1 Drugs and drug targets: an overview

1.1 What is a drug?

The medicinal chemist attempts to design and synthesize a pharmaceutical agent that has a desired biological effect on the human body or some other living system. Such a compound could also be called a 'drug', but this is a word that many scientists dislike because society views the term with suspicion. With media headlines such as 'Drugs Menace' or 'Drug Addiction Sweeps City Streets', this is hardly surprising. However, it suggests that a distinction can be drawn between drugs that are used in medicine and drugs that are abused. Is this really true? Can we draw a neat line between 'good drugs' like penicillin and 'bad drugs' like heroin? If so, how do we define what is meant by a good or a bad drug in the first place? Where would we place a so-called social drug like cannabis in this divide? What about nicotine or alcohol?

The answers we get depend on who we ask. As far as the law is concerned, the dividing line is defined in black and white. As far as the party-going teenager is concerned, the law is an ass. As far as we are concerned, the questions are irrelevant. Trying to divide drugs into two categories—safe or unsafe, good or bad—is futile and could even be dangerous.

First, let us consider the so-called 'good' drugs used in medicines. How 'good' are they? If a drug is to be truly 'good' it would have to do what it is meant to do, have no toxic or unwanted side effects, and be easy to take.

How many drugs fit these criteria?

The short answer is 'none'. There is no pharmaceutical compound on the market today that can completely satisfy all these conditions. Admittedly, some come quite close to the ideal. **Penicillin,** for example, has been one of the safest and most effective antibacterial agents ever discovered. Yet, it too has drawbacks. It cannot kill all known bacteria and, as the years have gone by, more and more bacterial strains have become resistant. Moreover, some individuals can experience severe allergic reactions to the compound.

Penicillin is a relatively safe drug, but there are some drugs that are distinctly dangerous. **Morphine** is one such example. It is an excellent analgesic, yet there are serious side effects, such as tolerance, respiratory depression, and addiction. It can even kill if taken in excess.

Barbiturates are also known to be dangerous. At Pearl Harbor, American casualties were given barbiturates as general anaesthetics before surgery. However, because of a poor understanding about how barbiturates are stored in the body, many patients received sudden and fatal overdoses. In fact, it is thought that more casualties died at the hands of the anaesthetists at Pearl Harbor than died of their wounds.

To conclude, the 'good' drugs are not as perfect as one might think.

What about the 'bad' drugs then? Is there anything good that can be said about them? Surely there is nothing we can say in defence of the highly addictive drug known as heroin?

Well, let us look at the facts about heroin. It is one of the best painkillers we know. In fact, it was named heroin at the end of the nineteenth century because it was thought to be the 'heroic' drug that would banish pain for good. Heroin went on the market in 1898, but five years later the true nature of its addictive properties became evident and the drug was speedily withdrawn from general distribution. However, heroin is still used in medicine today—under strict control, of course. The drug is called **diamorphine** and it is the drug of choice for treating patients dying of cancer. Not only does diamorphine reduce pain to acceptable levels, it also produces a euphoric effect that helps to counter the depression faced by patients close to death. Can we really condemn a drug which does that as being all 'bad'?

By now it should be evident that the division between good drugs and bad drugs is a woolly one and is not really relevant to our discussion of medicinal chemistry. All drugs have their good and bad points. Some have more good points than bad and vice versa, but, like people, they all have their own individual characteristics. So how are we to define a drug in general?

One definition could be to classify drugs as 'compounds which interact with a biological system to produce a biological response'. This definition covers all the drugs we have discussed so far, but it goes further. There are chemicals that we take every day and which have a biological effect on us. What are these everyday drugs?

One is contained in all the cups of tea, coffee, and cocoa that we consume. All of these beverages contain the stimulant **caffeine.** Whenever you take a cup of coffee, you are a drug user. We could go further. Whenever you crave a cup of coffee, you are a drug addict. Even children are not immune. They get their caffeine 'shot' from Coke or Pepsi. Whether you like it or not, caffeine is a drug. When you take it, you experience a change of mood or feeling.

So too, if you are a worshipper of the 'nicotine stick'. The biological effect is different. In this case you crave sedation or a calming influence, and it is the **nicotine** in the cigarette smoke which induces that effect.

There can be little doubt that **alcohol** is a drug and, as such, causes society more problems than all other drugs put together. One only has to study road accident statistics to appreciate that fact. If alcohol was discovered today, it would probably be restricted in exactly the same way as **cocaine.** Considered in a purely scientific way, alcohol is a most unsatisfactory drug. As many will testify, it is notoriously difficult to judge the correct dose required to gain the beneficial effect of 'happiness' without drifting into the higher dose levels that produce unwanted side effects, such as staggering down the street. Alcohol is also unpredictable in its biological effects. Either happiness or depression may result, depending on the user's state of mind. On a more serious note, **addiction** and **tolerance** in certain individuals have ruined the lives of addicts and relatives alike.

Our definition of a drug can also be used to include other compounds which may not be obvious as drugs, for example poisons and toxins. They too interact with a biological system and produce a biological response—a bit extreme, perhaps, but a response all the same. The idea of poisons acting as drugs may not appear so strange if we consider penicillin. We have no problem in thinking of penicillin as a drug, but if we were to look closely at how penicillin works, then it is really a poison. It interacts with bacteria (the biological system) and kills them (the biological response). Fortunately for us, penicillin has no such effect on human cells.

Even those drugs which do not act as poisons have the potential to become poisons—usually if they are taken in excess. We have already seen this with morphine. At low doses it is a painkiller; at high doses, it is a poison which kills by the suppression of breathing. Therefore, it is important that we treat all medicines as potential poisons and treat them with respect.

There is a term used in medicinal chemistry known as the **therapeutic index**, which indicates how safe a particular drug is. The therapeutic index is a measure of the drug's beneficial effects at a low dose versus its harmful effects at a high dose. To be more precise, the therapeutic index compares the dose level required to produce toxic effects in 50% of patients with the dose level required to produce the maximum therapeutic effects in 50% of patients. A high therapeutic index means that there is a large safety margin between beneficial and toxic doses. The values for cannabis and alcohol are 1000 and 10, respectively, which might imply that cannabis is safer and more predictable than alcohol. Indeed, a cannabis preparation **(nabiximols)** has now been approved to relieve the symptoms of multiple sclerosis. However, this does not suddenly make cannabis safe. For example, the favourable therapeutic index of cannabis does not indicate its potential toxicity if it is taken over a long period of time (chronic use). For example, the various side effects of cannabis include panic attacks, paranoid delusions, and hallucinations. Clearly, the safety of drugs is a complex matter and it is not helped by media sensationalism.

If useful drugs can be poisons at high doses or over long periods of use, does the opposite hold true? Can a poison be a medicine at low doses? In certain cases, this is found to be so.

Arsenic is well known as a poison, but arsenic-derived compounds are used as antiprotozoal and anticancer agents. **Curare** is a deadly poison which was used by the native people of South America to tip their arrows such that a minor arrow wound would be fatal, yet compounds based on the **tubocurarine** structure (the active principle of curare) are used in surgical operations to relax muscles. Under proper control and in the correct dosage, a lethal poison may well have an important medical role. Alternatively, lethal poisons can be the starting point for the development of useful drugs. For example, **ACE inhibitors** are important cardiovascular drugs that were developed, in part, from the structure of a snake venom.

As our definition covers any chemical that interacts with any biological system, we could include all pesticides and herbicides as drugs. They interact with bacteria, fungi, and insects, kill them, and thus protect plants.

Even food can act like a drug. Junk foods and fizzy drinks have been blamed for causing hyperactivity in children. It is believed that junk foods have high concentrations of certain amino acids which can be converted in the body to neurotransmitters. These are chemicals that pass messages between nerves. If an excess of these chemical messengers should accumulate, then too many messages are transmitted in the brain, leading to the disruptive behaviour observed in susceptible individuals. Allergies due to food additives and preservatives are also well recorded.

Some foods even contain toxic chemicals. Broccoli, cabbage, and cauliflower all contain high levels of a chemical that can cause reproductive abnormalities in rats. Peanuts and maize sometimes contain fungal toxins, and it is thought that fungal toxins in food were responsible for the biblical plagues. Basil contains over 50 compounds that are potentially carcinogenic, and other herbs contain some of the most potent carcinogens known. Carcinogenic compounds have also been identified in radishes, brown mustard, apricots, cherries, and plums. Such unpalatable facts might put you off your dinner, but take comfort—these chemicals are present in such small quantities that the risk is insignificant. Therein lies a great truth, which was recognized as long ago as the fifteenth century when it was stated that 'Everything is a poison, nothing is a poison. It is the dose that makes the poison'.

Almost anything taken in excess will be toxic. You can make yourself seriously ill by taking 100 aspirin tablets or a bottle of whisky or 9 kg of spinach. The choice is yours!

To conclude, drugs can be viewed as actual or potential poisons. An important principle is that of **selective toxicity**. Many drugs are effective because they are toxic to 'problem cells', but not normal cells. For example, antibacterial, antifungal, and antiprotozoal drugs are useful in medicine when they show a selective toxicity to microbial cells, rather than mammalian cells. Clinically effective anticancer agents show a selective toxicity for cancer cells over normal cells. Similarly, effective antiviral agents are toxic to viruses rather than normal cells.

Having discussed what drugs are, we shall now consider why, where, and how they act.

KEY POINTS

- Drugs are compounds that interact with a biological system to produce a biological response.
- No drug is totally safe. Drugs vary in the side effects they might have.
- The dose level of a compound determines whether it will act as a medicine or as a poison.
- The therapeutic index is a measure of a drug's beneficial effect at a low dose versus its harmful effects at higher dose. A high therapeutic index indicates a large safety margin between beneficial and toxic doses.
- The principle of selective toxicity means that useful drugs show toxicity against foreign or abnormal cells but not against normal host cells.

1.2 Drug targets

Why should chemicals, some of which have remarkably simple structures, have such an important effect on such a complicated and large structure as a human being? The answer lies in the way that the human body operates. If we could see inside our bodies to the molecular level, we would see a magnificent array of chemical reactions taking place, keeping the body healthy and functioning.

Drugs may be mere chemicals, but they are entering a world of chemical reactions with which they interact. Therefore, there should be nothing odd in the fact that they can have an effect. The surprising thing might be that they can have such *specific* effects. This is more a result of *where* they act in the body—the drug targets.

1.2.1 Cell structure

As life is made up of cells, then quite clearly drugs must act on cells. The structure of a typical mammalian cell is shown in Fig. 1.1. All cells in the human body contain a boundary wall called the **cell membrane** which encloses the contents of the cell—the **cytoplasm**. The cell membrane seen under the electron microscope consists of two identifiable layers, each of which is made up of an ordered row of phosphoglyceride molecules, such as **phosphatidylcholine** (**lecithin**) (Fig. 1.2). The outer layer of the membrane is made up of phosphatidylcholine, whereas the inner layer is made up of phosphatidylethanolamine, phosphatidylserine, and phosphatidylinositol. Each phosphoglyceride molecule consists of a small polar head-group and two long, hydrophobic (water-hating) chains.

In the cell membrane, the two layers of phospholipids are arranged such that the hydrophobic tails point towards each other and form a fatty, hydrophobic centre, while the ionic head-groups are placed at the inner and outer surfaces of the cell membrane (Fig. 1.3). This is a stable structure because the ionic, hydrophilic head-groups

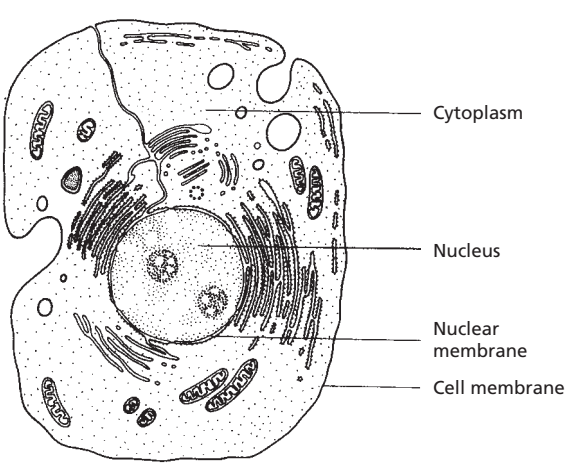


FIGURE 1.1 A typical mammalian cell. Taken from Mann, J. (1992) *Murder, Magic, and Medicine*. Oxford University Press, with permission.

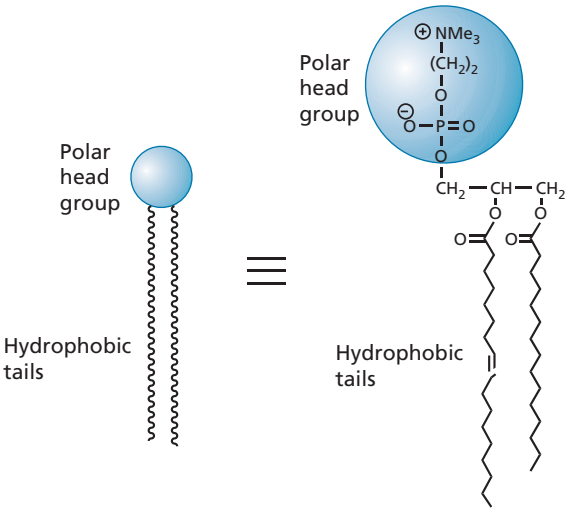


FIGURE 1.2 Phosphoglyceride structure.

interact with the aqueous media inside and outside the cell, whereas the hydrophobic tails maximize hydrophobic interactions with each other and are kept away from the aqueous environments. The overall result of this structure is to construct a fatty barrier between the cell's interior and its surroundings.

The membrane is not just made up of phospholipids, however. There are a large variety of proteins situated in the cell membrane (Fig. 1.3). Some proteins lie attached to the inner or the outer surface of the membrane. Others are embedded in the membrane with part of their structure exposed to one surface or both. The extent to which these proteins are embedded within the cell membrane structure depends on the types of amino acid present. Portions of protein that are embedded in the cell membrane have a large number of hydrophobic amino acids, whereas those portions that stick out from the surface have a large number of hydrophilic amino acids. Many surface proteins also have short chains of carbohydrates attached to them and are thus classed as **glycoproteins**.

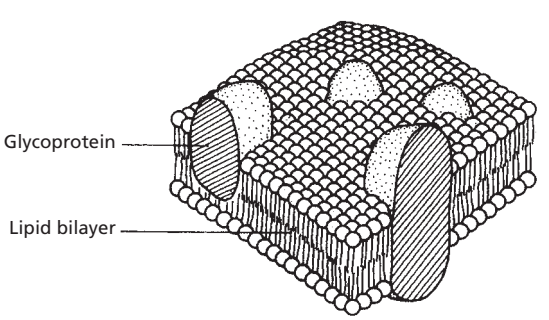


FIGURE 1.3 Cell membrane. Taken from Mann, J. (1992) *Murder, Magic, and Medicine*. Oxford University Press, with permission.

These carbohydrate segments are important in cell–cell recognition (section 10.7).

Within the cytoplasm there are several structures, one of which is the **nucleus**. This acts as the 'control centre' for the cell. The nucleus contains the genetic code—the DNA—which acts as the blueprint for the construction of all the cell's proteins. There are many other structures within a cell, such as the mitochondria, the Golgi apparatus, and the endoplasmic reticulum, but it is not the purpose of this book to look at the structure and function of these organelles. Suffice it to say that different drugs act on molecular targets at different locations in the cell.

1.2.2 Drug targets at the molecular level

We shall now move to the molecular level, because it is here that we can truly appreciate how drugs work. The main molecular targets for drugs are proteins (mainly enzymes, receptors, and transport proteins) and nucleic acids (DNA and RNA). These are large molecules (**macromolecules**) that have molecular weights measured in the order of several thousand atomic mass units. They are much bigger than the typical drug, which has a molecular weight in the order of a few hundred atomic mass units.

The interaction of a drug with a macromolecular target involves a process known as binding. There is usually a specific area of the macromolecule where this takes place, known as the **binding site** (Fig. 1.4). Typically, this takes the form of a hollow or canyon on the surface of the macromolecule allowing the drug to sink into the body of the larger molecule. Some drugs react with the binding site and become permanently attached via a covalent bond that has a bond strength of 200–400 kJ mol^{-1}. However, most drugs interact through weaker forms of interaction known as **intermolecular bonds**. These include electrostatic or ionic bonds, hydrogen bonds, van der Waals interactions, dipole–dipole interactions, and hydrophobic interactions. (It is also possible for these interactions to take place *within* a molecule, in which case they are called **intramolecular bonds**; see for example protein structure, sections 2.2 and 2.3.) None of these bonds is as strong as the covalent bonds that make up the skeleton of a molecule, and so they can be formed and then broken again. This means that an equilibrium takes place between the drug being bound and unbound to its target. The binding forces are strong enough to hold the drug for a certain period of time to let it have an effect on the target, but weak enough to allow the drug to depart once it has done its job. The length of time the drug remains at its target will then depend on the number of intermolecular bonds involved in holding it there. Drugs that have a large number of interactions are likely

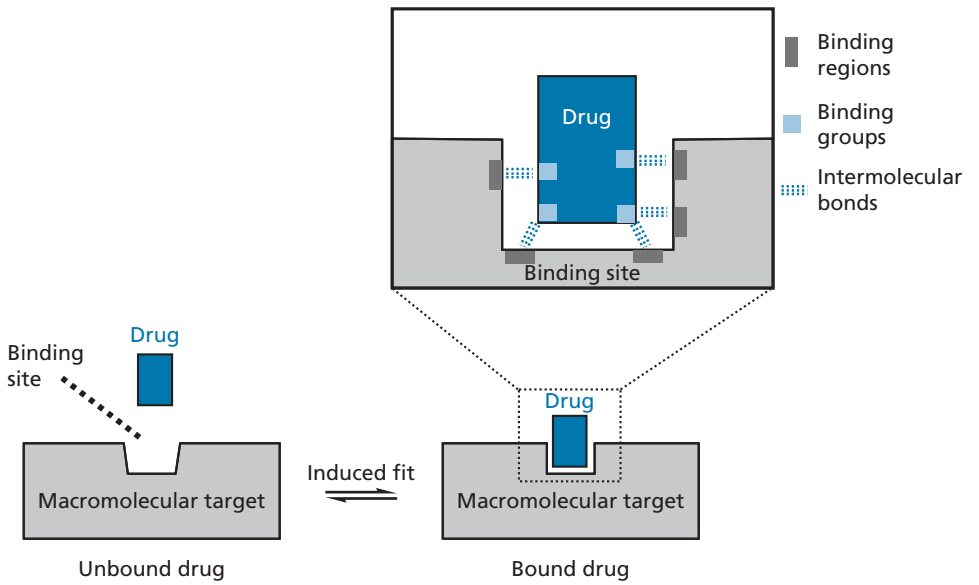


FIGURE 1.4 The equilibrium of a drug being bound and unbound to its target.

to remain bound longer than those that have only a few. The relative strength of the different intermolecular binding forces is also an important factor. Functional groups present in the drug can be important in forming intermolecular bonds with the target binding site. If they do so, they are called **binding groups**. However, the carbon skeleton of the drug also plays an important role in binding the drug to its target through van der Waals interactions. As far as the target binding site is concerned, it too contains functional groups and carbon skeletons which can form intermolecular bonds with 'visiting' drugs. The specific regions where this takes place are known as **binding regions**. The study of how drugs interact with their targets through binding interactions and produce a pharmacological effect is known as **pharmacodynamics**. Let us now consider the types of intermolecular bond that are possible.

1.3 Intermolecular bonding forces

There are several types of intermolecular bonding interactions, which differ in their bond strengths. The number and types of these interactions depend on the structure of the drug and the functional groups that are present (section 13.1 and Appendix 7). Thus, each drug may use one or more of the following interactions, but not necessarily all of them.

1.3.1 Electrostatic or ionic bonds

An ionic or electrostatic bond is the strongest of the intermolecular bonds (20–40 kJ mol^{-1}) and takes place between groups that have opposite charges, such as a carboxylate ion and an aminium ion (Fig. 1.5). The strength of the interaction is inversely proportional to the distance between the two charged atoms and it is also dependent on the nature of the environment, being stronger in hydrophobic environments than in polar environments. Usually, the binding sites of macromolecules are more hydrophobic in nature than the surface and so this enhances the effect of an ionic interaction. The drop-off in ionic bonding strength with separation is less than in other intermolecular interactions, so if an ionic interaction is possible, it is likely to be the most important initial interaction as the drug enters the binding site.

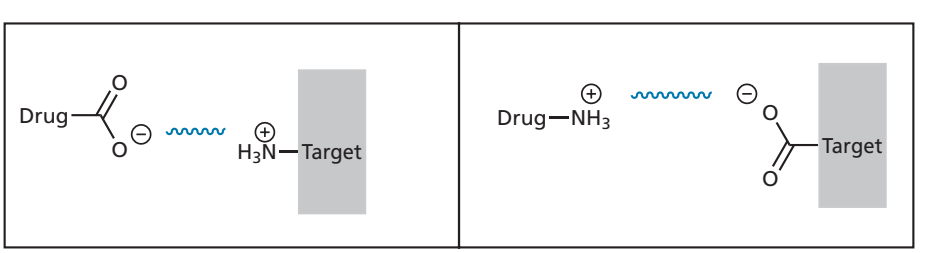


FIGURE 1.5 Electrostatic (ionic) interactions between a drug and the binding site.

FIGURE 1.6 Hydrogen bonding shown by a dashed line between a drug and a binding site (X, Y = oxygen or nitrogen; HBD = hydrogen bond donor, HBA = hydrogen bond acceptor).

1.3.2 Hydrogen bonds

A **hydrogen bond** can vary substantially in strength and normally takes place between an electron-rich heteroatom and an electron-deficient hydrogen (Fig. 1.6). The electron-rich heteroatom has to have a lone pair of electrons and is usually oxygen or nitrogen.

The electron-deficient hydrogen is usually linked by a covalent bond to an electronegative atom, such as oxygen or nitrogen. As the electronegative atom (X) has a greater attraction for electrons, the electron distribution in the covalent bond (X–H) is weighted towards the more electronegative atom and so the hydrogen gains its slight positive charge. The functional group containing this feature is known as a **hydrogen bond donor (HBD)** because it provides the hydrogen for the hydrogen bond. The functional group that provides the electron-rich atom to receive the hydrogen bond is known as the **hydrogen bond acceptor (HBA)**. Some functional groups can act both as hydrogen bond donors and hydrogen bond acceptors (e.g. OH, NH_2). When such a group is present in a binding site, it is possible that it might bind to one ligand as a hydrogen bond donor and to another as a hydrogen bond acceptor. This characteristic is given the term **hydrogen bond flip-flop**.

Hydrogen bonds have been viewed as a weak form of electrostatic interaction because the heteroatom is slightly negative and the hydrogen is slightly positive. However, there is more to hydrogen bonding than an attraction between partial charges. Unlike other intermolecular interactions, an interaction of orbitals takes place between the two molecules (Fig. 1.7). The orbital containing the lone pair of electrons on heteroatom (Y) interacts with the atomic orbitals normally involved in the covalent bond between X and H. This results in a weak form of sigma (σ) bonding and has an important directional consequence that is not evident in electrostatic bonds. The optimum orientation is where the X–H bond points directly to the lone pair on Y such that the angle formed between X, H, and Y is 180°. This is observed in very strong hydrogen bonds. However, the angle can vary between 130° and 180° for moderately strong hydrogen bonds, and can be as low as 90° for weak hydrogen bonds. The lone pair orbital of Y also has a directional property depending on its hybridization. For example, the nitrogen of a pyridine ring is sp^2 hybridized and so the lone pair points directly away from the ring and in the same plane (Fig. 1.8). The best location for a hydrogen bond donor would be the region of space indicated in the figure.

The strength of a hydrogen bond can vary widely, but most hydrogen bonds in drug–target interactions are moderate in strength, varying from 16 to 60 kJ mol^{-1}—approximately 10 times less than a covalent bond. The bond distance reflects this; hydrogen bonds are typically 1.5–2.2 Å compared with 1.0–1.5 Å for a covalent bond. The strength of a hydrogen bond depends on how strong the hydrogen bond acceptor and the hydrogen bond donor are. A good hydrogen bond acceptor has to be electronegative and have a lone pair of electrons. Nitrogen and oxygen are the most common atoms involved as hydrogen bond acceptors in biological systems. Nitrogen has one lone pair of electrons and can act as an acceptor for one hydrogen bond; oxygen has two lone pairs of electrons and can act as an acceptor for two hydrogen bonds (Fig. 1.9).

Several drugs and macromolecular targets contain a sulphur atom, which is also electronegative. However, sulphur is a weak hydrogen bond acceptor because its lone pairs are in third-shell orbitals that are larger and more

FIGURE 1.7 Orbital overlap in a hydrogen bond.

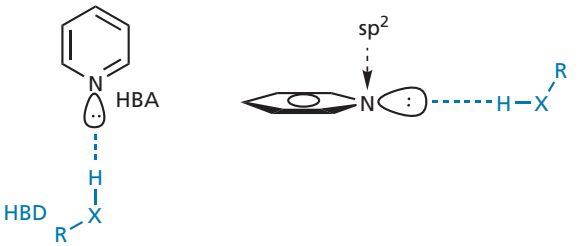


FIGURE 1.8 Directional influence of hybridization on hydrogen bonding.

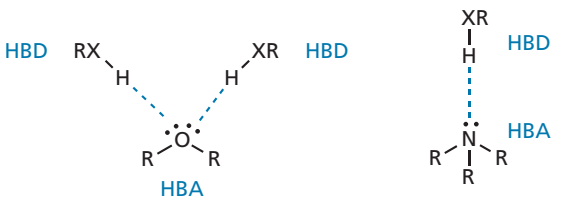


FIGURE 1.9 Oxygen and nitrogen acting as hydrogen bond acceptors (HBD = hydrogen bond donor, HBA = hydrogen bond acceptor).

diffuse. This means that the orbitals concerned interact less efficiently with the small 1s orbitals of hydrogen atoms.

Fluorine, which is present in several drugs, is more electronegative than either oxygen or nitrogen. It also has three lone pairs of electrons, which might suggest that it would make a good hydrogen bond acceptor. In fact, it is a weak hydrogen bond acceptor. It has been suggested that fluorine is so electronegative that it clings on tightly to its lone pairs of electrons, making them incapable of hydrogen bond interactions. This is in contrast to fluoride ions which are very strong hydrogen bond acceptors.

Any feature that affects the electron density of the hydrogen bond acceptor is likely to affect its ability to act as a hydrogen bond acceptor; the greater the electron density of the heteroatom, the greater its strength as a hydrogen bond acceptor. For example, the oxygen of a negatively charged carboxylate ion is a stronger hydrogen bond acceptor than the oxygen of the uncharged carboxylic acid (Fig. 1.10). Phosphate ions can also act as good hydrogen bond acceptors. Most hydrogen bond acceptors present in drugs and binding sites are neutral functional groups, such as ethers, alcohols, phenols, amides, amines, and ketones. These groups will form moderately strong hydrogen bonds.

It has been proposed that the pi (π) systems present in alkynes and aromatic rings are regions of high electron density and can act as hydrogen bond acceptors. However, the electron density in these systems is diffuse and so the hydrogen bonding interaction is much weaker than those involving oxygen or nitrogen. As a result, aromatic rings and alkynes are only likely to be significant hydrogen bond acceptors if they interact with a strong hydrogen bond donor, such as an alkylammonium ion (NHR_3^+).

More subtle effects can influence whether an atom is a good hydrogen bond acceptor or not. For example, the nitrogen atom of an aliphatic tertiary amine is a better hydrogen bond acceptor than the nitrogen of an amide or an aniline (Fig. 1.11). In the latter cases, the lone pair

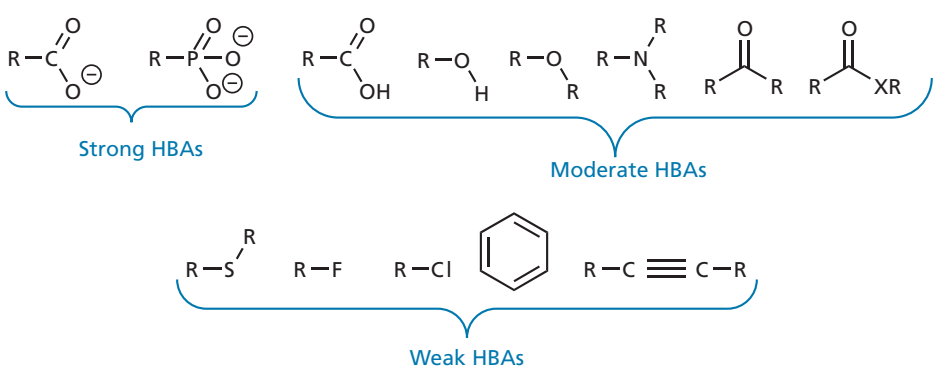


FIGURE 1.10 Relative strengths of hydrogen bond acceptors (HBAs).

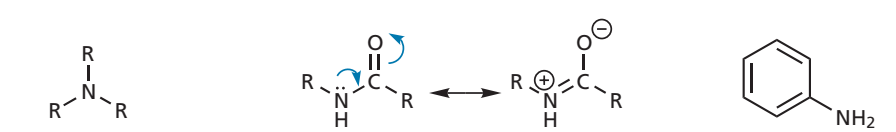


FIGURE 1.11 Comparison of different nitrogen containing functional groups as hydrogen bond acceptors (HBAs).

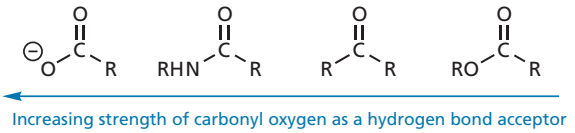


FIGURE 1.12 Comparison of carbonyl oxygens as hydrogen bond acceptors.

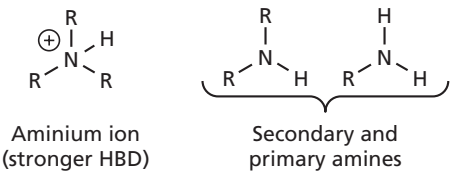


FIGURE 1.13 Comparison of hydrogen bond donors (HBDs).

of the nitrogen can interact with neighbouring π systems to form various resonance structures. As a result, it is less likely to take part in a hydrogen bond.

Similarly, the ability of a carbonyl group to act as a hydrogen bond acceptor varies depending on the functional group involved (Fig. 1.12).

It has also been observed that an sp^3 hybridized oxygen atom linked to an sp^2 carbon atom rarely acts as an HBA. This includes the alkoxy oxygen of esters and the oxygen atom present in aromatic ethers or furans.

Good hydrogen bond donors contain an electron-deficient proton linked to oxygen or nitrogen. The more electron-deficient the proton, the better it will act as a hydrogen bond donor. For example, a proton attached to a positively charged nitrogen atom acts as a stronger hydrogen bond donor than the proton of a primary or secondary amine (Fig. 1.13). Because the nitrogen is charged, it has a greater pull on the electrons surrounding it, making attached protons even more electron-deficient.

1.3.3 Van der Waals interactions

Van der Waals interactions are very weak interactions that are typically 2–4 kJ mol^{-1} in strength. They involve interactions between hydrophobic regions of different molecules, such as aliphatic substituents or the overall carbon skeleton. The electronic distribution in neutral, non-polar regions is never totally even or symmetrical, and there are always transient areas of high and low electron densities leading to temporary dipoles. The dipoles in one molecule can induce dipoles in a neighbouring molecule, leading to weak interactions between the two molecules (Fig. 1.14). Thus, an area of high electron density on one molecule can have an attraction for an area of low electron density on another molecule. The strength of these interactions falls off rapidly the further the two molecules are apart, decreasing to the seventh power of the separation. Therefore, the drug has to be close to the target binding site before the interactions become important. Van der Waals interactions are also referred to as **London forces**. Although the interactions are individually weak, there may be many such interactions between a drug and its target, and so the overall contribution of van der Waals interactions is often crucial to binding. Hydrophobic forces are also important when the non-polar regions of molecules interact (section 1.3.6).

1.3.4 Dipole–dipole and ion–dipole interactions

Many molecules have a permanent dipole moment resulting from the different electronegativities of the atoms and functional groups present. For example, a ketone has a dipole moment due to the different electronegativities of the carbon and oxygen making up the carbonyl bond. The binding site also contains functional groups, so it is inevitable that it too will have various local dipole moments. It is possible for the dipole moments of the drug and the binding site to interact as a drug approaches, aligning the drug such that the dipole moments are parallel and in opposite directions (Fig. 1.15). If this positions the drug such that other intermolecular interactions can take place between it and the target, the alignment is beneficial to both binding and activity. If not, then binding and activity may be weakened. An example of such an effect can be found in antiulcer drugs (section 25.2.8.3). The strength of dipole–dipole interactions reduces with the

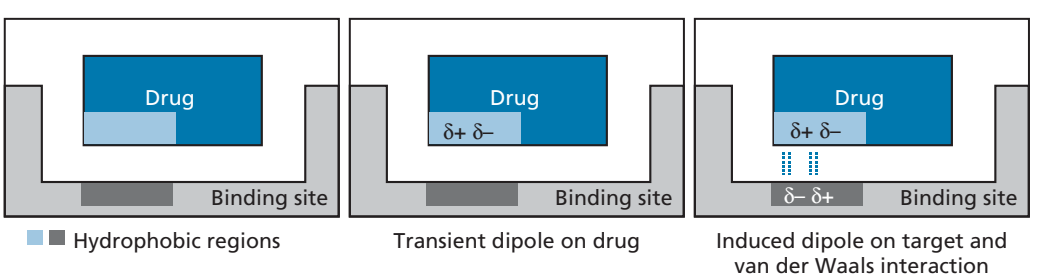


FIGURE 1.14 Van der Waals interactions between hydrophobic regions of a drug and a binding site.

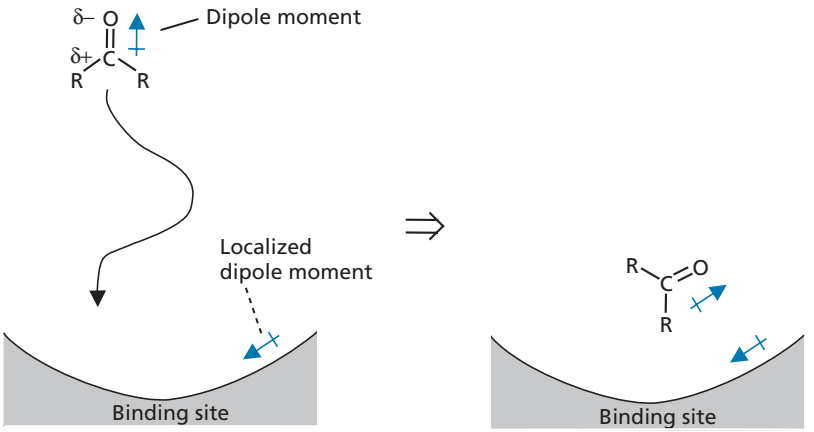


FIGURE 1.15 Dipole–dipole interactions between a drug and a binding site.

cube of the distance between the two dipoles. This means that dipole–dipole interactions fall away more quickly with distance than electrostatic interactions, but less quickly than van der Waals interactions.

An ion–dipole interaction is where a charged or ionic group in one molecule interacts with a dipole in a second molecule (Fig. 1.16). This is stronger than a dipole–dipole interaction and falls off less rapidly with separation (decreasing relative to the square of the separation).

Interactions involving an induced dipole moment have been proposed. There is evidence that an aromatic ring can interact with an ionic group such as a quaternary ammonium ion. Such an interaction is feasible if the positive charge of the quaternary ammonium group distorts the π electron cloud of the aromatic ring to produce a dipole moment where the face of the aromatic ring is electron-rich and the edges are electron-deficient (Fig. 1.17). This is also called a **cation-pi interaction**. An important neurotransmitter called **acetylcholine** forms this type of interaction with its binding site (section 22.5).

1.3.5 Repulsive interactions

So far we have concentrated on attractive forces, which increase in strength the closer the molecules approach each other. Repulsive interactions are also important.

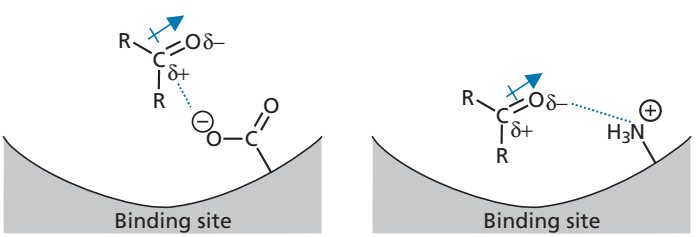


FIGURE 1.16 Ion–dipole interactions between a drug and a binding site.

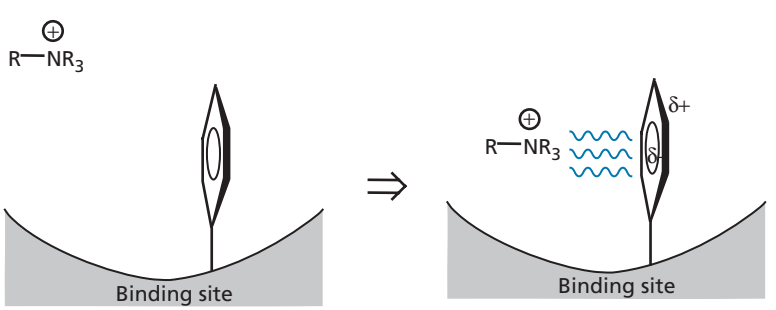


FIGURE 1.17 Induced dipole interaction between an alkylammonium ion and an aromatic ring.

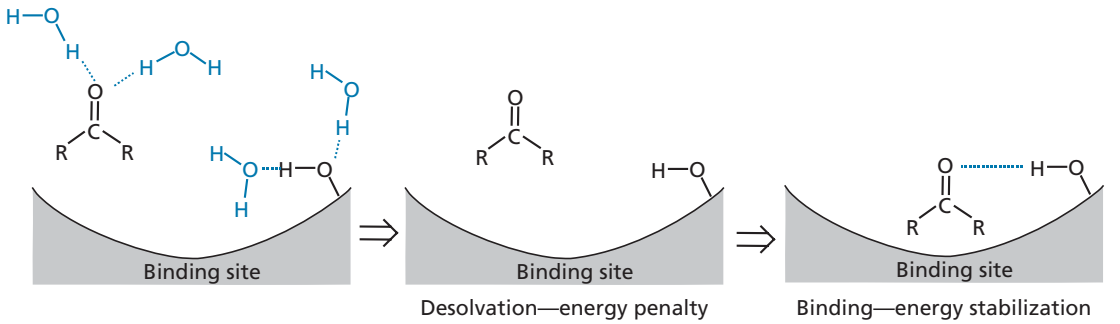


FIGURE 1.18 Desolvation of a drug and its target binding site prior to binding.

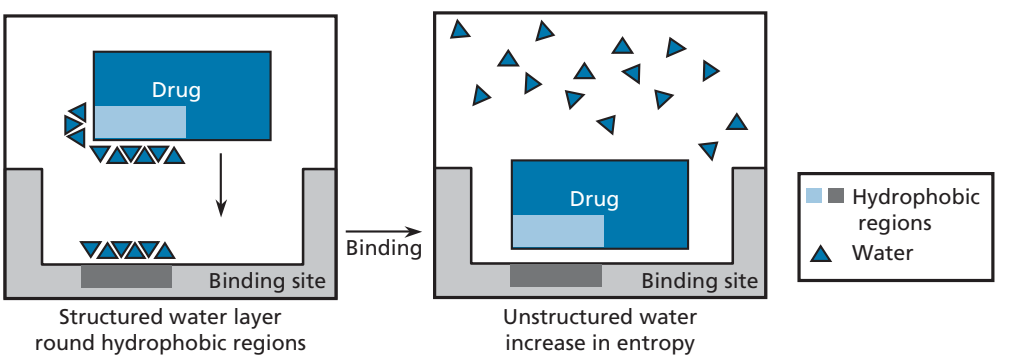


FIGURE 1.19 Hydrophobic interactions.

Otherwise, there would be nothing to stop molecules trying to merge with each other! If molecules come too close, their molecular orbitals start to overlap and this results in repulsion. Other forms of repulsion are related to the types of groups present in both molecules. For example, two charged groups of identical charge are repelled.

1.3.6 The role of water and hydrophobic interactions

A crucial feature that is often overlooked when considering the interaction of a drug with its target is the role of water. The macromolecular targets in the body exist in an aqueous environment and the drug has to travel through that environment in order to reach its target; therefore, both the drug and the macromolecule are solvated with water molecules before they meet each other. The water molecules surrounding the drug and the target binding site have to be stripped away before the interactions described above can take place (Fig. 1.18). This requires energy and if the energy required to desolvate both the drug and the binding site is greater than the stabilization energy gained by the binding interactions, then the drug may be ineffective. In certain cases, it has even proved beneficial to remove a polar binding group from a drug in order to lower its energy of desolvation. For example, this was carried out during the development of the antiviral drug **ritonavir** (section 20.7.4.4).

Sometimes polar groups are added to a drug to increase its water solubility. If this is the case, it is important that such groups are positioned in such a way that they protrude from the binding site when the drug binds; in other words, they are solvent-accessible or solvent-exposed. In this way, the water that solvates this highly polar group does not have to be stripped away and there is no energy penalty when the drug binds to its target (see section 21.6.2.1 and Case study 5).

It is not possible for water to solvate the non-polar or hydrophobic regions of a drug or its target binding site. Instead, the surrounding water molecules form stronger-than-usual interactions with each other, resulting in a more ordered layer of water next to the non-polar surface. This represents a negative entropy due to the increase in order. When the hydrophobic region of a drug interacts with a hydrophobic region of a binding site, these water molecules are freed and become less ordered (Fig. 1.19). This leads to an increase in entropy and a gain in binding energy.* The interactions involved are small at 0.1–0.2 kJ mol^{-1} for each $Å^2$ of hydrophobic surface, but overall they can be substantial. Sometimes, a hydrophobic region in the drug may not be sufficiently close to a hydrophobic

* The free energy gained by binding (ΔG) is related to the change in entropy (ΔS) by the equation $\Delta G = \Delta H - T\Delta S$. If entropy increases, ΔS is positive, which makes ΔG more negative. The more negative ΔG is, the more likely binding will take place.

region in the binding site and water may be trapped between the two surfaces. The entropy increase is not so substantial in that case and there is a benefit in designing a better drug that fits more snugly.

1.4 Pharmacokinetic issues and medicines

Pharmacodynamics is the study of how a drug binds to its target binding site and produces a pharmacological effect. However, a drug capable of binding to a particular target is not necessarily going to be useful as a clinical agent or medicine. For that to be the case, the drug not only has to bind to its target, it has to reach it in the first place. For an orally administered drug, that involves a long journey with many hazards to be overcome. The drug has to survive stomach acids then digestive enzymes in the intestine. It has to be absorbed from the gut into the blood supply and then it has to survive the liver where enzymes try to destroy it (drug metabolism). It has to be distributed round the body and not get mopped up by fat tissue. It should not be excreted too rapidly or else frequent doses will be required to maintain activity. However, it should not be excreted too slowly or its effects could linger on longer than required. The study of how a drug is absorbed, distributed, metabolized, and excreted (known as ADME in the pharmaceutical industry) is called **pharmacokinetics**. Pharmacokinetics has sometimes been described as 'what the body does to the drug' as opposed to pharmacodynamics—'what the drug does to the body'.

There are many ways in which medicinal chemists can design a drug to improve its pharmacokinetic properties, but the method by which the drug is formulated and administered is just as important. Medicines are not just composed of the active pharmaceutical agent. For example, a pill contains a whole range of chemicals that are present to give structure and stability to the pill, and also to aid the delivery and breakdown of the pill at the desired part of the gastrointestinal tract.

KEY POINTS

- Drugs act on molecular targets located in the cell membrane of cells or within the cells themselves.
- Drug targets are macromolecules that have a binding site into which the drug fits and binds.
- Most drugs bind to their targets by means of intermolecular bonds.
- Pharmacodynamics is the study of how drugs interact with their targets and produce a pharmacological effect.
- Electrostatic or ionic interactions occur between groups of opposite charge.
- Hydrogen bonds occur between an electron-rich heteroatom and an electron-deficient hydrogen.
- The functional group providing the hydrogen for a hydrogen bond is called the hydrogen bond donor. The functional group that interacts with the hydrogen in a hydrogen bond is called the hydrogen bond acceptor.
- Van der Waals interactions take place between non-polar regions of molecules and are caused by transient dipole–dipole interactions.
- Ion–dipole and dipole–dipole interactions are a weak form of electrostatic interaction.
- Hydrophobic interactions involve the displacement of ordered layers of water molecules which surround hydrophobic regions of molecules. The resulting increase in entropy contributes to the overall binding energy.
- Polar groups have to be desolvated before intermolecular interactions take place. This results in an energy penalty.
- The pharmacokinetics of a drug relate to its absorption, distribution, metabolism, and excretion in the body.

1.5 Classification of drugs

There are four main ways in which drugs might be classified or grouped.

By pharmacological effect Drugs can be classified depending on the biological or pharmacological effect that they have, for example analgesics, antipsychotics, antihypertensives, anti-asthmatics, and antibiotics. This is useful if one wishes to know the full scope of drugs available for a certain ailment, but it means that the drugs included are numerous and highly varied in structure. This is because there are a large variety of targets at which drugs could act in order to produce the desired effect. It is therefore not possible to compare different painkillers and expect them to look alike or to have some common mechanism of action.

The chapters on antibacterial, antiviral, anticancer, and anti-ulcer drugs (Chapters 19–21 and 25) illustrate the variety of drug structures and mechanisms of action that are possible when drugs are classified according to their pharmacological effect.

By chemical structure Many drugs which have a common skeleton are grouped together, for example penicillins, barbiturates, opiates, steroids, and catecholamines. In some cases, this is a useful classification as the biological activity and mechanism of action is the same for the structures involved, for example the antibiotic activity of penicillins. However, not all compounds with similar chemical structures have the same biological action. For example, steroids share a similar tetracyclic structure, but they have very different effects in the body. In this text, various groups of structurally related drugs are

discussed, for example penicillins, cephalosporins, sulphonamides, opioids, and glucocorticoids (sections 19.4 and 19.5, Chapter 24 and Case study 6). These are examples of compounds with a similar structure and similar mechanism of action. However, there are exceptions. Most sulphonamides are used as antibacterial agents, but there are a few which have totally different medical applications.

By target system Drugs can be classified according to whether they affect a certain target system in the body. An example of a target system is where a neurotransmitter is synthesized, released from its neuron, interacts with a protein target, and is either metabolized or reabsorbed into the neuron. This classification is a bit more specific than classifying drugs by their overall pharmacological effect. However, there are still several different targets with which drugs could interact in order to interfere with the system and so the drugs included in this category are likely to be quite varied in structure because of the different mechanisms of action that are involved. In Chapters 22 and 23 we look at drugs that act on target systems—the cholinergic and the adrenergic system respectively.

By target molecule Some drugs are classified according to the molecular target with which they interact. For example, anticholinesterases (sections 22.12–22.15) are drugs which act by inhibiting the enzyme **acetylcholinesterase**. This is a more specific classification as we have now identified the precise target at which the drugs act. In this situation we might expect some structural similarity between the agents involved and a common mechanism of action, although this is not an inviolable assumption. However, it is easy to lose the wood for the trees and to lose sight of why it is useful to have drugs which switch off a particular enzyme or receptor. For example, it is not intuitively obvious why an anticholinesterase agent could be useful in treating Alzheimer's disease or glaucoma.

1.6 Naming of drugs and medicines

The vast majority of chemicals that are synthesized in medicinal chemistry research never make it to the market place and it would be impractical to name them all. Instead, research groups label them with a code which usually consists of letters and numbers. The letters are specific to the research group undertaking the work and the number is specific for the compound. Thus, Ro31-8959, ABT-538, and MK-639 were compounds prepared by Roche, Abbott, and Merck pharmaceuticals respectively. If the compounds concerned show promise as therapeutic drugs they are taken into development and named. For example, the above compounds showed promise as anti-HIV drugs and were named **saquinavir, ritonavir**, and **indinavir** respectively. Finally, if the drugs prove successful and are marketed as medicines, they are given a proprietary, brand, or trade name, which only the company can use. For example, the above compounds were marketed as **Fortovase®, Norvir®** and **Crixivan®** respectively (note that brand names always start with a capital letter and have the symbol R or TM to indicate that they are registered brand names). The proprietary names are also specific for the preparation or formulation of the drug. For example, Fortovase® (or Fortovase™) is a preparation containing 200 mg of saquinavir in a gel-filled, beige-coloured capsule. If the formulation is changed, then a different name is used. For example, Roche sell a different preparation of saquinavir called **Invirase®** which consists of a brown/green capsule containing 200 mg of saquinavir as the mesylate salt. When a drug's patent has expired, it is possible for any pharmaceutical company to produce and sell that drug as a generic medicine. However, they are not allowed to use the trade name used by the company that originally invented it. European law requires that generic medicines are given a **recommended International Non-proprietary Name** (rINN), which is usually identical to the name of the drug. In the UK, such drugs were given a **British Approved Name** (BAN), but these have now been modified to fall in line with rINNs.

As the naming of drugs is progressive, early research articles in the literature may only use the original letter/number code as the name of the drug had not been allocated at the time of publication.

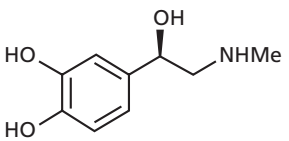
Throughout this text, the names of the active constituents are used rather than the trade names, although the trade name may be indicated if it is particularly well known. For example, it is indicated that **sildenafil** is **Viagra®** and that **paclitaxel** is **Taxol®**. If you wish to find out the trade name for a particular drug, these are listed in Appendix 6. If you wish to 'go the other way', Appendix 7 contains trade names and directs you to the relevant compound name. Only those drugs covered in the text are included and if you cannot find the drug you are looking for, you should refer to other textbooks or formularies such as the British National Formulary (see 'General further reading').

KEY POINTS

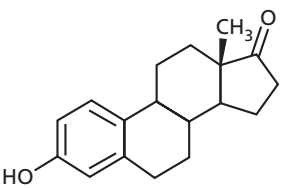
- Drugs can be classified by their pharmacological effect, their chemical structure, their effect on a target system, or their effect on a target structure.
- Clinically useful drugs have a trade (or brand) name, as well as a recommended international non-proprietary name.
- Most structures produced during the development of a new drug are not considered for the clinic. They are identified by simple codes that are specific to each research group.

QUESTIONS

1. The hormone adrenaline interacts with proteins located on the surface of cells and does not cross the cell membrane. However, larger steroid molecules, such as estrone, cross cell membranes and interact with proteins located in the cell nucleus. Why is a large steroid molecule able to cross the cell membrane when a smaller molecule such as adrenaline cannot?



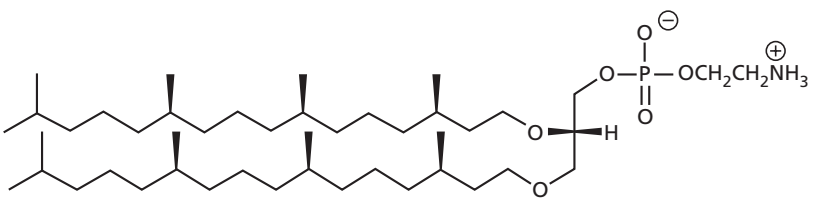
Adrenaline



Estrone

2. Valinomycin is an antibiotic which is able to transport ions across cell membranes and disrupt the ionic balance of the cell. Find out the structure of valinomycin and explain why it is able to carry out this task.

3. Archaea are microorganisms that can survive in extreme environments, such as high temperature, low pH, or high salt concentrations. It is observed that the cell membrane phospholipids in these organisms (see Structure I below) are markedly different from those in eukaryotic cell membranes. What differences are present and what function might they serve?

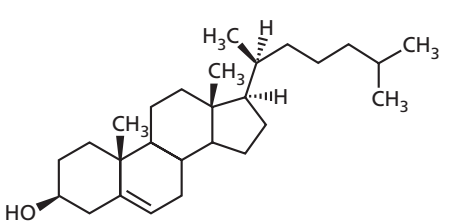


Structure I

4. Teicoplanin is an antibiotic which 'caps' the building blocks used in the construction of the bacterial cell wall such that they cannot be linked up. The cell wall is a barrier surrounding the bacterial cell membrane and the building blocks are anchored to the outside of this cell membrane prior to their incorporation into the cell wall. Teicoplanin contains a very long alkyl substituent which plays no role in the capping mechanism. However, if this substituent is absent, activity drops. What role do you think this alkyl substituent might serve?

5. The Ras protein is an important protein in signalling processes within the cell. It exists freely in the cell cytoplasm, but must become anchored to the inner surface of the cell membrane in order to carry out its function. What kind of modification to the protein might take place to allow this to happen?

6. Cholesterol is an important constituent of eukaryotic cell membranes and affects the fluidity of the membrane. Consider the structure of cholesterol (shown below) and suggest how it might be orientated in the membrane.



Cholesterol

7. Most unsaturated alkyl chains in phospholipids are *cis* rather than *trans*. Consider the *cis*-unsaturated alkyl chain in the phospholipid shown in Fig. 1.2. Redraw this chain to give a better representation of its shape and compare it with the shape of its *trans*-isomer. What conclusions can you make regarding the packing of such chains in the cell membrane and the effect on membrane fluidity?

8. The relative strength of carbonyl oxygens as hydrogen bond acceptors is shown in Fig. 1.12. Suggest why the order is as shown.

9. Consider the structures of adrenaline, estrone, and cholesterol and suggest what kind of intermolecular interactions are possible for these molecules and where they occur.

10. Using the index and Appendix 6, identify the structures and trade names for the following drugs—amoxicillin, ranitidine, gefitinib, and atracurium.

FURTHER READING

Hansch, C., Sammes, P. G., and Taylor, J. B. (eds) (1990) Classification of drugs. *Comprehensive Medicinal Chemistry*, Vol. 1, Chapter 3.1. Pergamon Press, ISBN 0-08-037057-8.

Howard, J. A. K., Hoy, V. J., O'Hagan, D., and Smith, G. T. (1996) How good is fluorine as a hydrogen bond acceptor? *Tetrahedron* **52**, 12613–12622.

Jeffrey, G. A. (1991) *Hydrogen Bonding in Biological Structures*. Springer-Verlag, London.

Kubinyi, H. (2001) Hydrogen bonding: The last mystery in drug design? In: Testa, B. (ed.) *Pharmacokinetic Optimisation in Drug Research*. Wiley, 513–24.

Mann, J. (1992) *Murder, Magic, and Medicine,* Chapter 1. Oxford University Press, Oxford.

Meyer, E. G., Botos, I., Scapozza, L., and Zhang, D. (1995) Backward binding and other structural surprises. *Perspectives in Drug Discovery and Design* **3**, 168–195.

Page, C., Curtis, M., Sutter, M., Walker, M., and Hoffman, B. (2002) Drug names and drug classification systems. *Integrated Pharmacology*, 2nd edn, Chapter 2. Mosby, St Louis, MO.

Titles for general further reading are listed on p.763.

2 Protein structure and function

The vast majority of drugs used in medicine are targeted to proteins, such as receptors, enzymes, and transport proteins. Therefore, it is important to understand protein structure in order to understand drug action on proteins. Proteins have four levels of structure: primary, secondary, tertiary, and quaternary.

2.1 The primary structure of proteins

The primary structure is the order in which the individual amino acids making up the protein are linked together through peptide bonds (Fig. 2.1). The 20 common amino acids found in humans are listed in Table 2.1, with the three- and one-letter codes often used to represent them. The structures of the amino acids are shown in Appendix 1. The primary structure of **Met-enkephalin** (one of the body's own painkillers) is shown in Fig. 2.2.

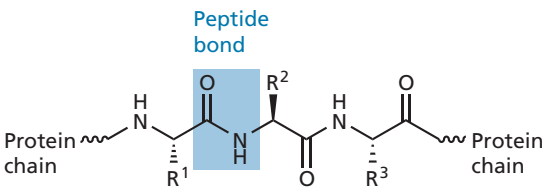


FIGURE 2.1 Primary structure of proteins (R^1, R^2, and R^3 = amino acid side chains).

The peptide bond in proteins is planar in nature as a result of the resonance structure shown in Fig. 2.3. This gives the peptide bond a significant double bond character, which prevents rotation. As a result, bond rotation in the protein backbone is only possible for the bonds on

TABLE 2.1 The 20 common amino acids found in humans

Synthesized in the human body			Essential to the diet		
Amino acid	Codes Three-letter	One-letter	Amino acid	Codes Three-letter	One-letter
Alanine	Ala	A	Histidine	His	H
Arginine	Arg	R	Isoleucine	Ile	I
Asparagine	Asn	N	Leucine	Leu	L
Aspartic acid	Asp	D	Lysine	Lys	K
Cysteine	Cys	C	Methionine	Met	M
Glutamic acid	Glu	E	Phenylalanine	Phe	F
Glutamine	Gln	Q	Threonine	Thr	T
Glycine	Gly	G	Tryptophan	Trp	W
Proline	Pro	P	Valine	Val	V
Serine	Ser	S			
Tyrosine	Tyr	Y			

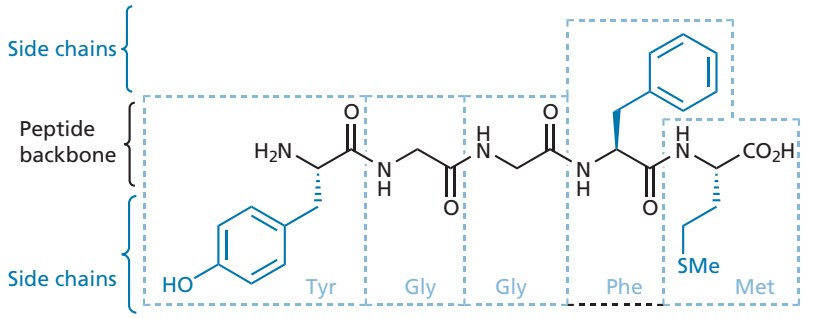


FIGURE 2.2 Met-enkephalin. The short hand notation for this peptide is H-Tyr-Gly-Gly-Phe-Met-OH or YGGFM.

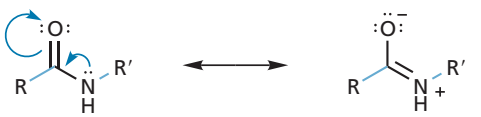


FIGURE 2.3 The planar peptide bond (bond rotation allowed for coloured bonds only).

trans conformation (favoured)

cis conformation (unfavoured)

FIGURE 2.4 *Trans* and *cis* conformations of the peptide bond.

either side of each peptide bond. This has an important consequence for protein tertiary structure.

There are two possible conformations for the peptide bond (Fig. 2.4). The *trans* conformation is the one that is normally present in proteins as the *cis* conformation leads to a steric clash between the residues. However, the *cis* conformation is possible for peptide bonds next to a proline residue.

2.2 The secondary structure of proteins

The secondary structure of proteins consists of regions of ordered structure adopted by the protein chain. In structural proteins, such as wool and silk, secondary structures are extensive and determine the overall shape and properties of such proteins. However, there are also regions of secondary structure in most other proteins. There are three main secondary structures: the α-helix, β-pleated sheet, and β-turn.

2.2.1 The α-helix

The α-helix results from coiling of the protein chain such that the peptide bonds making up the backbone are able to form hydrogen bonds between each other. These hydrogen bonds are directed along the axis of the helix, as shown in Fig. 2.5. The side chains of the component amino acids stick out at right angles from the helix, thus minimizing steric interactions and further stabilizing the structure. Other, less common, types of helices can occur in proteins, such as the 3(10)-helix, which is more stretched than the ideal α-helix, and the π-helix, which is more compact and extremely rare.

Test Your Understanding and Practise Your Molecular Modelling with Exercise 2.1.

2.2.2 The β-pleated sheet

The β-pleated sheet is a layering of protein chains one on top of another, as shown in Fig. 2.6. Here, too, the structure is held together by hydrogen bonds between the peptide chains. The side chains are situated at right angles to the sheets—once again to reduce steric interactions. The chains in β-sheets can run in opposite directions (antiparallel) or in the same direction (parallel) (Fig. 2.7).

2.2.3 The β-turn

A β-turn allows the polypeptide chain to turn abruptly and go in the opposite direction. This is important in allowing the protein to adopt a more globular compact shape. A hydrogen bonding interaction between the first and third peptide bond of the turn is important in stabilizing the turn (Fig. 2.8). Less abrupt changes in the direction of the polypeptide chain can also take place through longer loops, which are less regular in their structure, but often rigid and well defined.

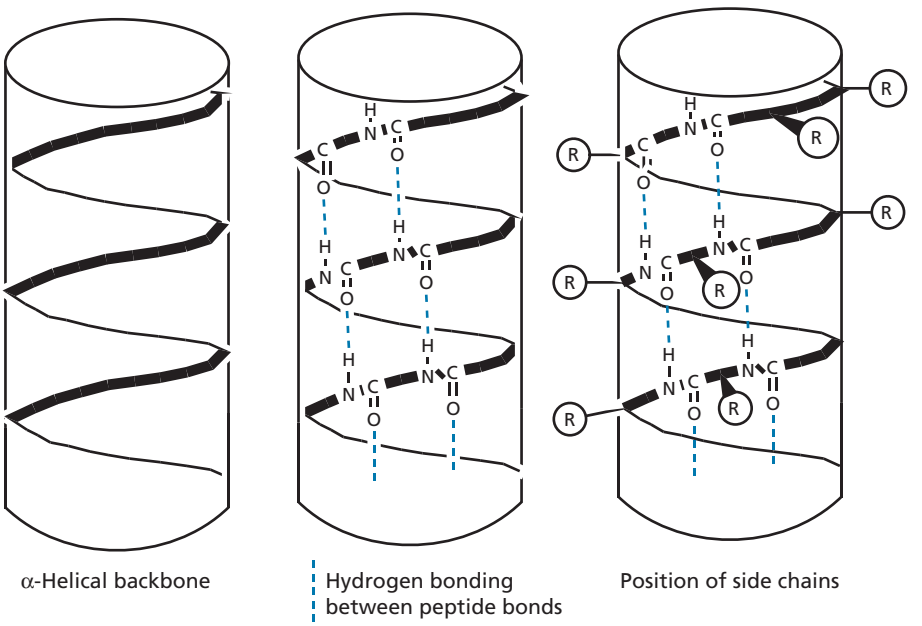


FIGURE 2.5 The α-helix for proteins showing intramolecular hydrogen bonds and the position of side chains.

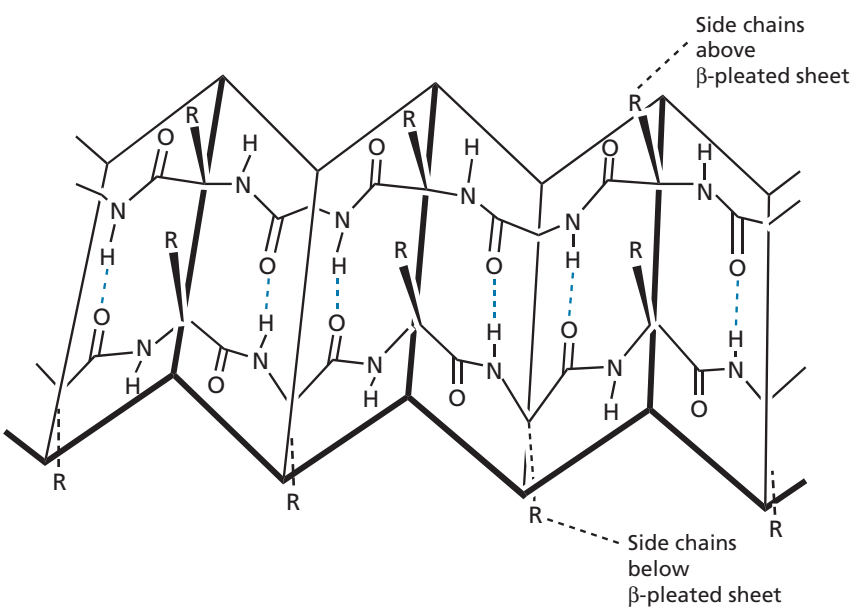


FIGURE 2.6 The β-pleated sheet (antiparallel arrangement).

2.3 The tertiary structure of proteins

The tertiary structure is the overall three-dimensional shape of a protein. Structural proteins are quite ordered in shape, whereas globular proteins, such as enzymes and receptors (Chapters 3 and 4), fold up to form more complex structures. The tertiary structure of enzymes and receptors is crucial to their function and also to their interaction with drugs; therefore, it is important to appreciate the forces that control tertiary structure.

Globular proteins often contain regions of ordered secondary structure, the extent of which varies from protein to protein. For example, **cyclin-dependent kinase 2** (a protein that catalyses phosphorylation reactions) has several regions of α-helices and β-pleated sheets

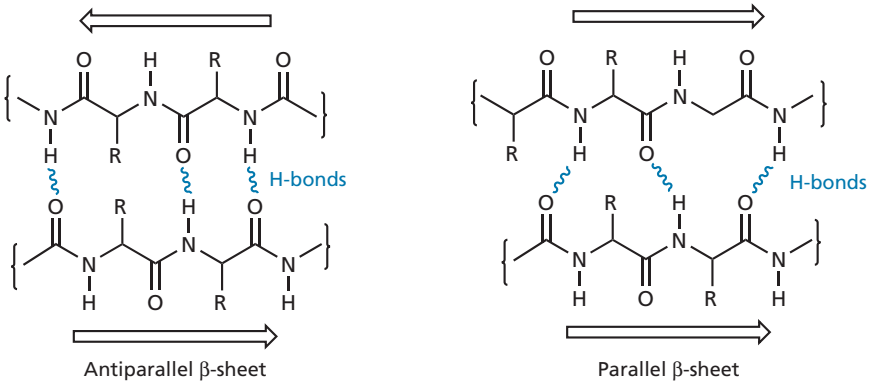


FIGURE 2.7 Hydrogen bonding in antiparallel and parallel β-sheets (the arrows are pointing to the *C*-terminal end of the chain).

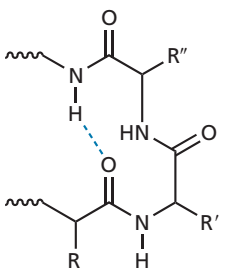


FIGURE 2.8 The β-turn showing hydrogen bonding between the first and third peptide bond.

(Fig. 2.9), whereas the digestive enzyme **chymotrypsin** has very little secondary structure. Nevertheless, the protein chains in both cyclin-dependent kinase 2 and chymotrypsin fold up to form a complex, but distinctive, globular shape. How does this come about?

At first sight, the three-dimensional structure of cyclin-dependent kinase 2 looks like a ball of string after a cat has been at it. In fact, the structure shown is a very precise shape which is taken up by every molecule of this protein, and which is determined by the protein's primary structure.* Indeed, it is possible to synthesize naturally occurring proteins in the laboratory which automatically adopt the same three-dimensional structure and function as the naturally occurring protein. The HIV-1 protease enzyme is an example (section 20.7.4.1).

This poses a problem. Why should a chain of amino acids take up such a precise three-dimensional shape? At first sight, it does not make sense. If we place a length of string on the table it does not fold itself up into a precise complex shape. So why should a chain of amino acids do such a thing?

* Some proteins contain species known as **cofactors** (e.g. metal ions or small organic molecules), which also have an effect on tertiary structure.

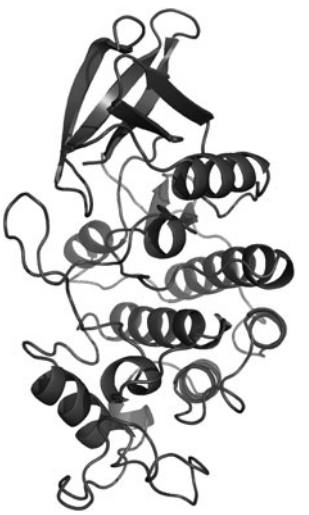


FIGURE 2.9 The pdb file (1hcl) for human cyclin-dependent kinase 2 (CDK2), where cylinders represent α-helices and arrows represent β-sheets. A pdb file contains the three-dimensional structural information for a protein and can be downloaded from the Brookhaven protein data bank. Each protein structure file is given a code, for example 1hcl.

Test Your understanding and practise your molecular modelling with Exercise 2.2.

The answer lies in the fact that a protein is not just a bland length of string. It contains a range of different chemical functional groups along its length—not only peptide links, but also the side chains of each amino acid. These can interact with each other such that there is either an attractive interaction or a repulsive interaction. Thus, the protein will twist and turn to minimize the unfavourable interactions and maximize the favourable ones until the most stable shape or conformation is found—the tertiary structure (Fig. 2.10).

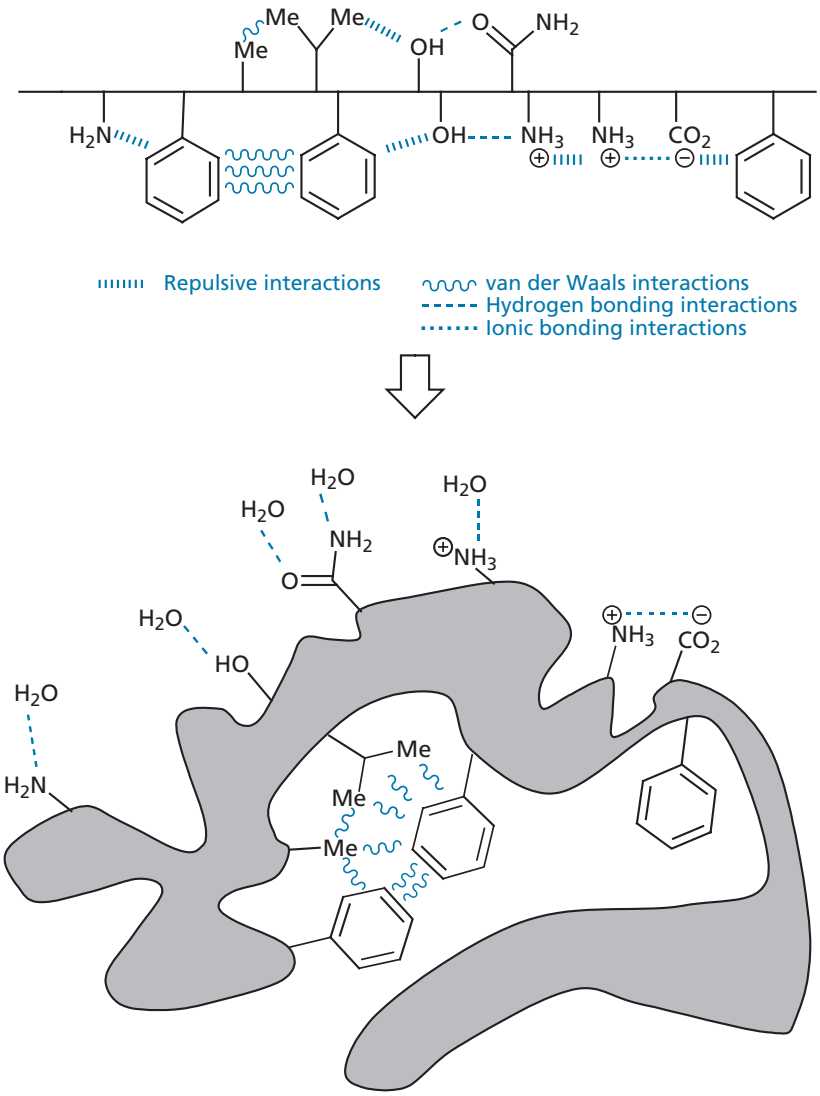


FIGURE 2.10 Tertiary structure formation as a result of intramolecular interactions.

With the exception of disulphide bonds, the bonding interactions involved in tertiary structure are the same as the **intermolecular bonds** described in section 1.3. The latter occur between different molecules, whereas the bonds controlling protein tertiary structure occur within the same molecule, and so they are called **intramolecular bonds**. Nevertheless, the principles described in section 1.3 are the same.

2.3.1 Covalent bonds—disulphide links

Cysteine has a residue containing a thiol group capable of forming a covalent bond in the protein tertiary structure. When two such residues are close together, a covalent disulphide bond can be formed as a result of oxidation. A covalent bridge is thus formed between two different parts of the protein chain (Fig. 2.11). It should be noted that the two cysteine residues involved in this bond formation may be far apart from each other in the primary structure of the protein, but are brought close together as a result of protein folding.

2.3.2 Ionic or electrostatic bonds

An ionic bond or salt bridge can be formed between the carboxylate ion of an acidic residue, such as aspartic acid or glutamic acid, and the aminium ion of a basic residue, such as lysine, arginine, or histidine (Fig. 2.12). This is the strongest of the intramolecular bonds.

2.3.3 Hydrogen bonds

Hydrogen bonds can be viewed as a weak form of ionic interaction as they involve interactions between atoms

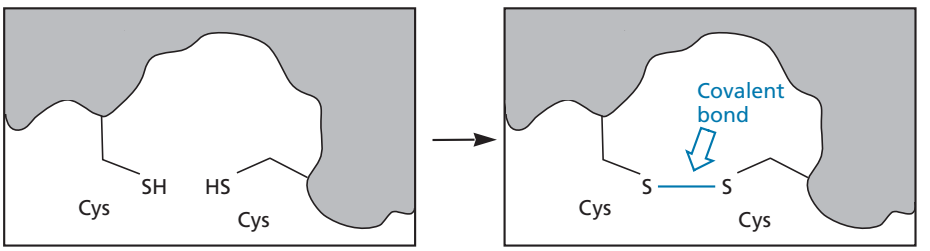


FIGURE 2.11 The formation of a disulphide covalent bond between two cysteine side chains.

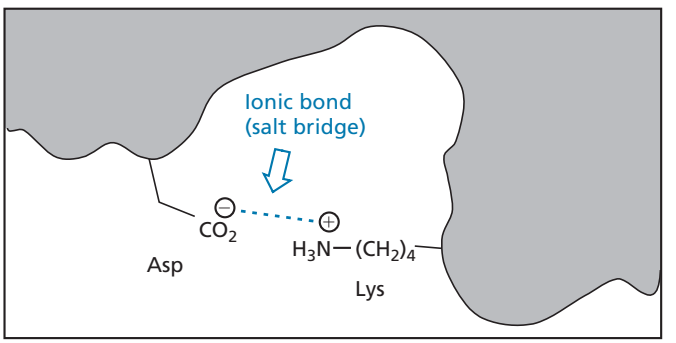


FIGURE 2.12 Ionic bonding between an aspartate side chain and a lysine side chain.

having partial charges. They can be formed between a large number of amino acid side chains, such as serine, threonine, aspartic acid, glutamic acid, glutamine, lysine, arginine, histidine, tryptophan, tyrosine, and asparagine. Two examples are shown in Fig. 2.13.

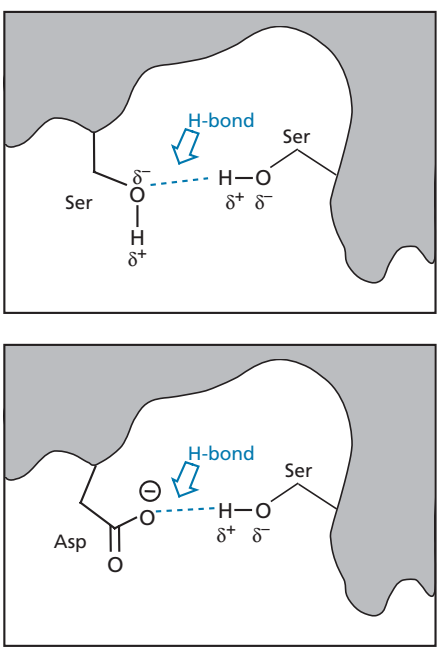


FIGURE 2.13 Hydrogen bonding between amino acid side chains.

2.3.4 Van der Waals and hydrophobic interactions

Van der Waals interactions are weaker interactions than hydrogen bonds and can take place between two hydrophobic regions of the protein. For example, they can take place between two alkyl groups (Fig. 2.14). The amino acids alanine, valine, leucine, isoleucine, phenylalanine, and proline all have hydrophobic side chains capable of interacting with each other by van der Waals interactions. The side chains of other amino acids, such as methionine, tryptophan, threonine, and tyrosine, contain polar functional groups, but the side chains also have a substantial hydrophobic character and so van der Waals interactions

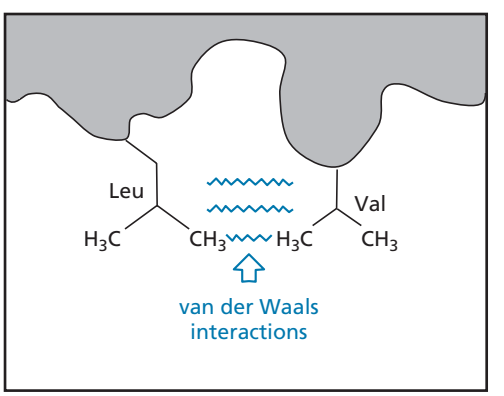


FIGURE 2.14 Van der Waals interactions between amino acid side chains.

are also possible for these amino acids. Hydrophobic interactions (section 1.3.6) are also important in the coming together of hydrophobic residues.

2.3.5 Relative importance of bonding interactions

We might expect the relative importance of the bonding interactions in protein tertiary structure to follow the same order as their strengths: covalent, ionic, hydrogen bonding, and, finally, van der Waals. In fact, the opposite is generally true. Usually, the most important bonding interactions are those due to van der Waals interactions and hydrogen bonding, while the least important interactions are those due to covalent and ionic bonding.

There are two reasons for this. Firstly, in most proteins there are more possible opportunities for van der Waals and hydrogen bonding interactions than for covalent or ionic bonding. We only need to consider the relative number of amino acids present in a typical globular protein to see why. The only amino acid that can form a covalent disulphide bond is cysteine, whereas there are many more amino acids that can interact with each other through hydrogen bonding and van der Waals interactions.

Having said that, there *are* examples of proteins with a large number of disulphide bridges, where the relative importance of the covalent link to tertiary structure is more significant. Disulphide links are also more significant in small polypeptides such as the peptide hormones **vasopressin** and **oxytocin** (Fig. 2.15). Nevertheless, in most proteins, disulphide links play a minor role in controlling tertiary structure.

As far as ionic bonding is concerned, there is only a limited number of amino acids with residues capable of forming ionic bonds, and so these, too, are outnumbered by the number of residues capable of forming hydrogen bonds or van der Waals interactions.

There is a second reason why van der Waals interactions are normally the most important form of bonding in tertiary structure. Proteins do not exist in a vacuum; they are surrounded by water. Water is a highly polar compound that interacts readily with polar, hydrophilic amino acid residues capable of forming hydrogen bonds

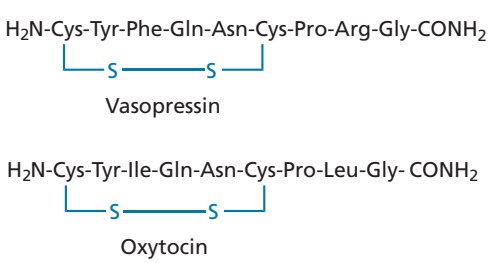


FIGURE 2.15 Vasopressin and oxytocin.

(Fig. 2.16). The remaining non-polar, hydrophobic amino acid residues cannot interact favourably with water, so the most stable tertiary structure will ensure that most of the hydrophilic groups are on the surface so that they interact with water and that most of the hydrophobic groups are in the centre so that they avoid water and interact with each other. As the hydrophilic amino acids form hydrogen bonds with water, the number of ionic and hydrogen bonds contributing to the tertiary structure is reduced leaving hydrophobic and van der Waals interactions to largely determine the three-dimensional shape of the protein.

For the reasons stated above, the centre of the protein must be hydrophobic and non-polar. This has important consequences. For example, it helps to explain why enzymes catalyse reactions that should be impossible in the aqueous environment of the human body. Enzymes contain a hollow, or canyon, on their surface called an **active site**. As the active site protrudes into the centre of the protein, it tends to be hydrophobic in nature and can provide a non-aqueous environment for the reaction taking place (Chapter 3).

Many other types of protein contain similar hollows or clefts that act as **binding sites** for natural ligands. They, too, are more hydrophobic than the surface and so van der Waals and hydrophobic interactions play an important role in the binding of the ligand. An understanding of these interactions is crucial to the design of effective drugs that will target these binding sites.

2.3.6 Role of the planar peptide bond

Planar peptide bonds indirectly play an important role in tertiary structure. Bond rotation in peptide bonds is hindered, with the *trans* conformation generally favoured, so the number of possible conformations that a protein can adopt is significantly restricted, making it more likely that a specific conformation is adopted. Polymers without peptide bonds do not fold into a specific conformation, because the entropy change required to form a highly ordered structure is extremely unfavourable. Peptide bonds can also form hydrogen bonds with amino acid side chains and play a further role in determining tertiary structure.

2.4 The quaternary structure of proteins

Only proteins that are made up of a number of protein subunits have quaternary structure. For example, **haemoglobin** is made up of four protein molecules—two identical alpha subunits and two identical beta subunits (not to be

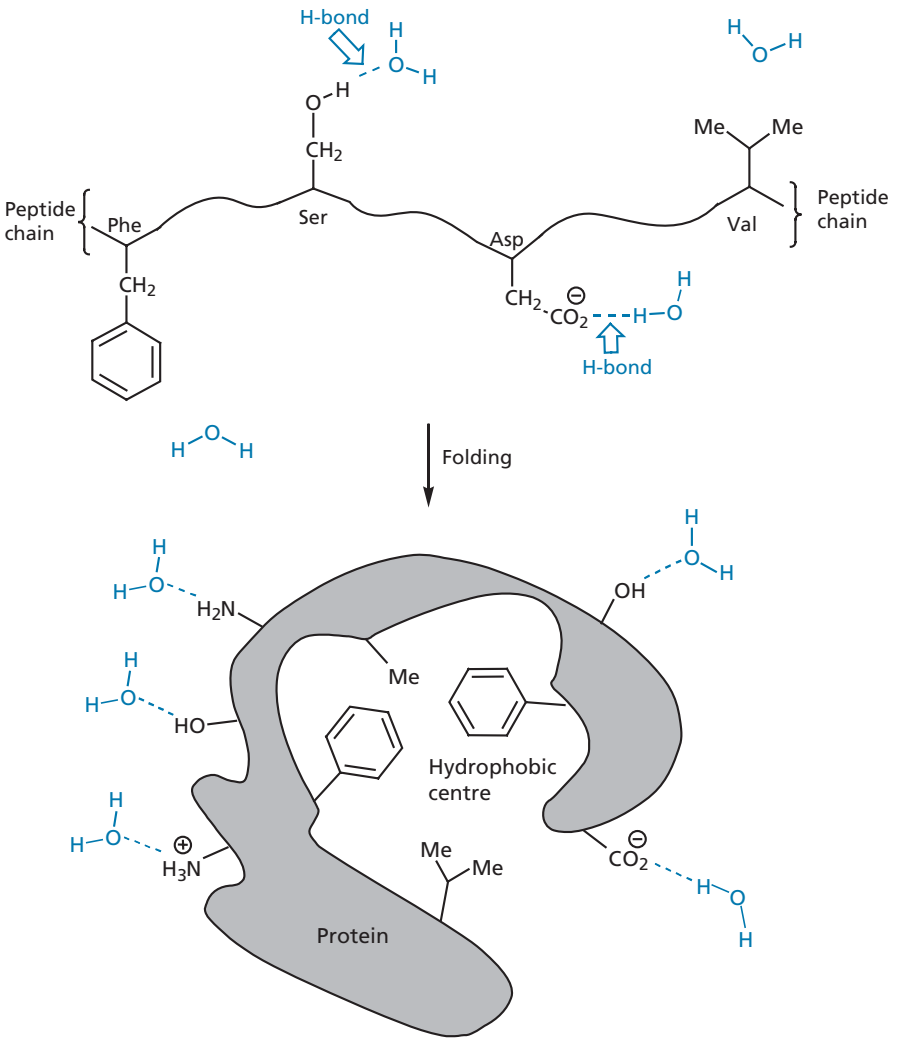


FIGURE 2.16 Bonding interactions with water.

confused with the alpha and beta terminology used in secondary structure). The quaternary structure of haemoglobin is the way in which these four protein units associate with each other.

As this must inevitably involve interactions between the exterior surfaces of proteins, ionic bonding can be more important to quaternary structure than it is to tertiary structure. Nevertheless, hydrophobic and van der Waals interactions have a role to play. It is not possible for a protein to fold up such that all its hydrophobic groups are placed towards the centre. Some of these groups may be stranded on the surface. If they form a small hydrophobic area on the protein surface, there is a distinct advantage for two protein molecules to form a dimer such that the two hydrophobic areas face each other rather than be exposed to an aqueous environment (Fig. 2.17). It is also

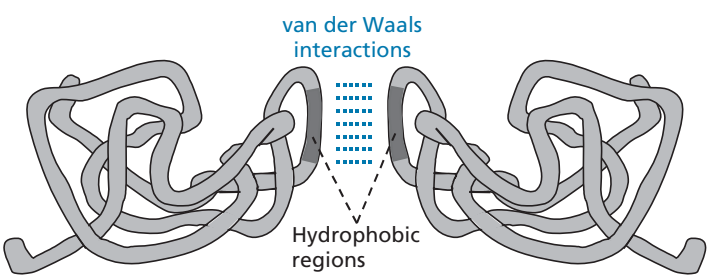


FIGURE 2.17 Quaternary structure involving two protein subunits.

possible for protein molecules to interlock in a quaternary structure (section 24.9.2).

2.5 Translation and post-translational modifications

The process by which a protein is synthesized in the cell is called **translation** (section 6.2.2). Many proteins are modified following translation (Fig. 2.18), and these modifications can have wide-ranging effects. For example, the *N*-terminals of many proteins are acetylated, making these proteins more resistant to degradation. Acetylation of proteins also has a role to play in the control of transcription, cell proliferation, and differentiation (section 21.7.3).

The fibres of **collagen** are stabilized by the hydroxylation of proline residues. Insufficient hydroxylation results in scurvy (caused by a deficiency of vitamin C). The glutamate residues of **prothrombin**, a clotting protein, are carboxylated to form γ-carboxyglutamate structures. In cases of vitamin K deficiency, carboxylation does not occur and excessive bleeding results. The serine, threonine, and tyrosine residues of many proteins are phosphorylated and this plays an important role in signalling pathways within the cell (sections 5.2–5.4).

Many of the proteins present on the surface of cells are linked to carbohydrates through asparagine residues. Such carbohydrates are added as post-translational modifications and are important to cell–cell recognition, disease processes, and drug treatments (section 10.7). The proteins concerned are called **glycoproteins** or **proteoglycans**, and are members of a larger group of molecules called **glycoconjugates.**

Several proteins are cleaved into smaller proteins or peptides following translation. For example, the **enkephalins** are small peptides which are derived from

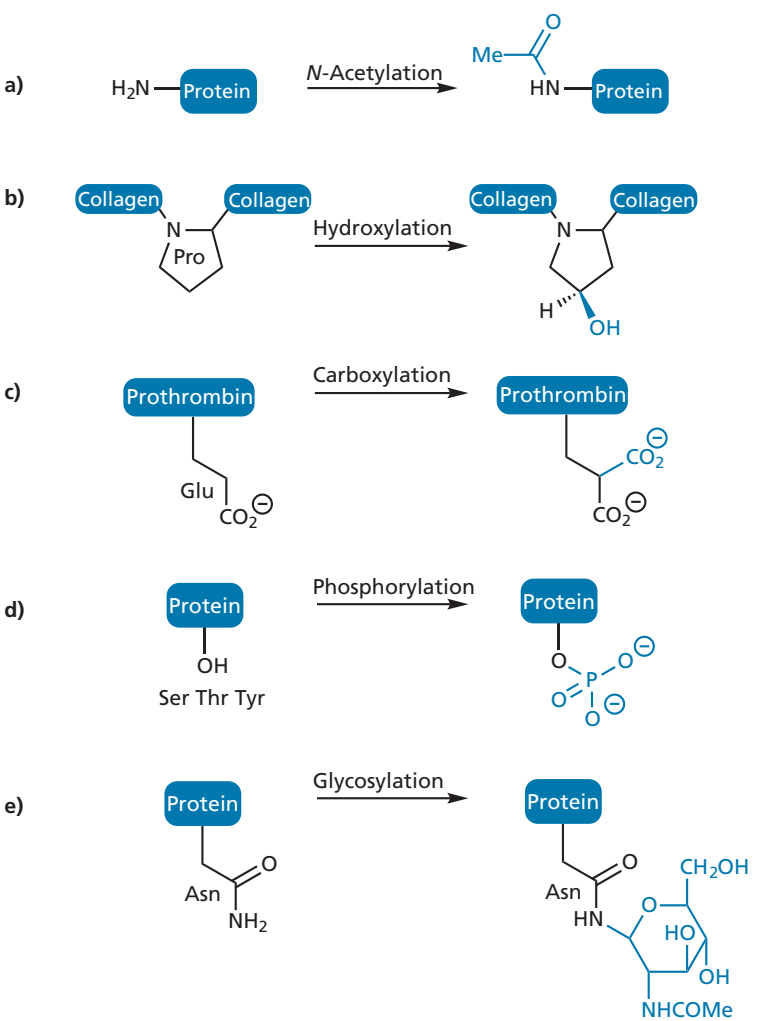


FIGURE 2.18 Examples of post-translational modifications carried out on proteins.

proteins in this manner (section 24.8). Active enzymes are sometimes formed by cleaving a larger protein precursor. Often, this serves to protect the cell from the indiscriminate action of an enzyme. For example, digestive enzymes are stored in the pancreas as inactive protein precursors and are only produced once the protein precursor is released into the intestine. In blood clotting, the soluble protein **fibrinogen** is cleaved to insoluble **fibrin** when the latter is required. Some polypeptide hormones are also produced from the cleavage of protein precursors. Finally, the cleavage of a viral polyprotein into constituent proteins is an important step in the life cycle of the HIV virus and has proved a useful target for several drugs currently used to combat AIDS (section 20.7.4).

2.6 Proteomics

A lot of publicity has been rightly accorded to the Human Genome Project, which has now been completed. The science behind this work is called **genomics** and it involves the identification of the genetic code in humans and other species. The success of this work has been hailed as a breakthrough that will lead to a new era in medicinal research. However, it is important to appreciate that this is only the start of a longer process. As we shall see in Chapter 6, DNA is the blueprint for the synthesis of proteins and so the task is now to identify all the proteins present in each cell of the body and, more importantly, how they interact with each other—an area of science known as **proteomics**. Proteomics is far more challenging than genomics because of the complexity of interactions that can take place between proteins (see Chapter 5). Moreover, the pattern and function of proteins present in a cell depend on the type of cell it is and this pattern can alter in the diseased state. Nevertheless, the race is now on to analyse the structure and function of proteins, many of which are completely new to science, and to see whether they can act as novel drug targets for the future. This is no easy task and it is made all the more difficult by the fact that it is not possible to simply derive the structure of proteins based on the known gene sequences. This is because different proteins can be derived from a single gene and proteins are often modified following their synthesis (section 2.5). There are roughly 40,000 genes, whereas a typical cell contains hundreds of thousands of different proteins. Moreover, knowing the structure of a protein does not necessarily suggest its function or interactions.

Identifying the proteins present in a cell usually involves analysing the contents of the cell and separating out the proteins using a technique known as two-dimensional gel electrophoresis. Mass spectrometry can then be used to study the molecular weight of each protein. Assuming a pure sample of protein is obtained, its primary structure can be identified by traditional sequencing techniques. The analysis of secondary and tertiary structures is trickier. If the protein can be crystallized, then it is possible to determine its structure by X-ray crystallography. Not all proteins can be crystallized, though, and even if they are, it is possible that the conformation in the crystal form is different from that in solution. In recent years nuclear magnetic resonance (NMR) spectroscopy has been successful in identifying the tertiary structure of some proteins.

There then comes the problem of identifying what role the protein has in the cell and whether it would serve as a useful drug target. If it does show promise as a target, the final problem is to discover or design a drug that will interact with it.

KEY POINTS

- The order in which amino acids are linked together in a protein is called the primary structure.
- The secondary structure of a protein refers to regions of ordered structure within the protein, such as α-helices, β-pleated sheets, or β-turns.
- The overall three-dimensional shape of a protein is called its tertiary structure.
- Proteins containing two or more subunits have a quaternary structure which defines how the subunits are arranged with respect to each other.
- Secondary, tertiary, and quaternary structures are formed to maximize favourable intramolecular and intermolecular bonds, and to minimize unfavourable interactions.
- Amino acids with polar residues are favoured on the outer surface of a protein because this allows hydrogen bonding interactions with water. Amino acids with non-polar residues are favoured within the protein because this maximizes van der Waals and hydrophobic interactions.
- Many proteins undergo post-translational modifications.
- Proteomics is the study of the structure and function of novel proteins discovered through genomics.

2.7 Protein function

We are now ready to discuss the various types of protein which act as drug targets.

2.7.1 Structural proteins

Structural proteins do not normally act as drug targets. However, the structural protein **tubulin** is an exception. Tubulin molecules polymerize to form small tubes called **microtubules** in the cell's cytoplasm (Fig 2.19). These

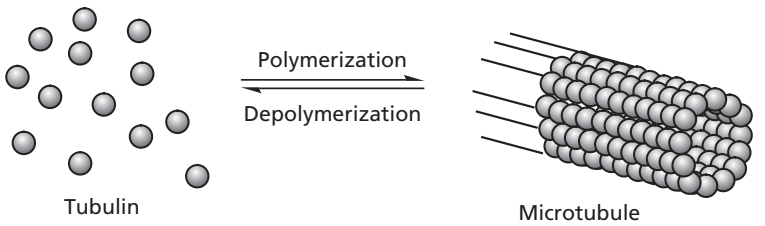


FIGURE 2.19 Polymerization of tubulin.

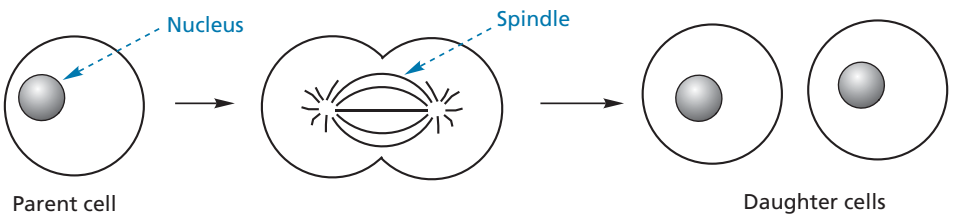


FIGURE 2.20 Cell division.

microtubules have various roles within the cell, including the maintenance of shape, exocytosis, and release of neurotransmitters. They are also involved in the mobility of cells. For example, inflammatory cells called **neutrophils** are mobile cells which normally protect the body against infection. However, they can also enter joints, leading to inflammation and arthritis.

Tubulin is also crucial to cell division. When a cell is about to divide, its microtubules depolymerize to give tubulin. The tubulin is then re-polymerized to form a structure called a **spindle** which then serves to push apart the two new cells and to act as a framework on which the chromosomes of the original cell are transferred to the nuclei of the daughter cells (Fig. 2.20). Drugs that target tubulin and inhibit this process are useful anticancer agents (section 10.2.2).

The structural proteins of viruses are important to the survival of the virus outside their host cell. Some of these proteins are proving to be interesting drug targets for the design of new antiviral agents and are discussed in more detail in sections 20.7.5 and 20.9.

2.7.2 Transport proteins

Transport proteins are present in the cell membrane and act as the cell's 'smugglers'—smuggling the important chemical building blocks of amino acids, sugars, and nucleic acid bases across the cell membrane such that the cell can synthesize its proteins, carbohydrates, and nucleic acids. They are also important in transporting important neurotransmitters (section 4.2) back into the neuron that released them so that the neurotransmitters only have a limited period of activity. But why is this smuggling operation necessary? Why can't these molecules pass through the membrane by themselves? Quite simply, the molecules concerned are polar structures and cannot pass through the hydrophobic cell membrane.

The transport proteins can float freely within the cell membrane because they have hydrophobic residues on their outer surface which interact favourably with the hydrophobic centre of the cell membrane. The portion of the transport protein that is exposed on the outer surface of the cell membrane contains a binding site that can bind a polar molecule, such as an amino acid, stow it away in a hydrophilic pocket, and ferry it across the membrane to release it on the other side (Fig. 2.21).

Transport proteins are not all identical; there are specific transport proteins for the different molecules that need to be smuggled across the membrane. The binding sites for these transport proteins vary in structure such that they can recognize and bind their specific guest. There are several important drugs which target transport proteins (section 10.1).

2.7.3 Enzymes and receptors

The most important drug targets in medicinal chemistry are enzymes and receptors. Chapters 3 and 4 are devoted to the structure and function of these proteins respectively.

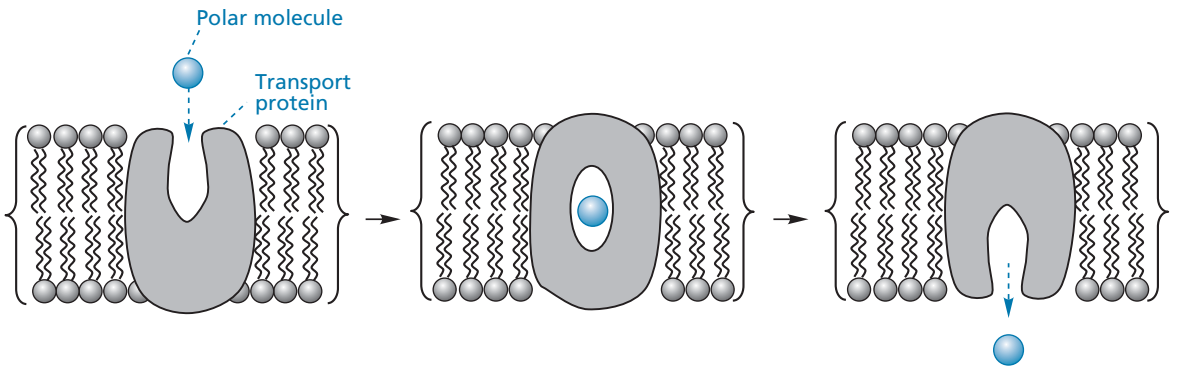


FIGURE 2.21 Transport proteins.

2.7.4 Miscellaneous proteins and protein–protein interactions

There are many situations in cell biology where proteins are required to interact with each other in order to produce a particular cellular effect. We have already seen an example of this in the polymerization of tubulin proteins in order to form microtubules (section 2.7.1). The structures of many important drug targets, such as ion channels, enzymes, and receptors consist of two or more protein subunits associated with each other. The signal transduction processes described in Chapter 5 show many instances where a variety of proteins, such as receptors, signal proteins, and enzymes, associate with each other in order to transmit a chemical signal into the cell. The actions of insulin are mediated through a protein–protein interaction (section 4.8.3). The control of gene expression involves the prior assembly of a variety of different proteins (section 4.9 and Box 8.2). An important part of the immune response involves proteins called antibodies interacting with foreign proteins (section 10.7.2). Cell–cell recognition involves protein–protein interactions—a process which is not only important in terms of the body's own proteins, but in the mechanism by which viruses invade human cells (sections 20.7.1, 20.8.1, and 20.9). Important processes that have an influence on tumour growth, such as angiogenesis and apoptosis (section 21.1), involve the association of proteins. Proteins called chaperones help to stabilize partially folded proteins during translation through protein–protein interactions. They are also important in the process by which old proteins are removed to the cell's recycling centre. Chaperones are particularly important when the cell experiences adverse environmental conditions which might damage proteins. It has been found that the synthesis of chaperones increases in tumour cells, which may reflect some of the stresses experienced in such cells, for example lack of oxygen, pH variation, and nutrient deprivation. Inhibiting chaperones could well lead to more damaged proteins and cell death. There are current studies looking into methods of inhibiting a chaperone protein called HSP90 (HSP stands for heat shock protein). Inhibition might prevent the synthesis of important receptors and enzymes involved in the process of cell growth and division and provide a new method of treating tumour cells (section 21.6.2.7). The inhibition of an enzyme acting as a chaperone protein is also being considered as a potential therapy for the treatment of Alzheimer's disease (section 22.15.2).

Protein–protein interactions are not limited to human biochemistry. Interfering with these interactions in other species could lead to novel antibacterial, antifungal, and antiviral agents. For example, HIV protease is an important enzyme in the life-cycle of the HIV virus and is an important target for antiviral agents (section 20.7.4). The enzyme consists of two identical proteins which bind together to produce the active site. Finding a drug that will prevent this association would be a novel method of inhibiting this enzyme.

To conclude, there is a lot of research currently underway looking at methods of inhibiting or promoting protein–protein interactions (section 10.5).

KEY POINTS

- Transport proteins, enzymes, and receptors are common drug targets.
- Transport proteins transport essential polar molecules across the hydrophobic cell membrane.
- Tubulin is a structural protein which is crucial to cell division and cell mobility.
- Many cell processes depend on the interactions of proteins with each other.

QUESTIONS

1. Draw the full structure of L-alanyl-L-phenylalanyl-glycine.
2. What is unique about glycine compared with other naturally-occurring amino acids?
3. Identify the intermolecular/intramolecular interactions that are possible for the side chains of the following amino acids; serine, phenylalanine, glycine, lysine, aspartic acid, and aspartate.
4. The chains of several cell membrane-bound proteins wind back and forth through the cell membrane, such that some parts of the protein structure are extracellular, some parts are intracellular, and some parts lie within the cell membrane. How might the primary structure of a protein help in distinguishing the portions of the protein embedded within the cell membrane from those that are not?
5. What problems might you foresee if you tried to synthesize L-alanyl-L-valine directly from its two component amino acids?
6. The tertiary structure of many enzymes is significantly altered by the phosphorylation of serine, threonine, or tyrosine residues. Identify the functional groups that are involved in these phosphorylations and suggest why phosphorylation affects tertiary structure.
7. What is the one-letter code for the polypeptide Glu–Leu–Pro–Asp–Val–Val–Ala–Phe–Lys–Ser–Gly–Gly–Thr?

FURTHER READING

Ball, P. (2009) Proteins unravelled. *Chemistry World*, December, 58–62.

Berg, C., Neumeyer, K., and Kirkpatrick, P. (2003) Teriparatide. *Nature Reviews Drug Discovery* **2**, 257–258.

Darby, N. J. and Creighton, T. E. (1993) *Protein Structure*. IRL Press, Oxford.

Dobson, C. M. (2003) Protein folding and disease: a view from the first Horizon symposium. *Nature Reviews Drug Discovery* **2**, 154–160.

Ezzell, C. (2002) Proteins rule. *Scientific American* **286**, 40–48.

Harris, J. M. and Chess, R. B. (2003) Effect of pegylation on pharmaceuticals. *Nature Reviews Drug Discovery* **2**, 214–221.

Jones, J. (1992) *Amino Acid and Peptide Synthesis*. Oxford University Press, Oxford.

Stevenson, R. (2002) Proteomic analysis honoured. *Chemistry in Britain* **38**(11).

Teague, S. J. (2003) Implications of protein flexibility for drug discovery. *Nature Reviews Drug Discovery* **2**, 527–541.

3 Enzymes: structure and function

In this chapter we discuss the structure and function of enzymes. Drug action at enzymes is discussed in Chapter 7 and in other chapters throughout the text.

3.1 Enzymes as catalysts

Enzymes are proteins which act as the body's catalysts—agents that speed up a chemical reaction without being consumed themselves. Without them, the cell's chemical reactions would either be too slow or not take place at all. An example of an enzyme-catalysed reaction is the reduction of **pyruvic acid** to **lactic acid**, which takes place when muscles are over-exercised, and is catalysed by an enzyme called **lactate dehydrogenase** (Fig. 3.1).

Note that the reaction is shown as an equilibrium. It is, therefore, more correct to describe an enzyme as an agent that speeds up the approach to equilibrium, because the enzyme speeds up the reverse reaction just as efficiently as the forward reaction. The final equilibrium concentrations of the starting materials and products are unaffected by the presence of an enzyme.

How do enzymes affect the rate of a reaction without affecting the equilibrium? The answer lies in the existence of a high-energy **transition state** that must be formed before the starting material (the **substrate**) can be converted to the product. The difference in energy between the transition state and the substrate is the **activation energy**, and it is the size of this activation energy that determines the rate of a reaction, rather than the difference in energy between the substrate and the product (Fig. 3.2). An enzyme acts to lower the activation energy by helping to stabilize the transition state. The energy of the substrate and products are unaffected, and therefore the equilibrium ratio of substrate to product is unaffected. We can relate energy to the rate and equilibrium constants with the following equations:

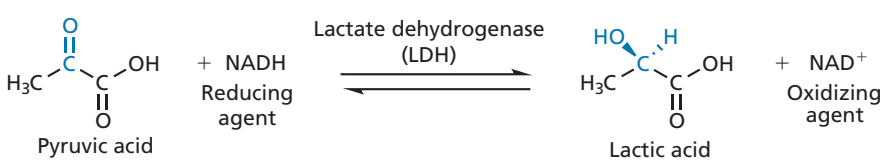


FIGURE 3.1 Reaction catalysed by lactate dehydrogenase.

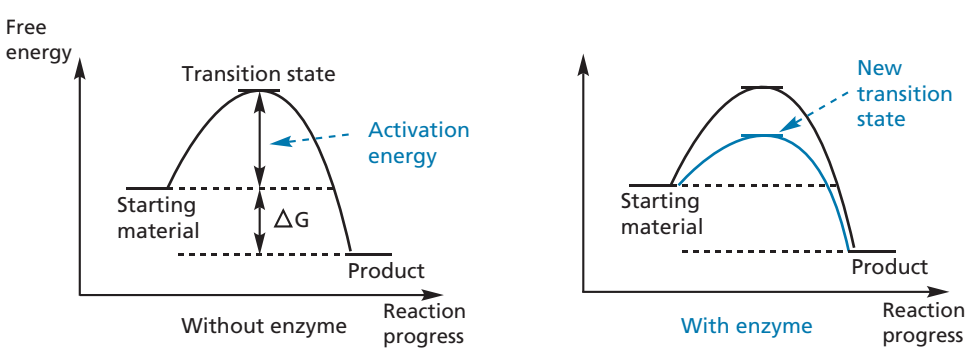


FIGURE 3.2 Graphs demonstrating the stabilization of a reaction's transition state by an enzyme.

$$\text{Energy difference} = \Delta G = -RT \ln K$$

where K is the equilibrium constant (= [products]/[reactants]), R is the gas constant (= 8.314 J mol^{-1} K^{-1}), and T is the temperature.

$$\text{Rate constant} = k = Ae^{-E/RT}$$

where E is the activation energy and A is the frequency factor.

Note that the rate constant k does not depend on the equilibrium constant K.

We have stated that enzymes catalyse reactions, but we have still to explain how.

3.2 How do enzymes catalyse reactions?

The factors involved in enzyme catalysis are summarized below and will be discussed in more detail in sections 3.2–3.5.

- Enzymes provide a reaction surface and a suitable environment.
- Enzymes bring reactants together and position them correctly so that they easily attain their transition-state configurations.
- Enzymes weaken bonds in the reactants.
- Enzymes may participate in the reaction mechanism.
- Enzymes form stronger interactions with the transition state than with the substrate or the product.

An enzyme catalyses a reaction by providing a surface to which a substrate can bind, resulting in the weakening of high-energy bonds. The binding also holds the substrate in the correct orientation to increase the chances of reaction. The reaction takes place, aided by the enzyme, to give a product which is then released (Fig. 3.3). Note again that it is a reversible process. Enzymes can catalyse both forward and backward reactions. The final equilibrium mixture will, however, be the same, regardless of whether we supply the enzyme with substrate or product. Substrates bind to, and react at, a specific area of the enzyme called the **active site**—usually quite a small part of the overall protein structure.

3.3 The active site of an enzyme

The active site of an enzyme (Fig. 3.4) has to be on or near the surface of the enzyme if a substrate is to reach it. However, the site could be a groove, hollow, or gully allowing the substrate to sink into the enzyme. Normally, the active site is more hydrophobic in character than the surface of the enzyme, providing a suitable environment for many reactions that would be difficult or impossible to carry out in an aqueous environment.

Because of the overall folding of the enzyme, the amino acid residues that are close together in the active site may be far apart in the primary structure. Several amino acids in the active site play an important role in enzyme function, which can be demonstrated by comparing the primary structures of the same enzyme from different organisms. Here, the primary structure differs from species to species as a result of **mutations** happening over millions of years. The variability is proportional to how far apart the organisms are on the evolutionary ladder. However, there are certain amino acids that remain constant, no matter the source of the enzyme. These are amino acids that are crucial to the enzyme's function and are often present in the active site. If one of these amino acids is altered through mutation, the enzyme could become useless and the cell bearing this mutation would have a poor chance of survival. Thus, the mutation would not be preserved. The only exception to this would be if the mutation introduced an amino acid which could either perform the same task as the original amino acid or improved substrate binding. This consistency of amino acids in the active site can often help scientists determine which amino acids are present in an active site, if this is not known already.

Amino acids present in the active site can have one of two roles:

- binding—the amino acid residue is involved in binding the substrate or a cofactor to the active site;

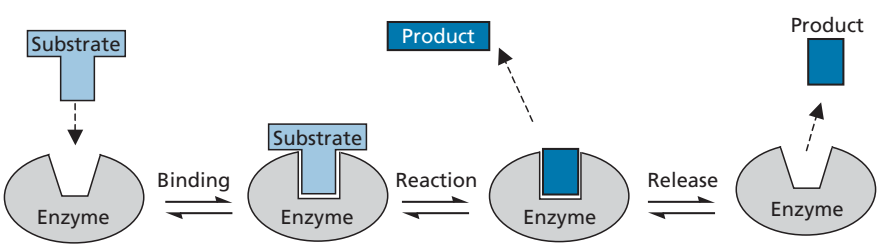


FIGURE 3.3 The process of enzyme catalysis.

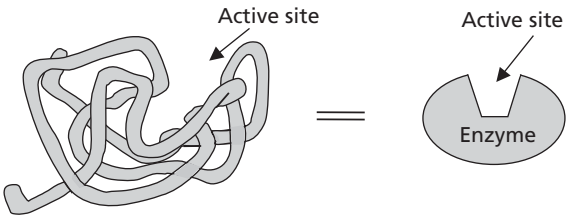


FIGURE 3.4 The active site of an enzyme.

- catalytic—the amino acid is involved in the mechanism of the reaction.

We shall study these in turn.

3.4 Substrate binding at an active site

The interactions which bind substrates to the active sites of enzymes include ionic bonds, hydrogen bonds, dipole–dipole, and ion–dipole interactions, as well as van der Waals and hydrophobic interactions (section 1.3). These binding interactions are the same bonding interactions responsible for the tertiary structure of proteins, but their relative importance differs. Ionic bonding plays a relatively minor role in protein tertiary structure compared with hydrogen bonding or van der Waals interactions, but it can play a crucial role in the binding of a substrate to an active site.

As intermolecular bonding forces are involved in substrate binding, it is possible to look at the structure of a substrate and postulate the probable interactions that it will have with its active site. As an example, consider **pyruvic acid**—the substrate for **lactate dehydrogenase** (Fig. 3.5).

If we look at the structure of pyruvic acid, we can propose three possible interactions by which it might bind to its active site—an ionic interaction involving the ionized carboxylate group, a hydrogen bond involving the ketonic oxygen, and a van der Waals interaction involving the methyl group. If these postulates are correct, it means that within the active site there must be **binding regions** containing suitable amino acids that can take part in these intermolecular interactions. Lysine, serine, and phenylalanine residues respectively would fit the bill. A knowledge of how a substrate binds to its active site is invaluable in designing drugs that will target specific enzymes (Chapter 7).

3.5 The catalytic role of enzymes

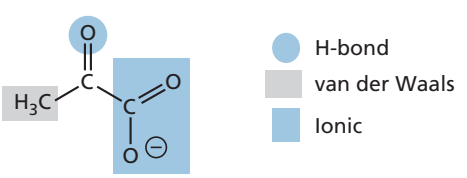
We now move on to consider the mechanism of enzymes and how they catalyse reactions. In general, enzymes catalyse reactions by providing binding interactions, acid/base catalysis, nucleophilic groups, and cofactors.

3.5.1 Binding interactions

In the past, it was thought that a substrate fitted its active site in a similar way to a key fitting a lock (**Fischer's lock and key hypothesis**). Both the enzyme and the substrate were seen as rigid structures, with the substrate (the key) fitting perfectly into the active site (the lock) (Fig. 3.6). However, this scenario does not explain how some enzymes can catalyse a reaction on a range of different substrates. It implies, instead, that an enzyme has an optimum substrate that fits it perfectly, whereas all other substrates fit less perfectly. This, in turn, would imply that the catalysed reaction is only efficient for the optimum substrate. As this is not the case for many enzymes, the lock and key analogy must be invalid.

It is now proposed that the substrate is not quite the ideal shape for the active site, and that it forces the active site to change shape when it enters—a kind of moulding process. This theory is known as **Koshland's theory of induced fit** as the substrate induces the active site to take up the ideal shape to accommodate it (Fig. 3.6).

For example, a substrate such as **pyruvic acid** might interact with specific binding regions in the active site of lactate dehydrogenase via one hydrogen bond, one ionic bond, and one van der Waals interaction (Fig. 3.7).



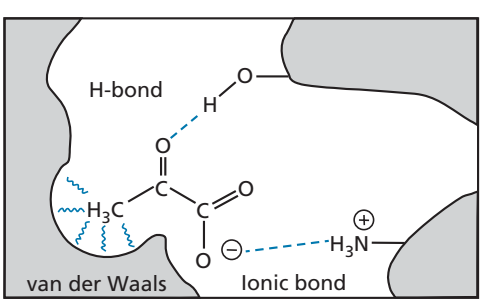


FIGURE 3.5 Binding interactions between pyruvic acid and lactate dehydrogenase.

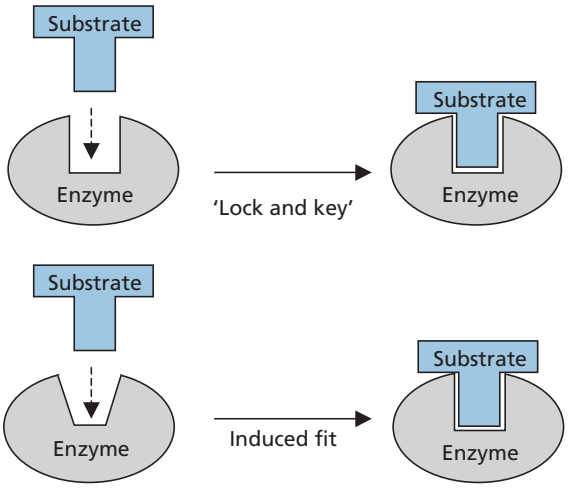


FIGURE 3.6 The 'lock and key' and 'induced fit' hypotheses for substrate–enzyme binding.

However, if the fit is not perfect, the three bonding interactions are not ideal either. For example, the binding groups may be slightly too far away from the corresponding binding regions in the active site. In order to maximize the strength of these bonds, the enzyme changes shape such that the amino acid residues involved in the binding move closer to the substrate.

This theory of induced fit helps to explain why some enzymes can catalyse reactions involving a wide range of substrates. Each substrate induces the active site into a shape that is ideal for it and, as long as the moulding process does not distort the active site so much that the reaction mechanism proves impossible, the reaction can proceed. The range of substrates that can be accepted depends on the substrates being the correct size to fit the active site and having the correct binding groups in the correct relative positions.

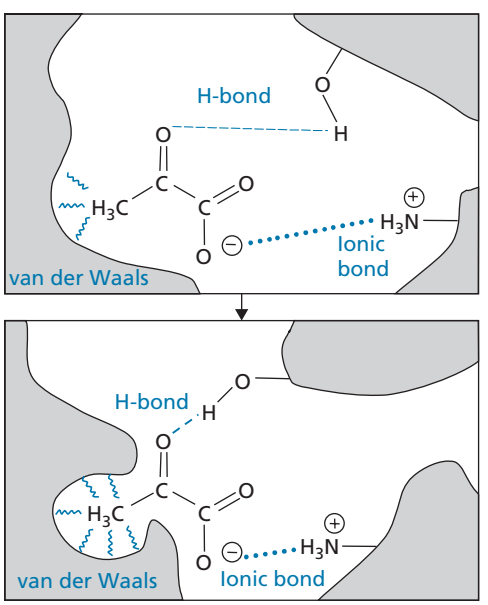


FIGURE 3.7 Example of an induced fit.

But note this. The substrate is not a passive spectator to the moulding process going on around it. As the enzyme changes shape to maximize bonding interactions, the same thing can happen to the substrate. It too may alter shape. Bond rotation may occur to fix the substrate in a particular conformation—and not necessarily the most stable one. Bonds may even be stretched and weakened. Consequently, this moulding process designed to maximize binding interactions may force the substrate into the ideal conformation for the reaction to follow and may also weaken the very bonds that have to be broken.

Once bound to an active site, the substrate is now held ready for the subsequent reaction. Binding has fixed the 'victim' (substrate) so that it cannot evade attack, and this same binding has weakened its defences (bonds) so that reaction is easier (a lower activation energy).

There is another point relating to substrate binding. The binding interactions with the active site must be sufficiently strong to hold the substrate for the subsequent reaction, but they cannot be too strong. If they were, the product might also be bound strongly and fail to depart the active site. This would block the active site of the enzyme and prevent it from catalysing another reaction. Therefore, a balance must be struck.

Finally, it is important to realize that the enzyme also binds the transition state involved in the enzyme-catalysed reaction. Indeed, the binding interactions involved are stronger than those binding the substrate, which means that the transition state is stabilized relatively more than the substrate. This results in a lower activation energy compared with the non-catalysed reaction.

3.5.2 Acid/base catalysis

Acid/base catalysis is often provided by the amino acid **histidine**, which contains an imidazole ring as part of its side chain. The imidazole ring acts as a weak base, which means that it exists in equilibrium between its protonated and free base forms (Fig. 3.8), allowing it to accept or donate protons during a reaction mechanism. This is important, as there are often very few water molecules present in an active site to carry out this role. Histidine is not the only amino acid residue that can provide acid/base catalysis. For example, a **glutamic acid** residue acts as a proton source in the reaction mechanism of the enzyme HMG-CoA reductase (Case study 1), while **aspartic acid** and **aspartate** residues act as proton donors and proton acceptors, respectively, in other enzyme-catalysed reactions (sections 7.4 and 20.7.4.1).

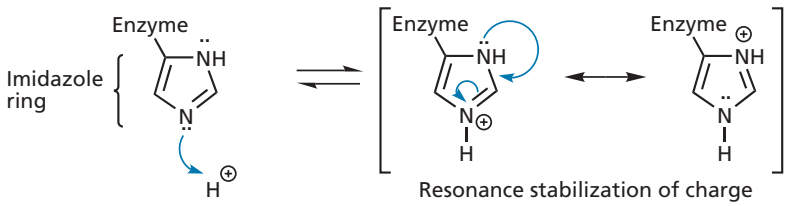


FIGURE 3.8 Histidine acting as a weak base.

Tyrosine acts as a proton source in the mechanism by which the enzyme 17β-hydroxysteroid type 1 catalyses the conversion of **estrone** to **estradiol**.

For additional material see Web article 1: steroids as novel anticancer agents.

3.5.3 Nucleophilic groups

The amino acids **serine** and **cysteine** are present in the active sites of some enzymes. These amino acids have nucleophilic residues (OH and SH respectively) which are able to participate in the reaction mechanism. They do this by reacting with the substrate to form intermediates that would not be formed in the uncatalysed reaction. These intermediates offer an alternative reaction pathway that may avoid a high-energy transition state and hence increase the rate of the reaction.

Normally, an alcoholic OH group, such as the one on serine, is not a good nucleophile. However, there is usually a histidine residue close by to catalyse the reaction. For example, the mechanism by which chymotrypsin hydrolyses peptide bonds (Fig. 3.9) involves a **catalytic triad** of amino acids—serine, histidine, and aspartic acid. Serine and histidine participate in the mechanism as a nucleophile and acid/base catalyst respectively. The aspartate group interacts with the histidine ring and serves to activate and orient it correctly for the mechanism.

The presence of a nucleophilic serine residue means that water is not required in the initial stages of the mechanism. This is important, firstly, because water is a

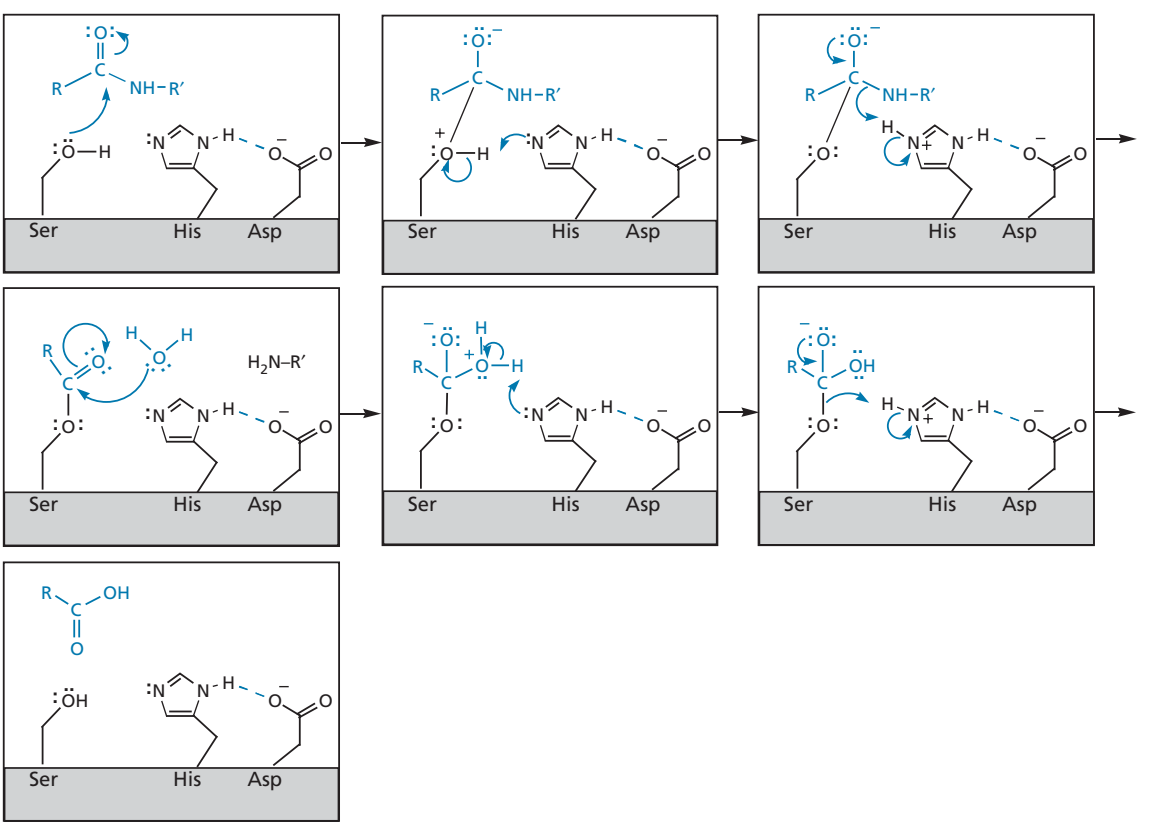


FIGURE 3.9 Hydrolysis of peptide bonds catalysed by the enzyme chymotrypsin.

poor nucleophile and may also find it difficult to penetrate the occupied active site. Secondly, a water molecule would have to drift into the active site and search out the carboxyl group before it could attack it. This would be something similar to a game of blind man's bluff. The enzyme, however, can provide a serine OH group positioned in exactly the right spot to react with the substrate. Therefore, the nucleophile has no need to search for its substrate: the substrate has been delivered to it.

Water is eventually required to hydrolyse the acyl group attached to the serine residue. However, this is a much easier step than the hydrolysis of a peptide link, as esters are more reactive than amides. Furthermore, the hydrolysis of the peptide link means that one half of the peptide can drift away from the active site and leave room for a water molecule to enter. A similar enzymatic mechanism is involved in the action of the enzyme **acetylcholinesterase** (section 22.12.3), **pancreatic lipase** (Box 7.2), and a viral protease enzyme carried by the hepatitis C virus (section 20.10).

The amino acid **lysine** has a primary amine group on its side chain which should make it a better nucleophilic group than serine or cysteine. However, the group is generally protonated at physiological pH, which precludes it acting as a nucleophile. Having said that, some enzymes have a lysine residue located in a hydrophobic pocket, which means that it is not protonated and can, indeed, act as a nucleophilic group.

3.5.4 Cofactors

Many enzymes require additional non-protein substances called cofactors for the reaction to take place. Deficiency of cofactors can arise from a poor diet resulting in the loss of enzyme activity and subsequent disease (e.g. scurvy). Cofactors are either metal ions (e.g. zinc) or small organic molecules called **coenzymes** (e.g. NAD^+, pyridoxal phosphate). Most coenzymes are bound by ionic bonds and other non-covalent bonding interactions, but some are bound covalently and are called **prosthetic groups**. Coenzymes are derived from water-soluble vitamins and act as the body's chemical reagents. For example, **lactate dehydrogenase** requires the coenzyme **nicotinamide adenine dinucleotide (NAD^+)** (Fig. 3.10) in order to catalyse the dehydrogenation of **lactic acid** to **pyruvic acid**. NAD^+ is bound to the active site along with lactic acid, and acts as the oxidizing agent. During the reaction it is converted to its reduced form (NADH) (Fig. 3.11). Conversely, NADH can bind to the enzyme and act as a reducing agent when the enzyme catalyses the reverse reaction.

$NADP^+$ and NADPH are phosphorylated analogues of NAD^+ and NADH, respectively, and carry out redox reactions by the same mechanism. NADPH is used almost exclusively for reductive biosynthesis, whereas NADH is used primarily for the generation of ATP.

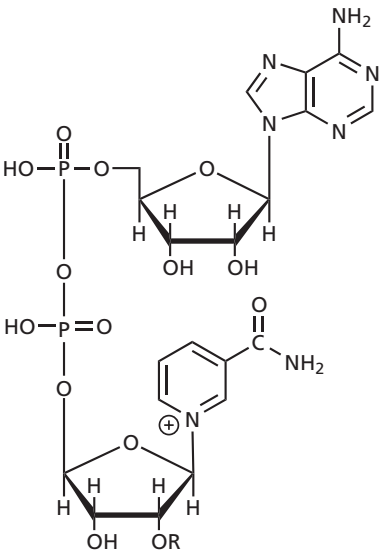


FIGURE 3.10 Nicotinamide adenine dinucleotide (R = H) and nicotinamide adenine dinucleotide phosphate (R = phosphate).

A knowledge of how the coenzyme binds to the active site allows the possibility of designing enzyme inhibitors that will fit the same region (see Case study 5 and section 21.6.2; see also Web article 1).

3.5.5 Naming and classification of enzymes

The name of an enzyme reflects the type of reaction it catalyses, and has the suffix '-ase' to indicate that it is an enzyme. For example, an **oxidase** enzyme catalyses an oxidation reaction. It is important to appreciate that enzymes can catalyse the forward and back reactions of an equilibrium reaction. This means that an oxidase enzyme can catalyse reductions, as well as oxidations. The reaction catalysed depends on the nature of the substrate, i.e. whether it is in the reduced or oxidized form.

Enzymes are classified according to the general class of reaction they catalyse and are coded with an EC number (Table 3.1).

3.5.6 Genetic polymorphism and enzymes

There are often subtle differences in the structure and properties of an enzyme between different individuals. This is owing to the fact that the DNA that codes for proteins (Chapter 6) is not identical from person to person. On average, there is a difference of one base pair in every thousand between individuals. This is known as **genetic polymorphism**. As the nucleic acid bases act as

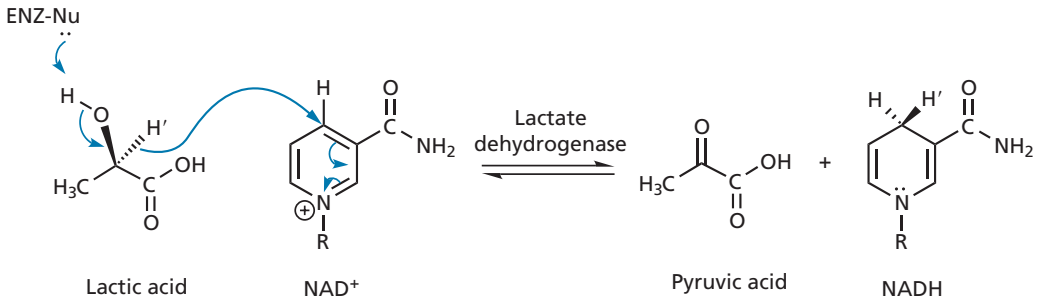


FIGURE 3.11 NAD^+ acting as a coenzyme.

TABLE 3.1 Classification of enzymes

EC number	Enzyme class	Type of reaction
E.C.1.x.x.x	Oxidoreductases	Oxidations and reductions
E.C.2.x.x.x	Transferases	Group transfer reactions
E.C.3.x.x.x	Hydrolases	Hydrolysis reactions
E.C.4.x.x.x	Lyases	Addition or removal of groups to form double bonds
E.C.5.x.x.x	Isomerases	Isomerizations and intramolecular group transfers
E.C.6.x.x.x	Ligases	Joining two substrates at the expense of ATP hydrolysis

Note: EC stands for Enzyme Commission, a body set up by the International Union of Biochemistry (as it then was) in 1955.

the code for amino acids in proteins, a difference at this level results in a different amino acid being introduced into the protein. Often, this has no observable effect on protein function, but not always. Some polymorphisms can adversely affect the proper functioning of an enzyme and lead to genetic disease. Others can have an influence on drug therapy. For example, individuals differ in their ability to metabolize drugs as a result of this phenomenon (section 11.5.6). Polymorphism can alter the sensitivity of an enzyme towards a drug, making the latter less effective. This is a particular problem in anticancer, antibacterial, and antiviral therapies where drug resistance can develop through the survival of cells containing less sensitive enzymes (Chapters 19–21).

3.6 Regulation of enzymes

Virtually all enzymes are controlled by agents which can either enhance or inhibit catalytic activity. Such control reflects the local conditions within the cell. For example, the enzyme **phosphorylase** *a* catalyses the breakdown of **glycogen** (a polymer of glucose monomers) to **glucose-1-phosphate** subunits (Fig. 3.12). It is stimulated by **adenosine 5′-monophosphate (AMP)** and inhibited by glucose-1-phosphate. Thus, rising levels of the product (glucose-1-phosphate) act as a self-regulating 'brake' on the enzyme.

But how does this control take place?

The answer is that many enzymes have a binding site which is separate from the active site called the **allosteric** binding site (Fig. 3.13). This is where the agents controlling the activity of the enzyme bind. When this occurs, an induced fit takes place which alters not only the allosteric binding site, but also the active site. Agents that inhibit the enzyme produce an induced fit that makes the active site unrecognizable to the substrate.

We might wonder why an agent inhibiting the enzyme has to bind to a separate, allosteric binding site and not to the active site itself. After all, if the agent could bind to the active site, it would directly block the natural substrate from entering. There are two explanations for this.

Firstly, many of the enzymes that are under allosteric control are at the start of a biosynthetic pathway (Fig. 3.14). A biosynthetic pathway involves a series of enzymes, all working efficiently to produce a final product. Eventually, the cell will have enough of the required material and will need to stop production. The most common control mechanism is known as **feedback control**, where the final product controls its own synthesis by inhibiting the first enzyme in the biochemical pathway. When there are low levels of final product in the cell, the first enzyme in the pathway is not inhibited and works normally. As the levels of final product increase, more and more of the enzyme is blocked and the rate of synthesis drops off in a graded fashion. Crucially, the final product has undergone many transformations from the original starting material and so it is no longer recognized by the active site of the first enzyme. A separate allosteric binding site is therefore needed which recognizes the final product. The biosynthesis of noradrenaline in section 23.4 is an example of a biosynthetic pathway under feedback control.

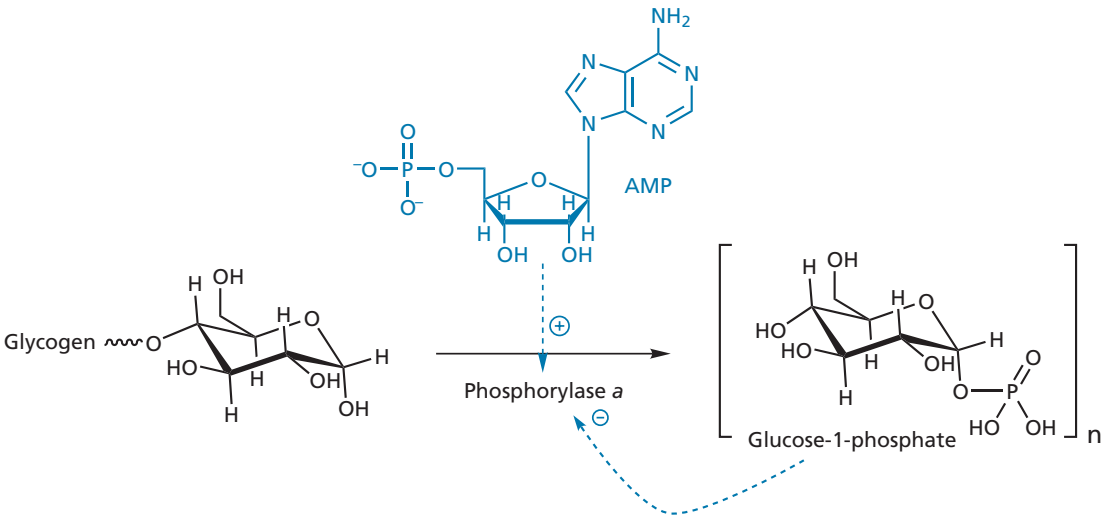


FIGURE 3.12 Internal control of the catalytic activity of phosphorylase *a* by glucose-1-phosphate and AMP.

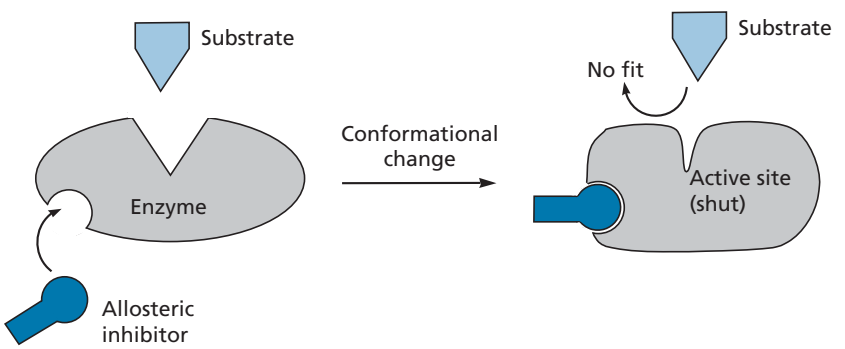


FIGURE 3.13 Allosteric inhibition of an enzyme.

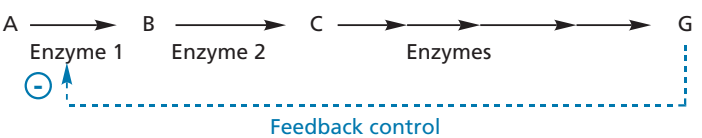


FIGURE 3.14 Feedback control of enzyme 1 by final product G.

Secondly, binding of the final product to the active site would not be a very efficient method of feedback control, as the product would have to compete with the enzyme's substrate. If levels of the latter increased, then the inhibitor would be displaced and feedback control would fail.

Many enzymes can also be regulated externally (Box 3.1). We shall look at this in more detail in Chapter 5, but, in essence, cells receive chemical messages from their environment which trigger a cascade of signals within the cell. In turn, these ultimately activate a set of enzymes known as **protein kinases**. The protein kinases play an important part in controlling enzyme activity within the cell by phosphorylating amino acids such as **serine, threonine**, or **tyrosine** in target enzymes—a covalent modification. For example, the hormone **adrenaline** is an external messenger which triggers a signalling sequence resulting in the activation of a protein kinase enzyme. Once activated, the protein kinase phosphorylates an inactive enzyme called **phosphorylase *b*** (Fig. 3.15). This enzyme now becomes active and is called **phosphorylase *a***. It catalyses the breakdown of glycogen and remains active until it is dephosphorylated back to phosphorylase ***b***.

In this case, phosphorylation of the target enzyme leads to activation. Other enzymes may be deactivated by phosphorylation. For example, **glycogen**

BOX 3.1 The external control of enzymes by nitric oxide

The external control of enzymes is usually initiated by external chemical messengers which do not enter the cell. However, there is an exception to this. It has been discovered that cells can generate the gas **nitric oxide** by the reaction sequence shown in Fig. 1, catalysed by the enzyme **nitric oxide synthase**.

Because nitric oxide is a gas, it can diffuse easily through cell membranes into target cells. There, it activates enzymes called **cyclases** to generate **cyclic GMP** from **GTP** (Fig. 2). Cyclic GMP then acts as a secondary messenger to influence other reactions within the cell. By this process, nitric oxide has an influence on a diverse range of physiological processes, including blood pressure, **neurotransmission**, and immunological defence mechanisms.

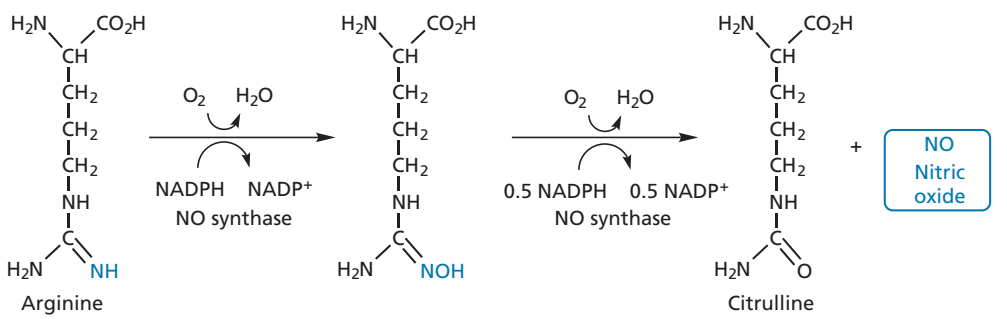


FIGURE 1 Synthesis of nitric oxide.

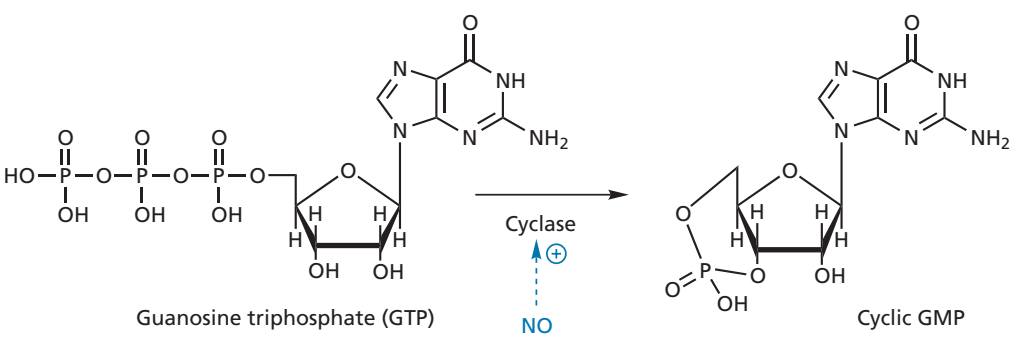


FIGURE 2 Activation of cyclase enzymes by nitric oxide (NO).

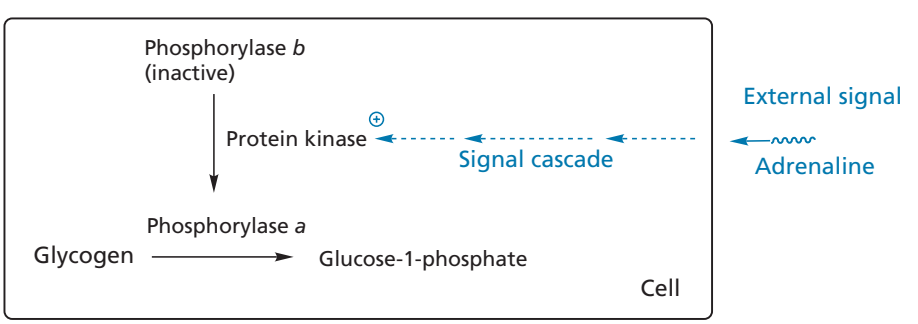


FIGURE 3.15 External control of phosphorylase *a*.

synthase—the enzyme that catalyses the *synthesis* of glycogen from glucose-1-phosphate—is inactivated by phosphorylation and activated by dephosphorylation. The latter is effected by the hormone **insulin**, which triggers a different signalling cascade from that of adrenaline.

Protein–protein interactions can also play a role in the regulation of enzyme activity. For example, signal proteins in the cell membrane are responsible for regulating the activity of membrane-bound enzymes (section 5.2).

3.7 Isozymes

Enzymes having a quaternary structure are made up of a number of polypeptide subunits. The combination of these subunits can differ in different tissues. Such variations are called **isozymes**. For example, there are five different isozymes of mammalian **lactate dehydrogenase** (LDH)—a tetrameric enzyme made up of four polypeptide subunits. There are two different types of subunits involved, which are labelled 'H' and 'M'. The former predominates in the LDH present in heart muscle, while the latter predominates in the LDH present in skeletal muscle. As there are two different types of subunit, five different isozymes are possible: HHHH, HHHM, HHMM, HMMM, and MMMM. Isozymes differ in their properties. For example, the M_4 isozyme in skeletal muscle catalyses the conversion of pyruvic acid to lactic acid and is twice as active as the H_4 isozyme in heart muscle. The H_4 isozyme catalyses the reverse reaction and is inhibited by excess pyruvic acid, whereas the M_4 isozyme is not.

KEY POINTS

- Enzymes are proteins that act as the body's catalysts by binding substrates and participating in the reaction mechanism.
- The active site of an enzyme is usually a hollow or cleft in the protein. There are important amino acids present in the active site that either bind substrates or participate in the reaction mechanism.
- Binding of substrate to an active site involves intermolecular bonds.
- Substrate binding involves an induced fit where the shape of the active site alters to maximize binding interactions. The binding process also orientates the substrate correctly and may weaken crucial bonds in the substrate to facilitate the reaction mechanism.
- The amino acid histidine is often present in active sites and acts as an acid/base catalyst. Glutamic acid, aspartic acid, and tyrosine also act as acid/base catalysts in some enzymes.
- The amino acids serine and cysteine act as nucleophiles in the reaction mechanisms of some enzymes. In some enzymes, lysine can act as a nucleophile.
- Cofactors are metal ions or small organic molecules (coenzymes) which are required by many enzymes. Coenzymes can be viewed as the body's chemical reagents.
- Prosthetic groups are coenzymes which are bound covalently to an enzyme.
- Enzymes are regulated by internal and/or external control.
- External control involves regulation initiated by a chemical messenger from outside the cell and which ultimately involves the phosphorylation of enzymes.
- Allosteric inhibitors bind to a different binding site from the active site and alter the shape of the enzyme such that the active site is no longer recognizable. Allosteric inhibitors are often involved in the feedback control of biosynthetic pathways.
- Isozymes are variations of the same enzyme. They catalyse the same reaction but differ in their primary structure, substrate specificity, and tissue distribution.
- The amino acid sequence in enzymes may differ between individuals as a result of genetic polymorphism. This may or may not result in a difference in enzyme activity.

3.8 Enzyme kinetics

3.8.1 The Michaelis-Menton equation

The Michaelis-Menten equation holds for an enzyme (E) which combines with its substrate (S) to form an enzyme–substrate complex (ES). The enzyme–substrate complex can then either dissociate back to E and S, or go on to form a product (P). It is assumed that formation of the product is irreversible.

$$E + S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_3} E + P$$

where k_1, k_2 *and* k_3 are rate constants.

For enzymes such as these, plotting the rate of enzyme reaction versus substrate concentration [S] gives a curve as shown in Fig. 3.16. At low substrate concentrations the rate of reaction increases almost proportionally to the substrate concentration, whereas at high substrate concentration the rate becomes almost constant and approaches a maximum rate ($rate_{max}$), which is independent of substrate concentration. This reflects a situation where there is more substrate present than active sites available; therefore, increasing the amount of substrate will have little effect.

The Michaelis-Menten equation relates the rate of reaction to the substrate concentration for the curve in Fig. 3.16.

$$\text{rate} = \text{rate}_{max} \frac{[S]}{[S] + K_M}$$

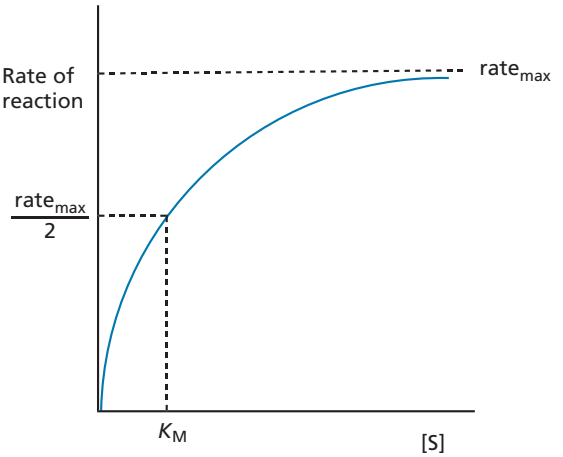


FIGURE 3.16 Reaction rate versus substrate concentration.

The derivation of this equation is not covered here, but can be found in most biochemistry textbooks. The constant K_M is known as the **Michaelis constant** and is equal to the substrate concentration at which the reaction rate is half of its maximum value. This can be demonstrated as follows. If Km = [S], then the Michaelis-Menten equation becomes:

$$\text{rate} = \text{rate}_{max}\frac{[S]}{[S]+[S]} = \text{rate}_{max}\frac{[S]}{2[S]} = \text{rate}_{max} \times \frac{1}{2}$$

The K_M of an enzyme is significant because it measures the concentration of substrate at which half the active sites in the enzyme are filled. This, in turn, provides a measure of the substrate concentration required for significant catalysis to occur.

K_M is also related to the rate constants of the enzyme-catalysed reaction:

$$K_M = \frac{k_2 + k_3}{k_1}$$

Consider now the situation where there is rapid equilibration between S and ES, and a slower conversion to product P. This means that the substrate binds to the active site and departs several times before it is finally converted to product.

$$E + S \underset{k_2 \text{ fast}}{\overset{k_1 \text{ fast}}{\rightleftharpoons}} ES \xrightarrow[\text{slow}]{k_3} E + P$$

Under these conditions, the dissociation rate (k_2) of ES is much greater than the rate of formation of product (k_3). k_3 now becomes insignificant relative to k_2 and the equation simplifies to:

$$K_M = \frac{k_2 + k_3}{k_1} = \frac{k_2}{k_1}$$

In this situation, K_M effectively equals the dissociation constant of ES and can be taken as a measure of how strongly the substrate binds to the enzyme.

$$[ES] \rightleftharpoons [E] + [S] \qquad \text{dissociation constant} = \frac{[E][S]}{[ES]}$$

A high value of K_M indicates weak binding because the equilibrium is pushed to the right; a low K_M indicates strong binding because the equilibrium is to the left. K_M is also dependent on the particular substrate involved and on environmental conditions, such as pH, temperature, and ionic strength.

The maximum rate is related to the total concentration of enzyme ($[E]_{total} = [E] + [ES]$) as follows:

$$\text{rate}_{max} = k_3[E]_{total}$$

A knowledge of the maximum rate and the enzyme concentration allows the determination of k_3. For example, the enzyme **carbonic anhydrase** catalyses the formation of hydrogen carbonate and does so at a maximum rate of 0.6 moles of hydrogen carbonate molecules formed per second for a solution containing 10^{-6} moles of the enzyme. Altering the above equation, k_3 can be determined as follows:

$$k_3 = \frac{\text{rate}_{max}}{[E]_{total}} = \frac{0.6}{10^{-6}}\frac{\text{M s}^{-1}}{\text{M}} = 600\,000\,\text{s}^{-1}$$

Therefore, each enzyme is catalysing the formation of 600,000 hydrogen carbonate molecules per second. The turnover number is the time taken for each catalysed reaction to take place, i.e. 1/600 000 = 1.7 µs.

3.8.2 Lineweaver-Burk plots

A problem related to Michaelis-Menton kinetics is the fact that there may not be sufficient data points to

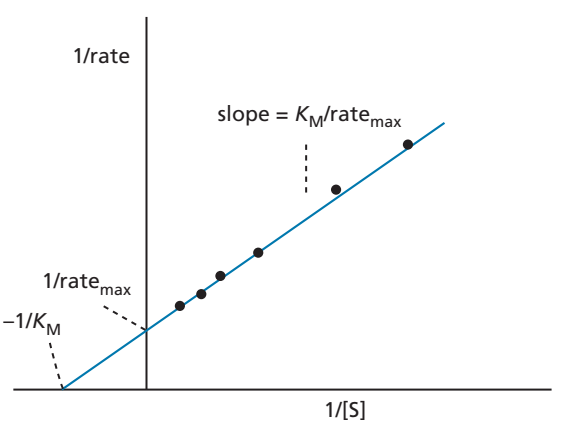


FIGURE 3.17 Lineweaver-Burk plot.

determine whether the curve of the Michaelis-Menton plot has reached a maximum value or not. This means that values for the maximum rate and K_M are likely to be inaccurate. More accurate values for these properties can be obtained by plotting the reciprocals of the rate and the substrate concentration to give a **Lineweaver-Burk plot** (Fig. 3.17):

$$\frac{1}{\text{rate}} = \frac{K_M}{\text{rate}_{max}} \cdot \frac{1}{[S]} + \frac{1}{\text{rate}_{max}} \qquad (y = m.x + c)$$

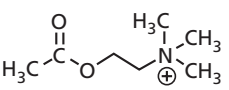
The maximum rate can then be obtained from the intersect of the line with the y-axis, while K_M can be obtained from the slope of the line or the intersect with the x-axis.

KEY POINTS

- The Michaelis-Menten equation relates the rate of an enzyme-catalysed reaction to substrate concentration.
- The Michaelis constant is equal to the substrate concentration at which the rate of the enzyme catalysed reaction is half of its maximum value.
- A Lineweaver-Burk plot provides more accurate values for the maximum rate and K_M.

QUESTIONS

1. Enzymes can be used in organic synthesis. For example, the reduction of an aldehyde is carried out using aldehyde dehydrogenase. Unfortunately, this reaction requires the use of the cofactor NADH, which is expensive and is used up in the reaction. If ethanol is added to the reaction, only catalytic amounts of cofactor are required. Why?
2. Acetylcholine is the substrate for the enzyme acetylcholinesterase. Suggest what sort of binding interactions could be involved in holding acetylcholine to the active site.



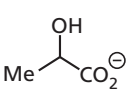
Acetylcholine

3. The ester bond of acetylcholine is hydrolysed by acetylcholinesterase. Suggest a mechanism by which the enzyme catalyses this reaction.
4. Suggest how binding interactions might make acetylcholine more susceptible to hydrolysis.
5. 17β-Hydroxysteroid dehydrogenase type 1 (17β-HSD1) is an enzyme that catalyses the conversion of estrone to estradiol in the presence of the cofactor NADH. The initial rate data for the enzyme-catalysed reaction in the absence of an inhibitor is as follows:

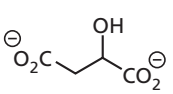
   Substrate concentration (10^{-2} mol dm^{-3}) 5 10 25 50 100

   Initial rate (10^{-1} mol dm^{-3} s^{-1}) 28.6 51.5 111 141 145

   Create a Michaelis Menton plot and a Lineweaver-Burk plot. Use both plots to calculate the values of K_M and the maximum rate of reaction. Identify which plot is likely to give the more accurate results and explain why this is the case.
6. Lactate dehydrogenase has a 1000-fold selectivity for lactate as a substrate over malate. However, if a mutation occurs that alters an active site glutamine residue to an arginine residue, the enzyme shows a 10,000-fold selectivity for malate over lactate. Explain this astonishing transformation.



Lactate



Malate

FURTHER READING

Broadwith, P. (2010) Enzymes do the twist. *Chemistry World*. Available at: http://www.rsc.org/chemistryworld/News/2010/January/06011001.asp (last accessed 14 June 2012).

Knowles, J. R. (1991) Enzyme catalysis: not different, just better. *Science* **350**, 121–124.

Maryanoff, B. E. and Maryanoff, C. A. (1992) Some thoughts on enzyme inhibition and the quiescent affinity label concept. *Advances in Medicinal Chemistry* **1**, 235–261.

Navia, M. A. and Murcko, M. A. (1992) Use of structural information in drug design. *Current Opinion in Structural Biology* **2**, 202–216.

Teague, S. J. (2003) Implications of protein flexibility for drug discovery. *Nature Reviews Drug Discovery* **2**, 527–541.

4 Receptors: structure and function

In this chapter we discuss the structure and function of receptors. Drug action at receptors is discussed in Chapter 8 and in other chapters throughout the text.

4.1 Role of the receptor

Receptors are proteins which are, by far, the most important drug targets in medicine. They are implicated in ailments such as pain, depression, Parkinson's disease, psychosis, heart failure, asthma, and many other problems. What are these receptors and what do they do?

In a complex organism there has to be a communication system between cells. After all, it would be pointless if individual heart cells were to contract at different times. The heart would then be a wobbly jelly and totally useless in its function as a pump. Communication is essential to ensure that all heart muscle cells contract at the same time. The same is true for all the organs and tissues of the body if they are to operate in a coordinated and controlled fashion.

Control and communication come primarily from the brain and spinal column (the **central nervous system**), which receives and sends messages via a vast network of nerves (Fig. 4.1). The detailed mechanism by which nerves transmit messages along their length need not concern us here (see Appendix 4). It is sufficient for our purposes to think of the message as being an electrical pulse which travels down the nerve cell (**neuron**) towards the target, whether that be a muscle cell or another neuron. If that was all there was to it, it would be difficult to imagine how drugs could affect this communication system. However, there is one important feature that is crucial to our understanding of drug action. Neurons do not connect directly to their target cells. They stop just short of the cell surface. The distance is minute, about 100 Å, but it is a space that the electrical 'pulse' is unable to jump.

Therefore, there has to be a method of carrying the message across the gap between the nerve ending and the target cell. The problem is solved by the release of a chemical messenger called a **neurotransmitter** from the nerve cell (Fig. 4.2). Once released, this chemical messenger diffuses across the gap to the target cell, where it binds and interacts with a specific protein (receptor) embedded in the cell membrane. This process of binding leads to a series or cascade of secondary effects, which results either in a flow of ions across the cell membrane or in the switching on (or off) of enzymes inside the target cell. A biological response then results, such as the contraction of a muscle cell or the activation of fatty acid metabolism in a fat cell.

The first person to propose the existence of receptors was Langley in 1905. Up until that point, it was thought that drugs acted to prevent the release of the neurotransmitter from the neuron, but Langley was able to show that certain target cells responded to the drug nicotine, even when the neurons supplying those cells were dead.

So far, we have talked about cellular communication involving neurons and neurotransmitters, but cells also receive chemical messages from circulating **hormones**. Once again, receptors are responsible for binding these messengers and triggering a series of secondary effects.

We shall consider these secondary effects and how they result in a biological action in Chapter 5, but, for the moment, the important thing to note is that the communication system depends crucially on a chemical messenger. As a chemical process is involved, it should be possible for other chemicals (drugs) to interfere or interact with the process.

4.2 Neurotransmitters and hormones

There are a large variety of messengers that interact with receptors and they vary significantly in structure and complexity. Some neurotransmitters are simple molecules, such as monoamines (e.g. **acetylcholine**, **noradrenaline**, **dopamine**, and **serotonin)** or amino acids (e.g.

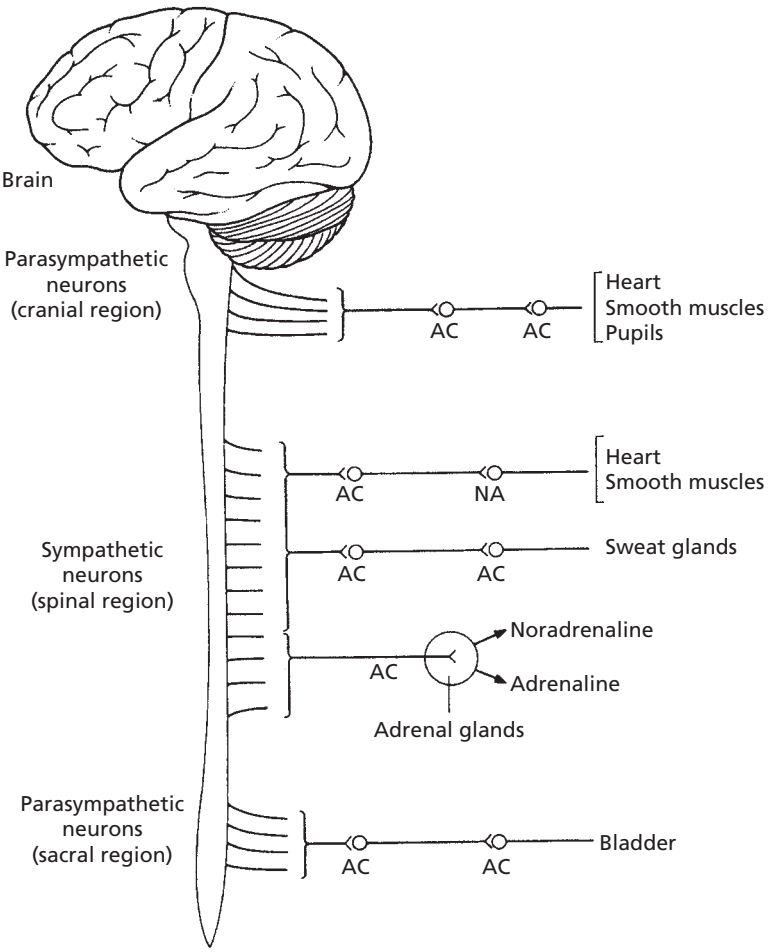


FIGURE 4.1 The central nervous system (AC = acetylcholine; NA = noradrenaline). Taken from Mann, J. (1992) *Murder, Magic, and Medicine*. Oxford University Press, with permission.

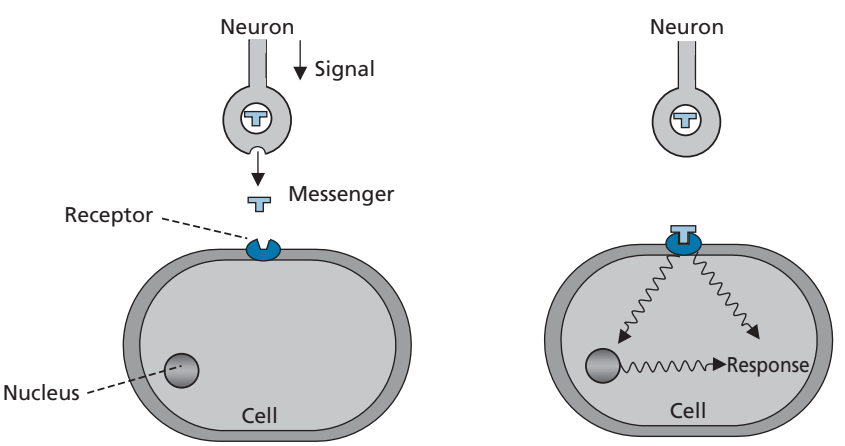


FIGURE 4.2 Neurotransmitters act as chemical messengers that bind to receptors and trigger reactions within a cell.

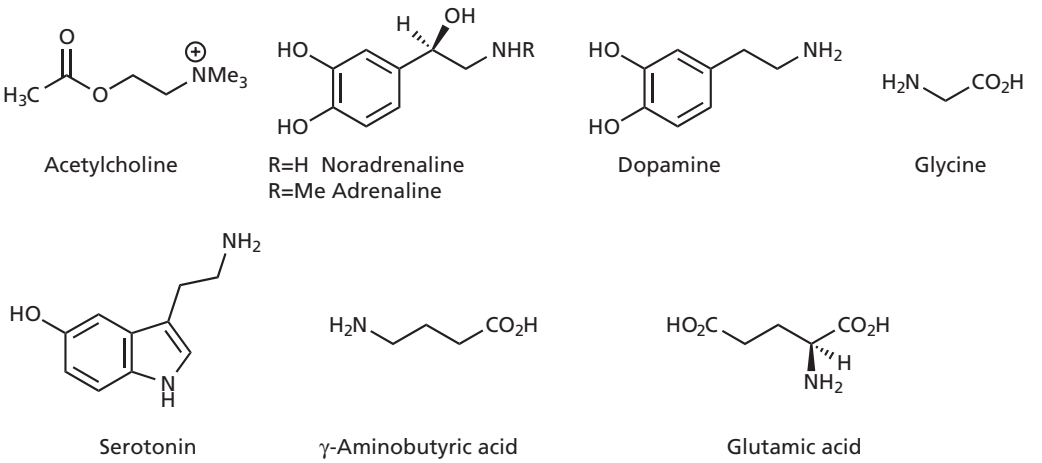


FIGURE 4.3 Examples of neurotransmitters and the hormone adrenaline.

γ-aminobutyric acid [GABA], **glutamic acid,** and **glycine**) (Fig. 4.3). Even the calcium ion can act as a chemical messenger. Other chemical messengers are more complex in structure and include lipids, such as **prostaglandins**; purines, such as **adenosine** or **ATP** (Chapter 6); **neuropeptides**, such as **endorphins** and **enkephalins** (section 24.8); peptide hormones, such as **angiotensin** or **bradykinin**; and even enzymes, such as **thrombin**.

In general, a neuron releases mainly one type of neurotransmitter, and the receptor which awaits it on the target cell will be specific for that messenger. However, that does not mean that the target cell has only one type of receptor protein. Each target cell has a large number of neurons communicating with it and they do not all use the same neurotransmitter (Fig. 4.4). Therefore, the target cell will have other types of receptors specific for those other neurotransmitters. It may also have receptors waiting to receive messages from chemical messengers that have longer distances to travel. These are the hormones released into the circulatory system by various glands in the body. The best known example of a hormone is **adrenaline**. When danger or exercise is anticipated, the adrenal medulla gland releases adrenaline into the bloodstream where it is carried round the body, preparing it for vigorous exercise.

Hormones and neurotransmitters can be distinguished by the route they travel and by the way they are released, but their action when they reach the target cell is the same. They both interact with a receptor and a message is received. The cell responds to that message and adjusts its internal chemistry accordingly, and a biological response results.

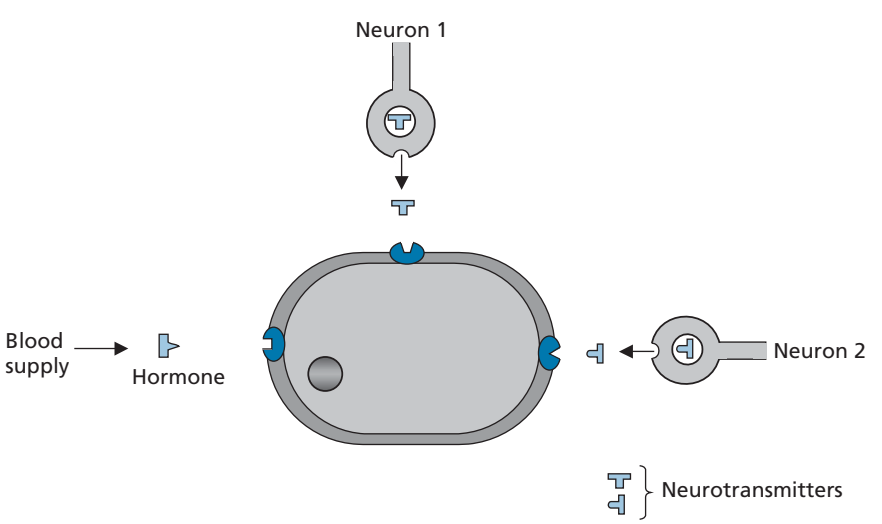


FIGURE 4.4 Target cell containing various receptors specific to different types of messenger.

4.3 Receptor types and subtypes

Receptors are identified by the specific neurotransmitter or hormone which activates them. Thus, the receptor activated by **dopamine** is called the **dopaminergic receptor**, the receptor activated by **acetylcholine** is called the **cholinergic receptor**, and the receptor activated by **adrenaline** or **noradrenaline** is called the **adrenergic receptor** or **adrenoceptor**.

However, not all receptors activated by the same chemical messenger are exactly the same throughout the body. For example, the adrenergic receptors in the lungs are slightly different from the adrenergic receptors in the heart. These differences arise from slight variations in amino acid composition; if the variations are in the binding site, it allows medicinal chemists to design drugs which can distinguish between them. For example, adrenergic drugs can be designed to be 'lung' or 'heart' selective. In general, there are various *types* of a particular receptor and various *subtypes* of these, which are normally identified by numbers or letters. Having said that, some of the early receptors that were discovered were named after natural products which bound to them, for example the muscarinic and nicotinic types of cholinergic receptor (section 22.4).

Some examples of receptor types and subtypes are given in Fig. 4.16. The identification of many of these subtypes is relatively recent and the current emphasis in medicinal chemistry is to design drugs that are as selective as possible for receptor types and subtypes so that the drugs are tissue selective and have fewer side effects.

4.4 Receptor activation

A receptor is a protein molecule usually embedded within the cell membrane with part of its structure exposed on the outside of the cell. The protein surface is a complicated shape containing hollows, ravines, and ridges. Somewhere within this complicated geography there is an area that has the correct shape to accept the incoming messenger. This area is known as the **binding site** and is analogous to the active site of an enzyme (section 3.3). When the chemical messenger fits into this site it 'switches on' the receptor molecule and a message is received (Fig. 4.5). However, there is an important difference between enzymes and receptors in that the chemical messenger does not undergo a chemical reaction. It fits into the binding site of the receptor protein, passes on its message, and then leaves unchanged. If no reaction takes place, what *has* happened? How does the chemical messenger tell the receptor its message and how is this message conveyed to the cell? The first thing to note is that when the messenger fits the binding site of the protein receptor it causes the binding site to change shape. This is known as an **induced fit**. This, in turn, has wider ramifications as there is a knock-on effect which causes the overall protein to change shape. But how does an induced fit happen and what is the significance of the receptor changing shape?

4.5 How does the binding site change shape?

As we have seen, the binding site of a receptor changes shape when a chemical messenger fits into it. This is not a moulding process in which the binding site wraps itself around the messenger. Instead, the induced fit is brought about by the intermolecular binding interactions that can take place between the messenger and the binding site. This is exactly the same process that occurs when a substrate binds to the active site of an enzyme (section 3.5.1), but, in this situation, no catalysed reaction follows binding.

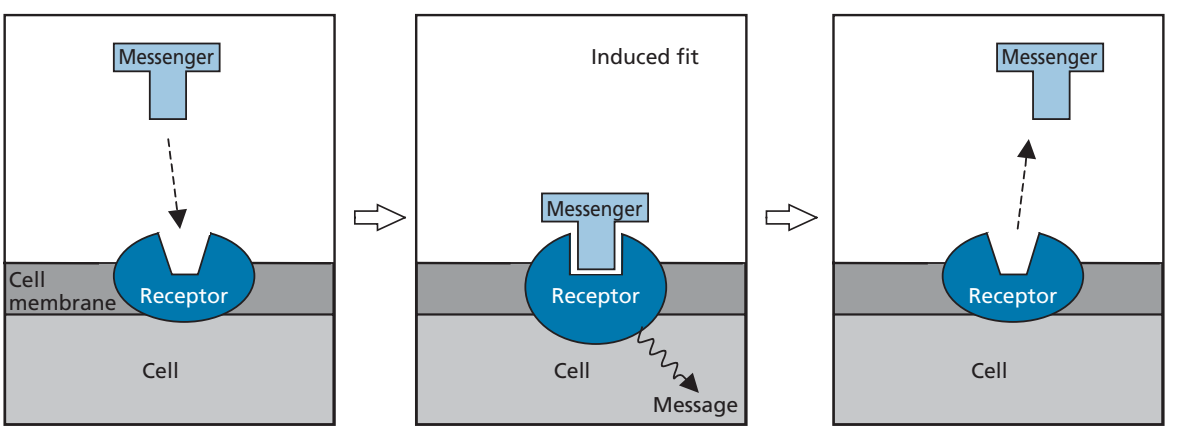


FIGURE 4.5 Binding of a chemical messenger to a protein receptor.

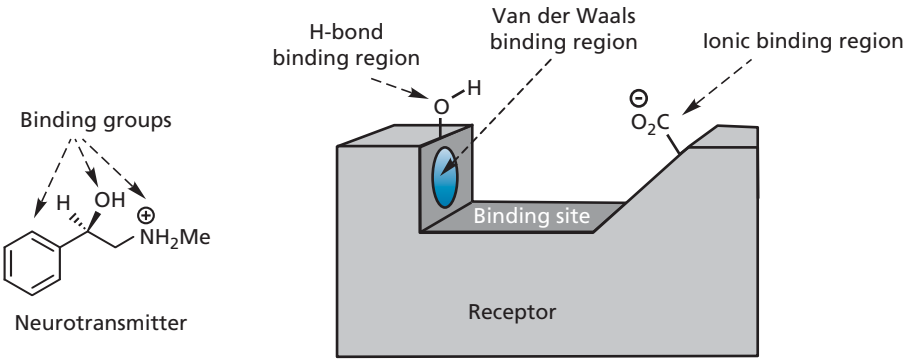


FIGURE 4.6 A hypothetical receptor and neurotransmitter.

To illustrate how binding interactions result in an induced fit, let us consider a hypothetical neurotransmitter and a hypothetical binding site as shown in Fig. 4.6. The neurotransmitter has an aromatic ring that can take part in van der Waals interactions, an alcohol OH group that can take part in hydrogen bonding interactions, and a charged nitrogen centre that can take part in ionic or electrostatic interactions. These functional groups are the messenger's **binding groups**.

The hypothetical binding site contains three **binding regions** which contain functional groups that are complimentary to the binding groups of the messenger. The messenger fits into the binding site such that intermolecular interactions take place between the messenger's binding groups and the receptor's binding regions (Fig. 4.7). However, the fit is not perfect. In the diagram, there are good van der Waals and hydrogen bond interactions, but the ionic interaction is not as strong as it could be. The ionic binding region is close enough to have a weak interaction with the messenger, but not close enough for the optimum interaction. The receptor protein therefore alters shape to bring the carboxylate group closer to the positively charged nitrogen and to obtain a stronger interaction. As a result, the shape of the binding site is altered and an induced fit has taken place.

The illustration shown here is a simplification of the induced fit process and, in reality, both the messenger and the binding site take up different conformations or shapes to maximize the bonding forces between them. As with enzyme–substrate binding, there is a fine balance involved in receptor–messenger binding. The bonding forces must be large enough to change the shape of the binding site, but not so strong that the messenger is unable to leave. Most neurotransmitters bind quickly to their receptors then 'shake themselves loose' once their message has been received.

We have now seen how a chemical messenger can cause an induced fit in the binding site of a receptor protein. However, this induced fit has a knock-on effect which alters the overall shape of the protein. It is this overall shape change that is crucial to the activation of a receptor and in its ability to trigger an amazing 'domino effect' which affects the cell's internal chemistry. This domino effect involves several different proteins and enzymes, and ultimately produces an observed biological effect. The process by which this takes place is called **signal transduction** and is covered in more detail in Chapter 5.

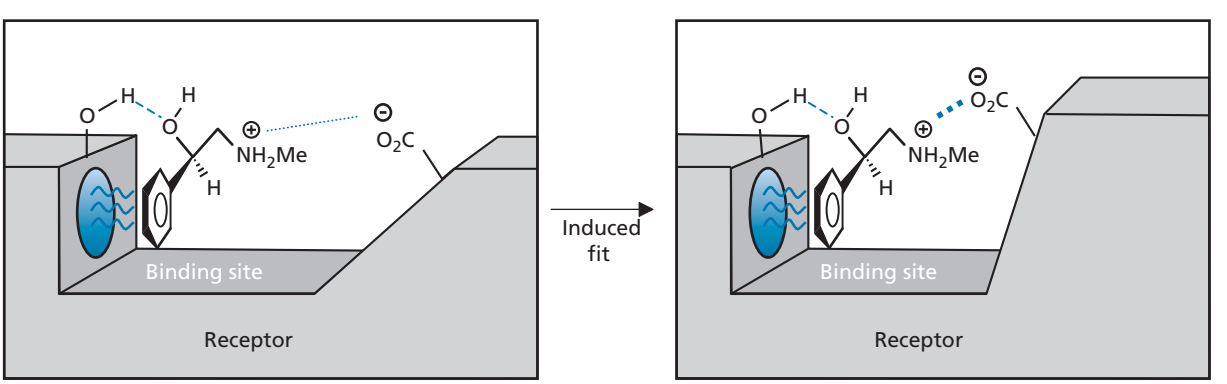


FIGURE 4.7 Binding of a hypothetical neurotransmitter to a binding site resulting in an induced fit.

Signal amplification is an important feature of this process as it means that a relatively small number of neurotransmitter molecules can have a dramatic effect on the cell's internal chemistry. In this chapter, we shall focus on the structure of different receptors and the process by which they are activated and trigger the signal transduction process.

There are three different types (or families) of membrane-bound receptors:

- ion channel receptors;
- G-protein-coupled receptors;
- kinase-linked receptors.

We shall consider each of these in turn in sections 4.6–4.8.

KEY POINTS

- Most receptors are membrane-bound proteins that contain an external binding site for hormones or neurotransmitters. Binding results in an induced fit that changes the receptor conformation. This triggers a series of events that ultimately results in a change in cellular chemistry.
- Neurotransmitters and hormones do not undergo a reaction when they bind to receptors. They depart the binding site unchanged once they have passed on their message.
- The interactions that bind a chemical messenger to the binding site must be strong enough to allow the chemical message to be received, but weak enough to allow the messenger to depart.
- Binding groups are the functional groups present on a messenger molecule which are used for binding it to the receptor binding site.
- Binding regions are regions of the receptor binding site which contain functional groups capable of forming intermolecular bonds to the binding groups of a messenger molecule.

4.6 Ion channel receptors

4.6.1 General principles

Some neurotransmitters operate by controlling ion channels. What are these ion channels and why are they necessary? Let us look again at the structure of the cell membrane.

As described in section 1.2.1, the membrane is made up of a bilayer of phospholipid molecules so the middle of the cell membrane is 'fatty' and hydrophobic. Such a barrier makes it difficult for polar molecules or ions to move in or out of the cell. Yet, it is important that these species should cross. For example, the movement of sodium and potassium ions across the membrane is crucial to the function of nerves (Appendix 4). It seems an intractable problem, but, once again, the ubiquitous proteins provide the answer by forming ion channels.

Ion channels are complexes made up of five protein subunits which traverse the cell membrane (Fig. 4.8). The centre of the complex is hollow and lined with polar amino acids to give a hydrophilic tunnel, or pore.

Ions can cross the fatty barrier of the cell membrane by moving through these hydrophilic channels or tunnels. But there has to be some control. In other words, there has to be a 'lock gate' that can be opened or closed as required. It makes sense that this lock gate should be controlled by a receptor protein sensitive to an external chemical messenger, and this is exactly what happens. In fact, the receptor protein is an integral part of the ion channel complex and is one or more of the constituent protein subunits. In the resting state, the ion channel is closed (i.e. the lock gate is shut). However, when a chemical messenger binds to the external binding site of the receptor protein, it causes an induced fit which causes the protein to change shape. This, in turn, causes the overall protein complex to change shape, opening up

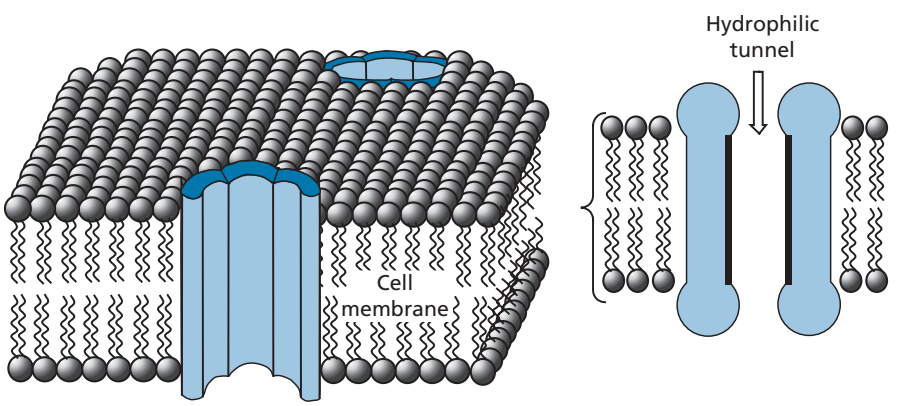


FIGURE 4.8 The structure of an ion channel. The bold lines show the hydrophilic sides of the channel.

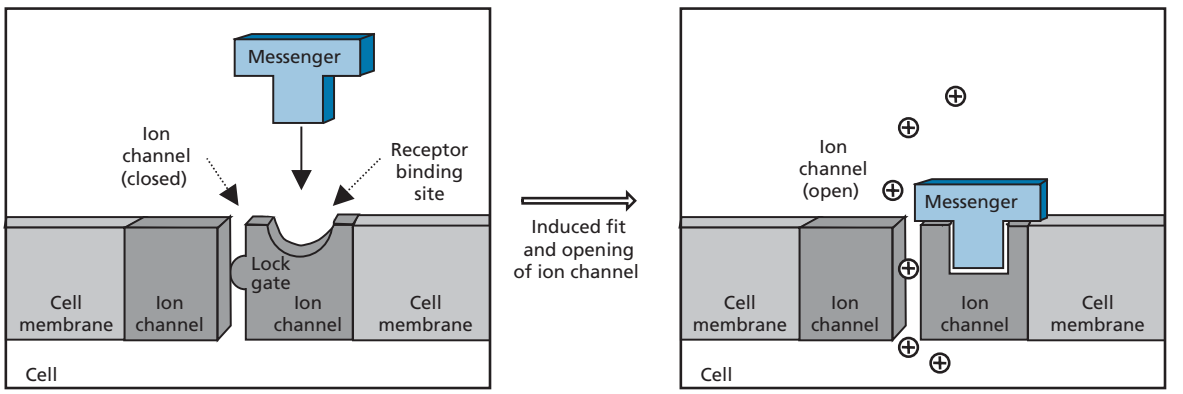


FIGURE 4.9 Lock-gate mechanism for opening ion channels.

the lock gate and allowing ions to pass through the ion channel (Fig. 4.9). We shall look at this in more detail in section 4.6.3.

The operation of an ion channel explains why the relatively small number of neurotransmitter molecules released by a neuron is able to have such a significant biological effect on the target cell. By opening a few ion channels, several thousand ions are mobilized for each neurotransmitter molecule involved. Moreover, the binding of a neurotransmitter to an ion channel results in a rapid response, measured in a matter of milliseconds. This is why the synaptic transmission of signals between neurons usually involves ion channels.

Ion channels are specific for certain ions. For example there are different cationic ion channels for sodium (Na^+), potassium (K^+), and calcium (Ca^{2+}) ions. There are also anionic ion channels for the chloride ion (Cl^-). The ion selectivity of different ion channels is dependent on the amino acids lining the ion channel. It is interesting to note that the mutation of just one amino acid in this area is sufficient to change a cationic-selective ion channel to one that is selective for anions.

4.6.2 Structure

The five protein subunits that make up an ion channel are actually **glycoproteins** (sections 2.5 and 10.7.1), but we will refer to them here as proteins. The protein subunits in an ion channel are not identical. For example, the ion channel controlled by the nicotinic cholinergic receptor is made up of five subunits of four different types [α (×2) β, γ, δ]; the ion channel controlled by the glycine receptor is made up of five subunits of two different types [α (×3), β (×2)] (Fig. 4.10).

The receptor protein in the ion channel controlled by glycine is the α-subunit. Three such subunits are present, all of which are capable of interacting with glycine. However, the situation is slightly more complex in the nicotinic ion channel controlled by the neurotransmitter acetylcholine. Most of the binding site is on the α-subunit, but there is some involvement from neighbouring subunits. In this case, the ion channel complex as a whole might be viewed as the receptor.

Let us now concentrate on the individual protein subunits. Although there are various types of these, they all fold up in a similar manner such that the protein chain traverses the cell membrane four times. This means that

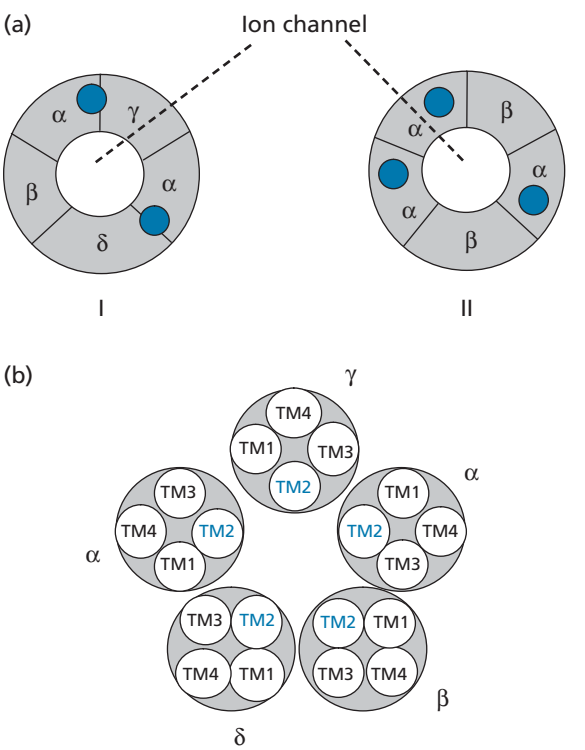


FIGURE 4.10 (**a**) Pentameric structure of ion channels (transverse view). I, ion channel controlled by a nicotinic cholinergic receptor; II, ion channel controlled by a glycine receptor. The coloured circles indicate ligand binding sites. (**b**) Transverse view of I, including transmembrane regions.

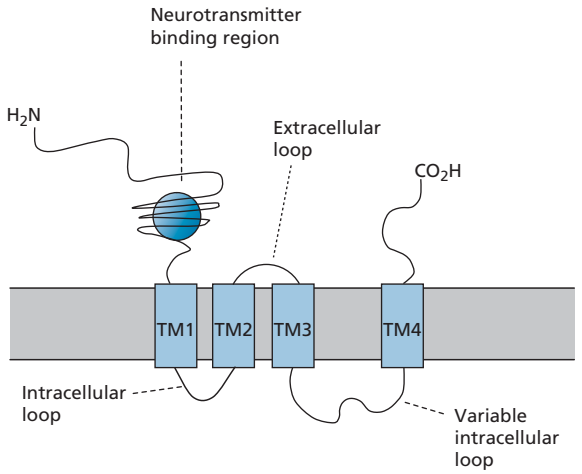


FIGURE 4.11 Structure of the four transmembrane (4-TM) receptor subunit.

each subunit has four transmembrane (TM) regions which are hydrophobic in nature. These are labelled TM1–TM4. There is also a lengthy *N*-terminal extracellular chain which (in the case of the α-subunit) contains the ligand-binding site (Fig. 4.11).

The subunits are arranged such that the second transmembrane region of each subunit faces the central pore of the ion channel (Fig. 4.10). We shall see the significance of this when we look at the next section.

4.6.3 Gating

When the receptor binds a ligand, it changes shape which has a knock-on effect on the protein complex, causing the ion channel to open—a process called **gating** (Fig. 4.12).

The binding of a neurotransmitter to its binding site causes a conformational change in the receptor, which eventually opens up the central pore and allows ions to flow. This conformational change is quite complex, involving several knock-on effects from the initial binding process. This must be so, as the binding site is quite far from the lock gate. Studies have shown that the lock gate is made up of five kinked α-helices where one helix (the 2-TM region) is contributed by each of the five protein subunits. In the closed state the kinks point towards each other. The conformational change induced by ligand binding causes each of these helices to rotate such that the kink points the other way, thus opening up the pore (Fig. 4.13).

4.6.4 Ligand-gated and voltage-gated ion channels

The ion channels that we have discussed so far are called **ligand-gated ion channels** as they are controlled by chemical messengers (**ligands**). There are other types of ion channel which are not controlled by ligands, but are instead sensitive to the potential difference that exists across a cell membrane—the **membrane potential**. These ion channels are present in the axons of excitable cells (i.e. neurons) and are called **voltage-gated ion channels.** They are crucial to the transmission of a signal along individual neurons and are important drug targets for local anaesthetics. A description of these ion channels is given in Appendix 4.

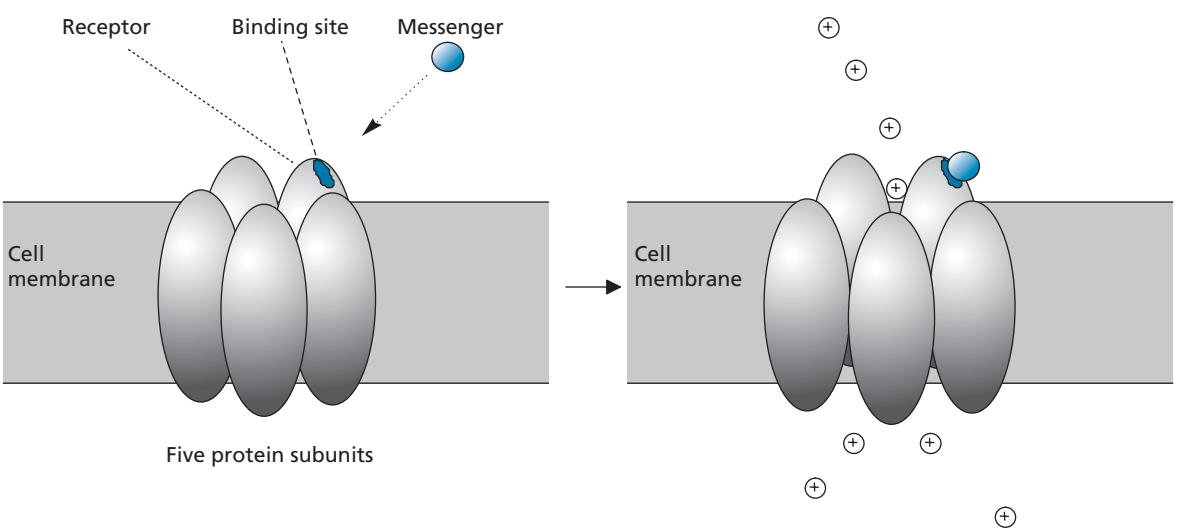


FIGURE 4.12 Opening of an ion channel (gating).

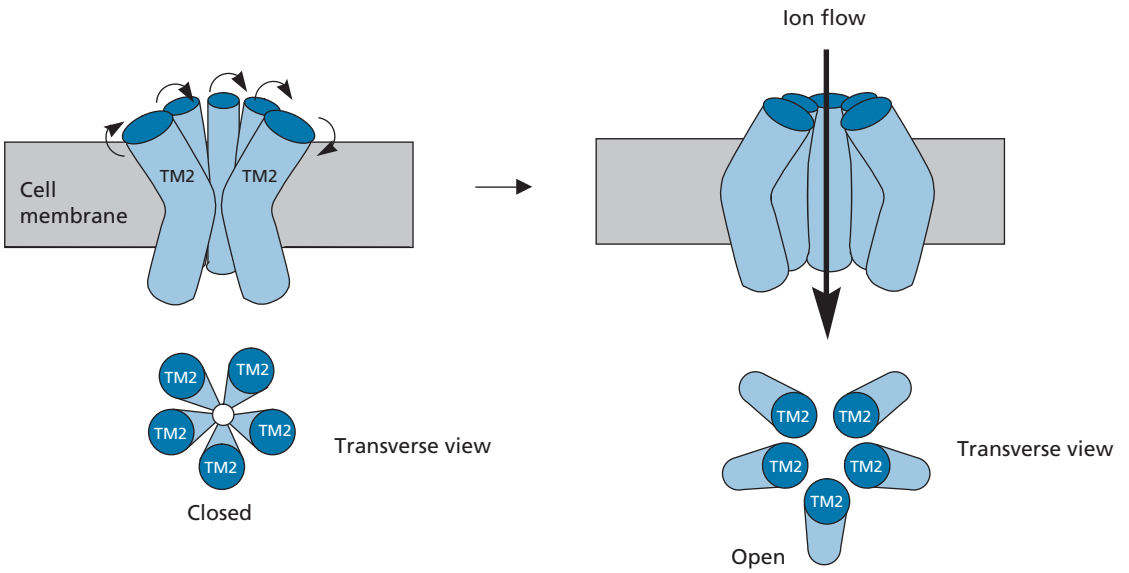


FIGURE 4.13 Opening of the 'lock gate' in an ion channel.

KEY POINTS

- Receptors controlling ion channels are an integral part of the ion channel. Binding of a messenger induces a change in shape, which results in the rapid opening of the ion channel.
- Receptors controlling ion channels are called ligand-gated ion channel receptors. They consist of five protein subunits with the receptor binding site being present on one or more of the subunits.
- Binding of a neurotransmitter to an ion channel receptor causes a conformational change in the protein subunits such that the second transmembrane domain of each subunit rotates to open the channel.

4.7 G-protein-coupled receptors

4.7.1 General principles

The **G-protein-coupled receptors** are some of the most important drug targets in medicinal chemistry. Indeed, some 30% of all drugs on the market act by binding to these receptors. In general, they are activated by hormones and slow-acting neurotransmitters. They include the **muscarinic receptor** (section 22.11), **adrenergic receptors** (section 23.2), and **opioid receptors** (section 24.4). The response from activated G-protein-coupled receptors is measured in seconds. This is slower than the response of ion channels, but faster than the response of kinase-linked receptors (section 4.8), which takes a matter of minutes. There are a large number of different G-protein-coupled receptors interacting with important neurotransmitters, such as acetylcholine, dopamine, histamine, serotonin, glutamate, and noradrenaline. Other G-protein-coupled receptors are activated by peptide and protein hormones, such as the enkephalins and endorphins.

G-protein-coupled receptors are membrane-bound proteins that are responsible for activating proteins called **G-proteins** (Fig. 4.14). These latter proteins act as **signal proteins** because they are capable of activating or deactivating membrane-bound enzymes (sections 5.1–5.2). Consequently, activation of the receptor by a chemical messenger influences the reactions that take place within the cell.

The receptor protein is embedded within the membrane, with the binding site for the chemical messenger exposed on the outer surface. On the inner surface, there is another binding site which is normally closed (Fig. 4.14, frame 1). When the chemical messenger binds to its binding site, the receptor protein changes shape, opening up the binding site on the inner surface. This new binding site is recognized by the G-protein, which then binds (Fig. 4.14, frame 2). The G-protein is attached to the inner surface of the cell membrane and is made up of three protein subunits, but once it binds to the receptor the complex is destabilized and fragments to a monomer and a dimer (Fig. 4.14, frame 3). These then interact with membrane-bound enzymes to continue the signal transduction process (sections 5.1–5.3).

There are several different G-proteins, which are recognized by different types of receptor. Some of the activated subunits from these G-proteins have an inhibitory effect on a membrane-bound enzyme, while others have a stimulatory effect. Nevertheless, the mechanism by which the G-protein is activated by fragmentation is the same.

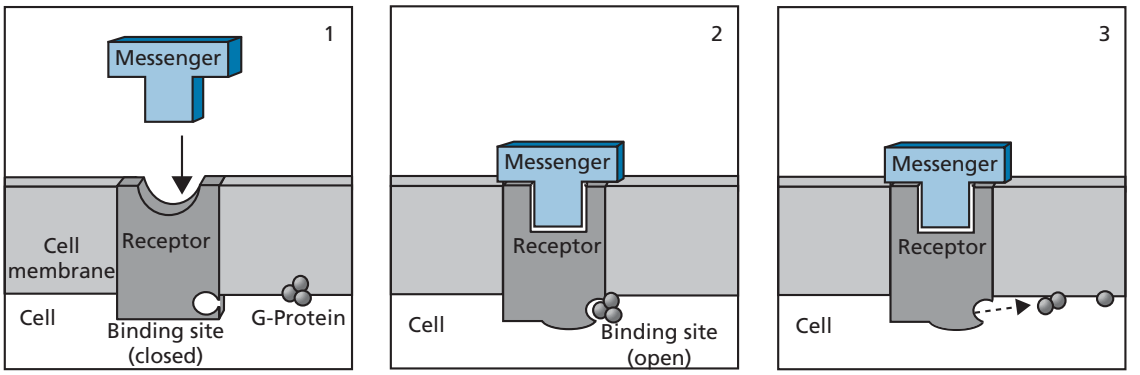


FIGURE 4.14 Activation of a G-protein-coupled receptor and G-protein.

There is a substantial amplification of the signal in this process, as one activated receptor activates several G-proteins.

4.7.2 Structure

The G-protein-coupled receptors fold up within the cell membrane such that the protein chain winds back and forth through the cell membrane seven times (Fig. 4.15). Each of the seven transmembrane sections is hydrophobic and helical in shape, and it is usual to assign these helices with roman numerals (I, II, etc.) starting from the *N*-terminus of the protein. Owing to the number of transmembrane regions, the G-proteins are also called **7-TM receptors**. The binding site for the G-protein is situated on the intracellular side of the protein and involves part of the *C*-terminal chain, as well as part of the variable intracellular loop (so called because the length of this loop varies between different types of receptor). As one might expect, the binding site for the neurotransmitter or hormone messenger is on the extracellular portion of the protein. The exact position of the binding site varies from receptor to receptor. For example, the binding site for the adrenergic receptor is in a deep binding pocket between the transmembrane helices, whereas the binding site for the glutamate receptor involves the *N*-terminal chain and is situated above the surface of the cell membrane.

4.7.3 The rhodopsin-like family of G-protein-coupled receptors

The G-protein-coupled receptors include the receptors for some of the best-known chemical messengers in medicinal chemistry (e.g. glutamic acid, GABA, noradrenaline, dopamine, acetylcholine, serotonin, prostaglandins, adenosine, endogenous opioids, angiotensin, bradykinin, and thrombin). Considering the structural variety of the chemical messengers involved, it is remarkable that the overall structures of the G-protein-coupled receptors

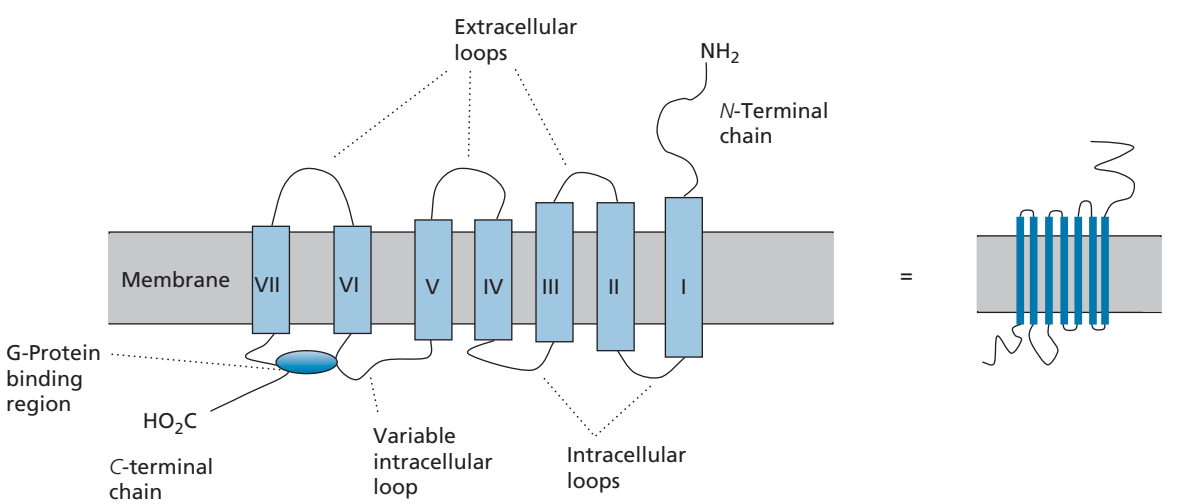


FIGURE 4.15 Structure of G-protein-coupled receptors.

are so similar. Nevertheless, despite their similar overall structure, the amino acid sequences of the receptors vary quite significantly. This implies that these receptors have evolved over millions of years from an ancient common ancestral protein. Comparing the amino acid sequences of the receptors allows us to construct an evolutionary tree and to group the receptors of this superfamily into various sub-families, which are defined as class A (rhodopsin-like receptors), class B (secretin-like receptors), and class C (metabotropic glutamate-like and pheromone receptors). The most important of these, as far as medicinal chemistry is concerned, is the rhodopsin-like family—so called because the first receptor of this family to be studied in detail was the rhodopsin receptor itself, a receptor involved in the visual process. A study of the evolutionary tree of rhodopsin-like receptors throws up some interesting observations (Fig. 4.16).

First of all, the evolutionary tree illustrates the similarity between different kinds of receptors based on their relative positions on the tree. Thus, the muscarinic, α-adrenergic, β-adrenergic, histamine, and dopamine receptors have evolved from a common branch of the evolutionary tree and have greater similarity to each other than to any receptors arising from an earlier evolutionary branch (e.g. the **angiotensin receptor**). Such receptor similarity may prove a problem in medicinal chemistry. Although the receptors are distinguished by different neurotransmitters or hormones in the body, a drug may not manage to make that distinction. Therefore, it is important to ensure that any new drug aimed at one kind of receptor (e.g. the dopamine receptor) does not interact with a similar kind of receptor (e.g. the muscarinic receptor).

Receptors have further evolved to give receptor *types* and *subtypes* which recognize the same chemical messenger, but are structurally different. For example, there are two types of adrenergic receptor (α and β), each of which has various subtypes (α_1, α_{2A}, α_{2B}, α_{2C}, β_1, β_2, β_3). There are two types of cholinergic receptor—nicotinic (an ion channel receptor) and muscarinic (a 7-TM receptor). Five subtypes of the muscarinic cholinergic receptor have been identified.

The existence of receptor subtypes allows the possibility of designing drugs that are selective for one receptor subtype over another. This is important, because one receptor subtype may be prevalent in one part of the body (e.g. the gut), while a different receptor subtype is prevalent in another part (e.g. the heart). Therefore, a drug that is designed to interact selectively with the receptor subtype in the gut is less likely to have side effects on the heart. Even if the different receptor subtypes are present in the same part of the body, it is still important to make drugs as selective as possible because different receptor subtypes frequently activate different signalling systems, leading to different biological results.

A closer study of the evolutionary tree reveals some curious facts about the origins of receptor subtypes. As one might expect, various receptor subtypes have diverged from a common evolutionary branch (e.g. the dopamine subtypes D2, D3, D4). This is known as **divergent evolution** and there should be close structural similarity between these subtypes. However, receptor subtypes are also found in separate branches of the tree. For example, the dopamine receptor subtypes ($D1_A$, $D1_B$, and D5) have developed from a different evolutionary branch. In other words, the ability of a receptor to bind dopamine has developed in different evolutionary branches—an example of **convergent evolution**.

Consequently, there may sometimes be greater similarities between receptors which bind different ligands but which have evolved from the same branch of the tree

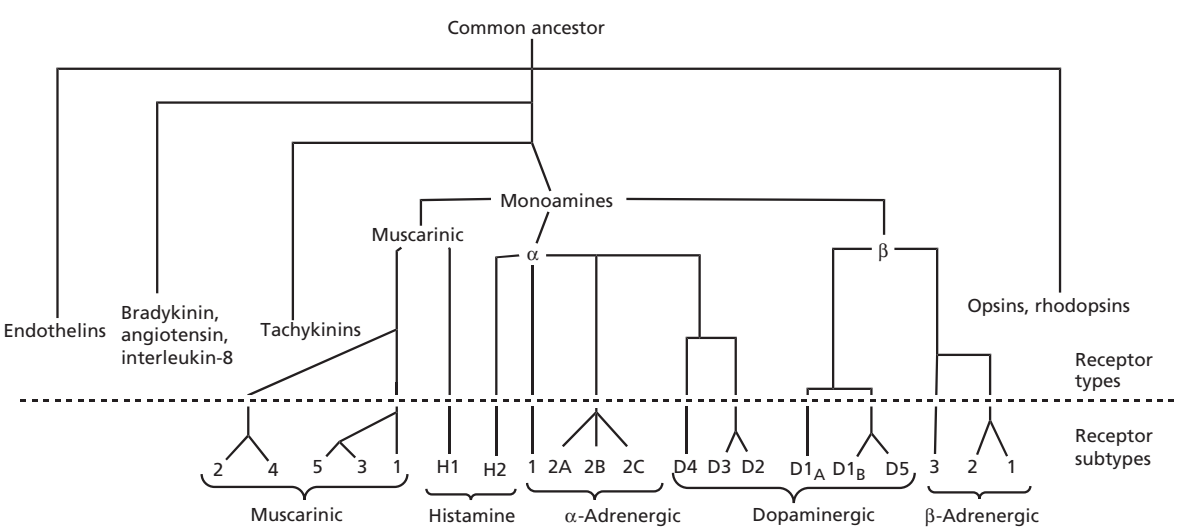


FIGURE 4.16 Evolutionary tree of G-protein-coupled receptors.

than there are between the various subtypes of receptors which bind the same ligand. For example, the histamine H_1 receptor resembles a muscarinic receptor more closely than it does the histamine H_2 receptor. Again, this has important consequences in drug design because there is an increased possibility that a drug aimed at a muscarinic receptor may also interact with a histamine H_1 receptor and lead to unwanted side effects.

As these receptors are membrane bound, it is not easy to crystallize them for X-ray crystallographic studies. However, the X-ray crystal structures of the β_2 and β_1 adrenoceptors have now been determined.

4.7.4 Dimerization of G-coupled receptors

There is strong evidence that some G-coupled receptors can exist as dimeric structures containing identical or different types of receptor—homodimers or heterodimers respectively. The presence of these receptor dimers appears to vary between different tissues and this has important consequences for drug design. An agent that is selective for one type of receptor would not normally affect other types. However, if receptor heterodimers are present, a 'communication' is possible between the component receptors such that an agent interacting with one half of the dimer may affect the activity of the other half. This is discussed further in section 24.9 with respect to opioid receptors.

KEY POINTS

- G-protein-coupled receptors activate signal proteins called G-proteins. Binding of a messenger results in the opening of a binding site for the signal protein. The latter binds and fragments, with one of the subunits departing to activate a membrane-bound enzyme.
- The G-protein-coupled receptors are membrane-bound proteins with seven transmembrane sections. The *C*-terminal chain lies within the cell and the *N*-terminal chain is extracellular.
- The location of the binding site differs between different G-protein-coupled receptors.
- The rhodopsin-like family of G-protein-coupled receptors includes many receptors that are targets for currently important drugs.
- Receptor types and subtypes recognize the same chemical messenger, but have structural differences, making it possible to design drugs that are selective for one type (or subtype) of receptor over another.
- Receptor subtypes can arise from divergent or convergent evolution.
- It is possible for some G-protein coupled receptors to exist as dimeric structures.

4.8 Kinase-linked receptors

4.8.1 General principles

Kinase-linked receptors are a superfamily of receptors which activate enzymes directly and do not require a G-protein (Fig. 4.17). **Tyrosine kinase receptors** are important examples of kinase-linked receptors and are proving to be highly important targets for novel anticancer drugs (section 21.6.2). In these structures, the protein concerned plays the dual role of receptor and enzyme. The receptor protein is embedded within the cell membrane, with part of its structure exposed on the outer surface of the cell and part exposed on the inner surface. The outer surface contains the binding site for the chemical messenger and the inner surface has an active site that is closed in the resting state. When a chemical messenger binds to the receptor it causes the protein to change shape. This results in the active site being opened up, allowing the protein to act as an enzyme within the cell. The reaction that is catalysed is a phosphorylation reaction where tyrosine residues on

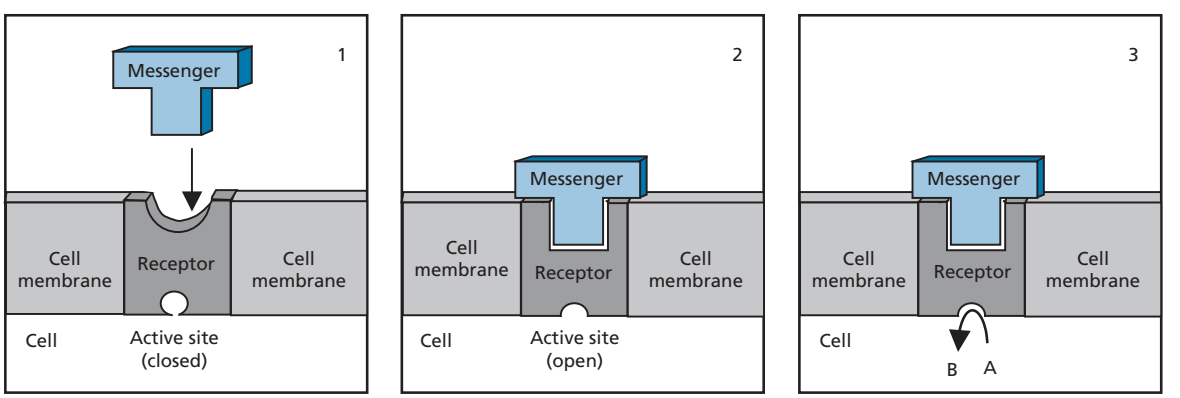


FIGURE 4.17 Enzyme activation.

a protein substrate are phosphorylated. An enzyme that catalyses phosphorylation reactions is known as a kinase enzyme and so the protein is referred to as a tyrosine kinase receptor. ATP is required as a cofactor to provide the necessary phosphate group. The active site remains open for as long as the messenger molecule is bound to the receptor, and so several phosphorylation reactions can occur, resulting in an amplification of the signal. A curiosity of this enzyme-catalysed reaction is that the substrate for the reaction is the receptor itself. This is explained more fully in section 4.8.3.

The kinase-linked receptors are activated by a large number of polypeptide hormones, growth factors, and cytokines. Loss of function of these receptors can lead to developmental defects or hormone resistance. Overexpression can result in malignant growth disorders.

4.8.2 Structure of tyrosine kinase receptors

The basic structure of a tyrosine kinase receptor consists of a single extracellular region (the *N*-terminal chain) that includes the binding site for the chemical messenger, a single hydrophobic region that traverses the membrane as an α-helix of seven turns (just sufficient to traverse the membrane), and a *C*-terminal chain on the inside of the cell membrane (Fig. 4.18). The *C*-terminal region contains the catalytic binding site. Examples of tyrosine kinase receptors include the receptor for **insulin,** and receptors for various **cytokines** and **growth factors**.

4.8.3 Activation mechanism for tyrosine kinase receptors

A specific example of a tyrosine kinase receptor is the receptor for a hormone called **epidermal growth factor** (EGF). EGF is a **bivalent ligand** which can bind to two receptors at the same time. This results in **receptor dimerization**, as well as activation of enzymatic activity. The dimerization process is important because the active site on each half of the receptor dimer catalyses the phosphorylation of accessible tyrosine residues on the other half (Fig. 4.19). If dimerization did not occur, no phosphorylation would take place. Note that these phosphorylations occur on the intracellular portion of the receptor protein chain. The relevance of these phosphorylation reactions will be explained in section 5.4.1. The important point to grasp at this stage is that an external chemical messenger has managed to convey its message to the interior of the cell without itself being altered or having to enter the cell.

Dimerization and auto-phosphorylation are common themes for receptors in this family. However, some of the receptors in this family already exist as dimers or tetramers, and only require binding of the ligand. For example, the **insulin** receptor is a heterotetrameric complex (Fig. 4.20).

4.8.4 Tyrosine kinase-linked receptors

Some kinase receptors bind ligands and dimerize in a similar fashion to the ones described above, but do not have inherent catalytic activity in their *C*-terminal chain. However, once they have dimerized, they can bind and activate a tyrosine kinase enzyme from the cytoplasm. The **growth hormone** (GH) receptor is an example of this type of receptor and is classified as a tyrosine kinase-linked receptor (Fig. 4.21).

KEY POINTS

- Kinase-linked receptors are receptors which are directly linked to kinase enzymes. Messenger binding results in the opening of the kinase-active site, allowing a catalytic reaction to take place.
- Tyrosine kinase receptors have an extracellular binding site for a chemical messenger and an intracellular enzymatic

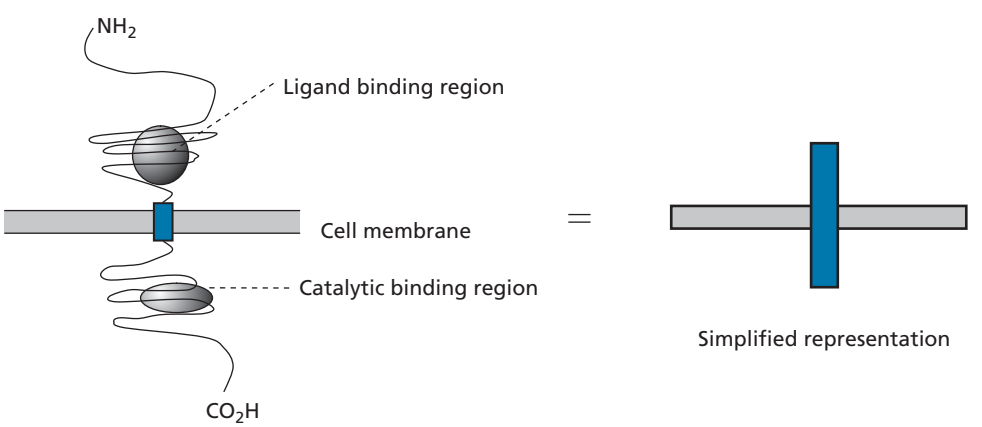


FIGURE 4.18 Structure of tyrosine kinase receptors.

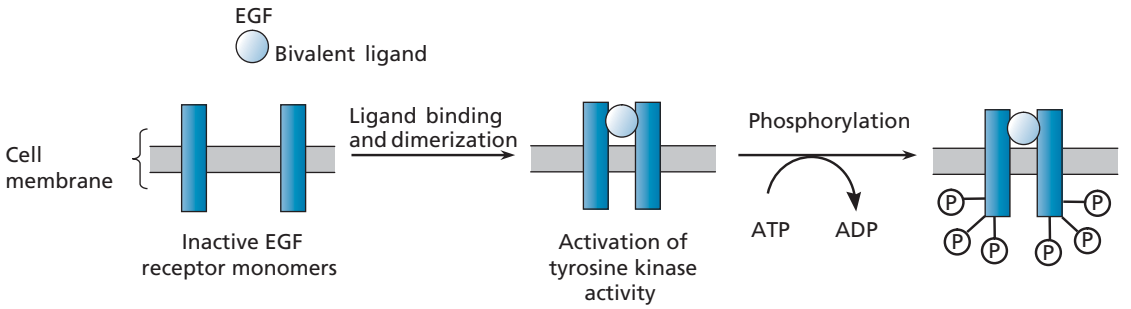


FIGURE 4.19 Activation mechanism for the epidermal growth factor (EGF) receptor.

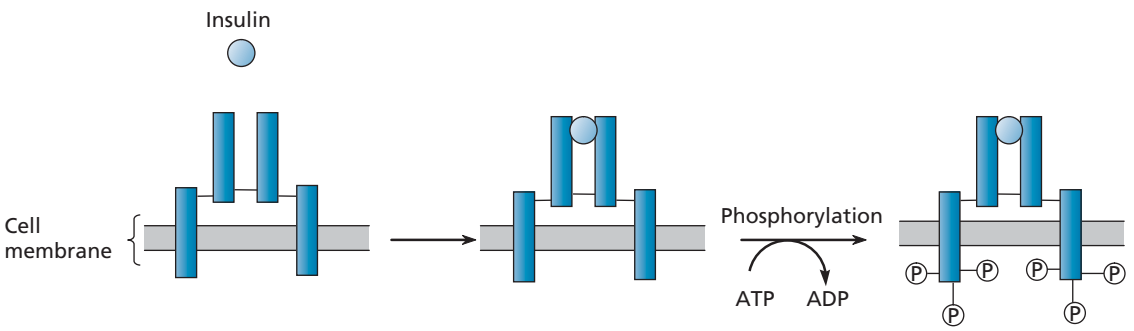


FIGURE 4.20 Ligand binding and activation of the insulin receptor.

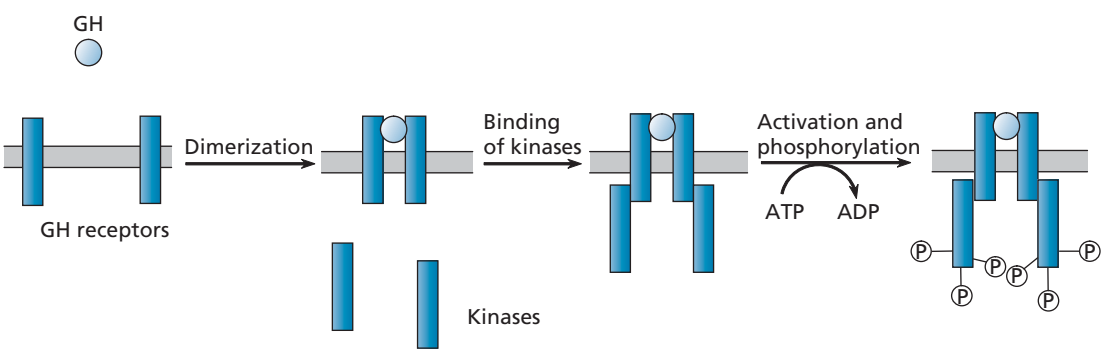


FIGURE 4.21 Activation of the growth hormone (GH) receptor.

active site which catalyses the phosphorylation of tyrosine residues in protein substrates.

- Ligand binding to the epidermal growth factor (EGF) receptor results in dimerization and opening of the active sites. The active site on one half of the dimer catalyses the phosphorylation of tyrosine residues present on the *C*-terminal chain of the other half.
- The insulin receptor is a preformed heterotetrameric structure which acts as a tyrosine kinase receptor.
- The growth hormone receptor dimerizes on binding its ligand, then binds and activates tyrosine kinase enzymes from the cytoplasm.

4.9 Intracellular receptors

Not all receptors are located in the cell membrane. Some receptors are within the cell and are defined as intracellular receptors. There are about 50 members of this group and they are particularly important in directly regulating gene transcription. As a result, they are often called **nuclear hormone receptors** or **nuclear transcription factors.** The chemical messengers for these receptors include steroid hormones, thyroid hormones, and retinoids. In all these cases, the messenger has to pass through the cell membrane in order to reach its receptor so it has to be hydrophobic in nature. The response time

resulting from the activation of the intracellular receptors is measured in hours or days, and is much slower than the response times of the membrane-bound receptors.

The intracellular receptors all have similar general structures. They consist of a single protein containing a ligand binding site at the *C*-terminus and a binding region for DNA near the centre (Fig. 4.22). The DNA binding region contains nine cysteine residues, eight of which are involved in binding two zinc ions. The zinc ions play a crucial role in stabilizing and determining the conformation of the DNA binding region. As a result, the stretches of protein concerned are called the **zinc finger domains.** The DNA binding region for each receptor can identify particular nucleotide sequences in DNA. For example, the zinc finger domains of the **estrogen receptor** recognize the sequence 5′-AGGTCA-3′, where A, G, C, and T are adenine, guanine, cytosine, and thymine.

The mechanism by which intracellular receptors work is also very similar (Fig. 4.23). Once the chemical messenger (ligand) has crossed the cell membrane, it seeks out its receptor and binds to it at the ligand binding site. An induced fit takes place which causes the receptor to change shape. This, in turn, leads to a dimerization of the ligand–receptor complex. The dimer then binds to a protein called a **co-activator** and, finally, the whole complex binds to a particular region of the cell's DNA. As there are two receptors in the complex and two DNA binding regions, the complex recognizes two identical sequences of nucleotides in the DNA separated by a short distance. For example, the estrogen ligand–receptor dimer binds to a nucleotide sequence of 5′-AGGTCANNNTGACCT-3′ where N can be any nucleic acid base. Depending on the complex involved, binding of the complex to DNA either triggers or inhibits the start of transcription, and affects the eventual synthesis of a protein.

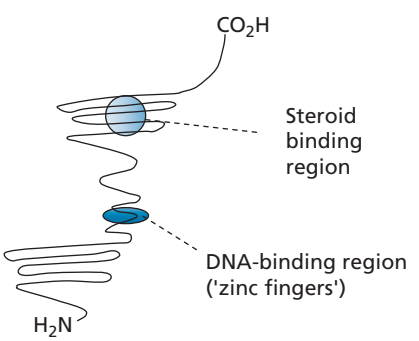


FIGURE 4.22 Structure of intracellular receptors.

4.10 Regulation of receptor activity

The role of allosteric binding sites in regulating the activity of enzymes was covered in section 3.6. Allosteric binding sites also play a role in regulating or modulating the activity of various receptors. These include ligand-gated ion channels, such as the nicotinic and the γ-aminobutyric acid receptors, and several G-protein-coupled receptors, such as the muscarinic, adenosine, and dopamine receptors. Structures that interact with these sites are called **allosteric modulators** and can either enhance or decrease the effect of the chemical messenger on the receptor (sections 8.2.7 and 8.3.2).

4.11 Genetic polymorphism and receptors

Genetic polymorphism was discussed in section 3.5.6 with respect to enzymes. Polymorphism is also responsible for receptors having subtle differences in structure and activity between individuals. In some cases, this can lead to diseases such as cancer (section 21.1.3).

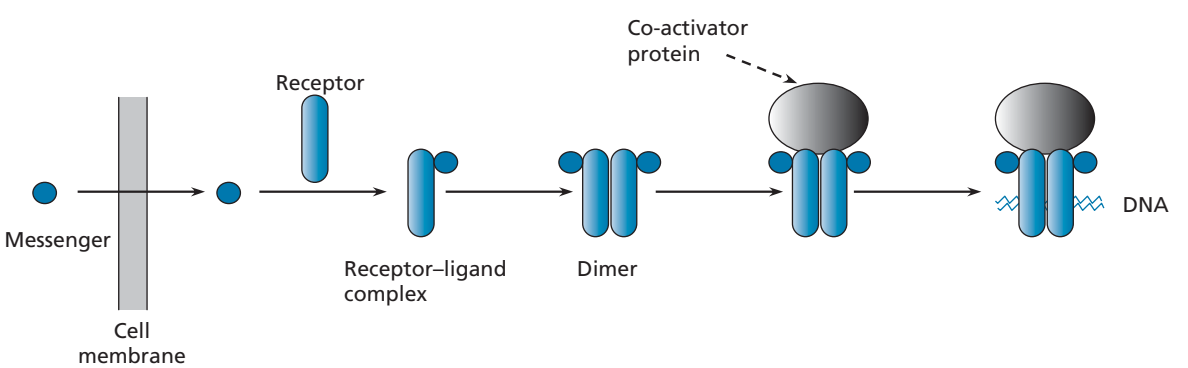


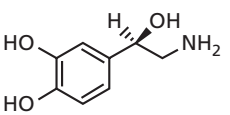
FIGURE 4.23 From messenger to control of gene transcription.

KEY NOTES

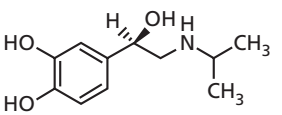
- Intracellular receptors are located within the cell and are important in controlling transcription.
- The chemical messengers for intracellular receptors must be sufficiently hydrophobic to pass through the cell membrane.
- The binding of a ligand with an intracellular receptor results in dimerization and the formation of a transcription factor complex which binds to a specific nucleotide sequence on DNA.

QUESTIONS

1. Explain the distinction between a binding site and a binding region.
2. Consider the structures of the neurotransmitters shown in Fig. 4.3 and suggest what type of binding interactions could be involved in binding them to a receptor binding site. Identify possible amino acids in the binding site which could take part in each of these binding interactions.
3. There are two main types of adrenergic receptor: the α and β–adrenoceptors. Noradrenaline shows slight selectivity for the α-receptor, whereas isoprenaline shows selectivity for the β-adrenoceptor. Adrenaline shows no selectivity and binds equally well to both the α- and β-adrenoceptors. Suggest an explanation for these differences in selectivity.



Noradrenaline



Isoprenaline

4. Suggest why the transmembrane regions of many membrane-bound proteins are α-helices.

FURTHER READING

Alexander, S. P. H., Mathie, A., and Peters, J. A. (2006) Guide to receptors and channels. *British Journal of Pharmacology* **147** (Suppl. 3), S1–S126.

Bikker, J. A., Trumpp-Kallmeyer, S., and Humblet, C. (1998) G-Protein coupled receptors: models, mutagenesis, and drug design. *Journal of Medicinal Chemistry* **41**, 2911–2927.

Chalmers, D. T. and Behan, D. P. (2002) The use of constitutively active GPCRs in drug discovery and functional genomics. *Nature Reviews Drug Discovery* **1**, 599–608.

Christopoulis, A. (2002) Allosteric binding sites on cell surface receptors: novel targets for drug discovery. *Nature Reviews Drug Discovery* **1**, 198–210.

Cohen, P. (2002) Protein kinases – the major drug targets of the twenty-first century? *Nature Reviews Drug Discovery* **1**, 309–315.

Kenakin, T. (2002) Efficacy at G-protein-coupled receptors. *Nature Reviews Drug Discovery* **1**, 103–110.

Kobilka, B. and Schertler, G. F. X. (2008) New G-protein-coupled receptor crystal structures: insights and limitations. *Trends in Pharmacological Sciences* **29**, 79–83.

Maehle, A-H., Prull, C-R., and Halliwell, R. F. (2002) The emergence of the drug receptor theory. *Nature Reviews Drug Discovery* **1**, 637–641.

Palczewski, K. (2010) Oligomeric forms of G protein-coupled receptors (GCPRs) *Trends in Biochemical Sciences* **35**, 595–600.

van Rijn, R. M., Whistler, J. L., and Waldhoer, M. (2010) Opioid-receptor-heteromer-specific trafficking and pharmacology. *Current Opinion in Pharmacology* **20**, 73–79.

Sansom, C. (2010) Receptive receptors. *Chemistry World* August, 52–55.

Tai, H. J., Grossmann, M., and Ji, I. (1998) G-Protein-coupled receptors. *Journal of Biological Chemistry* **273**, 17299–17302, 17979–17982.

Zhan-Guo, G. and Jacobson, K. A. (2006) Allosterism in membrane receptors. *Drug Discovery Today* **11**, 191–202.

5 Receptors and signal transduction

In Chapter 4, we discussed the structure and function of receptors. In this chapter, we consider what happens once a receptor has been activated. The interaction of a receptor with its chemical messenger is only the first step in a complex chain of events involving several secondary messengers, proteins, and enzymes that ultimately leads to a change in cell chemistry. These events are referred to as **signal transduction**. Unfortunately, a full and detailed account of these processes would fill a textbook in itself so the following account is focused mainly on the signal transduction processes that result from activation of G-protein-coupled receptors and kinase receptors. The signal transduction pathways following activation of G-protein-coupled receptors are of particular interest as 30% of all drugs on the market interact with these kinds of receptors. The transduction pathways for kinase receptors are also of great interest as they offer exciting new targets for novel drugs, particularly in the area of anticancer therapy (section 21.6.2). An understanding of the pathways and the various components involved helps to identify suitable drug targets.

5.1 Signal transduction pathways for G-protein-coupled receptors

G-protein-coupled receptors activate a signalling protein called a G-protein, which then initiates a signalling cascade involving a variety of enzymes. The sequence of events leading from the combination of receptor and ligand (the chemical messenger) to the final activation of a target enzyme is quite lengthy, so we shall look at each stage of the process in turn.

5.1.1 Interaction of the receptor–ligand complex with G-proteins

The first stage in the process is the binding of the chemical messenger or ligand to the receptor, followed by the binding of a G-protein to the receptor–ligand complex (Fig. 5.1). G-proteins are membrane-bound proteins situated at the inner surface of the cell membrane and are made up of three protein subunits (α, β, and γ). The α-subunit has a binding pocket which can bind guanyl nucleotides (hence the name G-protein) and which binds **guanosine diphosphate (GDP)** when the G-protein is in the resting state. There are several types of G-protein (e.g. Gs, Gi/Go, Gq/G_{11}) and several subtypes of these. Specific G-proteins are recognized by specific receptors. For example, G_s is recognized by the β-adrenoceptor, but not the α-adrenoceptor. However, in all cases, the G-protein acts as a molecular 'relay runner' carrying the message received by the receptor to the next target in the signalling pathway.

We shall now look at what happens in detail.

Firstly, the receptor binds its neurotransmitter or hormone (Fig. 5.1, frame 1). As a result, the receptor changes shape and exposes a new binding site on its inner surface (Fig. 5.1, frame 2). The newly exposed binding site now recognizes and binds a specific G-protein. Note that the cell membrane structure is a fluid structure and so it is possible for different proteins to 'float' through it. The binding process between the receptor and the G-protein causes the latter to change shape, which, in turn, changes the shape of the guanyl nucleotide binding site. This weakens the intermolecular bonding forces holding GDP and so GDP is released (Fig. 5.1, frame 3).

However, the binding pocket does not stay empty for long because it is now the right shape to bind **GTP** (**guanosine triphosphate**). Therefore, GTP replaces GDP (Fig. 5.1, frame 4).

Binding of GTP results in another conformational change in the G-protein (Fig. 5.1, frame 5), which weakens the links between the protein subunits such that the α-subunit (with its GTP attached) splits off from the β and γ-subunits (Fig. 5.1, frame 6). Both the α-subunit and the βγ-dimer then depart the receptor.

The receptor–ligand complex is able to activate several G-proteins in this way before the ligand departs and

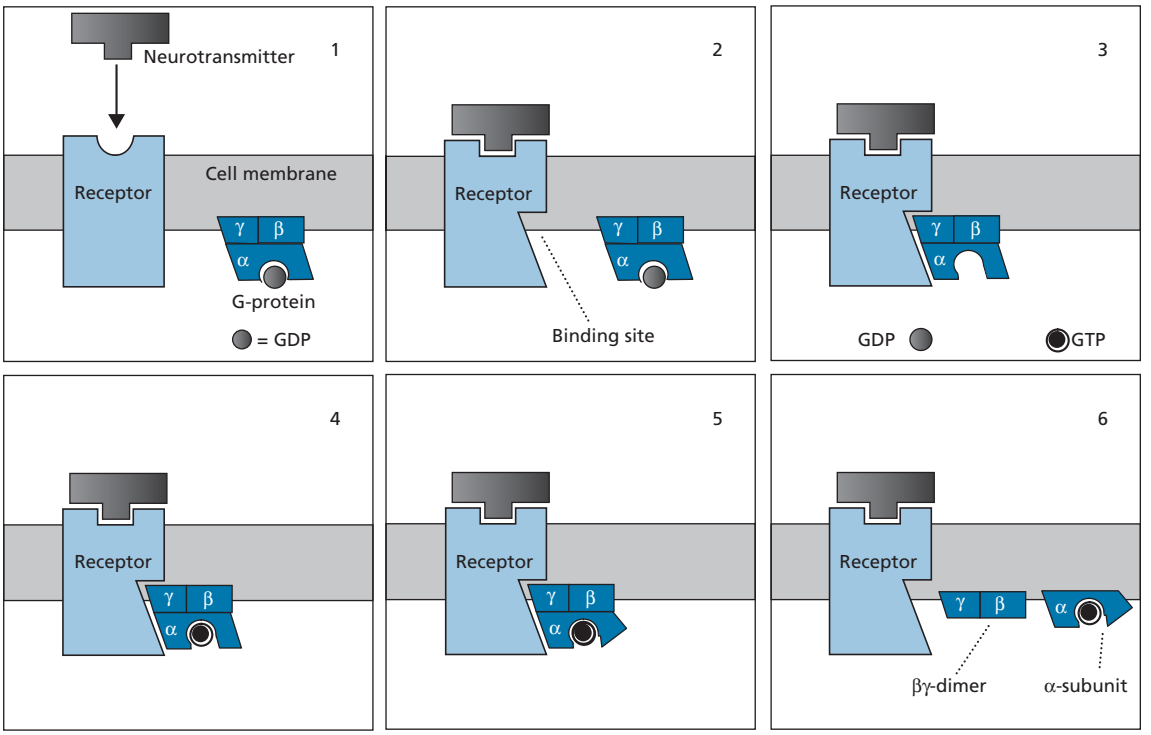


FIGURE 5.1 Activation of G-protein-coupled receptors and their interaction with G-proteins.

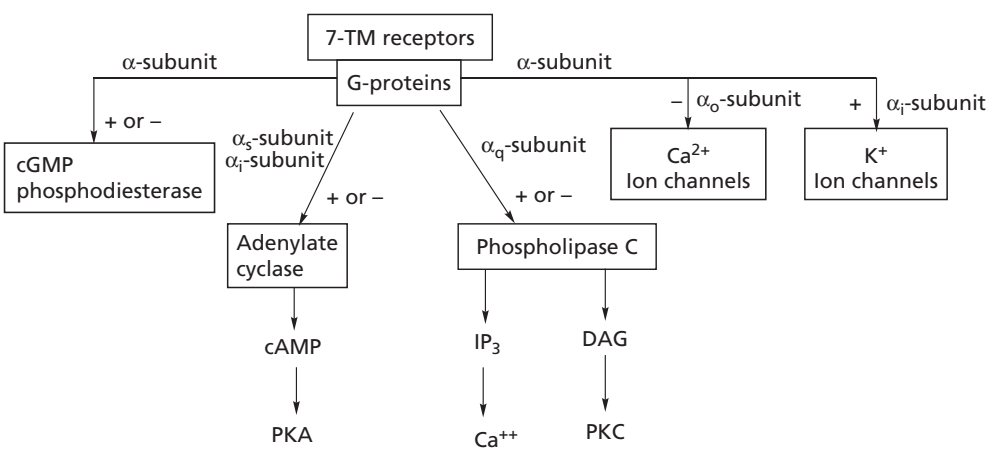


FIGURE 5.2 Signalling pathways arising from the splitting of different G-proteins.

switches off the receptor. This leads to an amplification of the signal.

Both the α-subunit and the βγ-dimer are now ready to enter the second stage of the signalling mechanism. We shall first consider what happens to the α-subunit.

5.1.2 Signal transduction pathways involving the α-subunit

The first stage of signal transduction (i.e. the splitting of a G-protein) is common to all of the 7-TM receptors. However, subsequent stages depend on what type of G-protein is involved and which specific α-subunit is formed (Fig. 5.2). Different α-subunits—there are at least 20 of them—have different targets and different effects:

- α_s stimulates adenylate cyclase;
- α_i inhibits adenylate cyclase and may also activate potassium ion channels;
- α_o activates receptors that inhibit neuronal calcium ion channels;
- α_q activates phospholipase C.

We do not have the space to study all these pathways in detail. Instead, we shall concentrate on two—the activation of **adenylate cyclase** and the activation of **phospholipase C**.

5.2 Signal transduction involving G-proteins and adenylate cyclase

5.2.1 Activation of adenylate cyclase by the α_s-subunit

The α_s-subunit binds to a membrane-bound enzyme called adenylate cyclase (or adenylyl cyclase) and 'switches' it on (Fig. 5.3). This enzyme now catalyses the synthesis of a molecule called cyclic AMP (cAMP) (Fig. 5.4). cAMP is an example of a **secondary messenger** which moves into the cell's cytoplasm and carries the signal from the cell membrane into the cell itself. The enzyme will continue to be active as long as the α_s-subunit is bound, resulting in the synthesis of several hundred cyclic AMP molecules—representing another substantial amplification of the signal. However, the α_s-subunit has intrinsic GTP-ase activity (i.e. it can catalyse the hydrolysis of its bound GTP to GDP) and so it deactivates itself after a certain time period and returns to the resting state. The α_s-subunit then departs the enzyme and recombines with the βγ-dimer to reform the G_s-protein while the enzyme returns to its inactive conformation.

5.2.2 Activation of protein kinase A

cAMP now proceeds to activate an enzyme called protein kinase A (PKA) (Fig. 5.5). PKA belongs to a group of enzymes called the **serine–threonine kinases** which catalyse the phosphorylation of serine and threonine residues in protein substrates (Fig. 5.6).

Protein kinase A catalyses the phosphorylation and activation of further enzymes with functions specific to the particular cell or organ in question, for example lipase enzymes in fat cells are activated to catalyse the breakdown of fat. The active site of a protein kinase has to be capable of binding the region of the protein substrate which is to be phosphorylated, as well as the cofactor ATP which provides the necessary phosphate group.

There may be several more enzymes involved in the signalling pathway between the activation of PKA and the activation (or deactivation) of the target enzyme. For example, the enzymes involved in the breakdown and synthesis of glycogen in a liver cell are regulated as shown in Fig. 5.7.

Adrenaline is the initial hormone involved in the regulation process and is released when the body requires immediate energy in the form of **glucose**. The hormone initiates a signal at the β-**adrenoceptor** leading to the

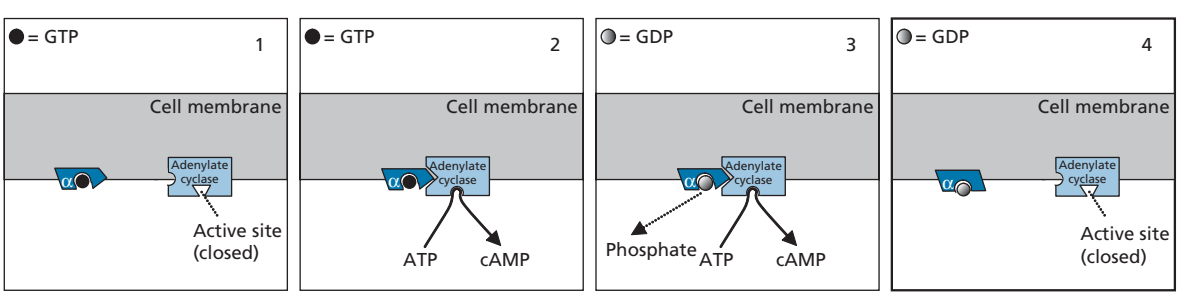


FIGURE 5.3 Interaction of α_s-subunit with adenylate cyclase and activation of the enzyme.

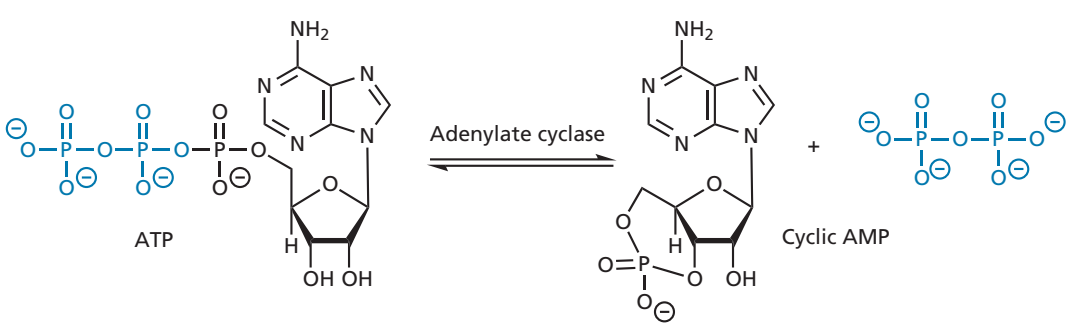


FIGURE 5.4 Synthesis of cyclic AMP.

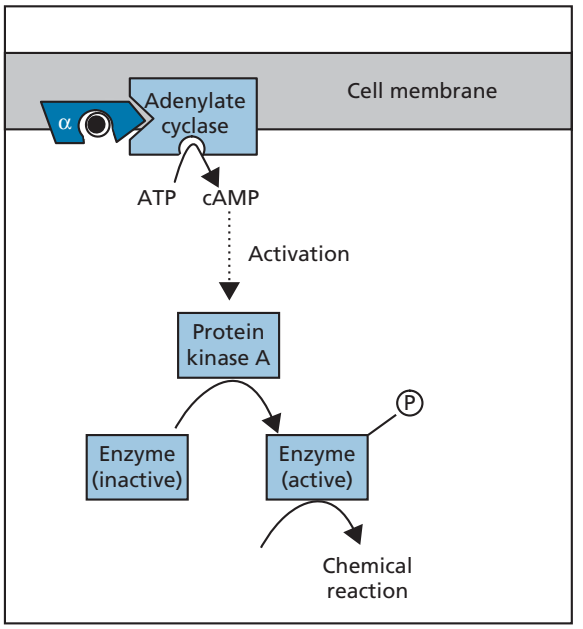


FIGURE 5.5 Activation of protein kinase A by cyclic AMP (Ⓟ = phosphate).

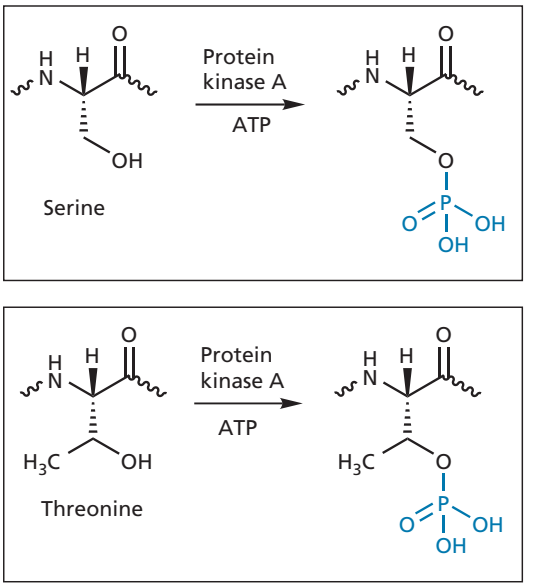


FIGURE 5.6 Phosphorylation of serine and threonine residues in protein substrates.

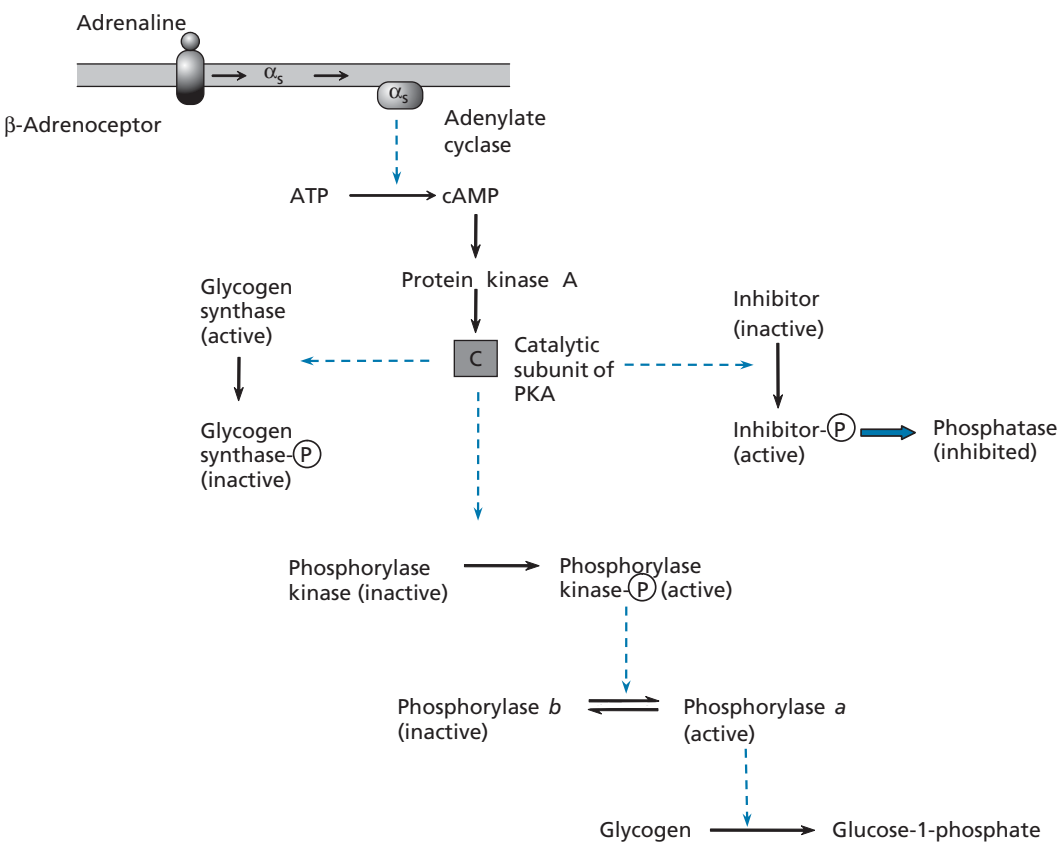


FIGURE 5.7 Regulation of glycogen synthesis and metabolism in a liver cell.

synthesis of cAMP and the activation of PKA by the mechanism already discussed. The catalytic subunit of PKA now phosphorylates three enzymes within the cell, as follows:

- an enzyme called **phosphorylase kinase** is phosphorylated and activated. This enzyme then catalyses the phosphorylation of an inactive enzyme called **phosphorylase *b*** which is converted to its active form, **phosphorylase *a***. Phosphorylase *a* now catalyses the breakdown of glycogen by splitting off glucose-1-phosphate units;
- **glycogen synthase** is phosphorylated to an inactive form, thus preventing the synthesis of glycogen;
- a molecule called **phosphorylase inhibitor** is phosphorylated. Once phosphorylated, it acts as an inhibitor for the **phosphatase** enzyme responsible for the conversion of phosphorylase *a* back to phosphorylase *b*. The lifetime of phosphorylase *a* is thereby prolonged.

The overall result of these different phosphorylations is a coordinated inhibition of glycogen synthesis and enhancement of glycogen metabolism to generate glucose in liver cells. Note that the effect of adrenaline on other types of cell may be quite different. For example, adrenaline activates β-adrenoceptors in fat cells leading to the activation of protein kinases, as before. This time, however, phosphorylation activates **lipase** enzymes which then catalyse the breakdown of fat to act as another source of glucose.

5.2.3 The G_i-protein

We have seen how the enzyme adenylate cyclase is activated by the α_s-subunit of the G_s-protein. Adenylate cyclase can also be inhibited by a different G-protein—the G_i-protein. The G_i-protein interacts with different receptors from those that interact with the G_s-protein, but the mechanism leading to inhibition is the same as that leading to activation. The only difference is that the α_i-subunit released binds to adenylate cyclase and inhibits the enzyme rather than activates it.

Receptors that bind G_i-proteins include the **muscarinic M_2 receptor** of cardiac muscle, **α_2-adrenoceptors** in smooth muscle, and **opioid receptors** in the central nervous system.

The existence of G_i- and G_s-proteins means that the generation of the secondary messenger cAMP is under the dual control of a brake and an accelerator, which explains the process by which two different neurotransmitters can have opposing effects at a target cell. A neurotransmitter which stimulates the production of cAMP forms a receptor–ligand complex which activates a G_s-protein, whereas a neurotransmitter which inhibits the production of cAMP forms a receptor–ligand complex which activates a G_i-protein. For example, **noradrenaline** interacts with the **β-adrenoceptor** to activate a G_s-protein, whereas **acetylcholine** interacts with the muscarinic receptor to activate a G_i-protein.

As there are various different types of receptor for a particular neurotransmitter, it is actually possible for that neurotransmitter to activate cAMP in one type of cell but inhibit it in another. For example, noradrenaline interacts with the β-adrenoceptor to activate adenylate cyclase because the β-adrenoceptor binds the G_s-protein. However, noradrenaline interacts with the α_2-adrenoceptor to inhibit adenylate cyclase because this receptor binds the G_i-protein. This example illustrates that it is the receptor that determines which G-protein is activated and not the neurotransmitter or hormone,

It is also worth pointing out that enzymes such as adenylate cyclase and the kinases are never fully active or inactive. At any one time, a certain proportion of these enzymes are active and the role of the G_s- and G_i-proteins is to either increase or decrease that proportion. In other words, the control is graded rather than all or nothing.

5.2.4 General points about the signalling cascade involving cyclic AMP

The signalling cascade involving the G_s-protein, cAMP and PKA appears very complex and you might wonder whether a simpler signalling process would be more efficient. There are several points worth noting about the process as it stands.

- Firstly, the action of the G-protein and the generation of a secondary messenger explains how a message delivered to the outside of the cell surface can be transmitted to enzymes within the cell—enzymes that have no direct association with the cell membrane or the receptor. Such a signalling process avoids the difficulties involved in a messenger molecule (which is commonly hydrophilic) having to cross a hydrophobic cell membrane.
- Secondly, the process involves a molecular 'relay runner' (the G-protein) and several different enzymes in the signalling cascade. At each of these stages, the action of one protein or enzyme results in the activation of a much larger number of enzymes. Therefore, the effect of one neurotransmitter interacting with one receptor molecule results in a final effect several factors larger than one might expect. For example, each molecule of **adrenaline** is thought to generate 100 molecules of cAMP and each cAMP molecule starts off an amplification effect of its own within the cell.
- Thirdly, there is an advantage in having the receptor, the G-protein, and adenylate cyclase as separate entities.

The G-protein can bind to several different types of receptor–ligand complexes. This means that different neurotransmitters and hormones interacting with different receptors can switch on the same G-protein leading to activation of adenylate cyclase. Therefore, there is an economy of organization involved in the cellular signalling chemistry, as the adenylate cyclase signalling pathway can be used in many different cells and yet respond to different signals. Moreover, different cellular effects will result depending on the type of cell involved (i.e. cells in different tissues will have different receptor types and subtypes, and the signalling system will switch on different target enzymes). For example, **glucagon** activates G_s-linked receptors in the liver leading to **gluconeogenesis**, **adrenaline** activates G_s-linked β_2-adrenoceptors in fat cells leading to **lipolysis**, and **vasopressin** interacts with G_s-linked vasopressin (V_2) receptors in the kidney to affect sodium/water resorption. Adrenaline acts on $G_{i/o}$-linked α_2-adrenoceptors leading to contraction of smooth muscle and **acetylcholine** acts on $G_{i/o}$-linked M_2 receptors leading to relaxation of heart muscle. All these effects are mediated by the cAMP signalling pathway.

- Finally, the dual control of 'brake/accelerator' provided by the G_s- and G_i-proteins allows fine control of adenylate cyclase activity.

5.2.5 The role of the βγ-dimer

If you've managed to follow the complexity of the G-protein signalling pathway so far, well done. Unfortunately, there's more! You may remember that when the G-protein binds to a receptor–ligand complex, it breaks up to form an α-subunit and a βγ-dimer. Until recently, the βγ-dimer was viewed merely as an **anchor** for the α-subunit to ensure that it remained bound to the inner surface of the cell membrane. However, it has now been found that the βγ-dimers from both the G_i- and the G_s-proteins can themselves activate or inhibit adenylate cyclase. There are actually six different types (or isozymes) of adenylate cyclase, and activation or inhibition depends on the isozyme involved. Moreover, adenylate cyclase is not the only enzyme that can be controlled by the βγ-dimer. The βγ-dimer is more promiscuous than the α-subunits and can affect several different targets, leading to a variety of different effects. This sounds like a recipe for anarchy. However, there is some advantage in the dimer having a signalling role, as it adds an extra subtlety to the signalling process. For example, it is found that higher concentrations of the dimer are required to result in any effect compared with the α-subunit. Therefore, regulation by the dimers becomes more important when a greater number of receptors are activated.

By now it should be clear that the activation of a cellular process is more complicated than the interaction of one type of neurotransmitter interacting with one type of receptor. In reality, the cell is receiving myriad signals from different chemical messengers via various receptors and receptor–ligand interactions. The final signal depends on the number and type of G-proteins activated at any one time, as well as the various signal transduction pathways that these proteins initiate.

5.2.6 Phosphorylation

As we have seen above, phosphorylation is a key reaction in the activation or deactivation of enzymes. Phosphorylation requires ATP as a source for the phosphate group and occurs on the phenolic group of tyrosine residues when catalysed by **tyrosine kinases,** and on the alcohol groups of serine and threonine residues when catalysed by **serine–threonine kinases**. These functional groups are all capable of participating in hydrogen bonding, but if a bulky phosphate group is added to the OH group, hydrogen bonding is disrupted. Furthermore, the phosphate group is usually ionized at physiological pH and so phosphorylation introduces two negatively charged oxygens. These charged groups can now form strong ionic bonds with a suitably positioned positively charged group in the protein causing the enzyme to change its tertiary structure. This change in shape results in the exposure or closure of the active site (Fig. 5.8).

Phosphorylation by kinase enzymes also accounts for the **desensitization** of G-protein-linked receptors. Phosphorylation of serine and threonine residues occurs on the intracellular *C*-terminal chain after prolonged ligand binding. As the *C*-terminal chain is involved in G-protein binding, phosphorylation changes the conformation of the protein in that region and prevents the G-protein from binding. Thus, the receptor–ligand complex is no longer able to activate the G-protein.

KEY POINTS

- G-proteins consist of three protein subunits, with the α-subunit bound to GDP. There are several types of G-protein.
- Receptor–ligand binding opens a binding site for the G-protein. On binding, GDP is exchanged for GTP, and the G-protein fragments into an α-subunit (bearing GTP) and a βγ-dimer.
- G-proteins are bound and split for as long as the chemical messenger is bound to the receptor, resulting in a signal amplification.
- An α_s-subunit binds to adenylate cyclase and activates it such that it catalyses the formation of cAMP from ATP. The reaction proceeds for as long as the α_s-subunit is bound, representing another signal amplification. An α_i-subunit inhibits adenylate cyclase.
- The α-subunits eventually hydrolyse bound GTP to GDP and depart adenylate cyclase. They combine with their respective βγ-dimers to reform the original G-proteins.

- cAMP acts as a secondary messenger within the cell and activates PKA. PKA catalyses the phosphorylation of serine and threonine residues in other enzymes leading to a biological effect determined by the type of cell involved.
- The signalling cascade initiated by receptor–ligand binding results in substantial signal amplification and does not require the original chemical messenger to enter the cell.
- The overall activity of adenylate cyclase is determined by the relevant proportions of G_s and G_i-proteins that are split, which, in turn, depends on the types of receptors that are being activated.
- The βγ-dimer of G-proteins has a moderating role on the activity of adenylate cyclase and other enzymes when it is present in relatively high concentration.
- Tyrosine kinases are enzymes which phosphorylate the phenol group of tyrosine residues in enzyme substrates. Serine–threonine kinases phosphorylate the alcohol groups of serine and threonine in enzyme substrates. In both cases, phosphorylation results in conformational changes that affect the activity of the substrate enzyme.
- Kinases are involved in the desensitization of receptors.

5.3 Signal transduction involving G-proteins and phospholipase C

5.3.1 G-protein effect on phospholipase C

Certain receptors bind G_s- or G_i-proteins and initiate a signalling pathway involving adenylate cyclase (section 5.2). Other 7-TM receptors bind a different G-protein called a **G_q-protein**, which initiates a different signalling pathway. This pathway involves the activation or deactivation of a membrane-bound enzyme called phospholipase C. The first part of the signalling mechanism is the interaction of the G-protein with a receptor–ligand complex as described previously in Fig. 5.1. This time, however, the G-protein is a G_q-protein rather than a G_s or G_i-protein, and so an α_q-subunit is released. Depending on the nature of the released α_q-subunit, phospholipase C is activated or deactivated. If activated, phospholipase C catalyses the hydrolysis of **phosphatidylinositol diphosphate** (PIP_2) (an integral part of the cell membrane structure) to generate the two secondary messengers **diacylglycerol** (DG) and **inositol triphosphate** (IP_3) (Figs 5.9 and 5.10).

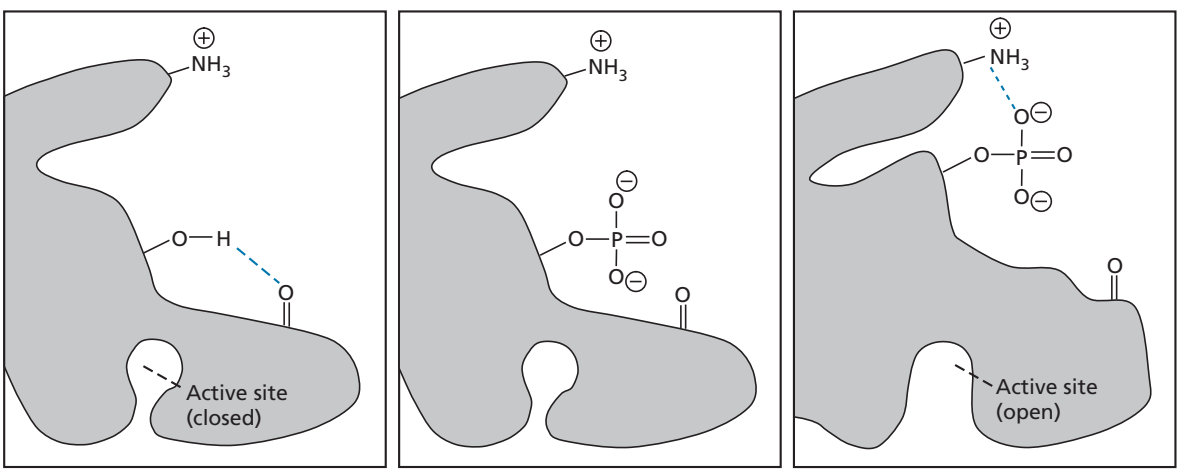


FIGURE 5.8 Conformational changes in a protein, induced by phosphorylation.

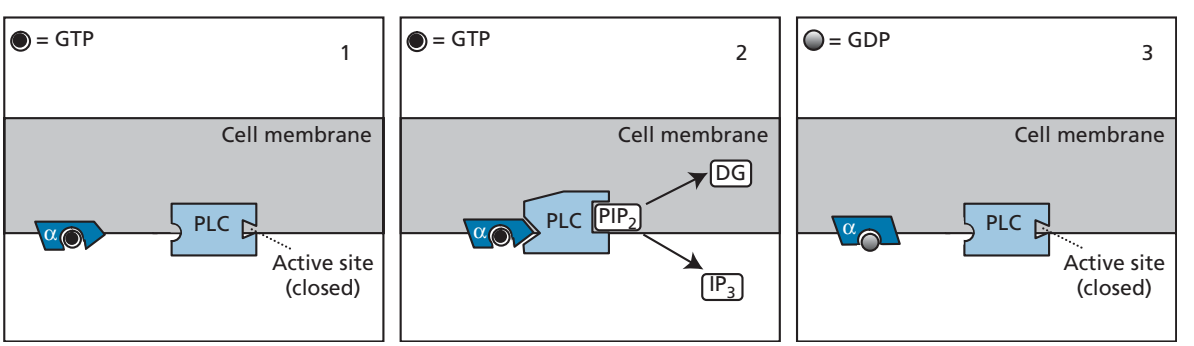


FIGURE 5.9 Activation of phospholipase C by an α_q-subunit.

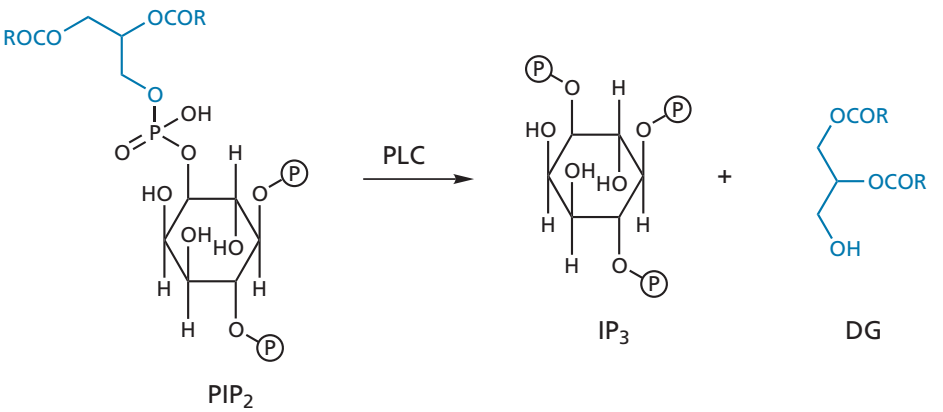


FIGURE 5.10 Hydrolysis of PIP_2 to inositol triphosphate (IP_3) and diacylglycerol (DG) (Ⓟ = phosphate).

5.3.2 Action of the secondary messenger: diacylglycerol

Diacylglycerol is a hydrophobic molecule and remains in the cell membrane once it is formed (Fig. 5.11). There, it activates an enzyme called **protein kinase C** (PKC) which moves from the cytoplasm to the cell membrane and then catalyses the phosphorylation of serine and threonine residues of enzymes within the cell. Once phosphorylated, these enzymes are activated and catalyse specific reactions within the cell. These induce effects such as tumour propagation, inflammatory responses, contraction or relaxation of smooth muscle, the increase or decrease of neurotransmitter release, the increase or decrease of neuronal excitability, and receptor desensitizations.

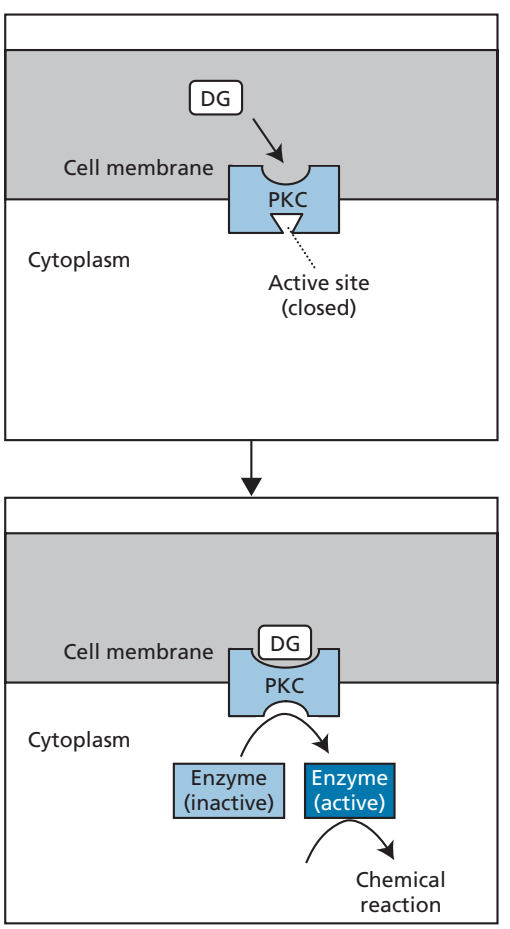


FIGURE 5.11 Activation of protein kinase C (PKC) by diacylglycerol (DG).

5.3.3 Action of the secondary messenger: inositol triphosphate

Inositol triphosphate is a hydrophilic molecule and moves into the cytoplasm (Fig. 5.12). This messenger works by mobilizing calcium ions from calcium stores in the endoplasmic reticulum. It does so by binding to a receptor and opening up a calcium ion channel. Once the ion channel is open, calcium ions flood the cell and activate calcium-dependent protein kinases which, in turn, phosphorylate and activate cell-specific enzymes. The released calcium ions also bind to a calcium binding protein called **calmodulin**, which then activates calmodulin-dependent protein kinases that phosphorylate and activate other cellular enzymes. Calcium has effects on contractile proteins and ion channels, but it is not possible to cover these effects in detail in this text. Suffice it to say that the release of calcium is crucial to a large variety of cellular functions including smooth muscle and cardiac muscle contraction, secretion from exocrine glands, transmitter release from nerves, and hormone release.

5.3.4 Re-synthesis of phosphatidylinositol diphosphate

Once IP_3 and DG have completed their tasks, they are recombined to form phosphatidylinositol diphosphate (PIP_2). Oddly enough, they cannot be linked directly and both molecules have to undergo several metabolic

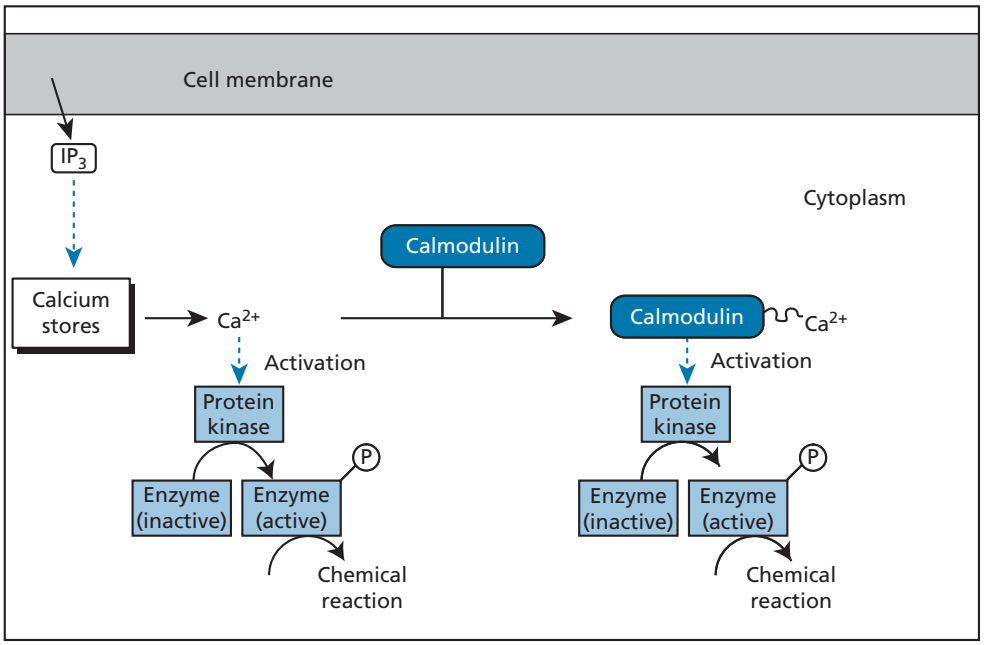


FIGURE 5.12 Signal transduction initiated by inositol triphosphate (IP_3). (Ⓟ = phosphate)

steps before re-synthesis can occur. For example, IP_3 is dephosphorylated in three steps to inositol which is then used as one of the building blocks for the re-synthesis of PIP_2 (Fig. 5.13). It is thought that **lithium salts** control the symptoms of manic depressive illness by interfering with this complex synthesis. They do so by inhibiting the monophosphatase enzyme responsible for the final dephosphorylation leading to inositol.

KEY POINTS

- G_q-proteins are split in a similar manner to G_s and G_i-proteins. The α_q-subunit affects the activity of phospholipase C which catalyses the hydrolysis of PIP_2 to form the secondary messengers IP_3 and DG.
- DG remains in the cell membrane and activates PKC, which is a serine–threonine kinase.
- IP_3 is a polar molecule which moves into the cytoplasm and mobilizes calcium ions. The latter activate protein kinases both directly and via the calcium binding protein calmodulin.
- IP_3 and DG are combined in a series of steps to reform PIP_2. Lithium salts are believed to interfere with this process.

5.4 Signal transduction involving kinase-linked receptors

5.4.1 Activation of signalling proteins and enzymes

We saw in section 4.8 that the binding of a chemical messenger to a kinase-linked receptor activates kinase activity such that a phosphorylation reaction takes place on the receptor itself. In the case of a tyrosine kinase, this involves the phosphorylation of tyrosine residues. We now continue that story.

Once phosphorylation has taken place, the phosphotyrosine groups and the regions around them act as binding sites for various signalling proteins or enzymes. Each phosphorylated tyrosine region can bind a specific signalling protein or enzyme. Some of these signalling proteins or enzymes become phosphorylated themselves once they are bound and act as further binding sites for yet more signalling proteins (Fig. 5.14).

Not all of the phosphotyrosine binding regions can be occupied by signalling proteins at one time so the type

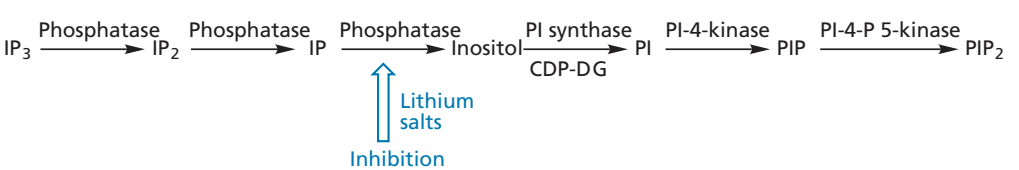


FIGURE 5.13 Re-synthesis of PIP_2 from IP_3 (CDP-DG = cytidine diphosphate-diacylglycerol).

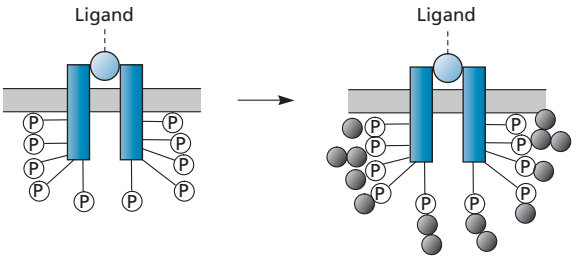


FIGURE 5.14 Binding of signalling proteins (indicated by dark circles) to activated kinase-linked receptors. (Ⓟ = phosphate)

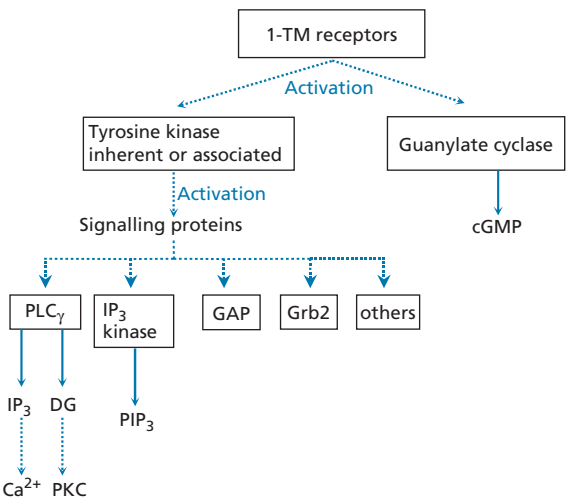


FIGURE 5.15 Signalling pathways from 1-TM receptors.

of signalling that results depends on which signalling proteins *do* manage to bind to the kinase receptors available. There is no room in an introductory text to consider what each and every signalling protein does, but most are the starting point for phosphorylation (kinase) cascades along the same principles as the cascades initiated by G-proteins (Fig. 5.15). Some growth factors activate a specific subtype of **phospholipase C** (PLCγ), which catalyses phospholipid breakdown leading to the generation of IP_3 and subsequent calcium release by the same mechanism as described in section 5.3.3. Other signalling proteins are chemical 'adaptors', which serve to transfer a signal from the receptor to a wide variety of other proteins, including many involved in cell division and differentiation. For example, the principal action of **growth factors** is to stimulate transcription of particular genes through a kinase signalling cascade (Fig. 5.16). A signalling protein called **Grb2** binds to a specific phosphorylated site of the receptor–ligand complex and becomes phosphorylated itself. A membrane protein called **Ras** (with a bound molecule of **GDP**) interacts with the receptor–ligand-signal protein complex and functions in a similar way to a G-protein (i.e. GDP is lost and GTP is gained). Ras is now activated and activates a serine–threonine kinase called **Raf**, initiating a serine–threonine kinase cascade which finishes with the activation of **mitogen-activated protein (MAP) kinase**. This phosphorylates and activates proteins called **transcription factors** which enter the nucleus and initiate gene expression resulting in various responses, including cell division. Many cancers can arise from malfunctions of this signalling cascade if the kinases involved become permanently activated, despite the absence of the initial receptor signal. Alternatively, some cancer cells over-express kinases and, as a result, the cell becomes super-sensitive to signals that stimulate growth and division. Consequently, inhibiting the kinase receptors or targeting the signalling pathway is proving to be an important method of designing new drugs for the treatment of cancer (section 21.6).

5.4.2 Small G-proteins

The Ras signal protein described in section 5.4.1 is an example of a class of signal proteins called the **small G-proteins**, so called because they are about two-thirds the size of the G-proteins described in sections 5.1–5.3. There are several subfamilies of small G-proteins (Ras, Rho, Arf, Rab, and Ran) and they can be viewed as being similar to the α-subunit of the larger G-proteins. Like the α-subunits, they are able to bind either GDP in the resting state or GTP in the activated state. Unlike their larger cousins, the small G-proteins are not activated by direct interaction with a receptor, but are activated downstream of receptor activation through intermediary proteins, which are classed as **guanine nucleotide exchange factors** (GEF). For example, activation of Ras (as shown in Fig. 5.16) requires the prior involvement of the protein Grb2 following receptor activation. Like the α-subunits, small G-proteins can autocatalyse the hydrolysis of bound GTP to give bound GDP, resulting in a return to the resting state. However, this process can be accelerated by helper proteins known as **GTPase activating proteins** (GAPs). This means that the level of activity of small G-proteins is under simultaneous brake and accelerator control involving GAP and GEF respectively.

The small G-proteins are responsible for stimulating cell growth and differentiation through different signal transduction pathways. Many cancers are associated with defects in small G-proteins, such as the Ras protein. *Ras* is the gene coding for the Ras protein and is one of the genes most commonly mutated in human tumours. There are three Ras proteins in mammalian cells: H-, K-, and N-Ras. Mutations which result in the inability of these proteins to autocatalyse the hydrolysis of bound GTP can occur. As a result, they remain permanently activated, leading, in turn, to permanent cell growth and division (see also section 21.6.1).

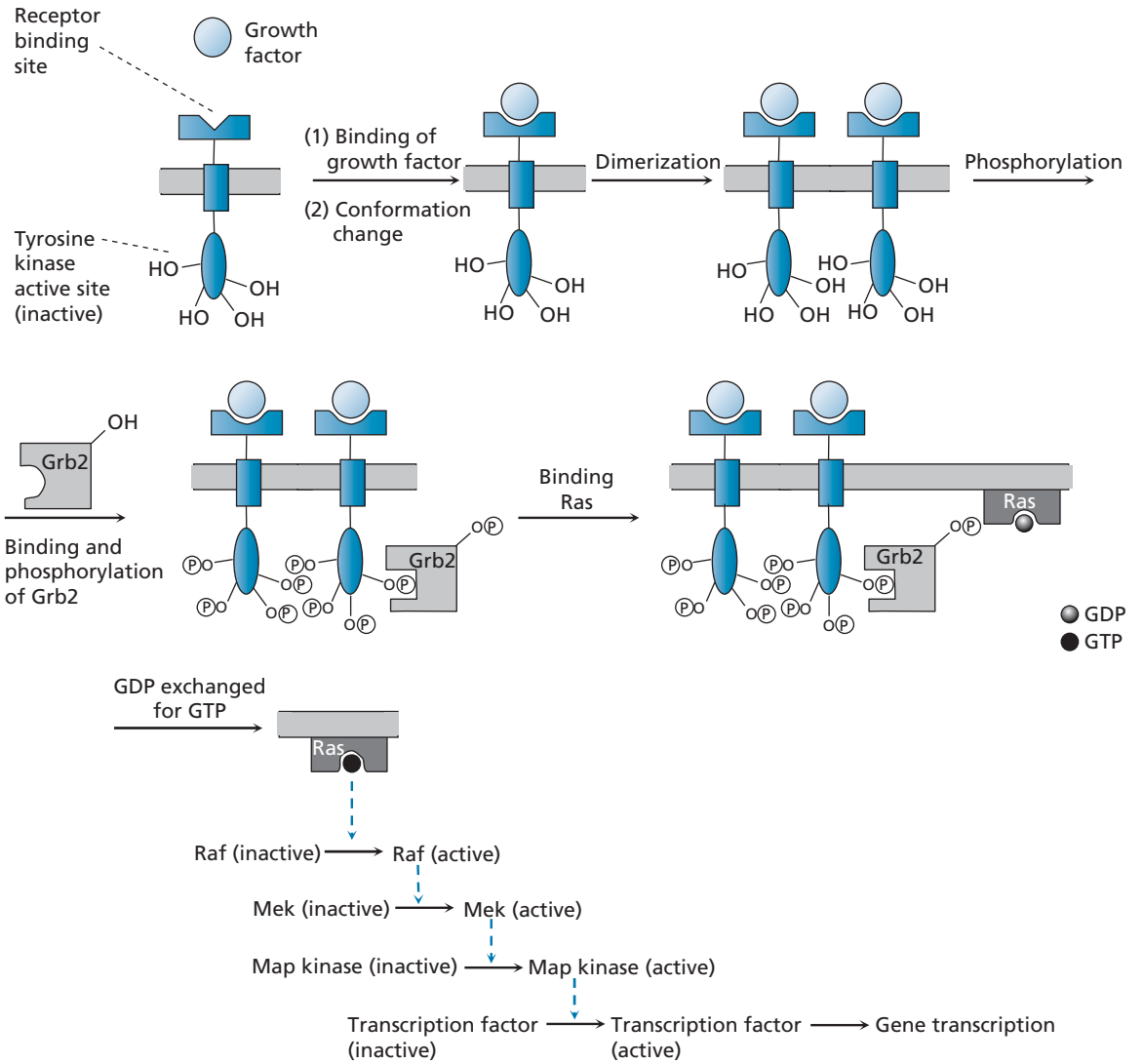


FIGURE 5.16 From growth factor to gene transcription. (Ⓟ = phosphate)

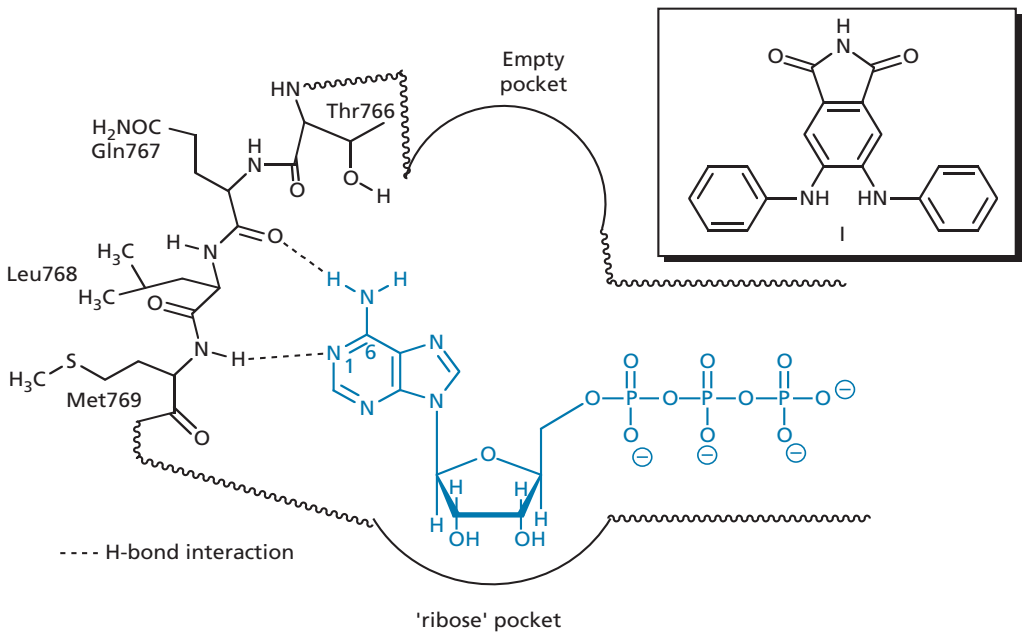
5.4.3 Activation of guanylate cyclase by kinase receptors

Some kinase receptors have the ability to catalyse the formation of **cyclic GMP** from **GTP**. Therefore, they are both receptor and enzyme (guanylate cyclase). The membrane-bound receptor/enzyme spans the cell membrane and has a single transmembrane segment. It has an extracellular receptor binding site and an intracellular guanylate cyclase active site. Its ligands are **α-atrial natriuretic peptide** and **brain natriuretic peptide**. Cyclic GMP appears to open sodium ion channels in the kidney, promoting the excretion of sodium.

KEY POINTS

- The phosphorylated tyrosine residues on activated kinase receptors act as binding sites for various signalling proteins and enzymes which are activated in turn.
- Small G-proteins are similar in nature to G-proteins, binding GDP in the resting state, and GTP in the activated state. They are single proteins activated by guanine nucleotide exchange factors.
- Some kinase receptors have an intracellular active site capable of catalysing the formation of cyclic GMP from GTP.

QUESTIONS



1. A model binding site for ATP was created for endothelial growth factor (EGF) receptor kinase, which demonstrates how ATP is bound (see above). Structure I is known to inhibit the binding of ATP. Suggest how structure I might bind.
2. Small G-proteins like Ras have an autocatalytic property. What does this mean and what consequences would there be (if any) should that property be lost?
3. Farnesyl transferase is an enzyme which catalyses the attachment of a long hydrophobic chain to the Ras protein. What do you think the purpose of this chain is and what would be the effect if the enzyme was inhibited?
4. Consider the signal transduction pathways shown in Fig. 5.16 and identify where signal amplification takes place.
5. The enzyme cAMP phosphodiesterase hydrolyses cAMP to AMP. What effect would an inhibitor of this enzyme have on glucose-1-phosphate production (Fig. 5.7)?
6. An enzyme was produced by genetic engineering where several of the serine residues were replaced by glutamate residues. The mutated enzyme was permanently active, whereas the natural enzyme was only active in the presence of a serine–threonine protein kinase. Give an explanation.
7. Suggest why tyrosine kinases phosphorylate tyrosine residues in protein substrates, but not serine or threonine residues.
8. Antibodies have been generated to recognize the extracellular regions of growth factor receptors. Binding of the antibody to the receptor should block the growth factor from reaching its binding site and block its signal. However, it has been observed that antibodies can sometimes trigger the same signal as the growth factor. Why should this occur? Consult section 10.7.2 to see the structure of an antibody.

FURTHER READING

Alexander, S., Mead, A., and Peters, J. (eds) (1998) TiPS receptor and ion channel nomenclature. *Trends in Pharmacological Sciences* **19**(Suppl. 1), 1–98.

Bikker, J. A., Trumpp-Kallmeyer, S., and Humblet, C. (1998) G-Protein coupled receptors: models, mutagenesis, and drug design. *Journal of Medicinal Chemistry* **41**, 2911–2927.

Cohen, P. (2002) Protein kinases—the major drug targets of the twenty-first century? *Nature Reviews Drug Discovery* **1**, 309–315.

Flower, D. (2000) Throwing light on GPCRs. *Chemistry in Britain* November, 25.

Foreman, J. C. and Johansen, T. (eds) (1996) *Textbook of Receptor Pharmacology*. CRC Press, Boca Raton, FL.

George, S. R., O'Dowd, B. F., and Lee, S. P. (2002) G-Protein-coupled receptor oligomerization and its potential for drug discovery. *Nature Reviews Drug Discovery* **1**, 808–820.

Kenakin, T. (2002) Efficacy at G-protein-coupled receptors. *Nature Reviews Drug Discovery*, **1**, 103–110.

Neubig, R. R. and Siderovski, D. P. (2002) Regulators of G-protein signalling as new central nervous system drug targets. *Nature Reviews Drug Discovery* **1**, 187–197.

Schwarz, M. K. and Wells, T. N. C. (2002) New therapeutics that modulate chemokine networks. *Nature Reviews Drug Discovery* **1**, 347–358.

Tai, H. J., Grossmann, M., and Ji, I. (1998) G-Protein-coupled receptors. *Journal of Biological Chemistry* **273**, 17299–17302, 17979–17982.

Takai, Y., Sasaki, T., and Matozaki, T. (2001) Small GTP-binding proteins. *Physiological Reviews* **81**, 153–208.

Vlahos, C. J., McDowell, S. A., and Clerk, A. (2003) Kinases as therapeutic targets for heart failure. *Nature Reviews Drug Discovery* **2**, 99–113.

Titles for general further reading are listed on p.763.