mediated Ca^{2+} entry. The various positive and negative feedback mechanisms that regulate $[\text{Ca}^{2+}]_i$ give rise to a variety of temporal and spatial oscillatory patterns ([Fig. 4.2B](#)) that are responsible for spontaneous rhythmic activity in smooth muscle and nerve cells (see [Berridge, 2008](#)).

OTHER SECOND MESSENGERS

- Two intracellular metabolites, cyclic ADP-ribose (cADPR) and nicotinic acid adenine dinucleotide phosphate (NAADP) formed from the ubiquitous coenzymes nicotinamide adenine dinucleotide (NAD) and NAD phosphate, also affect Ca^{2+} signalling (see [Morgan et al., 2015](#); [Parrington et al., 2015](#)). cADPR acts by increasing the sensitivity of RyRs to Ca^{2+} , thus increasing the ‘gain’ of the CICR effect, whereas NAADP has been proposed to release Ca^{2+} from lysosomes by activating two-pore domain calcium channels ([Fig. 4.1](#)).

The levels of these messengers in mammalian cells may be regulated mainly in response to changes in the metabolic status of the cell, although the details are not yet clear. Abnormal Ca^{2+} signalling is involved in many pathophysiological conditions, such as ischaemic cell death, endocrine disorders and cardiac dysrhythmias, where the roles of cADPR and NAADP, and their interaction with other mechanisms that regulate $[\text{Ca}^{2+}]_i$, are the subject of much current work.

THE ROLE OF MITOCHONDRIA

- Under normal conditions, mitochondria accumulate Ca^{2+} passively as a result of the intramitochondrial potential, which is strongly negative with respect to the cytosol. This negativity is maintained

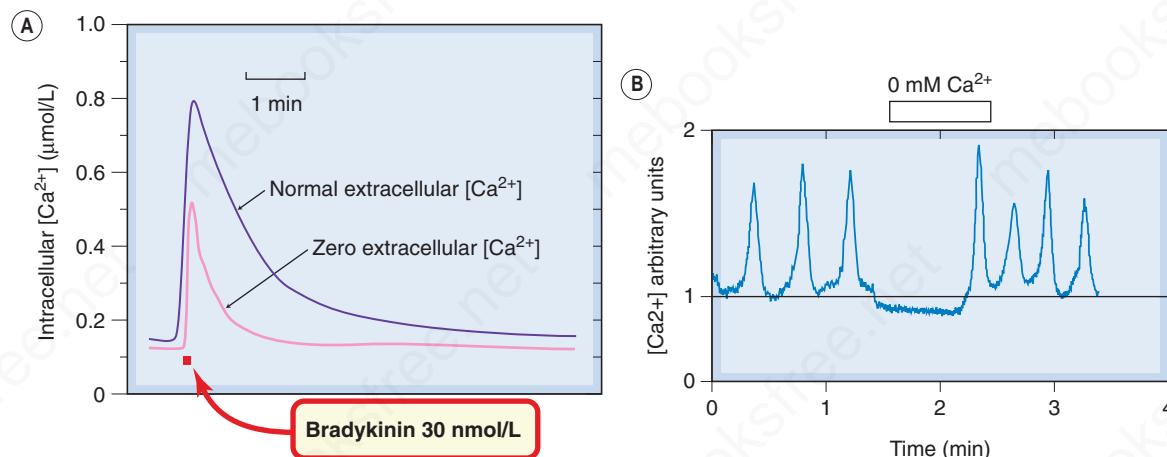


Fig. 4.2 (A) Increase in intracellular free calcium concentration in response to receptor activation. The records were obtained from a single rat sensory neuron grown in tissue culture. The cells were loaded with the fluorescent Ca²⁺ indicator Fura-2, and the signal from a single cell monitored with a fluorescence microscope. A brief exposure to the peptide bradykinin, which causes excitation of sensory neurons (see Ch. 43), causes a transient increase in [Ca²⁺]_i from the resting value of about 150 nmol/L. When Ca²⁺ is removed from the extracellular solution, the bradykinin-induced increase in [Ca²⁺]_i is still present but is smaller and briefer. The response in the absence of extracellular Ca²⁺ represents the release of stored intracellular Ca²⁺ resulting from the intracellular production of inositol triphosphate. The difference between this and the larger response when Ca²⁺ is present extracellularly is believed to represent Ca²⁺ entry through store-operated ion channels in the cell membrane. (Figure kindly provided by G. M. Burgess and A. Forbes, Novartis Institute for Medical Research.) (B) Spontaneous intracellular calcium oscillations in pacemaker cells from the rabbit urethra that regulate the rhythmic contractions of the smooth muscle. The signals cease when external Ca²⁺ is removed, showing that activation of membrane Ca²⁺ channels is involved in the mechanism. (From McHale, N., et al., 2006. J. Physiol. 570, 23–28.)

by active extrusion of protons, and is lost – thus releasing Ca²⁺ into the cytosol – if the cell runs short of ATP, for example under conditions of hypoxia. This only happens *in extremis*, and the resulting Ca²⁺ release contributes to the cytotoxicity associated with severe metabolic disturbance. Cell death resulting from brain ischaemia or coronary ischaemia (see Chs 22 and 41) involves this mechanism, along with others that contribute to an excessive rise in [Ca²⁺]_i.

CALMODULIN

Calcium exerts its control over cell functions by virtue of its ability to regulate the activity of many different proteins, including enzymes (particularly kinases and phosphatases), channels, transporters, transcription factors, synaptic vesicle proteins and many others either by binding directly to these proteins or through a Ca²⁺-binding protein that serves as an intermediate between Ca²⁺ and the regulated functional protein, the best known such binding protein being the ubiquitous *calmodulin*. This regulates at least 40 different functional proteins – indeed a powerful fixer. Calmodulin is a dumbbell-shaped protein with a globular domain at either end, each with two Ca²⁺ binding sites. When all are occupied, the protein undergoes a conformational change, exposing a ‘sticky’ hydrophobic domain that lures many proteins into association, thereby affecting their functional properties.

EXCITATION

Excitability describes the ability of a cell to show a regenerative all-or-nothing electrical response to depolarisation of its membrane, this membrane response being known as

Calcium regulation

Intracellular Ca²⁺ concentration, [Ca²⁺]_i, is critically important as a regulator of cell function.

- Intracellular Ca²⁺ is determined by (a) Ca²⁺ entry; (b) Ca²⁺ extrusion; and (c) Ca²⁺ exchange between the cytosol, endoplasmic or sarcoplasmic reticulum (ER, SR), lysosomes and mitochondria.
- Calcium entry occurs by various routes, including voltage- and ligand-gated calcium channels and Na⁺–Ca²⁺ exchange.
- Calcium extrusion depends mainly on an ATP-driven Ca²⁺ pump.
- Calcium ions are actively taken up and stored by the ER/SR, from which they are released in response to various stimuli.
- Calcium ions are released from ER/SR stores by (a) the second messenger inositol triphosphate (IP₃) acting on IP₃ receptors; or (b) increased [Ca²⁺]_i itself acting on ryanodine receptors, a mechanism known as Ca²⁺-induced Ca²⁺ release.
- Other second messengers, cyclic ADP-ribose and nicotinic acid dinucleotide phosphate, also promote the release of Ca²⁺ from Ca²⁺ stores.
- Depletion of ER/SR Ca²⁺ stores promotes Ca²⁺ entry through the plasma membrane, via store-operated channels.
- Calcium ions affect many aspects of cell function by binding to proteins such as calmodulin, which in turn bind other proteins and regulate their function.

an action potential. It is a characteristic of most neurons and muscle cells (including skeletal, cardiac and smooth muscle) and of many endocrine gland cells. In neurons and muscle cells, the ability of the action potential, once initiated, to propagate to all parts of the cell membrane, and often to spread to neighbouring cells, explains the importance of membrane excitation in intra- and intercellular signalling. In the nervous system and in skeletal muscle, action potential propagation is the mechanism responsible for communication over long distances at high speed, indispensable for large, fast-moving creatures. In cardiac and smooth muscle, as well as in some central neurons, spontaneous rhythmic activity occurs. In gland cells, the action potential, where it occurs, serves to amplify the signal that causes the cell to secrete. In each type of tissue, the properties of the excitation process reflect the special characteristics of the ion channels that underlie the process. The molecular nature of ion channels, and their importance as drug targets, is considered in Chapter 3; here we discuss the cellular processes that depend primarily on ion channel function. For more detail, see Hille (2001).

THE 'RESTING' CELL

The resting cell is not resting at all but very busy controlling the state of its interior, and it requires a continuous supply of energy to do so. In relation to the topics discussed in this chapter, the following characteristics are especially important:

- membrane potential
- permeability of the plasma membrane to different ions
- intracellular ion concentrations, especially $[Ca^{2+}]_i$

Under resting conditions, all cells maintain a negative internal potential between about -30 and -80 mV, depending on the cell type. This arises because (a) the membrane is relatively impermeable to Na^+ , and (b) Na^+ ions are actively extruded from the cell in exchange for K^+ ions by an energy-dependent transporter, the Na^+ pump (or Na^+-K^+ -ATPase). The result is that the intracellular K^+ concentration, $[K^+]_i$, is higher, and $[Na^+]_i$ is lower, than the respective extracellular concentrations. In many cells, other ions, particularly Cl^- , are also actively transported and unequally distributed across the membrane. In many cases (e.g. in neurons), the membrane permeability to K^+ is relatively high, and the membrane potential settles at a value of -60 to -80 mV, close to the equilibrium potential for K^+ (Fig. 4.3). In other cells (e.g. smooth muscle), anions play a larger part, and the membrane potential is generally lower (-30 to -50 mV) and less dependent on K^+ .

ELECTRICAL AND IONIC EVENTS UNDERLYING THE ACTION POTENTIAL

Our present understanding of electrical excitability rests firmly on the work of Hodgkin, Huxley and Katz on squid axons, published in 1949–1952. Their experiments (see Katz, 1966) revealed the existence of voltage-gated ion channels (see pp. 60–61) and showed that the action potential is generated by the interplay of two processes:

1. A rapid, transient increase in Na^+ permeability that occurs when the membrane is depolarised beyond about -50 mV.
2. A slower, sustained increase in K^+ permeability.

Because of the inequality of Na^+ and K^+ concentrations on the two sides of the membrane, an increase in Na^+

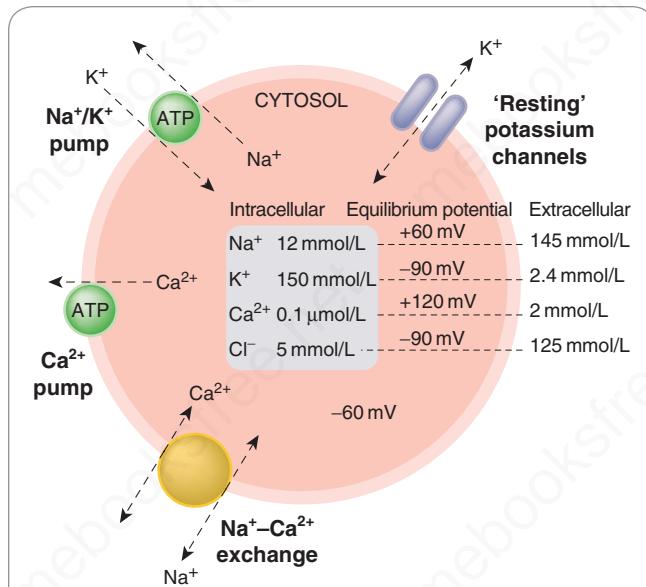


Fig. 4.3 Simplified diagram showing the ionic balance of a typical 'resting' cell. The main transport mechanisms that maintain the ionic gradients across the plasma membrane are the ATP-driven Na^+-K^+ and Ca^{2+} pumps and the Na^+-Ca^{2+} exchange transporter. The membrane is relatively permeable to K^+ , because some types of potassium channel are open at rest, but impermeable to other cations. The unequal ion concentrations on either side of the membrane give rise to the 'equilibrium potentials' shown. The resting membrane potential, typically about -60 mV but differing between different cell types, is determined by the equilibrium potentials and the permeabilities of the various ions involved, and by the 'electrogenic' effect of the transporters. For simplicity, anions and other ions, such as protons, are not shown, although these play an important role in many cell types.

permeability causes an inward (depolarising) current of Na^+ ions, whereas an increase in K^+ permeability causes an outward (repolarising) current. The separability of these two currents can be most clearly demonstrated by the use of drugs blocking sodium and potassium channels, as shown in Fig. 4.4. During the physiological initiation or propagation of a nerve impulse, the first event is a small depolarisation of the membrane, produced either by transmitter action or by the approach of an action potential passing along the axon. This opens sodium channels, allowing an inward current of Na^+ ions to flow, which depolarises the membrane still further. The process is thus a regenerative one, and the increase in Na^+ permeability is enough to bring the membrane potential towards E_{Na} . The increased Na^+ conductance is transient, because the channels inactivate rapidly and the membrane returns to its resting state.

In many types of cell, including most nerve cells, repolarisation is assisted by the opening of voltage-dependent K^+ channels. These function in much the same way as sodium channels, but their activation kinetics are about 10 times slower and they do not inactivate appreciably. This means that the potassium channels open later than the sodium channels, contributing to the rapid termination of the action potential and to the after-hyperpolarisation that follows the depolarising phase. The behaviour of the sodium and

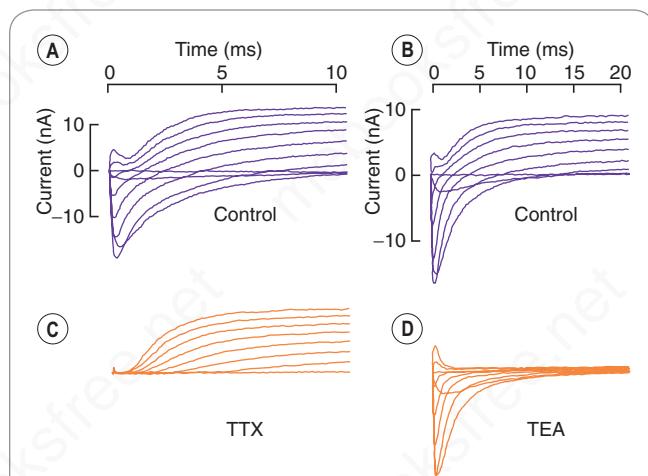


Fig. 4.4 Separation of sodium and potassium currents in the nerve membrane. Voltage clamp records from the node of Ranvier of a single frog nerve fibre. At time 0, the membrane potential was stepped to a depolarised level, ranging from -60 mV (lower trace in each series) to $+60\text{ mV}$ (upper trace in each series) in 15-mV steps. (A and B) Control records from two fibres. (C) Effect of tetrodotoxin (TTX), which abolishes Na^+ currents. (D) Effect of tetraethylammonium (TEA), which abolishes K^+ currents. (From Hille, B., 1970. Ionic channels in nerve membranes. *Prog. Biophys. Mol. Biol.* 21, 1–32.)

through voltage-gated calcium channels plays a key role in intracellular signalling, as described on pp. 52–56.

CHANNEL FUNCTION

The discharge patterns of excitable cells vary greatly. Skeletal muscle fibres are quiescent unless stimulated by the arrival of a nerve impulse at the neuromuscular junction. Cardiac muscle fibres discharge spontaneously at a regular rate (see Ch. 22). Neurons may be normally silent, or they may discharge spontaneously, either regularly or in bursts; smooth muscle cells show a similar variety of firing patterns. The frequency at which different cells normally discharge action potentials also varies greatly, from 100 Hz or more for fast-conducting neurons, down to about 1 Hz for cardiac muscle cells. These very pronounced functional variations reflect the different characteristics of the ion channels expressed in different cell types. Rhythmic fluctuations of $[\text{Ca}^{2+}]_i$ underlie the distinct firing patterns that occur in different types of cell (see Berridge, 2016).

Drugs that alter channel characteristics, either by interacting directly with the channel itself or indirectly through second messengers, affect the function of many organ systems, including the nervous, cardiovascular, endocrine, respiratory and reproductive systems, and are a frequent theme in this book. Here we describe some of the key mechanisms involved in the regulation of excitable cells.

In general, action potentials are initiated by membrane currents that cause depolarisation of the cell. These currents may be produced by synaptic activity, by an action potential approaching from another part of the cell, by a sensory stimulus or by spontaneous *pacemaker* activity. The tendency of such currents to initiate an action potential is governed by the *excitability* of the cell, which depends mainly on the state of (a) the voltage-gated sodium and/or calcium channels, and (b) the potassium channels of the resting membrane. Anything that increases the number of available sodium or calcium channels, or reduces their activation threshold, will tend to increase excitability, whereas increasing the resting K^+ conductance reduces it. Agents that do the reverse, by blocking channels or interfering with their opening, will have the opposite effect. Some examples are shown in Figs 4.6 and 4.7. Inherited mutations of channel proteins are responsible for a wide variety of neurological and other genetic disorders (see Imbrici et al., 2016).

USE DEPENDENCE AND VOLTAGE DEPENDENCE

▼ Voltage-gated channels can exist in three functional states (Fig. 4.8): *resting* (the closed state that prevails at the normal resting potential), *activated* (the open state favoured by brief depolarisation) and *inactivated* (the blocked state resulting from a trapdoor-like occlusion of the open channel by a floppy intracellular appendage of the channel protein). After the action potential has passed, many sodium channels are in the inactivated state; after the membrane potential returns to its resting value, the inactivated channels take time to revert to the resting state and thus become available for activation once more. In the meantime, the membrane is temporarily *refractory*. Each action potential causes the channels to cycle through these states. The duration of the refractory period determines the maximum frequency at which action potentials can occur. Drugs that block sodium channels, such as local anaesthetics (Ch. 44), antidysrhythmic drugs (Ch. 22) and antiepileptic drugs (Ch. 46), commonly show a selective affinity for one or other of these functional states of the channel, and in their presence the proportion of channels in the high-affinity state is increased. Of particular importance are drugs that bind most strongly to the inactivated state of the channel and thus favour the adoption of this state, prolonging the refractory period and reducing the maximum frequency at which action potentials can

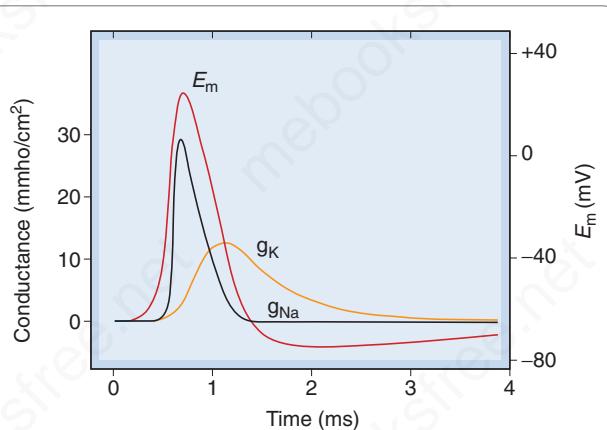


Fig. 4.5 Behaviour of sodium and potassium channels during a conducted action potential. Rapid opening of sodium channels occurs during the action potential upstroke. Delayed opening of potassium channels, and inactivation of sodium channels, causes repolarisation. E_m , membrane potential; g_K , membrane conductance to K^+ ; g_{Na} , membrane conductance to Na^+ .

potassium channels during an action potential is shown in Fig. 4.5.

The foregoing account, based on Hodgkin and Huxley's work 65 years ago, involves only Na^+ and K^+ channels. Subsequently (see Hille, 2001), voltage-gated calcium channels (see Fig. 4.1) were discovered. These function in basically the same way as sodium channels, if on a slightly slower timescale; they contribute to action potential generation in many cells, particularly cardiac and smooth muscle cells, but also in neurons and secretory cells. Ca^{2+} entry

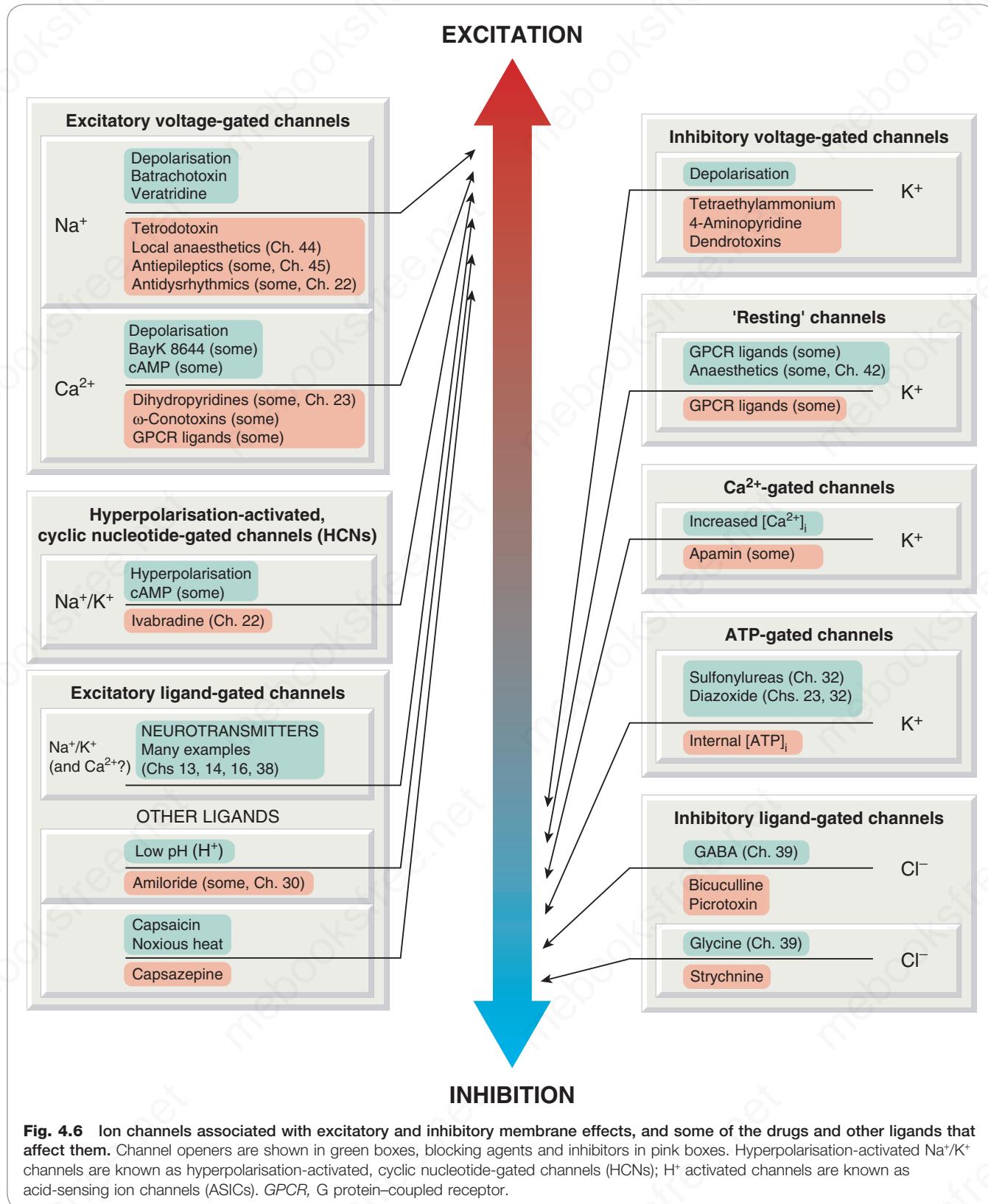


Fig. 4.6 Ion channels associated with excitatory and inhibitory membrane effects, and some of the drugs and other ligands that affect them. Channel openers are shown in green boxes, blocking agents and inhibitors in pink boxes. Hyperpolarisation-activated Na⁺/K⁺ channels are known as hyperpolarisation-activated, cyclic nucleotide-gated channels (HCNs); H⁺ activated channels are known as acid-sensing ion channels (ASICs). GPCR, G protein-coupled receptor.

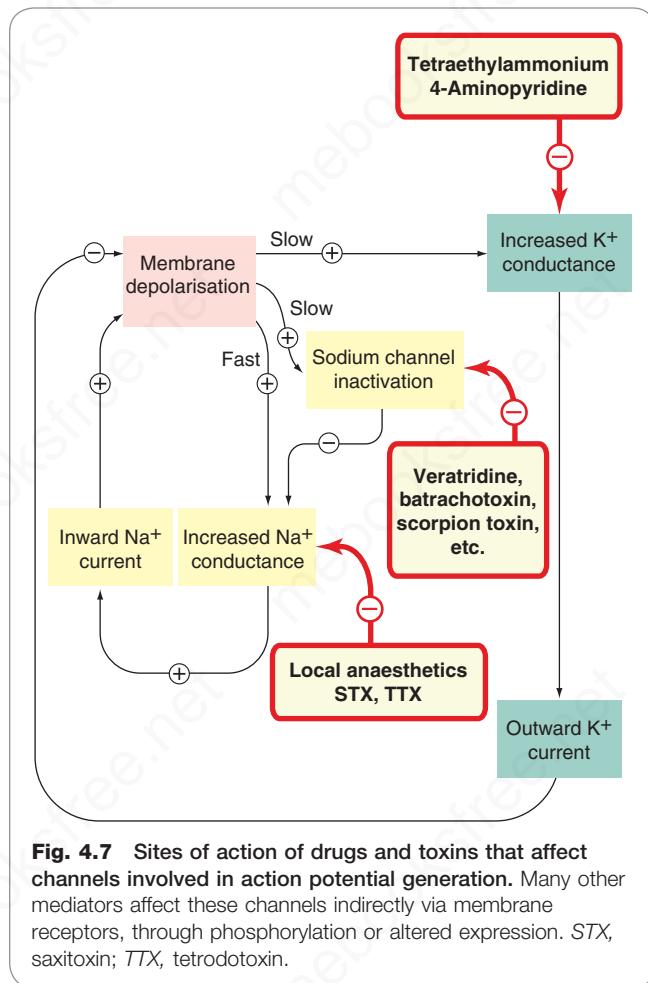


Fig. 4.7 Sites of action of drugs and toxins that affect channels involved in action potential generation. Many other mediators affect these channels indirectly via membrane receptors, through phosphorylation or altered expression. STX, saxitoxin; TTX, tetrodotoxin.

be generated. This type of block is called *use dependent*, because the binding of such drugs increases as a function of the rate of action potential discharge, which governs the rate at which inactivated – and therefore drug-sensitive – channels are generated. This is important for some antidysrhythmic drugs (see Ch. 22) and for antiepileptic drugs (Ch. 46), because high-frequency discharges can be inhibited without affecting excitability at normal frequencies. Drugs that readily block sodium channels in their resting state (e.g. local anaesthetics, Ch. 44) prevent excitation at low as well as high frequencies.

Most sodium channel-blocking drugs are cationic at physiological pH and are therefore affected by the voltage gradient across the cell membrane. They block the channel from the inside, so that their blocking action is favoured by depolarisation. This phenomenon, known as *voltage dependence*, is also of relevance to the action of antidysrhythmic and antiepileptic drugs, because the cells that are the seat of dysrhythmias or seizure activity are generally somewhat depolarised and therefore more strongly blocked than healthy cells. Similar considerations apply also to drugs that block potassium or calcium channels, but we know less about the importance of use and voltage dependence for these than we do for sodium channels.

SODIUM CHANNELS

In most excitable cells, the regenerative inward current that initiates the action potential results from activation of voltage-gated sodium channels. The early voltage clamp studies by Hodgkin and Huxley on the squid giant axon, described on p. 57, revealed the essential functional properties of these channels. Later, advantage was taken of the potent and highly selective blocking action of **tetrodotoxin**

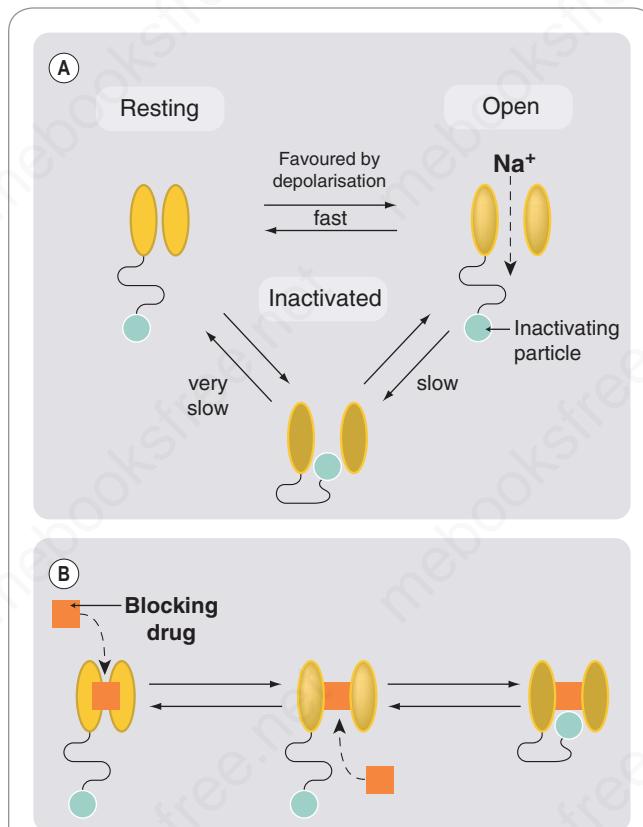


Fig. 4.8 Resting, activated and inactivated states of voltage-gated channels, exemplified by the sodium channel. (A) Membrane depolarisation causes a rapid transition from the resting (closed) state to the open state. The inactivating particle (part of the intracellular domain of the channel protein) is then able to block the channel. With prolonged depolarisation below the threshold for opening, channels can go directly from resting to inactivated without opening. (B) Some blocking drugs (such as tetrodotoxin) block the channel from the outside like a plug, whereas others (such as local anaesthetics and antiepileptic drugs) enter from the inside of the cell and often show preference for the open or inactivated states, and thus affect the kinetic behaviour of the channels, with implications for their clinical application.

(TTX, see Ch. 44) to label and purify the channel proteins, and subsequently to clone them. Sodium channels consist of a central, pore-forming α subunit (shown in Fig. 3.20) and two auxiliary β subunits. Nine α -subunits ($\text{Na}_v1.1$ through $\text{Na}_v1.9$) and four β subunits have been identified in mammals. The α subunits contain four similar domains, each comprising six membrane-spanning helices (see Caterall & Swanson, 2015). One of these helices, S4, which contains several basic amino acids and forms the voltage sensor, moves outwards, thus opening the channel, when the membrane is depolarised. One of the intracellular loops is designed to swing across and block the channel when S4 is displaced, thus inactivating the channel.

It was known from physiological studies that the sodium channels of heart and skeletal muscle differ in various ways from those of neurons. In particular, cardiac sodium channels (and those of some sensory neurons) are relatively insensitive to TTX and slower in their kinetics, compared with

most neuronal sodium channels. This is explained by the relative insensitivity of some α subunits ($\text{Na}_v1.5$, $\text{Na}_v1.8$ and $\text{Na}_v1.9$) to tetrodotoxin. Changes in the level of expression of some sodium channel subunits is thought to underlie the hyperexcitability of sensory neurons in different types of neuropathic pain (see Ch. 43).

In addition to channel-blocking compounds such as tetrodotoxin, other compounds affect sodium channel gating. For example, the plant alkaloid **veratridine** and the frog skin poison **batrachotoxin** cause persistent activation, while various scorpion toxins prevent inactivation, mechanisms resulting in enhanced neuronal excitability.

POTASSIUM CHANNELS

In a typical resting cell (see p. 57, Fig. 4.3), the membrane is selectively permeable to K^+ and the membrane potential (about -60 mV) is somewhat positive to the K^+ equilibrium (about -90 mV). This resting permeability comes about because some potassium channels are open. If more potassium channels open, the membrane hyperpolarises and the cell is inhibited, whereas the opposite happens if potassium channels close. As well as affecting excitability in this way, potassium channels also play an important role in regulating the duration of the action potential and the temporal patterning of action potential discharges; altogether, these channels play a central role in regulating cell function. As mentioned in Chapter 3, the number and variety of potassium channel subtypes is extraordinary, implying that evolution has been driven by the scope for biological advantage to be gained from subtle variations in the functional properties of these channels. A recent résumé lists over 60 different pore-forming subunits, plus another 20 or so auxiliary subunits. An impressive evolutionary display, maybe, but hard going for most of us.

▼ Potassium channels fall into three main classes (Table 4.2),⁴ of which the structures are shown in Fig. 3.20.

- *Voltage-gated potassium channels*, which possess six membrane-spanning helices, one of which serves as the voltage sensor, causing the channel to open when the membrane is depolarised. Included in this group are channels of the shaker family, accounting for most of the voltage-gated K^+ currents familiar to electrophysiologists, and others such as Ca^{2+} -activated potassium channels and two subtypes that are important in the heart, HERG and LQT channels. Many of these channels are blocked by drugs such as **tetraethylammonium** and **4-aminopyridine**.
- *Inwardly rectifying potassium channels*, so called because they allow K^+ to pass inwards much more readily than outwards. These have two membrane-spanning helices and a single pore-forming loop (P loop). These channels are regulated by interaction with G proteins (see Ch. 3) and mediate the inhibitory effects of many agonists acting on G protein-coupled receptors. Certain types are important in the heart, particularly in regulating the duration of the cardiac action potential (Ch. 22); others are the target for the action of **sulfonylureas** (antidiabetic drugs that stimulate insulin secretion by blocking

⁴Potassium channel terminology is confusing, to put it mildly. Electrophysiologists have named K^+ currents prosaically on the basis of their functional properties (I_{KV} , I_{KCav} , I_{KATP} , I_{KIR} , etc.); geneticists have named genes somewhat fancifully according to the phenotypes associated with mutations ('shaker', 'ether-a-go-go', etc.), while molecular biologists have introduced a rational but unmemorable nomenclature on the basis of sequence data (KCNK, KCNQ, etc., with numerical suffixes). The rest of us must make what we can of the unlovely jargon of labels such as HERG (which – don't blink – stands for 'Human Ether-a-go-go Related Gene'), TWIK, TREK and TASK.

them; see Ch. 32) and smooth muscle relaxant drugs, such as **minoxidil** and **diazoxide**, which open them (see Ch. 23).

- *Two-pore domain potassium channels*, with four helices and two P loops. These show outward rectification and therefore exert a strong repolarising influence, opposing any tendency to excitation. They may contribute to the resting K^+ conductance in many cells and are susceptible to regulation via G proteins; certain subtypes have been implicated in the action of volatile anaesthetics such as **isoflurane** (Ch. 42).

For more details, and information on potassium channels and the various drugs and toxins that affect them, see [Jenkinson \(2006\)](#) and [Alexander et al. \(2015\)](#).

Inherited abnormalities of potassium channels (channelopathies) contribute to a rapidly growing number of cardiac, neurological and other diseases. These include the *long QT syndrome* associated with mutations in cardiac voltage-gated potassium channels, causing episodes of ventricular arrest that can result in sudden death. Drug-induced prolongation of the *QT* interval is an unwanted side effect of several drugs (see Ch. 22), including **methadone** and various antipsychotic agents. Nowadays, new drugs are screened for this property at an early stage in the development process (see Ch. 60). Certain familial types of deafness and epilepsy are associated with mutations in voltage-gated potassium channels ([Imbrici et al., 2016](#)).

Ion channels and electrical excitability



- Excitable cells generate an all-or-nothing action potential in response to membrane depolarisation. This occurs in most neurons and muscle cells, and in some gland cells. The ionic basis and time course of the response varies between tissues.
- The regenerative response results from the depolarising current associated with opening of voltage-gated cation channels (mainly Na^+ and Ca^{2+}). It is terminated by inactivation of these channels accompanied by opening of K^+ channels.
- These voltage-gated channels exist in many molecular varieties, with specific functions in different types of cell.
- The membrane of the 'resting' cell is relatively permeable to K^+ but impermeable to Na^+ and Ca^{2+} . Drugs or mediators that open K^+ channels reduce membrane excitability, as do inhibitors of Na^+ or Ca^{2+} channel function. Blocking K^+ channels or activating Na^+ or Ca^{2+} channels increases excitability.
- Cardiac muscle cells, some neurons and some smooth muscle cells generate spontaneous action potentials whose amplitude, rate and rhythm are affected by drugs that affect ion channel function.

MUSCLE CONTRACTION

Effects of drugs on the contractile machinery of smooth muscle are the basis of many therapeutic applications, for smooth muscle is an important component of most physiological systems, including blood vessels and the gastrointestinal, respiratory and urinary tracts. For many decades, smooth muscle pharmacology with its trademark

Table 4.2 Types and functions of K⁺ channels

Structural class ^a	Functional subtypes ^b	Functions	Drug effects	Notes
Voltage-gated (6T, 1P)	Voltage-gated K ⁺ channels	Action potential repolarisation Limits maximum firing frequency	Blocked by tetraethylammonium, 4-aminopyridine Certain subtypes blocked by dendrotoxins (from mamba snake venom)	Subtypes in the heart include HERG and LQT channels, which are involved in congenital and drug-induced dysrhythmias Other subtypes may be involved in inherited forms of epilepsy
	Ca ²⁺ -activated K ⁺ channels	Inhibition following stimuli which increase [Ca ²⁺] _i	Certain subtypes blocked by apamin (from bee venom), and charybdotoxin (from scorpion venom)	
Inward rectifying (2T, 1P)	G protein-activated	Mediate effects of Gi/Go-coupled GPCRs which cause inhibition by increasing K ⁺ conductance	GPCR agonists and antagonists Some are blocked by tertiapin (from honey bee venom)	Other inward rectifying K ⁺ channels important in kidney
	ATP-sensitive	Found in many cells Channels open when [ATP] is low, causing inhibition Important in control of insulin secretion in the pancreas	Association of one subtype with the sulfonylurea receptor (SUR) results in modulation by sulfonylureas (e.g. glibenclamide) which close channel, and by K ⁺ channel openers (e.g. diazoxide, minoxidil) which relax smooth muscle	
Two-pore domain (4T, 2P)	Several subtypes identified (TWIK, TRAAK, TREK, TASK, etc.)	Most are voltage-insensitive; some are normally open and contribute to the 'resting' K ⁺ conductance Modulated by GPCRs	Certain subtypes are activated by volatile anaesthetics (e.g. isoflurane) No selective blocking agents	The nomenclature is misleading, especially when they are incorrectly referred to as two-pore channels

^aK⁺ channel structures (see Fig. 3.20) are defined according to the number of transmembrane helices (T) and the number of pore-forming loops (P) in each α subunit. Functional channels contain several subunits (often four) which may be identical or different, and they are often associated with accessory (β) subunits.

^bWithin each functional subtype, several molecular variants have been identified, often restricted to particular cells and tissues. The physiological and pharmacological significance of this heterogeneity is not yet understood.

technology – the isolated organ bath – held the centre of the pharmacological stage, and neither the subject nor the technology shows any sign of flagging, even though the stage has become much more crowded. Cardiac and skeletal muscle contractility are also the targets of important drug effects.

Although in each case the basic molecular basis of contraction is similar, namely an interaction between actin and myosin, fuelled by ATP and initiated by an increase in [Ca²⁺]_i, there are differences between these three kinds of muscle that account for their different responsiveness to drugs and chemical mediators.

These differences (Fig. 4.9) involve (a) the linkage between membrane events and increase in [Ca²⁺]_i, and (b) the mechanism by which [Ca²⁺]_i regulates contraction.

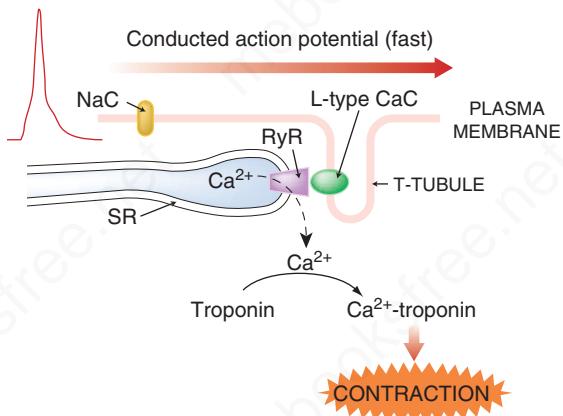
SKELETAL MUSCLE

Skeletal muscle possesses an array of transverse T-tubules extending into the cell from the plasma membrane. The

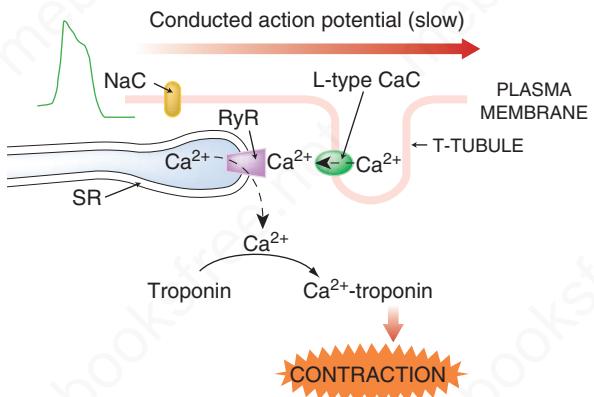
action potential of the plasma membrane depends on voltage-gated sodium channels, as in most nerve cells, and propagates rapidly from its site of origin, the motor endplate (see Ch. 14), to the rest of the fibre. The T-tubule membrane contains voltage-gated calcium channels termed dihydropyridine receptors (DHPRs),⁵ that respond to membrane depolarisation conducted passively along the T-tubule when the plasma membrane is invaded by an action potential. DHPRs are located extremely close to RyRs (see Ch. 3) in the adjacent SR membrane and activation of these RyRs causes release of Ca²⁺ from the SR. Direct coupling between the DHPRs of the T-tubule and the RyRs of the SR (as

⁵Although these are, to all intents and purposes, just a form of L-type calcium channel the term dihydropyridine receptor (DHPR) is used to reflect that they are not identical to the L-type channels in neurons and cardiac muscle.

(A) Skeletal muscle



(B) Cardiac muscle



(C) Smooth muscle

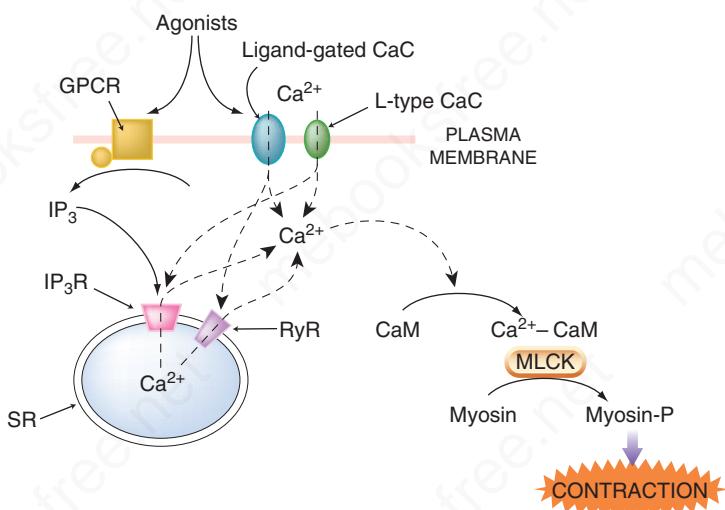


Fig. 4.9 Comparison of excitation–contraction coupling in (A) skeletal muscle, (B) cardiac muscle and (C) smooth muscle.

Skeletal and cardiac muscle differ mainly in the mechanism by which membrane depolarisation is coupled to Ca²⁺ release. The calcium channel (CaC) and ryanodine receptor (RyR) are very closely positioned in both types of muscle. In cardiac muscle, Ca²⁺ entry via voltage-gated calcium channels initiates Ca²⁺ release through activation of the Ca²⁺-sensitive RyRs, whereas in skeletal muscle the sarcolemmal calcium channels activate the RyRs through a voltage-dependent physical interaction. The control of intracellular Ca²⁺ in smooth muscle cells may vary depending upon the type of smooth muscle. In general terms, smooth muscle contraction is largely dependent upon inositol trisphosphate (IP₃)-induced Ca²⁺ release from SR stores through IP₃ receptors (IP₃R). Smooth muscle contraction can also be produced either by Ca²⁺ entry through voltage- or ligand-gated calcium channels. The mechanism by which Ca²⁺ activates contraction is different, and operates more slowly, in smooth muscle compared with in skeletal or cardiac muscle. CaC, calcium channel; CaM, calmodulin; GPCR, G protein-coupled receptor; MLCK, myosin light-chain kinase; NaC, voltage-gated sodium channel; RyR, ryanodine receptor; SR, sarcoplasmic reticulum.

shown in Fig. 4.9) causes the opening of the RyRs on membrane depolarisation. Through this link, depolarisation rapidly activates the RyRs, releasing a short puff of Ca²⁺ from the SR into the sarcoplasm. The Ca²⁺ binds to troponin, a protein that normally blocks the interaction between actin and myosin. When Ca²⁺ binds, troponin moves out of the

way and allows the contractile machinery to operate. Ca²⁺ release is rapid and brief, and the muscle responds with a short-lasting ‘twitch’ response. This is a relatively fast and direct mechanism compared with the arrangement in cardiac and smooth muscle (see later), and consequently less susceptible to pharmacological modulation.

CARDIAC MUSCLE

Cardiac muscle differs from skeletal muscle in several important respects. The nature of the cardiac action potential, the ionic mechanisms underlying its inherent rhythmicity and the effects of drugs on the rate and rhythm of the heart are described in Chapter 22. The cardiac action potential varies in its configuration in different parts of the heart, but commonly shows a plateau lasting several hundred milliseconds following the initial rapid depolarisation. T-tubules in cardiac muscle contain L-type calcium channels, which open during this plateau and allow Ca^{2+} to enter. This Ca^{2+} entry acts on RyRs (a different molecular type from those of skeletal muscle) to release Ca^{2+} from the SR (see Fig. 4.9). With minor differences, the subsequent mechanism by which Ca^{2+} activates the contractile machinery is the same as in skeletal muscle. Ca^{2+} -induced Ca^{2+} release via RyRs may play a role in some forms of cardiac arrhythmia. Mutations of RyRs are implicated in various disorders of skeletal and cardiac muscle function (see Priori & Napolitano, 2005).

SMOOTH MUSCLE

The properties of smooth muscle vary considerably in different organs, and the mechanisms linking membrane events and contraction are correspondingly variable and more complex than in other kinds of muscle. Spontaneous rhythmic activity occurs in many organs, by mechanisms producing oscillations of $[\text{Ca}^{2+}]_i$ (see Fig. 4.2B). The action potential of smooth muscle is generally a rather lazy and vague affair compared with the more military behaviour of skeletal and cardiac muscle, and it propagates through the tissue much more slowly and uncertainly. The action potential is, in most cases, generated by L-type calcium channels rather than by voltage-gated sodium channels, and this is one important route of Ca^{2+} entry. In addition, many smooth muscle cells possess P2X receptors, ligand-gated cation channels, which allow Ca^{2+} entry when activated by ATP released from autonomic nerves (see Ch. 13). Smooth muscle cells also store Ca^{2+} in the ER, from which it can be released when the IP₃R is activated (see Ch. 3). IP₃ is generated by activation of many types of G protein-coupled receptor. Thus in contrast to skeletal and cardiac muscle, Ca^{2+} release and contraction can occur in smooth muscle when such receptors are activated without necessarily involving depolarisation and Ca^{2+} entry through the plasma membrane. RyRs are also present in many smooth muscle cells and calcium-induced Ca^{2+} release via these channels may play a role in generating muscle contraction (see Fig. 4.9) or couple to plasma membrane calcium-activated K⁺ channels resulting in cell hyperpolarisation, thereby reducing Ca^{2+} entry through voltage-gated calcium channels (Fig. 4.10).

The contractile machinery of smooth muscle is activated when the *myosin light chain* undergoes phosphorylation, causing it to become detached from the actin filaments. This phosphorylation is catalysed by a kinase, *myosin light-chain kinase* (MLCK), which is activated when it binds to Ca^{2+} -calmodulin (see p. 56, Fig. 4.9). A second enzyme, *myosin phosphatase*, reverses the phosphorylation and causes relaxation. The activity of MLCK and myosin phosphatase thus exerts a balanced effect, promoting contraction and relaxation, respectively. Both enzymes are regulated by cyclic nucleotides (cAMP and cGMP; see Ch. 3), and many drugs that cause smooth muscle contraction or relaxation

mediated through G protein-coupled receptors or through guanylyl cyclase-linked receptors act in this way. Fig. 4.10 summarises the main mechanisms by which drugs control smooth muscle contraction. The complexity of these control mechanisms and interactions explains why pharmacologists have been entranced for so long by smooth muscle. Many therapeutic drugs work by contracting or relaxing smooth muscle, particularly those affecting the cardiovascular, respiratory and gastrointestinal systems, as discussed in later chapters, where details of specific drugs and their physiological effects are given.

Muscle contraction



- Muscle contraction occurs in response to a rise in $[\text{Ca}^{2+}]_i$.
- In skeletal muscle, depolarisation causes rapid Ca^{2+} release from the sarcoplasmic reticulum (SR); in cardiac muscle, Ca^{2+} enters through voltage-gated channels, and this initial entry triggers further release from the SR; in smooth muscle, the Ca^{2+} signal is due partly to Ca^{2+} entry and partly to inositol trisphosphate (IP₃)-mediated release from the SR.
- In smooth muscle, contraction can occur without action potentials, for example, when agonists at G protein-coupled receptors lead to IP₃ formation.
- Activation of the contractile machinery in smooth muscle involves phosphorylation of the myosin light chain, a mechanism that is regulated by a variety of second messenger systems.

RELEASE OF CHEMICAL MEDIATORS

Much of pharmacology is based on interference with the body's own chemical mediators, particularly neurotransmitters, hormones and inflammatory mediators. Here we discuss some of the common mechanisms involved in the release of such mediators, and it will come as no surprise that Ca^{2+} plays a central role. Drugs and other agents that affect the various control mechanisms that regulate $[\text{Ca}^{2+}]_i$ will therefore also affect mediator release, and this accounts for many of the physiological effects that they produce.

Chemical mediators that are released from cells fall into two main groups (Fig. 4.11):

- Mediators that are preformed and packaged in storage vesicles – sometimes called storage granules – from which they are released by *exocytosis*. This large group comprises all the conventional neurotransmitters and neuromodulators (see Chs 13 and 38), and many hormones. It also includes secreted proteins such as cytokines and various growth factors (Ch. 19).
- Mediators that are produced on demand and are released by diffusion or by membrane carriers.⁶ This group includes nitric oxide (Ch. 21) and many lipid mediators (e.g. prostaglandins, Ch. 18) and

⁶Carrier-mediated release can also occur with neurotransmitters that are stored in vesicles but is quantitatively less significant than exocytosis.

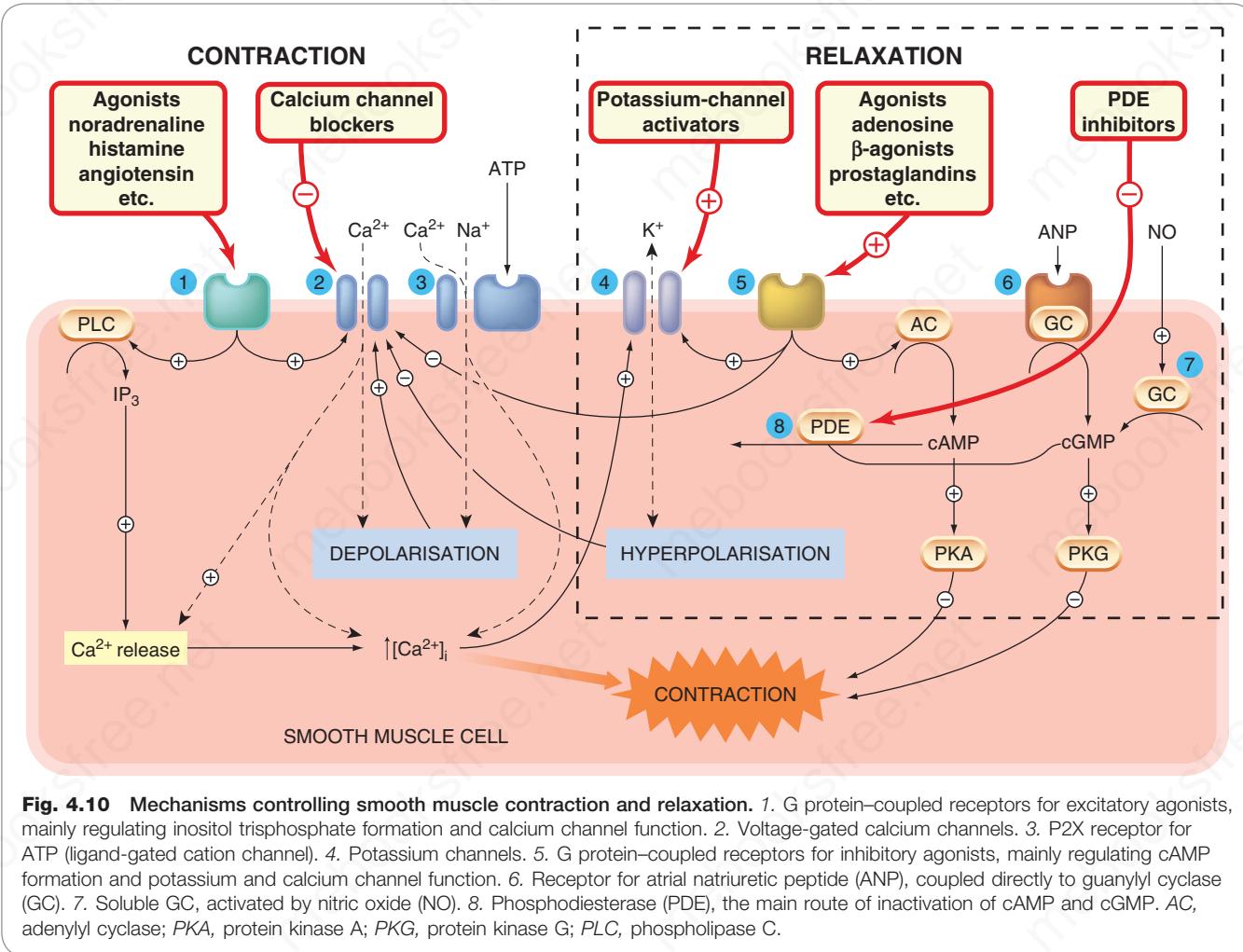


Fig. 4.10 Mechanisms controlling smooth muscle contraction and relaxation. 1. G protein-coupled receptors for excitatory agonists, mainly regulating inositol trisphosphate formation and calcium channel function. 2. Voltage-gated calcium channels. 3. P2X receptor for ATP (ligand-gated cation channel). 4. Potassium channels. 5. G protein-coupled receptors for inhibitory agonists, mainly regulating cAMP formation and potassium and calcium channel function. 6. Receptor for atrial natriuretic peptide (ANP), coupled directly to guanylyl cyclase (GC). 7. Soluble GC, activated by nitric oxide (NO). 8. Phosphodiesterase (PDE), the main route of inactivation of cAMP and cGMP. AC, adenyl cyclase; PKA, protein kinase A; PKG, protein kinase G; PLC, phospholipase C.

endocannabinoids (Ch. 20), which are released from the postsynaptic cell to act retrogradely on nerve terminals.

Calcium ions play a key role in both cases, because a rise in [Ca²⁺]_i initiates exocytosis and is also the main activator of the enzymes responsible for the synthesis of diffusible mediators.

In addition to mediators that are released from cells, some are formed from precursors in the plasma, two important examples being *kinins* (Ch. 19) and *angiotensin* (Ch. 23), which are peptides produced by protease-mediated cleavage of circulating proteins.

EXOCYTOSIS

Exocytosis, occurring in response to an increase of [Ca²⁺]_i, is the principal mechanism of transmitter release (see Fig. 4.11) in the peripheral and central nervous systems, as well as in endocrine cells and mast cells. The secretion of enzymes and other proteins by gastrointestinal and exocrine glands and by vascular endothelial cells is also basically similar. Exocytosis (see Thorn et al., 2016) involves fusion between the membrane of synaptic vesicles and the inner surface of the plasma membrane. The vesicles are preloaded with stored transmitter, and release occurs in discrete packets,

or quanta, each representing the contents of a single vesicle. The first evidence for this came from the work of Katz and his colleagues in the 1950s, who recorded spontaneous ‘miniature endplate potentials’ at the frog neuromuscular junction, and showed that each resulted from the spontaneous release of a packet of the transmitter, acetylcholine. They also showed that release evoked by nerve stimulation occurred by the synchronous release of several hundred such quanta and was highly dependent on the presence of Ca²⁺ in the bathing solution. Unequivocal evidence that the quanta represented vesicles releasing their contents by exocytosis came from electron microscopic studies, in which the tissue was rapidly frozen in mid-release, revealing vesicles in the process of extrusion, and from elegant electrophysiological measurements showing that membrane capacitance (reflecting the area of the presynaptic membrane) increased in a stepwise manner as each vesicle fused and then gradually returned as the vesicle membrane was recovered from the surface. There is also biochemical evidence showing that, in addition to the transmitter, other constituents of the vesicles are released at the same time.

▼ In nerve terminals specialised for fast synaptic transmission, Ca²⁺ enters through voltage-gated calcium channels, mainly of the N and P/Q type (see p. 52 and Table 4.1), and the synaptic vesicles are ‘docked’ at active zones – specialised regions of the presynaptic

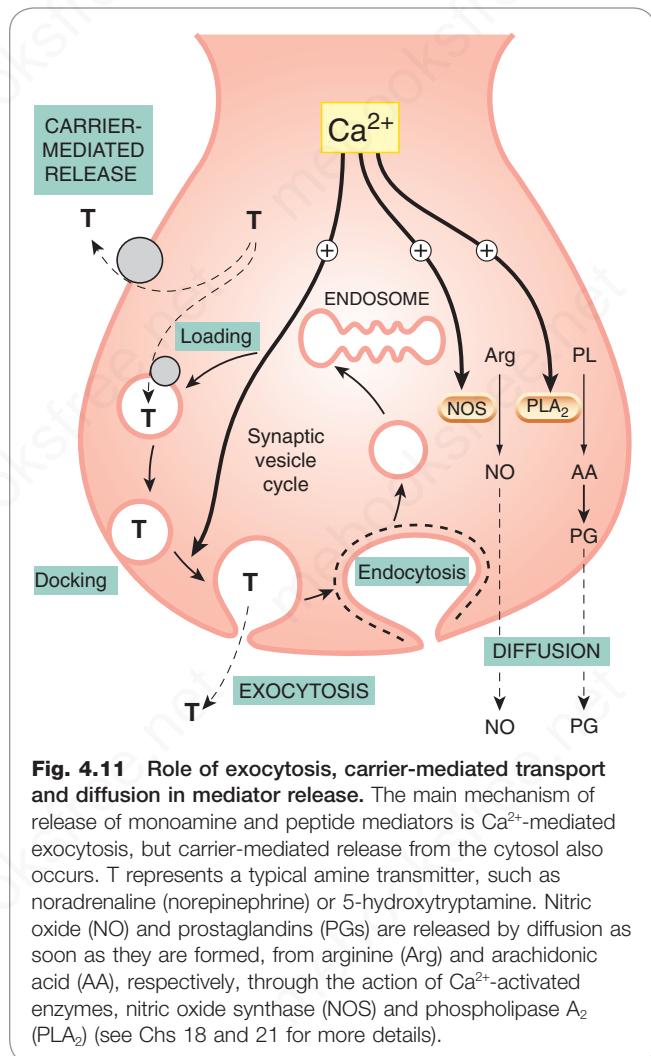


Fig. 4.11 Role of exocytosis, carrier-mediated transport and diffusion in mediator release. The main mechanism of release of monoamine and peptide mediators is Ca^{2+} -mediated exocytosis, but carrier-mediated release from the cytosol also occurs. T represents a typical amine transmitter, such as noradrenaline (norepinephrine) or 5-hydroxytryptamine. Nitric oxide (NO) and prostaglandins (PGs) are released by diffusion as soon as they are formed, from arginine (Arg) and arachidonic acid (AA), respectively, through the action of Ca^{2+} -activated enzymes, nitric oxide synthase (NOS) and phospholipase A₂ (PLA₂) (see Chs 18 and 21 for more details).

membrane from which exocytosis occurs, situated close to the relevant calcium channels and opposite receptor-rich zones of the postsynaptic membrane. Elsewhere, where speed is less critical, Ca^{2+} may come from intracellular stores and the spatial organisation of active zones is less clear. It is common for secretory cells, including neurons, to release more than one mediator (for example, a 'fast' transmitter such as glutamate and a 'slow' transmitter such as a neuropeptide) from different vesicle pools (see Ch. 13). The fast transmitter vesicles are located close to active zones, while the slow transmitter vesicles are further away. Release of the fast transmitter, because of the tight spatial organisation, occurs as soon as the neighbouring calcium channels open, before the Ca^{2+} has a chance to diffuse throughout the terminal, whereas release of the slow transmitter requires the Ca^{2+} to diffuse more widely. As a result, release of fast transmitters occurs impulse by impulse, even at low stimulation frequencies, whereas release of slow transmitters builds up only at higher stimulation frequencies. The release rates of the two therefore depend critically on the frequency and patterning of firing of the presynaptic neuron (Fig. 4.12). In non-excitable cells (e.g. most exocrine and endocrine glands), the slow mechanism predominates and is activated mainly by Ca^{2+} release from intracellular stores.

Calcium causes exocytosis by binding to the vesicle-bound protein *synaptotagmin*, and this favours association between a second vesicle-bound protein, *synaptobrevin*, and a related protein, *synaptotaxin*, on the inner surface of the plasma membrane. This association brings the vesicle membrane into close apposition with the plasma membrane, causing membrane fusion. This group of proteins, known collectively as SNAREs, plays a key role in exocytosis.

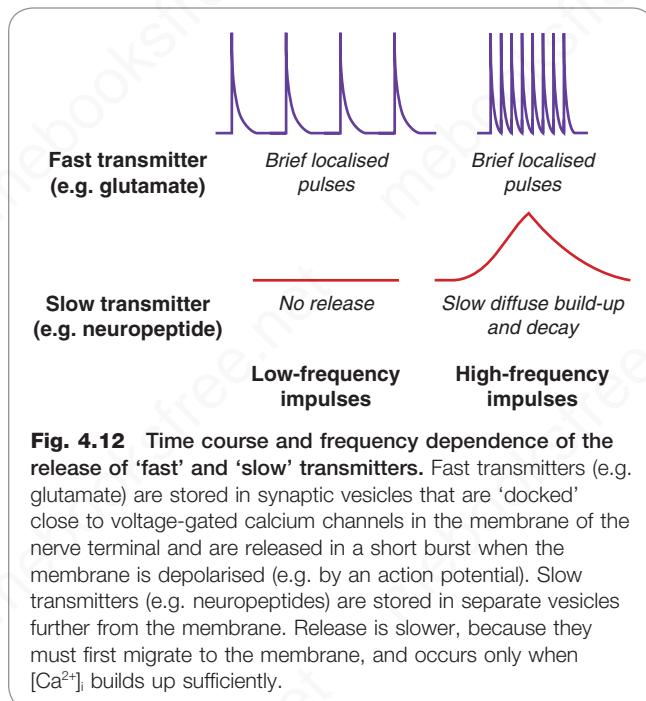


Fig. 4.12 Time course and frequency dependence of the release of 'fast' and 'slow' transmitters. Fast transmitters (e.g. glutamate) are stored in synaptic vesicles that are 'docked' close to voltage-gated calcium channels in the membrane of the nerve terminal and are released in a short burst when the membrane is depolarised (e.g. by an action potential). Slow transmitters (e.g. neuropeptides) are stored in separate vesicles further from the membrane. Release is slower, because they must first migrate to the membrane, and occurs only when $[\text{Ca}^{2+}]_i$ builds up sufficiently.

Having undergone exocytosis, the empty vesicle⁷ is recaptured by endocytosis and returns to the interior of the terminal, where it fuses with the larger endosomal membrane. The endosome buds off new vesicles, which take up transmitter from the cytosol by means of specific transport proteins and are again docked on the presynaptic membrane. This sequence, which typically takes several minutes, is controlled by various trafficking proteins associated with the plasma membrane and the vesicles, as well as cytosolic proteins. So far, there are few examples of drugs that affect transmitter release by interacting with synaptic proteins, although the botulinum neurotoxins (see Ch. 14) produce their effects by proteolytic cleavage of SNARE proteins.

NON-VESICULAR RELEASE MECHANISMS

If this neat and tidy picture of transmitter packets ready and waiting to pop obediently out of the cell in response to a puff of Ca^{2+} seems a little too good to be true, rest assured that the picture is not quite so simple. Acetylcholine, noradrenaline (norepinephrine) and other mediators can leak out of nerve endings from the cytosolic compartment, independently of vesicle fusion, by utilising carriers in the plasma membrane (see Fig. 4.11). Drugs such as **amphetamines**, which release amines from central and peripheral nerve terminals (see Chs 15 and 40), do so by displacing the endogenous amine from storage vesicles into the cytosol, whence it escapes via the monoamine transporter in the plasma membrane, a mechanism that does not depend on Ca^{2+} .⁸

Nitric oxide (see Ch. 21), arachidonic acid metabolites (e.g. prostaglandins; Ch. 18) and endocannabinoids (see

⁷The vesicle contents may not always discharge completely. Instead, vesicles may fuse transiently with the cell membrane and release only part of their contents before becoming disconnected (termed *kiss-and-run exocytosis*).

⁸Some cheeses may have high levels of the trace amino acid tyramine, which can act akin to amphetamines and release noradrenaline (particularly in those being treated with monoamine oxidase (MAO) inhibitors, see Ch. 48) giving a dramatic sympathetic episode known as a 'cheese effect'.

Ch. 20) are important examples of mediators that are released from the cytosol by diffusion across the membrane or by carrier-mediated extrusion, rather than by exocytosis. The mediators are not stored but escape from the cell as soon as they are synthesised. In each case, the synthetic enzyme(s) is activated by Ca^{2+} , and the moment-to-moment control of the rate of synthesis depends on $[\text{Ca}^{2+}]_i$. This kind of release is necessarily slower than the classic exocytic mechanism, but in the case of nitric oxide is fast enough for it to function as a true transmitter (see Fig. 13.5 and Ch. 21).

Mediator release

- Most chemical mediators are packaged into storage vesicles and released by exocytosis. Some are not stored but synthesised on demand and released by diffusion or the operation of membrane carriers.
- Exocytosis occurs in response to increased $[\text{Ca}^{2+}]_i$ as a result of a Ca^{2+} -mediated interaction between proteins of the synaptic vesicle and the plasma membrane, causing the membranes to fuse.
- After releasing their contents, vesicles are recycled and reloaded with transmitter.
- Many secretory cells contain more than one type of vesicle, loaded with different mediators and secreted independently.
- Stored mediators (e.g. neurotransmitters) may be released directly from the cytosol independently of Ca^{2+} and exocytosis, by drugs that interact with membrane transport mechanisms.
- Non-stored mediators, such as prostanooids, nitric oxide and endocannabinoids are released by increased $[\text{Ca}^{2+}]_i$, which activates the enzymes responsible for their synthesis.

EPITHELIAL ION TRANSPORT

Fluid-secreting epithelia include the renal tubule, salivary glands, gastrointestinal tract and airways epithelia. In each case, epithelial cells are arranged in sheets separating the interior (blood-perfused) compartment from the exterior lumen compartment, into which, or from which, secretion takes place. Fluid secretion involves two main mechanisms, which often co-exist in the same cell and indeed interact with each other. The two mechanisms (Fig. 4.13) are concerned, respectively, with Na^+ transport and Cl^- transport.

In the case of Na^+ transport, secretion occurs because Na^+ enters the cell passively at one end and is pumped out actively at the other, with water following passively. Critical to this mechanism is a class of highly regulated epithelial sodium channels (ENaCs) that allow Na^+ entry.

ENaCs (see Hanukoglu & Hanukoglu, 2016) are widely expressed, not only in epithelial cells but also in neurons and other excitable cells, where their function is largely unknown. They are regulated mainly by aldosterone, a hormone produced by the adrenal cortex that enhances Na^+ reabsorption by the kidney (see Ch. 30). Aldosterone, like other steroid hormones, exerts its effects by regulating

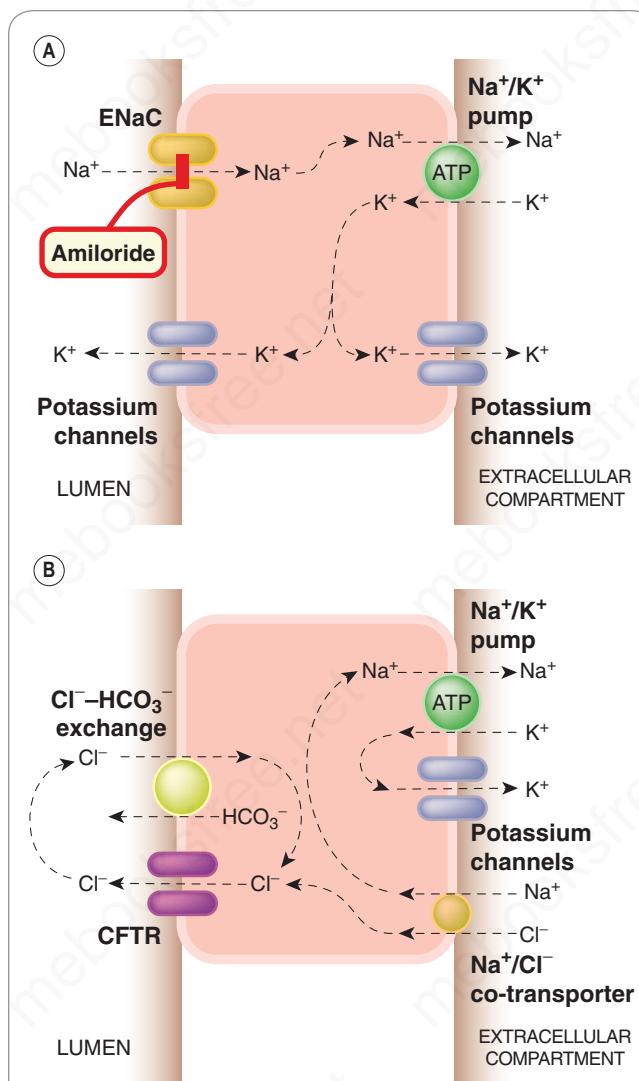


Fig. 4.13 Generalised mechanisms of epithelial ion transport. Such mechanisms are important in renal tubules (see Ch. 30 for more details) and in many other situations, such as the gastrointestinal and respiratory tracts. The exact mechanism may vary from tissue to tissue depending upon channel and pump expression and location. (A) Sodium transport. A special type of epithelial sodium channel (ENaC) controls entry of Na^+ into the cell from the luminal surface, the Na^+ being actively pumped out at the apical surface by the Na^+-K^+ exchange pump. K^+ moves passively via potassium channels. (B) Chloride transport. Cl^- leaves the cell via a special membrane channel, the cystic fibrosis transmembrane conductance regulator (CFTR), after entering the cell either from the apical surface via the Na^+/Cl^- co-transporter, or at the luminal surface via the $\text{Cl}^-/\text{HCO}_3^-$ co-transporter.

gene expression (see Ch. 3), and causes an increase in ENaC expression, thereby increasing the rate of Na^+ and fluid transport. ENaCs are selectively blocked by certain diuretic drugs, notably **amiloride** (see Ch. 30), a compound that is widely used to study the functioning of ENaCs in other situations.

Chloride transport is particularly important in the airways and gastrointestinal tract. In the airways it is essential for

fluid secretion, whereas in the colon it mediates fluid reabsorption, the difference being due to the different arrangement of various transporters and channels with respect to the polarity of the cells. The simplified diagram in Fig. 4.13B represents the situation in the pancreas, where secretion depends on Cl^- transport. The key molecule in Cl^- transport is the *cystic fibrosis transmembrane conductance regulator* (CFTR), so named because early studies on the inherited disorder cystic fibrosis showed it to be associated with impaired Cl^- conductance in the membrane of secretory epithelial cells, and the CFTR gene, identified through painstaking genetic linkage studies and isolated in 1989, was found to encode a Cl^- -conducting ion channel. Severe physiological consequences follow from CFTR mutations and the resulting impairment of secretion, particularly in the airways but also in many other systems, such as sweat glands and pancreas. Studies on the disease-associated mutations of the CFTR gene have revealed much about the molecular mechanisms involved in Cl^- transport (Wang et al., 2014).

Both Na^+ and Cl^- transport are regulated by intracellular messengers, notably by Ca^{2+} and cAMP, the latter exerting its effects by activating protein kinases and thereby causing phosphorylation of channels and transporters. CFTR itself is activated by cAMP. In the gastrointestinal tract, increased cAMP formation causes a large increase in the rate of fluid secretion, an effect that leads to the copious diarrhoea produced by cholera infection (see Ch. 3) and by inflammatory conditions in which prostaglandin formation

is increased (see Ch. 18). Activation of G protein-coupled receptors, which cause release of Ca^{2+} , also stimulates secretion, possibly also by activating CFTR. Many examples of therapeutic drugs that affect epithelial secretion by activating or blocking G protein-coupled receptors appear in later chapters.

Epithelial ion transport



- Many epithelia (e.g. renal tubules, exocrine glands and airways) are specialised to transport specific ions.
- This type of transport depends on a special class of epithelial sodium channels (ENaCs) which allow Na^+ entry into the cell at one surface, coupled to active extrusion of Na^+ , or exchange for another ion, from the opposite surface.
- Anion transport depends on a specific chloride channel (the cystic fibrosis transmembrane conductance regulator [CFTR]), mutations of which result in cystic fibrosis.
- The activity of channels, pumps and exchange transporters is regulated by various second messengers and nuclear receptors, which control the transport of ions in specific ways.

REFERENCES AND FURTHER READING

General references

- Alexander, S.P., Catterall, W.A., Kelly, E., et al., 2015. The concise guide to pharmacology 2015/16: Voltage-gated ion channels. *Br. J. Pharmacol.* 172, 5904–5941. (*Contains a brief description of a range of ion channels and the drugs that interact with them*)
- Berridge, M.J., 2014. Cell Signalling Biology. Portland Press. doi:10.1042/csb0001002. (*Free ebook available online at www.cellsignallingbiology.org; a regularly updated resource that covers various aspects of cell signalling in a highly readable format*)
- Kandel, E.R., Schwartz, J.H., Jessel, T.M., Siegelbaum, S.A., Hudspeth, A.J., 2012. Principles of Neural Science. McGraw-Hill, New York. (*Excellent, well-written textbook of neuroscience*)
- Katz, B., 1966. Nerve, Muscle and Synapse. McGraw-Hill, New York. (*A classic account of the ground-breaking electrophysiological experiments that established the basis of nerve and muscle function*)
- Berridge, M.J., 2016. The inositol trisphosphate/calcium signaling pathway in health and disease. *Physiol. Rev.* 96, 1261–1296. (*Clear and readable up-to-date account of the mechanisms and versatility of calcium signaling and how it may be altered in disease states*)
- Morgan, A.J., Davis, L.C., Ruas, M., Galione, A., 2015. TPC: the NAADP discovery channel? *Biochem. Soc. Trans.* 43, 384–389. (*Interesting discussion of this controversial topic from the laboratory of one of the main protagonists*)
- Parrington, J., Lear, P., Hachem, A., 2015. Calcium signals regulated by NAADP and two-pore channels—their role in development, differentiation and cancer. *Int. J. Dev. Biol.* 59, 341–355. (*Review of the potential pathophysiological importance*)
- Prakriya, M., Lewis, R.S., 2015. Store-operated calcium channels. *Physiol. Rev.* 95, 1383–1436. (*Comprehensive review of this topic*)

Excitation and ion channels

- Catterall, W.A., Swanson, T.M., 2015. Structural basis for pharmacology of voltage-gated sodium and calcium channels. *Mol. Pharmacol.* 88, 141–150. (*Useful, authoritative review article*)

- Hille, B., 2001. Ionic Channels of Excitable Membranes. Sinauer Associates, Sunderland. (*Still probably the best, clear and detailed account of the basic principles of ion channels, with emphasis on their biophysical properties*)
- Imbrici, P., Liantonio, A., Camerino, G.M., et al., 2016. Therapeutic approaches to genetic ion channelopathies and perspectives in drug discovery. *Front. Pharmacol.* 7, 121. (*Detailed review of the importance of channelopathies in multiple disorders*)

- Jenkinson, D.H., 2006. Potassium channels – multiplicity and challenges. *Br. J. Pharmacol.* 147 (Suppl. 1), 63–71. (*Useful short article on the many types of K^+ channel*)

Muscle contraction

- Berridge, M.J., 2008. Smooth muscle cell calcium activation mechanisms. *J. Physiol.* 586, 5047–5061. (*Excellent review article describing the various mechanisms by which calcium signals control activity in different types of smooth muscle – complicated but clear*)
- Priori, S.G., Napolitano, C., 2005. Cardiac and skeletal muscle disorders caused by mutations in the intracellular Ca^{2+} release channels. *J. Clin. Invest.* 115, 2033–2038. (*Focuses on RyR mutations in various inherited diseases*)
- Van Petegem, F., 2012. Ryanodine receptors: structure and function. *J. Biol. Chem.* 287 (31), 31624–31632.

Secretion and exocytosis

- Hanukoglu, I., Hanukoglu, A., 2016. Epithelial sodium channel (ENaC) family: phylogeny, structure-function, tissue distribution, and associated inherited diseases. *Gene* 579, 95–132. (*Extensive review of the function and importance of these channels in health and disease*)
- Thorn, P., Zorec, R., Rettig, J., Keating, D.J., 2016. Exocytosis in non-neuronal cells. *J. Neurochem.* 137, 849–859.
- Wang, Y., Wrennall, J.A., Cai, Z., Li, H., Sheppard, D.N., 2014. Understanding how cystic fibrosis mutations disrupt CFTR function: from single molecules to animal models. *Int. J. Biochem. Cell Biol.* 52, 47–57.

How drugs act: biopharmaceuticals and gene therapy

OVERVIEW

In this chapter, we discuss the properties of a group of therapeutic agents known collectively as **biopharmaceuticals**. These are relatively recent additions to our therapeutic armoury, but they have already made a major impact on treatment of diseases such as rheumatoid arthritis and cancer. The number of biopharmaceuticals approved for clinical use is growing and the sector will assume more significance in the future. In this chapter, we first introduce protein- and oligonucleotide-based biopharmaceuticals, highlight the major differences with 'conventional' small molecule drugs and explain how they are manufactured, how they work and how they are metabolised. We then introduce the central concepts of **gene therapy**, discuss the promise and problems associated with this therapeutic modality and highlight some limited successes.

INTRODUCTION

This chapter deals with the general pharmacological characteristics of protein and nucleic acid-based pharmaceuticals produced using genetic engineering techniques (i.e. biotechnology, as distinct from synthetic chemistry). Such agents currently account for 20%–25% of new registrations, and are increasingly important therapeutically. Individual agents are discussed in later chapters.

▼ Annoyingly for authors of textbooks and their readers, there is no consensus on what actually constitutes a 'biopharmaceutical' as opposed to a conventional drug. An apparently obvious distinguishing feature is whether it is predominately 'chemical' in nature (like almost all the small molecule drugs in this book) or of 'biological' origin (like insulin or growth hormone, for example). Unfortunately, this simplistic distinction breaks down rather quickly when we consider that some 'small molecule' drugs (such as **morphine** or **penicillin**) are plant or fungal products whilst other 'biological molecules' such as short peptides or antisense oligonucleotides are produced by synthetic organic chemistry techniques.

A further problem is the stance adopted by the main drug regulatory agencies. The FDA and their European counterparts use slightly different definitions when classifying 'biopharmaceuticals' and this has profound effects on the companies that manufacture them, affecting their regulatory obligations, their business models, patent filings, investment funding and even public relations. As one commentator (Rader, 2008) put it, 'The result is a Babel-like situation with terminological chaos and anarchy confounding communication, comparative and industry analyses, understanding and regulation'.

To simplify the situation, we will adopt a largely pragmatic definition in this chapter; there is no doubt that biopharmaceuticals¹ are different in many respects from conventional small molecule drugs

(including their pharmacology) and we will use these as our criteria for distinguishing between them. We will begin with a discussion of protein and oligonucleotide biopharmaceuticals.

PROTEIN AND OLIGONUCLEOTIDE BIOPHARMACEUTICALS

The use of proteins as therapeutic agents is not a novel idea; insulin, extracted from animal pancreas tissue (Ch. 32), and human growth hormone, extracted at one time from human cadaver pituitary glands (Ch. 34), were among the first therapeutic proteins to be used and, for many years, such purified extracts provided the only option for treating protein hormone deficiency disorders. However, there were problems. Technical difficulties in extraction of the hormone from tissue often led to disappointing yields. Administration of animal hormones (e.g. pig insulin) to humans could evoke an immune response and there was another insidious danger – transmission of infectious agents across species or between people. This was highlighted in the 1970s, when cases of *Creutzfeldt-Jakob disease* (see Ch. 41) were seen in patients treated with human growth hormone obtained from cadavers. This serious problem was later traced to contamination of the donor pituitary glands with infectious *prions* (Ch. 41). The advent of 'genetic engineering' techniques offered a new way to deal with these perennial problems.

Biopharmaceuticals and gene therapy: definition and potential uses



- **Biopharmaceuticals** include proteins and antibodies (and oligonucleotides) used as drugs:
 - First-generation biopharmaceuticals are mainly copies of endogenous proteins or antibodies, produced by recombinant DNA technology.
 - Second-generation biopharmaceuticals have been 'engineered' to improve the performance of the protein, antibody or antisense agent.
- **Applications:**
 - therapeutic monoclonal antibodies
 - recombinant hormones
 - controlling gene expression (oligonucleotides)
- **Gene therapy** is the addition of genetic material to cells to prevent, alleviate or cure disease.
- **Potential applications:**
 - radical cure of monogenic diseases (e.g. cystic fibrosis, haemoglobinopathies);
 - amelioration of diseases with or without a genetic component, including many malignant, neurodegenerative and infectious diseases.

¹There is a tiresome terminological issue here too: such drugs are often known simply as 'biologics', but this term is also used to refer to *any* biological reagents (e.g. antibody-based laboratory tests, plasma expanders and so on).

PROTEINS AND POLYPEPTIDES

The biopharmaceuticals in use today are generally classified as first- or second-generation agents. *First-generation* biopharmaceuticals are usually straightforward copies of human hormones or other proteins, prepared by *transfected* the human gene into a suitable *expression system* (a cell line that produces the protein in good yield), harvesting and purifying the *recombinant protein* for use as a drug. The first agent to be produced in this way was recombinant human insulin in 1982.

Second-generation biopharmaceuticals are those that have been *engineered*; that is to say, either the gene has been deliberately altered prior to transfection such that the structure of the recombinant protein is changed, or some alteration is made to the purified end product. Such changes are generally made to improve some aspect of the protein's activity profile. Human recombinant insulins designed to act faster or last longer were among the first in this class to be marketed; **Table 5.1** contains other examples.

Production methods

▼ There are several problems associated with the manufacture of any type of recombinant protein, and one of the most pressing is the choice of expression system. Many recombinant proteins are expressed in bacterial systems (*Escherichia coli*, for example), which is useful because cultures grow quickly and are generally easy to manipulate. Disadvantages include the fact that the product may contain bacterial endotoxins, which must be removed before administration to patients, and also that bacterial cells differ from mammalian cells in patterns of *post-translational processing* (e.g. glycosylation) of proteins, which may affect the protein's biological action. To circumvent these problems, mammalian (e.g. Chinese hamster ovary [CHO]) cells can also be used as expression systems, although such cells require more careful culture, grow more slowly than bacteria and produce less product, all of which contributes to the cost of the final medicine.

A number of emergent technologies are set to transform the production process. The use of plants to produce recombinant proteins has attracted considerable interest (see [Melnik & Stoger, 2013](#)). Several species have shown promise, including the tobacco plant. Human genes of interest can readily be transfected into the plant using tobacco mosaic virus as a vector; the crop grows rapidly (yields a high *biomass*) and offers a number of other advantages. Edible plants

Table 5.1 Some examples of biopharmaceuticals

Class	Type	Biopharmaceutical	Change	Target	Indication	Reason for change
First generation	Protein	Human insulin	None	Insulin receptor	Diabetes	N/A
	Protein	Human growth hormone	None	Growth hormone receptor (agonist)	Pituitary dwarfism, Turner's syndrome	N/A
Second generation	Protein	Insulin	AA sequence	Insulin receptor	Diabetes	Faster acting hormone
	Protein	Interferon analogue	AA sequence	Viral replication	Viral infection	Superior antiviral activity
	Protein	Glucocerebrosidase enzyme	Carbohydrate residue	Glucocerebrosides	Gaucher's disease	Promotes phagocytic uptake
	Protein	Erythropoietin analogue	Carbohydrate residue	Erythropoietin receptor	Anaemia	Prolongs half-life
	Protein	Human growth hormone	AA sequence, prosthetic group	Growth hormone receptor (antagonist)	Acromegaly	Converts agonist into antagonist with long duration of action
	Protein	Adalimumab ^a	Humanised mAb	Tumour necrosis factor	Rheumatoid disease	Persists in circulation
	Protein	Omalizumab	Humanised mAb	IgE	IgE-mediated asthma	Persists in circulation
	Antisense oligonucleotide	Mipomersin	Modified nucleotides	Apolipoprotein B gene	Familial hypercholesterolaemia	Stability
	Antisense oligonucleotide	Eteplirsen	Modified nucleotides	Dystrophin gene	Duchenne muscular dystrophy	Stability
	Antisense oligonucleotide	Nusinersen	Modified nucleotides	Survival motor neuron protein (SMN 1)	Spinal muscular atrophy	Stability

^aTherapeutic monoclonal antibody names all end in '-mab', prefixed by an indication of their species nature: -umab (human), -omab (mouse), -ximab (chimera), -zumab (humanised).

AA, amino acids; IgE, immunoglobulin E; mAb, monoclonal antibody.

(Source: [Walsh, 2004](#); The British National Formulary; and others.)

such as lettuce and bananas could be used to produce some orally active proteins, such as vaccines, which could then be consumed directly without the need for prior purification. Several such proteins have already been produced in plants, and have entered clinical trials (Kwon et al., 2013).

Another technology that could dramatically increase the yield of human recombinant proteins is the use of transgenic cattle. A dairy cow can produce some 10,000 litres of milk per year, and recombinant proteins introduced into the genome and under the control of promoters that regulate production of other milk proteins, can generate yields as high as 1 g/L (see Brink et al., 2000).

Engineered proteins

There are several ways in which proteins can be altered prior to expression. Alteration of the nucleotide sequence of the coding gene can be used to change single amino acids or, indeed, whole regions of the polypeptide chain. There are good reasons why it may be an advantage to engineer proteins in this way. These include:

- modification of pharmacokinetic properties
- creation of novel *fusion* or other proteins
- reducing immunogenicity, for example, by *humanising* the protein

It is often useful to modify the pharmacokinetic properties of recombinant proteins. Changes in the structure of human insulin, for example, provided a form of the hormone that did not self-associate during storage and was thus faster acting and easier to manage. The half-life of proteins in the blood can often be extended by *PEGylation* (see Ch. 11), the addition of polyethylene glycol to the molecule. This *post-translational engineering* approach has been applied to some human hormones, such as recombinant growth hormone, interferons and others. This is not merely a convenience to patients; it also reduces the overall cost of the treatment, an important factor in the adoption of this type of therapy.

Fusion proteins comprise two or more proteins engineered to be expressed as one single polypeptide chain, sometimes joined by a short linker. An example is **etanercept**, an anti-inflammatory drug used in the treatment of rheumatoid arthritis and other conditions (see Ch. 27). Etanercept consists of the ligand-binding domain taken from the tumour necrosis factor (TNF) receptor, joined to the Fc domain of a human immunoglobulin G antibody. The receptor moiety sequesters endogenous TNF ligand in a complexed inactive form, while the immunoglobulin increases persistence of the drug in the blood. Reduction of immunogenicity through bioengineering is discussed later.

MONOCLONAL ANTIBODIES

Although antisera preparations can be used to confer *passive immunity*, there are a number of inherent disadvantages that limit their utility. Conventionally, antisera are produced from the blood of immunised humans or animals. Antisera containing high levels of specific antibodies (e.g. to tetanus toxin or snake venom) is prepared from the plasma and this can then be used therapeutically to neutralise pathogens or other dangerous substances in the blood of the patient.

Such preparations are comprised of *polyclonal antibodies* – that is, a *polyvalent* mixture of antibodies from all the plasma cell clones that reacted to that particular antigen. The actual composition and efficacy of these varies over time, and obviously there is a limit to how much plasma can be collected on any one occasion. However, in 1975,

Milstein and Köhler² discovered a method of producing from immunised mice an immortalised *hybridoma*, a fusion of one particular lymphocytic clone with an immortalised tumour cell. This furnished a method of producing *monoclonal antibodies* (*mAbs*) – a single species of monovalent antibody – at high abundance *in vitro*. The hybridoma cell line could be retained and expanded indefinitely while preserving the integrity of its product.

mAbs can be classified into first- or second-generation reagents along similar lines to the other therapeutic proteins discussed above. First-generation *mAbs* were simply murine monoclonals (or fragments thereof) but had several drawbacks. As mouse proteins, they provoked an immune response in 50%–75% of all human recipients, had a short half-life in the human circulation and were unable to activate human complement.

Most of these problems can now be surmounted by using either *chimeric* or *humanised* *mAbs*. These two terms refer to the degree to which they have been engineered. Fig. 5.1 shows how this is done; the antibody molecule consists of a *constant* domain (Fc) and the antibody-binding domain (Fab), with *hypervariable* regions that recognise and bind to the antigen in question. The genes for chimeric *mAbs* are engineered to contain the cDNA of the *murine* Fab domain coupled with the *human* Fc domain sequences. This

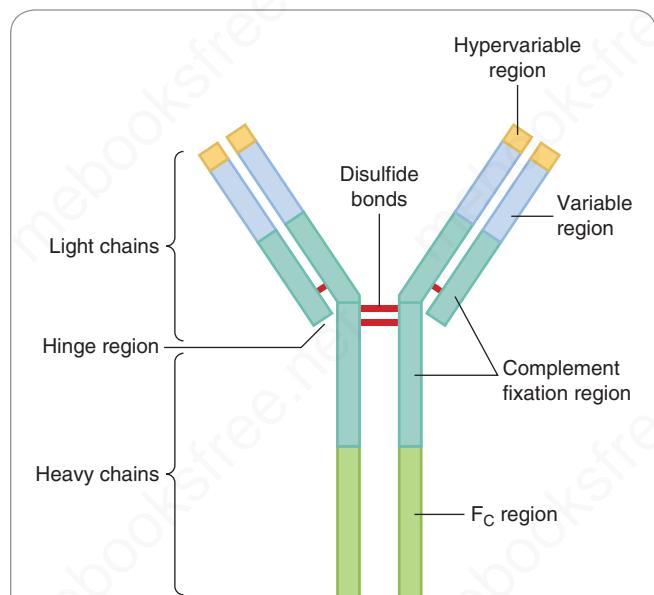


Fig. 5.1 Production of engineered 'chimeric' and 'humanised' monoclonal antibodies. The Y-shaped antibody molecule consists of two main domains: the Fc (constant) domain and the Fab (antigen-binding) domain. At the tip of the Fab regions (on the arms of the 'Y') are the hypervariable regions that actually bind the antigen. Chimeric antibodies are produced by replacing the murine Fc region with its human equivalent by altering and splicing the gene. For humanised antibodies, only the murine hypervariable regions are retained, the remainder of the molecule being human in origin. (After Walsh, 2004.)

²They won the 1984 Nobel Prize for Physiology or Medicine for this work.

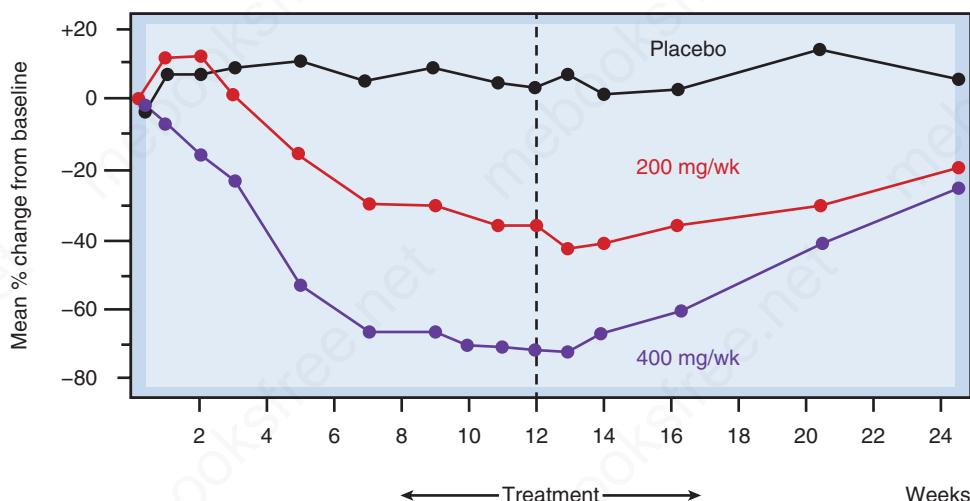


Fig. 5.2 Using antisense oligonucleotides to correct mild-moderate hyperlipidaemia. The antisense oligonucleotide mipomersen was administered to 50 patients for 13 weeks. The data shows the mean reduction in low-density lipoprotein (LDL) cholesterol expressed in percentage terms from the baseline readings at day 1 with doses of 200 mg/week (red) and 400 mg/week (blue) compared with a placebo (black). The reduction in expression of apolipoprotein B caused by the drug exactly paralleled the LDL cholesterol data. After discontinuation of the treatment (indicated by dotted line) the blood levels showed signs of returning to baseline values, but were still depressed by 20%–30% at the conclusion of the study at these doses (Redrawn from Geary et al., 2015).

greatly (around five-fold) extends the plasma half-life because whilst most plasma proteins turn over quite rapidly immunoglobulins are an exception (it is easy to see why this provides a selective advantage to the host). Incorporation of human Fc sequences also improves the functionality of the antibody in human medicine. A further development (and now the preferred approach) is to replace the entire Fc and Fab region with the human equivalent with the exception of the hypervariable regions, giving a molecule which, while essentially human in nature, contains just the minimal murine antibody-binding sites. The anticancer monoclonal **Herceptin** (**trastuzumab**; see Ch. 57) is an example of such an antibody, and some others (together with an explanation of the tongue-twisting nomenclature system) are given in Table 5.1.

OLIGONUCLEOTIDES

We turn next to another type of biopharmaceutical, this time based upon oligonucleotide structures. These offer an alternative way of modifying genetic material that is far less problematic than delivering entire genes (see later). Amongst the most useful approaches is the use of *antisense oligonucleotides*. These are short oligonucleotides that are complementary to part of a gene or gene product that one wishes to modify or suppress. The oligomer needs to be at least 15 bases long to confer specificity and tight binding to its target sequence (most antisense oligos are 15–25 mers). These snippets of genetic material can be designed to suppress the expression of a harmful gene either by forming a triplex (three-stranded helix) with a regulatory component of chromosomal DNA, or by complexing a region of mRNA as a duplex. Unlike entire gene constructs, oligonucleotides can cross plasma and nuclear membranes by endocytosis as well as by direct diffusion, despite their molecular size and charge. To avoid destruction by nucleotidases which

are ubiquitous in biological fluids, enzyme-resistant *methylphosphorate*, *phosphothiorate* or other analogues have been developed.

Following parenteral administration, such oligomers distribute widely throughout the body (although not to the CNS) and work in part by interfering with the transcription of mRNA and in part by stimulating its breakdown by ribonuclease H. **Mipomersen**, the first ever licensed antisense therapeutic (2013), is a phosphothiorate analogue that suppresses the expression of apolipoprotein B, acts through this mechanism. It can be used to treat a rare form of hypercholesterolaemia (Fig. 5.2). Another oligonucleotide drug recently (2016) approved by the FDA, **etioplenersen**, targets a specific region of the *dystrophin* gene that is mutated in Duchenne muscular dystrophy, changing the reading frame and causing the faulty exon to be removed, with the result that the final translated product is a partially functional version of the protein.

After a slow start with only two agents actually reaching the market prior to 2017, at the time of writing, several candidate drugs are being tested for use in the treatment of viral diseases including HIV, cytomegalovirus and haemorrhagic virus infections as well as cancer, spinal muscular atrophy and other disorders.

A related approach (see Castanotto & Rossi, 2009), which provides more efficient gene silencing than antisense oligonucleotides, is the use of *short interfering RNA* (siRNA),³ whereby short lengths of double-stranded RNA recruit an enzyme complex, known as RISC (RNA-Induced Silencing Complex), which selectively degrades the corresponding

³Discovered when plant scientists found, to their surprise, that introducing RNA which encoded the colour-producing enzyme in petunias made the flowers *less* colourful, not more so. Subsequently siRNA has emerged as an important physiological mechanism for controlling gene expression and was recognised by the award in 2006 of the Nobel Prize to Craig Mello and Andrew Fire.

Table 5.2 Differences between biopharmaceuticals and conventional small molecule drugs

Property	Conventional drug	Biopharmaceutical
Size	Generally <500 kDa (10^2)	Generally >5000 kDa (10^3), e.g. oligonucleotides, 10^3 ; small proteins 10^3 – 10^4 ; mAbs 10^5 ; genes $>10^6$.
Synthesis	Easy to synthesise identical batches	Most biopharmaceuticals are unique (except small peptides and short oligonucleotides which can be synthesised chemically)
Relationship between dose and effect	Usually a predictable relationship between dose and effect	Complex mechanisms of action, usually high-affinity binding, slow on- and off-rates, unusual shaped D/R curves
Pharmacokinetics	Often oral administration, variable absorption and bioavailability, phase 1 and phase 2 metabolism, excretion of drug in urine or faeces	Usually parenteral administration, bioavailability high, long half-life, atypical biodistribution and removal mechanisms
Toxicology and adverse effects	Variable, possible drug interactions	Immunogenicity, few drug interactions, generally fewer adverse effects

mRNA produced by the cell, thereby blocking expression. These silencing RNA sequences can also be efficiently produced in a cell infected with a virus that is engineered to express the right sequence. Clinical trials of siRNA therapeutics are in progress.

PHARMACOLOGY OF PROTEIN AND OLIGONUCLEOTIDE PHARMACEUTICALS

There are important differences between the pharmacological properties of protein and oligonucleotide biopharmaceuticals and those of conventional small molecule drugs (Table 5.2), attributable in part to their difference in molecular mass. Most conventional drugs have molecular masses of less than 1000 and are usually less than 500 – in fact, it is thought that this factor is important in achieving optimal distribution in the body and for the biological activity of the drug. In contrast, even the smallest protein biopharmaceutical, insulin, has a molecular mass of almost 6000. Antibodies usually weigh in at about 150,000 and oligonucleotides about 2000–3000. Their size obviously affects the absorption and bioavailability of biopharmaceuticals.

Another distinguishing factor is a consequence of their production. Conventional drugs and short oligonucleotides are produced by total (occasionally partial) chemical synthesis, with identical characteristics wherever the compound is made. However, this is not the case with many protein-based biopharmaceuticals. The gene expression system used to produce therapeutically active proteins differs from one company to another – deliberately so in most cases because, unlike genes themselves, proprietary gene constructs and expression systems can be patented, enabling pharma to protect its intellectual property. Each expression system produces a slightly different product in terms of its purity, post-translational modifications and protein ‘fingerprint’⁴. This has important consequences for drug regulation, because unlike synthetic small molecule drugs, each biopharmaceutical is unique and a common saying in the biotech industry is that ‘the product is the process’. Compared to the first-in-the-field drug, subsequent

entry biologics (SEBs) or follow-on biologics (FOBs), may be bioequivalents (i.e. drugs that are interchangeable with, and therapeutically equivalent to, the original preparation); more often, while still therapeutically effective, they have different clinical properties.⁵ However, each requires separate regulatory approval.

Another manufacturing issue concerns the number of steps required to prepare biopharmaceuticals. With chemical synthesis, one can assess the exact purity of the final product, but preparations of biopharmaceuticals may not be homogenous and could contain mixtures of different glycoforms of the protein or possibly bacterial proteins or endotoxins. This means that there is a requirement for greatly enhanced quality control and this obviously has profound implications for ease of manufacture and final cost (Revers & Furzcon, 2010).

Some actions of biopharmaceuticals resemble those of conventional drugs: for example, insulin or growth hormone have identical actions to the native hormone. But there are differences too. Some mAbs immunoneutralise unwanted substances: for example, **infliximab** directly neutralises the cytokine TNF to produce its therapeutic effect. However, another mAb, **rituximab**, binds to CD20 on lymphocytes and causes actual destruction of the cells to diminish an unwanted immune response. **Ibritumomab tiuxetan** also binds CD20, but delivers ⁹⁰Y to kill the cells.

Because of these different modes of action, the relationship between dose and effect, so beloved of pharmacologists, is much less clear cut. Agoram (2009) highlights some of the problems. In the case of mAbs, high-affinity binding is usual (sometimes a mixture of specific and non-specific binding), slow on- and off-rates are common and the mAb may be internalised, thus modifying its target cell. Dose-response relationships are sometimes bell-shaped or U-shaped. Human recombinant erythropoietin has a bell-shaped dose-response and, in the case of many mAbs, there is a single optimal dose at which effective immunoneutralisation

⁴This ‘biological variation’ using cells to produce a drug is obviously not inherent in more exact medicinal chemistry processes.

⁵More bewildering terminology: *biosimilars* are generic drugs with a similar function to the original but with different pharmacology or toxicology; *biobetters* are generic drugs with a similar function to the original but with superior pharmacology or toxicology.

Table 5.3 A comparison of pharmacokinetics between two conventional small molecule drugs and some biopharmaceuticals

Type	Drug	Route ^a	Dosing frequency	T _{max}	T _{1/2}	Bioavailability	V
Conventional drug	20 mg simvastatin	p.o.	1 per day	0.7 h	1.5 h	<5%	215 L/kg
	75 mg indometacin	p.o.	1–2 per day	2–3 h	2–3 h	>90%	1.0 L/kg
Biopharmaceutical	25 mg etanercept	i.m.	1–2 per week	69 h	102 h	58%	6–11 L/kg
	40 mg adalimumab	i.m.	1 per 2 weeks	131 h	10–20 days	64%	4.7–6.0 L/kg
	75 mg omalizumab	i.m.	1 per month	7–8 days	26 days	62%	5.5 L/kg

^aRoute of administration: *i.m.*, intramuscular; *p.o.*, by mouth. All data approximated from information from manufacturers.

T_{max}, time to maximum plasma concentration; T_{1/2}, half-life; V, volume of distribution.

occurs, instead of the proportional effects that we are more accustomed to when dealing with small molecule drugs. In the case of oligonucleotides, there are also several different mechanisms of action, as we saw previously.

The differences in the nature and size of most biopharmaceuticals when compared with conventional drugs also have implications for their pharmacokinetic properties. Because proteins do not usually survive oral administration, most are administered subcutaneously, intramuscularly or intravenously and so bioavailability is typically high compared with many small molecule drugs, often in the region of 80%–100%. But, except in the case of intravenous administration, absorption from the injection site is usually slow and the time taken to achieve attain the T_{max} reflects this. Once in the circulation however, the half-life is typically long. Because antibodies bind to their target with high affinity, the volume of distribution is often small, but transcellular and unusual trafficking may redistribute the drug to other tissues (Zhao et al., 2012). A comparison of the pharmacokinetics of conventional small molecule drugs with several biopharmaceuticals is shown in Table 5.3.

Biopharmaceuticals are not removed from the body following metabolic transformation and excretion of the type described in Chapter 10 and indeed, some antibodies can persist in the circulation for weeks. Instead, uptake of large biopharmaceuticals by the lymphatic system is the usual first step, followed by lysosomal degradation. However, some ‘small’ (<69 kDa) mAbs may be eliminated directly by the kidney. Immunogenicity is an issue that plagued early development of proteins as drugs and whilst this has been largely overcome by ‘humanisation’ of antibodies and proteins, it is still important because it alters the pharmacokinetic properties of the drug (Richter et al., 1999) by increasing its clearance from the circulation.

Drug interactions are less of an issue with biopharmaceuticals, as are general toxicity problems and adverse side effects, an advantage that is reflected in their relatively rapid approval by regulatory agencies. In part, this is due to their extraordinary specificity. In fact, few drugs come closer to the idea of a ‘magic bullet’⁶ than biopharmaceuticals which, because of the specificity of the immune system,

can inactivate single targets with an extraordinary degree of precision. Ironically, this can cause major problems when testing these drugs.

▼ In 2006, for example, a UK clinical trial of a new mAb (TGN 1412) designed to activate T cells (see Ch. 7) and thus treat B-cell lymphocytic leukaemia went badly wrong. All six participants became severely ill following a ‘cytokine storm’ and suffered lasting damage. The incident provoked wide media publicity⁷ and, while the subsequent investigation blamed an ‘unpredictable’ biological reaction, it caused many to think hard about how such trials should be conducted in the future (see Muller & Brennan, 2009). Highly specific reagents, such as monoclonals intended for human use, pose particular problems as they may not cross-react with the corresponding proteins of other species, thus evading detection in the usual preclinical animal safety screens. It may be the case that ‘surrogate’ mAbs which are species-specific will have to be developed to test in animal models of the disease.

GENE THERAPY

Astonishingly, the first study to demonstrate the theoretical feasibility of gene transfer took place in 1944 when Avery and his colleagues showed that a virulence factor could be transferred between two strains of pneumococcus and identified the factor as (what we now call) DNA, which was not even recognised at that time as the genetic material. Following the molecular biology revolution in the 1980s however, the significance of this experiment became clear and the notion that one could replace faulty or missing genes became a thrilling – if distant – prospect. Despite the high hopes and intensive research efforts in the intervening years, the full potential of gene therapy is still unrealised. However, the idea commands such appeal that vast resources (both public and private) have been committed to its development. There are several reasons why it is so attractive. First, it is a (deceptively) simple approach to a radical cure of single-gene diseases such as *cystic fibrosis* and the *haemoglobinopathies*, which are collectively responsible for much misery throughout the world. Second, many other more common conditions, including malignant, neurodegenerative and infectious diseases, have a large genetic component. Conventional treatment of such disorders is (as readers of later chapters will appreciate) far from ideal, so the promise of a completely new approach has enormous allure.

⁶It was Weber’s opera, *Der Freischütz* (The Sharpshooter, 1821), that introduced the idea of a ‘magic bullet’ which, once fired, always found its mark. Ehrlich liked the idea and thought that it was a good description of a highly specific drug. The term, and the concept, has haunted our discipline ever since.

⁷One tabloid headline read: ‘We saw human guinea pigs explode’ (quoted by Stobart et al., 2007).

Table 5.4 Characteristics of some delivery systems for gene therapy

Vector	Advantages	Disadvantages	Utilisation of system ^a
Liposomes	Virus-free, cheap to produce	Low efficiency, sometimes cytotoxic	6%
DNA cassettes	Virus-free	Low efficiency, expression temporary	18%
Herpes simplex virus type I	Highly infective, persistent expression	No integration with host DNA, cytotoxic, difficult to handle	3%
Adenovirus	Highly infective in epithelia	Immunogenic and transient, requires repeated administration	23%
Adeno-associated virus	Stable	Low capacity	5%
Retrovirus	Efficient, permanent	Low capacity, unstable, must integrate into host DNA, requires dividing cells	22%

^aThe approximate percentage of trials employing this type of delivery system. (After Wolf & Jenkins, 2002; and with data from Wirth et al., 2013.)

The gurus are emphatic that ‘the conceptual part of the gene therapy revolution has indeed occurred ...’ – so where are the therapies? The devil, of course, is in the detail: in this case, the details of:

- pharmacokinetics: delivery of the gene to the interior of appropriate target cells (especially those in the CNS);
- pharmacodynamics: the controlled expression of the gene in question;
- safety;
- clinical efficacy and long-term practicability.

There is a broad consensus that the *Weismann barrier*⁸ should not be breached and so a moratorium has been agreed on making alterations to the DNA of germ cells (which could influence future generations) and gene therapy trials have focused on somatic cells only.

Here we focus first on the main problems and approaches being used to transform gene therapy into useful medicines, and conclude with a final section on the limited success achieved so far.

GENE DELIVERY

The transfer of large sections of recombinant nucleic acid into target cells is critical to the success of gene therapy. In the words of one commentator (Galun, quoted in Bender, 2016), ‘Gene therapy is actually three things: delivery, delivery and delivery’. To overcome this first and most fundamental hurdle, techniques borrowed from viruses, which are masters of the sort of molecular hijacking that is required to introduce functional genes into mammalian cells, are often used in gene therapy research. The constructs must pass from the extracellular space across the plasma and nuclear membranes, and be incorporated into the chromosomes. Because DNA is negatively charged and single genes have molecular weights around 10^4 times greater than conventional drugs, the problem is of a different order from the equivalent stage of routine drug development.

There are several important considerations in choosing a gene delivery system; these include:

- the *capacity* of the system (e.g. how much DNA it can carry);
- the *transfection efficiency* (its ability to enter and become utilised by cells);
- the *lifetime* of the transfected material (determined by the lifetime of the targeted cells);
- the *safety issue*, especially important in the case of viral delivery systems.

Various approaches have been developed (Table 5.4) in an attempt to produce the optimal system.

There are two main gene therapy strategies. Using the *in vivo* technique, the vector containing the therapeutic gene is injected into the patient, either intravenously (in which case some form of organ or tissue targeting is required) or directly into the target tissue (e.g. the retina). When using the *ex vivo* strategy, cells are removed from the patient (e.g. stem cells from bone marrow or circulating blood, or myoblasts from a biopsy of striated muscle), and treated with the vector in the laboratory. The genetically altered cells are injected back into the patient they came from, thus avoiding any immune rejection (autologous rather than allogenic).

An ideal vector should be *safe*, highly *efficient* (i.e. insert the therapeutic gene into a high proportion of target cells and under control of the appropriate promoter) and *selective* in that it should lead to expression of the therapeutic protein in the target cells but *not* to the expression of other viral proteins. Ideally, and provided that the cell into which it is inserted is itself long-lived, the vector should cause persistent expression, avoiding the need for repeated treatment. The latter consideration can be a problem in some tissues. In the autosomal recessive disorder *cystic fibrosis*, for example, the airway epithelium malfunctions because it lacks a membrane Cl^- transporter known as the *cystic fibrosis transport regulator* (CFTR). Epithelial cells in the airways are continuously dying and being replaced, so even if the unmutated CFTR gene could be stably transfected into the epithelium, there would still be a periodic need for further treatment unless the gene could be inserted into the progenitor (stem) cells. Similar problems are anticipated in other cells that turn over continuously, such as gastrointestinal epithelium and skin.

⁸Named after August Weismann (1834–1914), who formulated the concept that inheritance utilises only germ, and not somatic, cells.

VIRAL VECTORS

Many contemporary gene delivery strategies aim to capitalise on the capacity of viruses to subvert the transcriptional machinery of the cells they infect and their ability (in some cases) to fuse with the host genome. While seemingly simple, there remain substantial practical problems with this *viral vector* approach. As viruses have evolved the means to invade human cells, so humans have evolved immune responses and other protective countermeasures. Although limiting in some respects, this is not all bad news from the point of view of safety. As many of the viruses used for vectors are pathogenic, they are usually modified such that they are ‘replication defective’ to avoid toxicity.

Retroviruses

▼ If introduced into stem cells, *retroviral vectors* have long-lasting effects because they are incorporated into, and replicate along with, host DNA, and so the ‘therapeutic’ gene is passed down to each daughter cell during division. Against this, the *retroviral integrase* inserts the construct into chromosomes randomly, so it may cause damage. Also, retroviruses could infect germ or non-target cells and produce undesired effects if administered *in vivo*. For this reason, retroviruses have been used mainly for *ex vivo* gene therapy. The life cycle of naturally occurring retroviruses may be exploited to create useful vectors for gene therapy (Fig. 5.3).

Many viruses are equipped to infect specific cell types, though not necessarily the target cell of interest. It is possible to alter the retroviral envelope to alter specificity, such that the vector could be administered

systemically but would target only the desired cell population. An example of this approach with a *lentivirus* (a type of retrovirus) is the substitution of the envelope protein of a non-pathogenic vector (e.g. mouse leukaemia virus) with the envelope protein of human vesicular stomatitis virus, to specifically target human epithelial cells. Most retrovirus vectors are unable to penetrate the nuclear envelope, and because this dissolves during cell division, they only infect dividing cells rather than non-dividing cells (such as adult neurons).

Adenovirus

▼ *Adenovirus vectors* are popular because of the high transgene expression that can be achieved. They transfer genes to the nucleus of the host cell, but (unlike retroviruses) these are not inserted into the host genome and so do not produce effects that outlast the lifetime of the transfected cell. This property also obviates the risk of disturbing the function of other cellular genes and the theoretical risks of carcinogenicity and germ cell transfection. Because of these favourable properties, adenovirus vectors have been used for *in vivo* gene therapy. Engineered deletions in the viral genome render it unable to replicate or cause widespread infection in the host while at the same time creating space in the viral genome for the therapeutic transgene to be inserted.

One of the first adenoviral vectors lacked part of a growth-controlling region called *E_I*, while incorporating the desired transgene. This vector gave excellent results, demonstrating gene transfer to cell lines and animal models of disease, but it proved disappointing as a treatment for cystic fibrosis in human trials. Low doses (administered by aerosol to patients with this disease) produced only a very low-efficiency transfer, whereas higher doses caused inflammation, a host immune response and short-lived gene expression. Furthermore, treatment

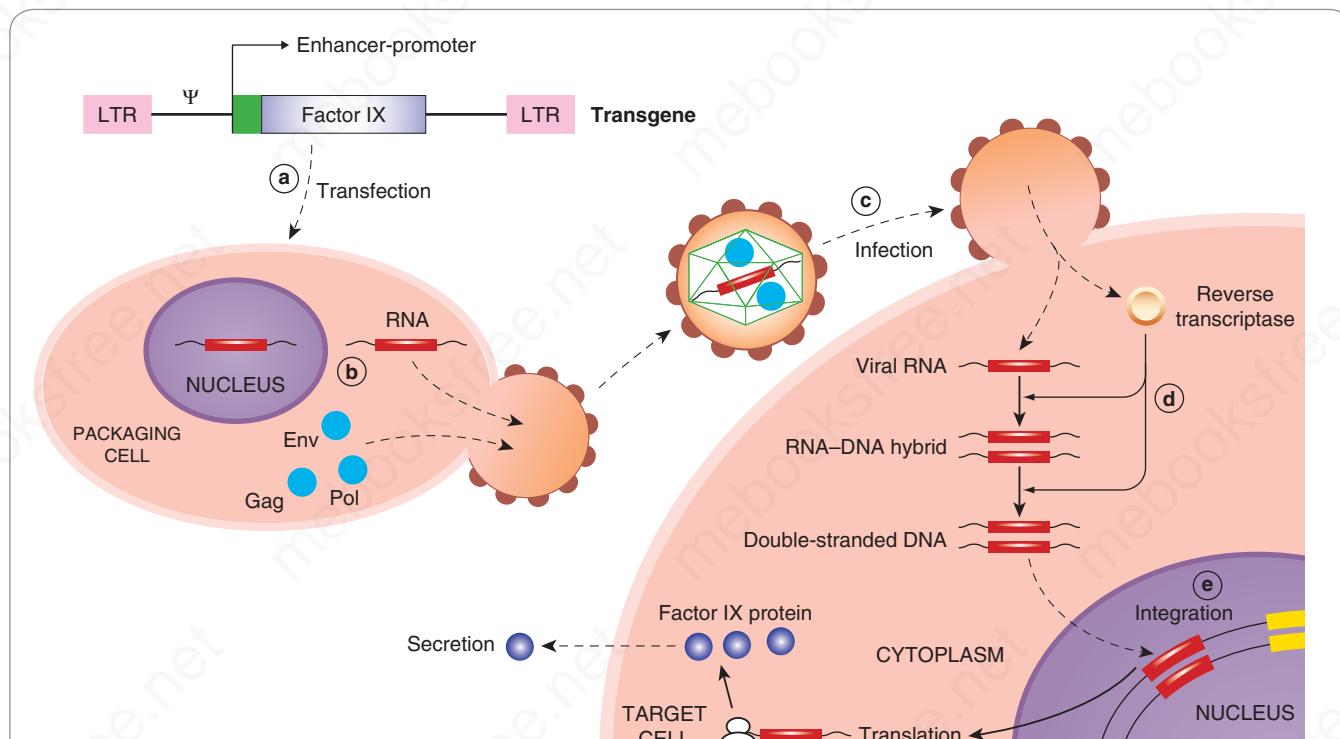


Fig. 5.3 Strategy for making retroviral vectors. The transgene (the example shows the gene for factor IX) in a vector backbone is introduced (a) into a packaging cell, where it is integrated into a chromosome in the nucleus, and (b) transcribed to make vector mRNA, which is packaged into the retroviral vector and shed from the packaging cell. It then infects the target cell (c). Virally encoded reverse transcriptase (d) converts vector RNA into an RNA–DNA hybrid, and then into double-stranded DNA, which is integrated (e) into the genome of the target cell. It can then be transcribed and translated to make (in this case) factor IX protein. ‘Env’, ‘Gag’ and ‘Pol’ represent components of the retroviral vector. LTR, long terminal repeat. (Redrawn from Verma & Somia, 1997.)

could not be repeated because of the appearance, in the circulation, of neutralising antibodies. This has led to attempts to manipulate adenoviral vectors to mutate or remove the genes that are most strongly immunogenic.

Other viral vectors

▼ Other potential viral vectors under investigation include *adenovirus*, *herpes virus* and disabled versions of *human immunodeficiency virus* (HIV). Adeno-associated virus associates with host DNA but is not activated unless the cell is infected with an adenovirus. It is less immunogenic than other vectors but is difficult to mass produce and cannot be used to carry large transgenes. Herpes virus does not associate with host DNA but is very long lived in nervous tissue (so could have a specific application in treating neurological disease). HIV, unlike most other retroviruses, can infect non-dividing cells such as neurons. It is possible to remove the genes from HIV that control replication and substitute other genes. Alternatively, it may prove possible to transfer to other non-pathogenic retroviruses those genes that permit HIV to penetrate the nuclear envelope.

NON-VIRAL VECTORS

To reduce the problems associated with viral vectors, a variety of other substances have been used to deliver genes and other material. These are often collectively known as *nanocarriers*. The list includes the following (but see also Xu et al., 2014):

Liposomes

▼ Non-viral vectors include a variant of liposomes (Ch. 9). Plasmids (diameter up to approximately 2 µm) are too big to package in regular liposomes (diameter 0.025–0.1 µm), but larger particles can be made from positively charged lipids ('lipoplexes'), which interact with both negatively charged cell membranes and DNA, improving delivery into the cell nucleus and incorporation into the host chromosome. Such particles have been used to deliver the genes for HLA-B7, interleukin-2 and CFTR. They are much less efficient than viruses, and attempts are currently under way to improve this by incorporating various viral signal proteins (membrane fusion proteins, for example) in their outer coat. Direct injection of these complexes into solid tumours (e.g. melanoma, breast, kidney and colon cancers) can, however, achieve high local concentrations within the tumour.

Microspheres

▼ Biodegradable microspheres made from polyanhydride co-polymers of fumaric and sebamic acids (see Ch. 9) can be loaded with plasmid DNA. A plasmid with bacterial β-galactosidase activity formulated in this way and given by mouth to rats has resulted in systemic absorption and expression of the bacterial enzyme in the rat liver, raising the possibility of oral gene therapy.

Plasmid DNA

▼ Surprisingly, plasmid DNA itself ('naked DNA') enters the nucleus of some cells and is expressed, albeit much less efficiently than when it is packaged in a vector. Such DNA carries no risk of viral replication and is not usually immunogenic, but it cannot be targeted precisely. There is considerable interest in the possibility of using naked DNA for vaccines, which has several theoretical advantages and numerous trials are using the technique (Liu, 2011).

CONTROLLING GENE EXPRESSION

To realise the full potential of gene therapy, it is not enough to transfer the gene selectively to the desired target cells and maintain acceptable expression of its product – difficult though these goals are. It is also essential that the activity of the gene is controlled. Historically, it was the realisation of the magnitude of this task that diverted attention from the haemoglobinopathies (which were the first projected targets of gene therapy). Correction of these disorders

demands an appropriate balance of normal α- and β-globin chain synthesis to be effective, and for this, and many other potential applications, precisely controlled gene expression is essential.

▼ It has not yet proved possible to control transgenes precisely in human recipients, but there are techniques that may eventually enable us to achieve this goal. One hinges on the use of an inducible expression system. This is a fairly standard laboratory technique whereby the inserted gene also includes a *doxycycline*-inducible promoter such that expression of the gene can be switched on or off by treatment with, or withdrawal of, doxycycline.

The control of transfected genes is important in gene targeting as well. By splicing the gene of interest with a tissue-specific promoter, it should be possible to restrict expression of the gene to the target tissue. Such an approach has been used in the design of gene therapy constructs for use in ovarian cancer, the cells of which express several proteins at high abundance, including the proteinase inhibitor SLP1. In combination with the SLP1 promoter, plasmids carrying various genes were successfully and selectively expressed in ovarian cancer cell lines (Wolf & Jenkins, 2002).

SAFETY AND SOCIETAL ISSUES

Experiments or protocols involving the transfer of genetic material tend to provoke deep unease in some sectors of society – witness the genetically modified (GM) crop debate (and see Freier et al., 2014). Partly, this may be traced to ignorance or prejudice but it is nevertheless a problem that can hinder the introduction of new agents. Societal issues aside, the technique does raise a number of specific concerns that generally relate to the use of viral vectors. These are usually selected because they are non-pathogenic, or modified to render them innocuous, but there is a concern that such agents might still acquire virulence during use. Retroviruses, which insert randomly into host DNA, could damage the genome and interfere with the protective mechanisms that normally regulate the cell cycle (see Ch. 6), and if they happen to disrupt essential cellular functions, this could increase the risk of malignancy.⁹

Another problem is that immunogenic viral proteins may elicit an inflammatory response, and this could be harmful in some situations (e.g. in the airways of patients with cystic fibrosis). Initial clinical experience was reassuring, but the death of Jesse Gelsinger, an 18-year-old volunteer in a gene therapy trial for the non-fatal disease *ornithine decarboxylase deficiency* (which can be controlled, albeit tediously, by diet and drugs anyway), led to the appreciation that safety concerns related to immune-mediated responses to vectors are very real (see Marshall, 1999).

THERAPEUTIC APPLICATIONS

Despite the plethora of technical problems and safety concerns, there have been some encouraging successes and interest – and confidence – in the area is still very strong with some 3000 new publications appearing each year.

⁹This risk is more than a theoretical possibility; several children treated for *severe combined immunodeficiency* (SCID) with a retrovirus vector developed a leukaemia-like illness (Woods et al., 2006). The retroviral vector was shown to have inserted itself into a gene called LMO-2, mutations of which are associated with childhood cancers.



Gene delivery and expression

- Gene delivery is the main hurdle to practical gene therapy.
- Recombinant genes are transferred using a *vector*, often a suitably modified virus.
- There are two main strategies for delivering genes into patients:
 - *in vivo* injection of the vector directly into the patient (e.g. into a malignant tumour);
 - *ex vivo* treatment of cells from the patient (e.g. stem cells from marrow or circulating blood), which are then returned to the patient.
- An ideal vector should be safe, efficient, selective and produce long-lasting expression of the therapeutic gene.
- Viral vectors include retroviruses, adenoviruses, adeno-associated virus, herpesvirus and disabled HIV:
 - *Retroviruses* infect many different types of dividing cells and become incorporated randomly into host DNA.
 - *Adenoviruses* are genetically modified to prevent replication and accommodate the therapeutic transgene. They transfer genes to the nucleus but not to the genome of the host cell. Problems include a
- strong host immune response, inflammation and short-lived expression. Treatment can be compromised by neutralising antibodies.
- *Adeno-associated virus* associates with host DNA and is non-immunogenic but is hard to mass produce and has a small capacity.
- *Herpesvirus* does not associate with host DNA but persists in nervous tissue and may be useful in treating neurological disease.
- Disabled versions of HIV differ from most other retroviruses in that they infect non-dividing cells, including neurons.
- Non-viral vectors include:
 - a variant of liposomes, made using positively charged lipids and called ‘lipoplexes’;
 - biodegradable microspheres, which may offer orally active gene therapy;
 - plasmid DNA (‘naked DNA’), which can be used as a vaccine.
- A *tetracycline-inducible expression system* or similar technique can control the activity of the therapeutic gene.

According to a recent reviewer (Tani, 2016) 2210 gene therapy trials had been approved by regulatory agencies around the world by 2015, with six gene therapies already approved by a limited number of countries. The first was **Gendicine**, a treatment for replacing the faulty p53 protein causing head and neck cancer, which was licensed in China in 2003. The European Medicines Agency granted its first license for a gene therapy product, **Glybera**, in 2012¹⁰. This is an adeno-associated virus construct that delivers a correct copy of lipoprotein lipase to patients lacking this enzyme (a very rare disorder that causes severe pancreatitis) and in 2016, **Strimvelis** was also approved in Europe. This is an *ex vivo* gene therapy approach to replace adenosine deaminase which is absent in children with a rare (~15 patients per year in Europe) type of SCID. At the time of writing, no gene therapy-based therapeutics had been approved in the United States although intense interest continues.

Recently, a system of gene editing originally discovered in bacteria is promising to transform gene therapy (garnering nearly 7000 publications over the last few years). This bears the rather complex name of *Clustered Regulatory Interspersed Short Palindromic Repeats* (CRISPR) and can target nucleases (notably Cas9) to precisely edit genes of interest. Viruses can deliver the CRISPR-Cas9 components, thereby acting as delivery vehicles for the biochemical machinery required to repair faulty genes in humans (see for example, Gori et al., 2015; Gee et al., 2017). To date,

no successful gene therapy using CRISPR-Cas9 has been approved for therapeutic use, but several human trials are underway¹¹, and are very close to getting this therapy into the clinic.

▼ Target diseases eliciting interest from companies specialising in gene therapy applications include:

Single-Gene Defects

Single-gene (*monogenic*), often rare, disorders were the obvious starting point for gene therapy trials and haemoglobinopathies were the first projected targets, but early attempts (in the 1980s) were put ‘on hold’ because of the problem (mentioned previously) of controlling precisely the expression of the genes encoding the different polypeptide chains of the haemoglobin molecule. Recent trials have proved encouraging in the treatment of thalassaemia (the commonest monogenic disease) and sickle cell disease (see Rai & Malik, 2016) although no products have yet been approved.

Another early target was cystic fibrosis, but progress here has been disappointing (see Kim et al., 2016 for details) largely because of the biological barriers that must be penetrated. There have been other successes though. For example, *X-linked chronic granulomatous disease* has been successfully treated using a retroviral technique to deliver a functional version of the mutated NADPH oxidase protein (Ott et al., 2006 and Fig. 5.4) and a form of inherited blindness, *Leber's congenital amaurosis*, associated with a mutation in a gene that produces retinal pigment, has been rectified using an adeno-associated virus vector bearing a cDNA coding for the intact gene (Maguire et al., 2009). Several other ocular conditions also seem to be promising candidates for gene therapy approaches (see Borras, 2017; Boye et al., 2013). Williams and Thrasher (2014) have reviewed the general

¹⁰With an annual cost of at least US\$1 million per treatment, Glybera has been called the most expensive medicine in the world. No wonder that only one patient has been treated so far! (see Regolado, 2016)

¹¹Allogenic transplants in childhood leukaemias, autologous immune cell PD-1 knock-out in lung cancer patients, and OCT4’s role in human embryo development are all recent examples of breakthroughs using CRISPR-Cas9 in human cells.

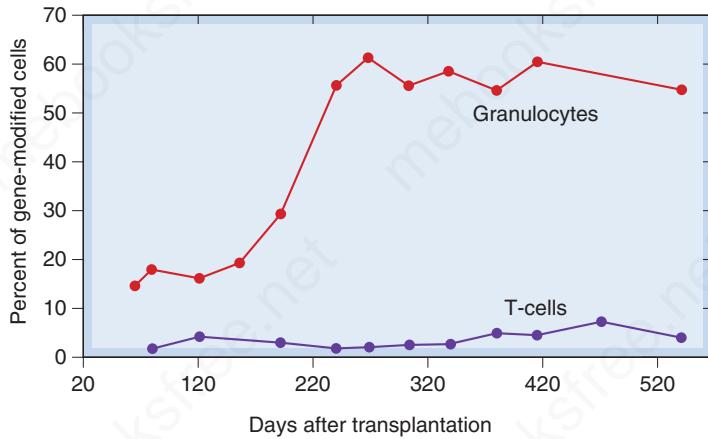


Fig. 5.4 Correcting an inherited defect using gene therapy. In this clinical trial, two patients with X-linked chronic granulomatous disease were transfused with GM-CSF (Granulocyte-Macrophage Colony Stimulating Factor)-treated peripheral blood cells that had been genetically modified with a retroviral vector bearing the intact *gp91phox* gene ('in vitro protocol' – see text). The graph shows that the number of gene-modified peripheral blood leukocytes remained high for well over a year and this was accompanied by good levels of superoxide production in these cells – a clinical 'cure'. (Data redrawn from Ott et al., 2006.)

Safety issues for gene therapy



- There are those safety concerns that are specific to any particular therapy (e.g. polycythaemia from overexpression of **erythropoietin**) and also additional general concerns relating, for example, to the nature of the vectors used.
- Viral vectors:
 - might acquire virulence during use
 - contain viral proteins, which may be immunogenic
 - can elicit an inflammatory response
 - could damage the host genome and interfere with the cell cycle, provoking malignancy.
- The limited clinical experience to date has not so far provided evidence of insurmountable problems.

problems associated with gene therapy in the treatment of monogenic immunodeficiency diseases.

Cancer

Gene therapy for cancer and related diseases currently comprise almost 70% of gene therapy trials. Several therapeutic approaches (see Barar & Omidi, 2012) are under investigation, including:

- restoring 'protective' proteins, such as the tumour suppressor gene (see Ch. 6);
- inactivating oncogene expression (e.g. by using a retroviral vector bearing an antisense transcript RNA to the K-Ras oncogene);
- delivering a gene to malignant cells that renders them sensitive to cytotoxic drugs (e.g. thymidylate kinase, which activates **ganciclovir**) – the so-called 'suicide gene' approach;
- delivery of proteins to healthy host cells, which, for example, renders them resistant to chemotherapy (e.g. addition of the multidrug resistance channel to bone marrow cells ex vivo);

- tagging cancer cells with genes expressing proteins that render malignant cells more visible to the immune system (e.g. for antigens such as HLA-B7 or cytokines such as granulocyte-macrophage colony-stimulating factor and interleukin-2).

- Recent progress in research in some of these areas has been reviewed by Gilham et al. (2015).

Gene Therapy and Infectious Disease

In addition to DNA vaccines mentioned previously, there is considerable interest in the potential of gene therapy for HIV and other viral infections. The aim is to render stem cells (which differentiate into immune cells) resistant to HIV before they mature. For an account of the strategies under investigation, see Chung et al. (2013).

Gene Therapy and Cardiovascular Disease

Gene therapy trials for treating cardiovascular diseases are reviewed by Bradshaw and Baker (2013). Vascular gene transfer is attractive not least because cardiologists and vascular surgeons routinely perform invasive studies that offer the opportunity to administer gene therapy vectors ex vivo (e.g. to a blood vessel that has been removed to use as an autograft) or locally in vivo (e.g. by injection through a catheter directly into a diseased coronary or femoral artery). The nature of many vascular disorders, such as restenosis following angioplasty (stretching a narrowed artery using a balloon that can be inflated via a catheter), is such that transient gene expression might be all that is required therapeutically. Extension of vein graft patency by gene therapy approaches has been reviewed by Chandiwal et al. (2005). This seems to be a promising area (see Hammond & McKirnan, 2001; Ghosh et al., 2008) although Hammer and Steiner (2013) concluded that most trials had proved disappointing.

CONCLUDING REMARKS

Whilst protein and oligonucleotide biopharmaceuticals share some of the characteristics of other drugs described in this book, the same cannot be said for gene therapy. Is a gene a drug? Is a virus a drug? You could argue that it satisfies the broad definition that we posited in on page 1

of this book in that ‘administration to a living organism produces a biological effect’, but it does not seem sensible to discuss the pharmacology of gene therapy as such and most would consider it beyond the scope of the subject. A gene has no inherent pharmacodynamic or pharmacokinetic properties, most of the toxicity and adverse effects mentioned here are due to the vector or carrier and not the

gene itself. And how do you assess the dose of a ‘drug’ that is self-replicating? Having said that, we make no apologies for including gene therapy in this section. There is little doubt that it will become a major therapeutic modality in the future and that physicians and pharmacologists alike will be called on to assess and comment on the biological effects produced.

REFERENCES AND FURTHER READING

General reviews on biopharmaceuticals, gene therapy and utilities

- Agoram, B.M., 2009. Use of pharmacokinetic/ pharmacodynamic modelling for starting dose selection in first-in-human trials of high-risk biologics. *Br. J. Clin. Pharmacol.* 67, 153–160. (*Interesting review of the PK problems of biopharmaceuticals*)
- Bender, E., 2016. Gene therapy: Industrial strength. *Nature* 537, S57–S59. (*Short commentary on recent progress*)
- Brink, M.F., Bishop, M.D., Pieper, F.R., 2000. Developing efficient strategies for the generation of transgenic cattle which produce biopharmaceuticals in milk. *Theriogenology* 53, 139–148. (*A bit specialised, as it focuses mainly on the husbandry of transgenic cattle, but interesting nonetheless*)
- Castanotto, D., Rossi, J.J., 2009. The promises and pitfalls of RNA-interference-based therapeutics. *Nature* 457, 426–433. (*Useful review of the mechanism, current status and potential applications of RNAi as a means of controlling gene expression*)
- Guttmacher, A.E., Collins, F.S., 2002. Genomic medicine: a primer. *N. Engl. J. Med.* 347, 1512–1520. (*First in a series on genomic medicine*)
- Kwon, K.C., Verma, D., Singh, N.D., Herzog, R., Daniell, H., 2013. Oral delivery of human biopharmaceuticals, autoantigens and vaccine antigens bioencapsulated in plant cells. *Adv. Drug Deliv. Rev.* 65, 782–799. (*The title is self-explanatory*)
- Liu, M.A., 2011. DNA vaccines: an historical perspective and view to the future. *Immunol. Rev.* 239, 62–84.
- Melnik, S., Stoger, E., 2013. Green factories for biopharmaceuticals. *Curr. Med. Chem.* 20, 1038–1046. (*Another paper on the use of plants to produce biologics*)
- Rader, R.A., 2008. (Re)defining biopharmaceutical. *Nat. Biotechnol.* 26, 743–751. (*Short commentary dealing with the vexed problem of defining what biopharmaceuticals are*)
- Regolado, A., 2016. The world’s most expensive medicine is a bust. *MIT Technology Review*. May issue.
- Revers, L., Furczon, E., 2010. An introduction to biologics and biosimilars. Part II: Subsequent entry biologics: biosame or biodifferent? *Can. Pharm. J.* 143, 184–191. (*Useful review on regulatory issues surrounding the introduction of follow-on biopharmaceuticals*)
- Tani, K., 2016. Current status of ex vivo gene therapy for hematological disorders: a review of clinical trials in Japan around the world. *Int. J. Hematol.* 104, 42–72.
- Verma, I.M., Somia, N., 1997. Gene therapy – promises, problems and prospects. *Nature* 389, 239–242. (*The authors, from the Salk Institute, describe the principle of inserting corrective genetic material into cells to alleviate disease, the practical obstacles to this and the hopes that better delivery systems will overcome them*)
- Walsh, G., 2004. Second-generation biopharmaceuticals. *Eur. J. Pharm. Biopharm.* 58, 185–196. (*Excellent overview of therapeutic proteins and antibodies; some good tables and figures*)
- Wirth, T., Parker, N., Yla-Herttuala, S., 2013. History of gene therapy. *Gene* 525, 162–169. (*An excellent review of the area from its very beginnings. Highly recommended*)
- Xu, H., Li, Z., Si, J., 2014. Nanocarriers in gene therapy: a review. *J. Biomed. Nanotechnol.* 10, 3483–3507.
- Zhao, L., Ren, T.H., Wang, D.D., 2012. Clinical pharmacology considerations in biologics development. *Acta Pharmacol. Sin.* 33, 1339–1347. (*Review article stressing the differences between conventional drugs and biopharmaceuticals. Good place to start reading*)

Problems

- Check, E., 2002. A tragic setback. *Nature* 420, 116–118. (*News feature describing efforts to explain the mechanism underlying a leukaemia-like illness in a child previously cured of SCID by gene therapy*)

- Freire, J.E., Medeiros, S.C., Lopes Neto, A.V., et al., 2014. Bioethical conflicts of gene therapy: a brief critical review. *Rev. Assoc. Med. Bras.* (1992) 60, 520–524. (*The title is self-explanatory. Easy to read*)
- Kim, N., Duncan, G.A., Hanes, J., Suk, J.S., 2016. Barriers to inhaled gene therapy of obstructive lung diseases: A review. *J. Control. Release* 240, 465–488. (*Describes the rather disappointing results obtained in the cystic fibrosis trials. See also Pickett et al., 2013 below*)
- Hammer, A., Steiner, S., 2013. Gene therapy for therapeutic angiogenesis in peripheral arterial disease - a systematic review and meta-analysis of randomized, controlled trials. *Vasa* 42, 331–339.
- Marshall, E., 1999. Gene therapy death prompts review of adenovirus vector. *Science* 286, 2244–2245. (*Deals with the tragic ‘Gelsinger affair’*)
- Muller, P.Y., Brennan, F.R., 2009. Safety assessment and dose selection for first-in-human clinical trials with immunomodulatory monoclonal antibodies. *Clin. Pharmacol. Ther.* 85, 247–258. (*A sober and, at times, rather technical assessment of the safety procedures required for the ‘first-in-man’ testing of therapeutic monoclonals. Written in the wake of the TGN 1412 affair*)
- Richter, W.F., Gallati, H., Schiller, C.D., 1999. Animal pharmacokinetics of the tumor necrosis factor receptor-immunoglobulin fusion protein lenercept and their extrapolation to humans. *Drug Metab. Dispos.* 27, 21–25.
- Stobart, L., Murtagh, M.J., Rapley, T., et al., 2007. We saw human guinea pigs explode. *BMJ* 334, 566–567. (*Analysis of the, often grotesque, press coverage of the above clinical trial*)
- Woods, N.B., Bottero, V., Schmidt, M., von Kalle, C., Verma, I.M., 2006. Gene therapy: therapeutic gene causing lymphoma. *Nature* 440, 1123.
- Williams, D.A., Thrasher, A.J., 2014. Concise review: lessons learned from clinical trials of gene therapy in monogenic immunodeficiency diseases. *Stem Cells Transl. Med.* 3, 636–642.

Therapeutic uses

- Barar, J., Omidi, Y., 2012. Translational approaches towards cancer gene therapy: hurdles and hopes. *Bioimpacts* 2, 127–143.
- Borras, T., 2017. The pathway from genes to gene therapy in glaucoma: a review of possibilities for using genes as glaucoma drugs. *Asia Pac. J. Ophthalmol. (Phila.)* 6, 80–93.
- Boye, S.E., Boye, S.L., Lewin, A.S., Hauswirth, W.W., 2013. A comprehensive review of retinal gene therapy. *Mol. Ther.* 21, 509–519.
- Bradshaw, A.C., Baker, A.H., 2013. Gene therapy for cardiovascular disease: perspectives and potential. *Vasc. Pharm.* 58, 174–181.
- Chandni, A., Balasubramanian, V., Baldwin, Z.K., Conte, M.S., Schwartz, L.B., 2005. Gene therapy for the extension of vein graft patency: a review. *Vasc. Endovasc. Surg.* 39, 1–14.
- Chung, J., DiGiusto, D.L., Rossi, J.J., 2013. Combinatorial RNA-based gene therapy for the treatment of HIV/AIDS. *Expert Opin. Biol. Ther.* 13 (3), 437–445. (*Review on prophylactic gene therapy approaches for HIV*)
- Gee, P., Xu, H., Hotta, A., 2017. Cellular reprogramming, genome editing, and alternative CRISPR Cas9 technologies for precise gene therapy of Duchenne muscular dystrophy. *Stem Cells Int.* 2017, 8765154. (*A rather technical paper but gives a good idea of the power of the technique*)
- Geary, R.S., Baker, B.F., Crooke, S.T., 2015. Clinical and preclinical pharmacokinetics and pharmacodynamics of mipomersen (Kynamro[®]): a second-generation antisense oligonucleotide inhibitor of apolipoprotein B. *Clin. Pharmacokinet.* 54, 133–146.
- Ghosh, R., Walsh, S.R., Tang, T.Y., Noorani, A., Hayes, P.D., 2008. Gene therapy as a novel therapeutic option in the treatment of peripheral vascular disease: systematic review and meta-analysis. *Int. J. Clin. Pract.* 62, 1383–1390.
- Gori, J.L., Hsu, P.D., Maeder, M.L., Shen, S., Welstead, G.G., Bumcrot, D., 2015. Delivery and specificity of CRISPR-Cas9 genome editing

- technologies for human gene therapy. *Hum. Gene Ther.* 26, 443–451.
- Hammond, H.K., McKirnan, M.D., 2001. Angiogenic gene therapy for heart disease: a review of animal studies and clinical trials. *Cardiovasc. Res.* 49, 561–567. (*Comprehensive review spanning animal and human trials of gene therapy for myocardial ischaemia*)
- Maguire, A.M., High, K.A., Auricchio, A., et al., 2009. Age-dependent effects of RPE65 gene therapy for Leber's congenital amaurosis: a phase 1 dose-escalation trial. *Lancet* 374, 1597–1605.
- Nathwani, A.C., Davidoff, A.M., Linch, D.C., 2005. A review of gene therapy for haematological disorders. *Br. J. Haematol.* 128, 3–17. (*The title is self-explanatory; easy to read and comprehensive in scope*)
- Ott, M.G., Schmidt, M., Schwarzwälder, K., et al., 2006. Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. *Nat. Med.* 12, 401–409. (*Clinical trial of gene therapy to correct hereditary neutrophil dysfunction*)
- Rai, P., Malik, P., 2016. Gene therapy for hemoglobin disorders - a mini-review. *J. Rare Dis. Res. Treat.* 1, 25–31.
- Prickett, M., Jain, M., 2013. Gene therapy in cystic fibrosis. *Transl. Res.* 161, 255–264. (*Cystic fibrosis was one of the first monogenic disorders identified as a candidate for gene therapy. This review explains how and why things haven't quite worked out the way it was hoped. See also Kim et al., 2016, above*)
- Wolf, J.K., Jenkins, A.D., 2002. Gene therapy for ovarian cancer (review). *Int. J. Oncol.* 21, 461–468. (*Excellent review and broad introduction to gene therapy in general*)