2 Proteins: structure and function

The vast majority of drugs used in medicine are targeted on 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.

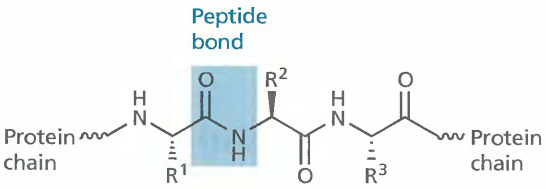


FIGURE 2.1 Primary structure of proteins (R^1, R^2 and R^3 = amino acid residues).

2.1 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 three-letter 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.

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

TABLE 2.1 The 20 common amino acids found in humans

Synthesized in the human body			Essential to the diet		
Amino acid	Codes		Amino acid	Codes	
	3-letter	1-letter		3-letter	1-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			

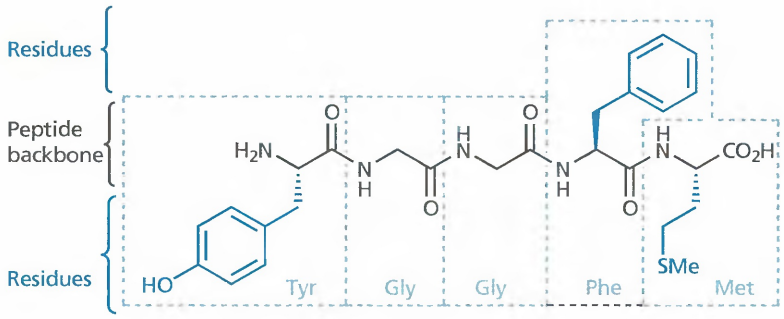


FIGURE 2.2 Met-enkephalin. The short-hand notation for this peptide is H-Tyr-Gly-Gly-Phe-Met-OH or YGGFM.

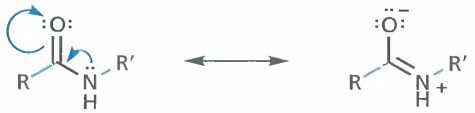
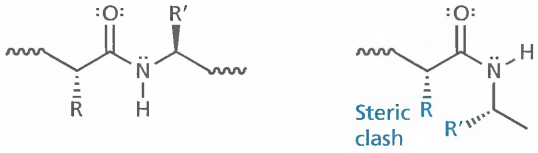


FIGURE 2.3 The planar peptide bond (bond rotation allowed for coloured bonds only).



trans conformation (favoured)

cis conformation (unfavoured)

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

which prevents rotation. As a result, bond rotation in the protein backbone is only possible for the bonds on 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, because 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 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 regions of secondary structure in most other proteins as well. 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 residues 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.

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 residues 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 that are less regular in their structure, but are often rigid and well defined.

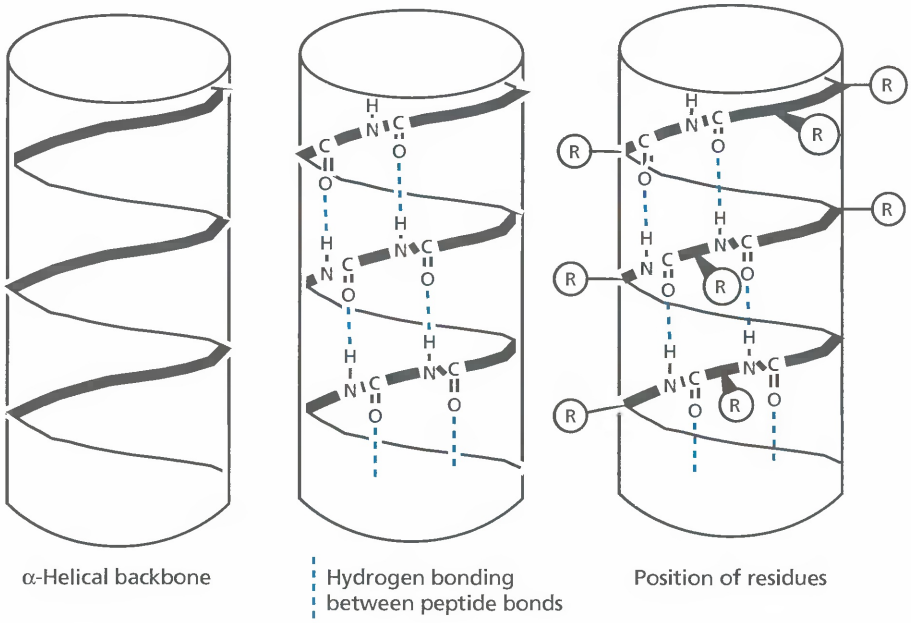


FIGURE 2.5 The α-helix for proteins showing intramolecular hydrogen bonds and the position of residues.

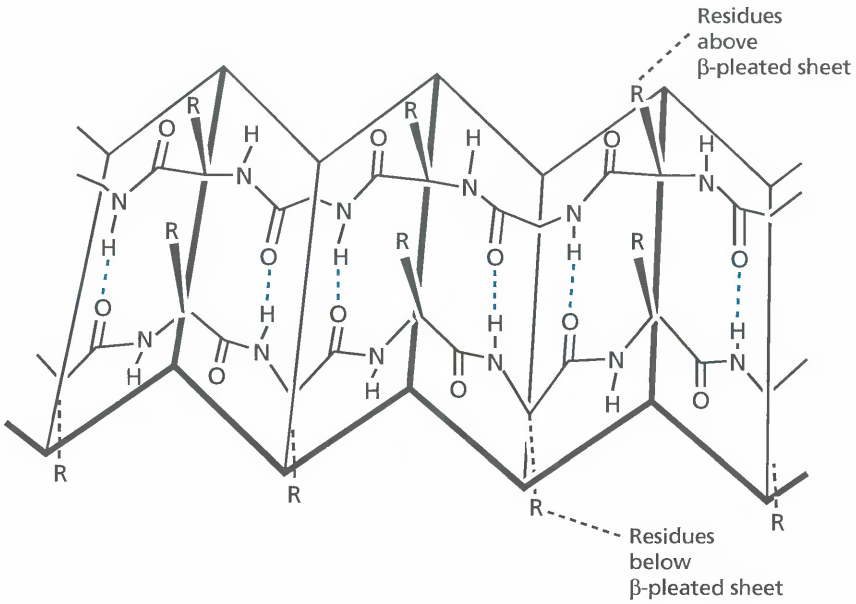


FIGURE 2.6 The β-pleated sheet (antiparallel arrangement).

2.3 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 & 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, so 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

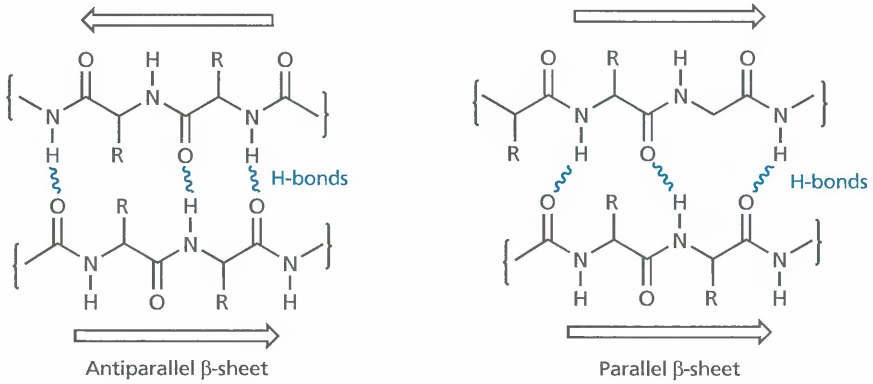


FIGURE 2.7 Hydrogen bonding in antiparallel and parallel β-sheets (the arrows are pointing to the *C*-terminal end of the chain).

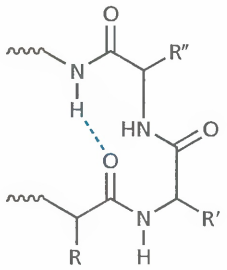


FIGURE 2.8 The β-turn showing hydrogen bonding between the first and third peptide bond.

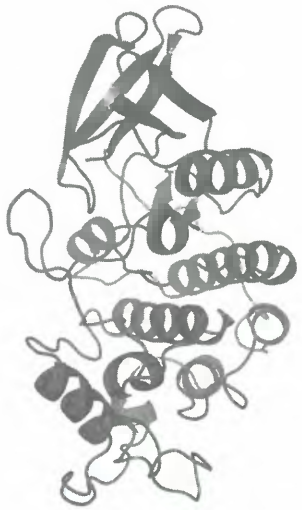


FIGURE 2.9 The pdb file (1hcl) for human cyclin-dependent kinase 2 (CDK2) where cylinders represent α-helices and arrows represent β-sheets. A pdb file contains the three-dimensional structural information for a protein, and can be downloaded from the Brookhaven protein data bank. Each protein structure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

(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 the 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.[1] 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?

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 the peptide links, but also the residues 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).

With the exception of disulfide bonds, the bonding interactions involved in tertiary structure are the same as the **intermolecular bonds** described in section 1.3. The

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

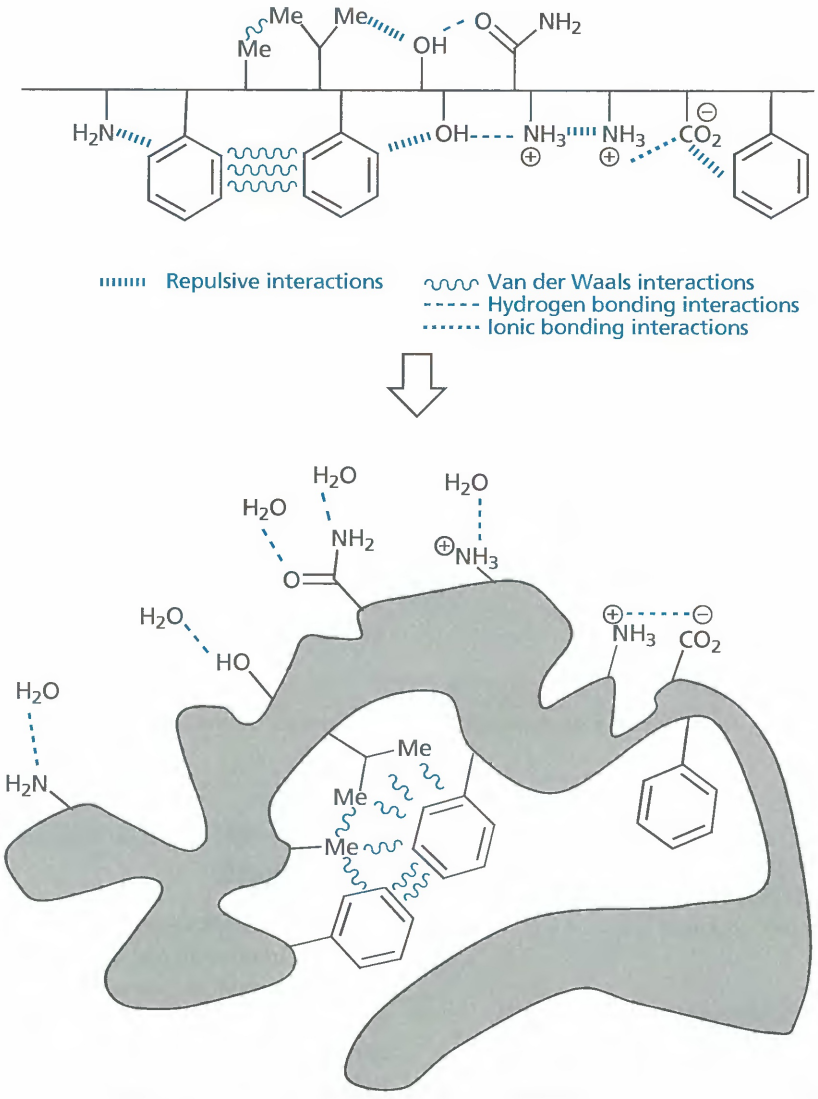


FIGURE 2.10 Tertiary structure formation as a result of intramolecular interactions.

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: disulfide links

Cysteine has a residue containing a thiol group capable of forming a covalent bond in protein tertiary structure. When two such residues are close together, a covalent disulfide 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 ammonium 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 since they involve interactions between atoms having partial charges. They can be formed between

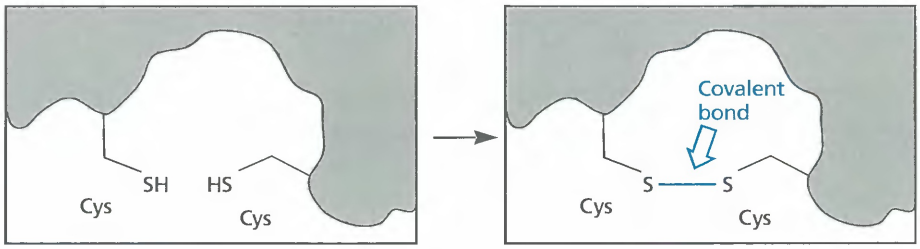


FIGURE 2.11 The formation of a disulfide covalent bond between two cysteine residues.

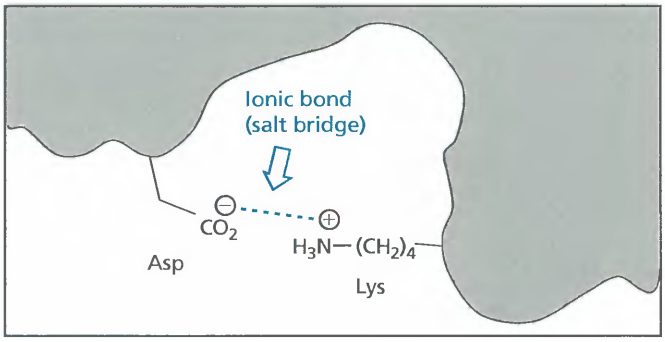


FIGURE 2.12 Ionic bonding between an aspartate residue and a lysine residue.

a large number of amino acid residues such as serine, threonine, aspartic acid, glutamic acid, glutamine, lysine, arginine, histidine, tryptophan, tyrosine, and asparagine. Two examples are shown in Fig. 2.13.

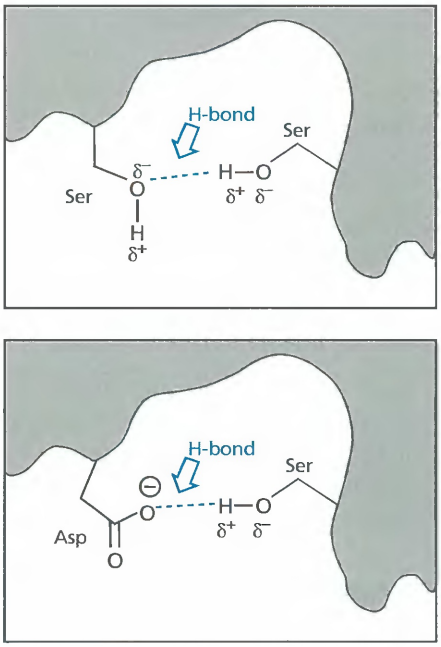


FIGURE 2.13 Hydrogen bonding between amino acid residues.

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 residues capable of interacting with each other by van der Waals interactions. The residues of other amino acids such as methionine, tryptophan, threonine, and tyrosine contain polar functional groups, but the residues also have a substantial

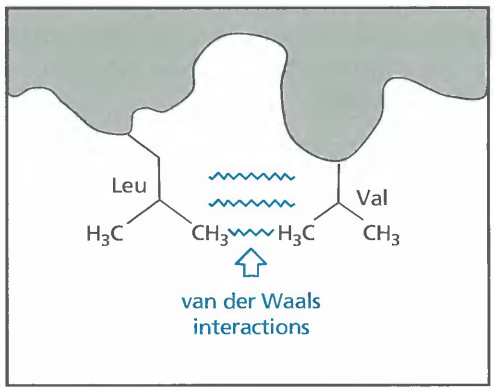


FIGURE 2.14 van der Waals interactions between amino acid residues.

hydrophobic character and so van der Waals interactions are possible for these amino acids as well. 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 described above to follow the same order as their strengths: covalent, ionic, hydrogen bonding, and finally van der Waals. In fact, the opposite is usually true. Usually the most important bonding interactions in tertiary structure are those due to van der Waals interactions and hydrogen bonding, and the least important interactions are those due to covalent and ionic bonding.

There are two reasons for this. First, 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 types of amino acids in any typical globular protein to see why. The only covalent bond that can contribute to tertiary structure is a disulfide bond. Only cysteine can form such a bond, whereas there are several 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 disulfide bridges, where the relative importance of the covalent link to tertiary structure is more significant. Disulfide links are also more significant in small polypeptides such as the peptide hormones **vasopressin** and **oxytocin** (Fig. 2.15). Nevertheless, in most proteins, disulfide 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 (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 most of the hydrophobic groups are in the centre so that they avoid water and interact with each other. Since the hydrophilic amino acids form hydrogen bonds with water, the number of ionic and hydrogen bonds contributing to the tertiary structure is reduced and this leaves 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 as far as the action of enzymes is concerned and helps to explain why reactions that should be impossible in an aqueous environment can take place in the presence of enzymes. The enzyme can provide a non-aqueous environment for the reaction to take place (chapter 3).

There are also important consequences for drug design. Drugs interact with proteins in binding sites which are normally hollows or canyons on the protein surface. These sites are also more hydrophobic in character than the surface, and so van der Waals and hydrophobic interactions play an important role in the binding of a drug to its target, and consequently to its activity.

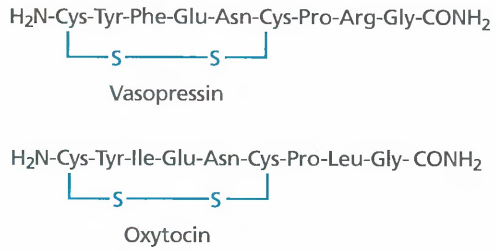


FIGURE 2.15 Vasopressin and oxytocin.

2.3.6 Role of the planar peptide bond

Planar peptide bonds indirectly play an important role in tertiary structure. Since bond rotation is hindered in peptide bonds with the *trans* conformation generally favoured, 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 this restriction do not fold into a specific conformation, because the entropy change required to go from a highly disordered structure to an ordered structure is highly unfavourable. Peptide bonds can also form hydrogen bonds with amino acid residues, and play a role in determining tertiary structure.

2.4 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 confused with the alpha and beta terminology

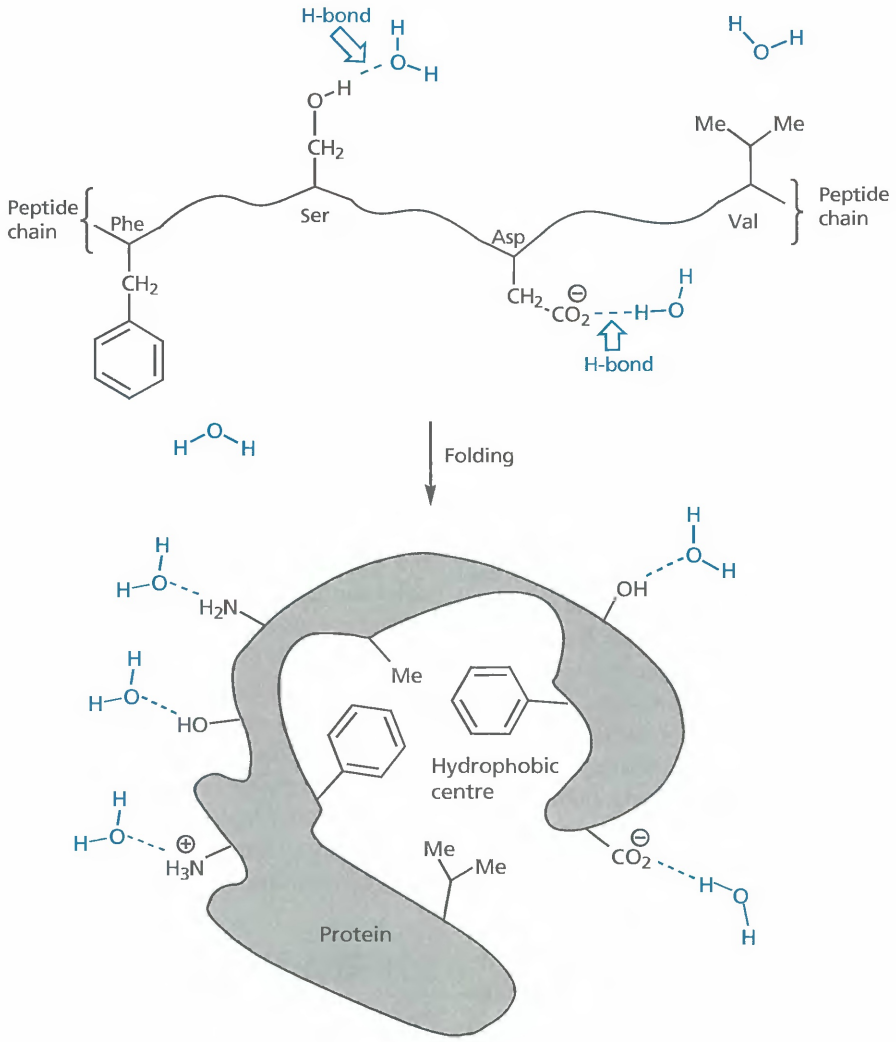


FIGURE 2.16 Bonding interactions with water.

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

Since 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).

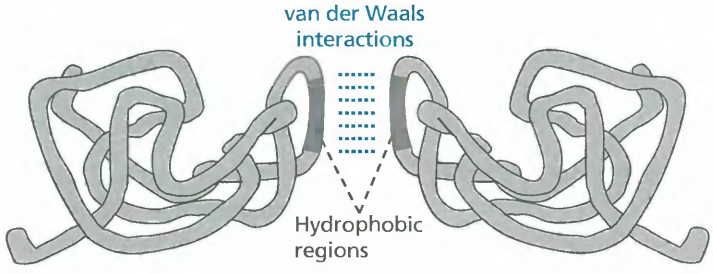


FIGURE 2.17 Quaternary structure involving two protein subunits.

2.5 Translation and post-translational modifications

The process by which a protein is synthesized in the cell is called **translation** (section 6.2). 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 21.7.4).

The fibres of **collagen** are stabilized by the hydroxylation of proline residues, and 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

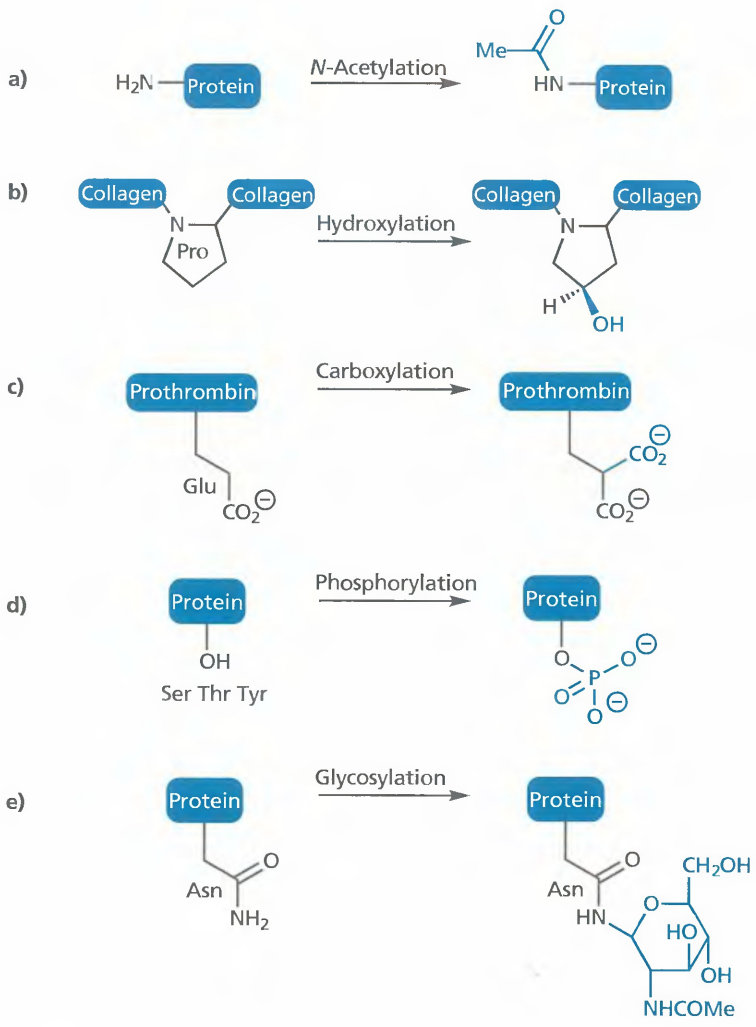


FIGURE 2.18 Examples of post-translational modifications carried out on proteins.

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 accorded rightly to the Human Genome Project, which has now been completed. The science behind this work is called **genomics** and involves the identification of the genetic code not only in humans but in other species as well. 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 this 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 (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 structure 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 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 repolymerized 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. Individual chapters are devoted to the structure and function of these proteins—chapters 3 and 4 respectively.

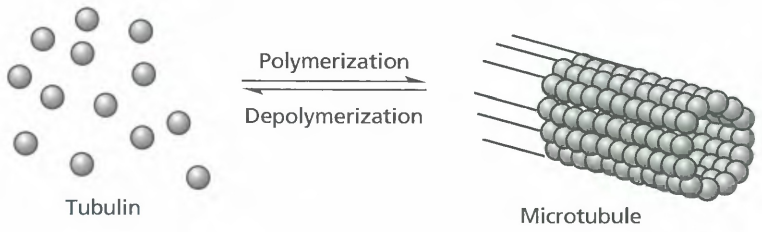


FIGURE 2.19 Polymerization of tubulin.

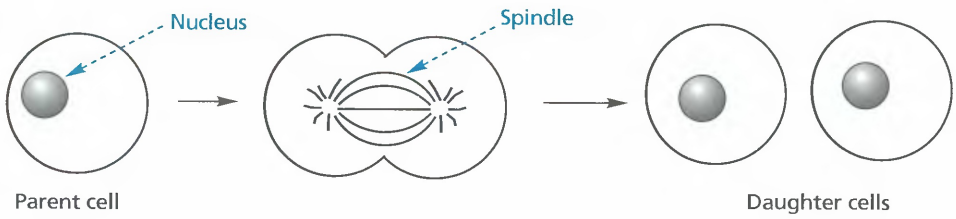


FIGURE 2.20 Cell division.

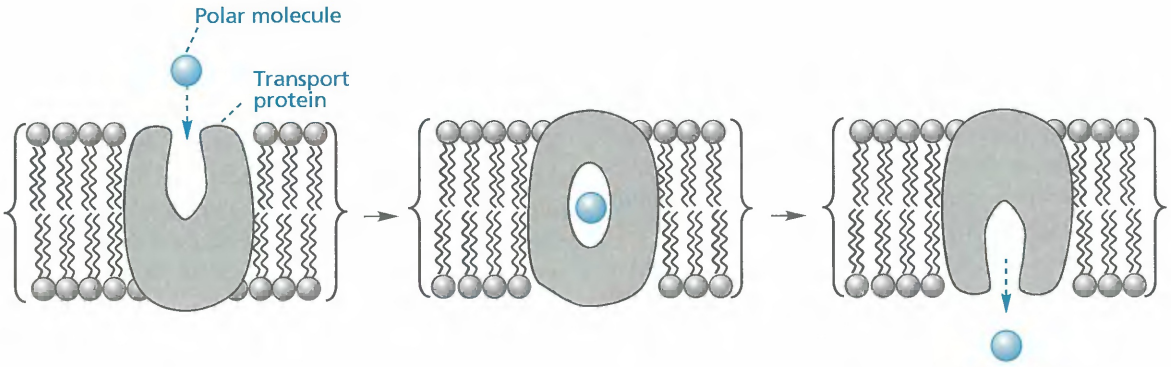


FIGURE 2.21 Transport proteins.

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, and this 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).

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 the 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 to other naturally occurring amino acids?
3. Identify the intermolecular/intramolecular interactions that are possible for the residues of the following amino acids; serine, phenylalanine, glycine, lysine, aspartic acid and aspartate.
4. The chains of several cell membrane-bound proteins are known to 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 such a protein help in distinguishing the portions of the protein embedded within the cell membrane from those positioned intracellularly or extracellularly?
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 should have such an effect on 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

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*, April, 7–33 (proteomics).

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*, November, 21–23.

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