1 – THE CHEMISTRY OF LIFE

1.1 – Introduction

NASA scientists have possibly found fossilized traces of early life on Mars. Oceans of ice have been discovered on Europa, one of the moons of Jupiter, suggesting that primitive life forms could have evolved there. The SETI programme, which attempts to make contact with life elsewhere in the universe, has beamed out the structure of DNA to inform other intelligent life forms of the genetic basis of life on Earth. These events make headline news, demonstrating our fascination with the origins of life. Recent generations have undergone a psychological frameshift; we have seen the Earth – 'the blue marble' – from outside the planet.

Figure 1.1: The Earth as seen from space

The predominance of the water environment is dramatically evident from space. Life began in the oceans and the chemistry of life is that chemistry which takes place in water under mild conditions.

Water – Life's Matrix

What makes water so vital for life? Animals and plants living in 'normal' environments cannot survive without a regular supply of water. Without water life could never have evolved. Some of the most significant biological roles of water are summarized below.

© University of Cambridge International Examinations 2006

www.theallpapers.com

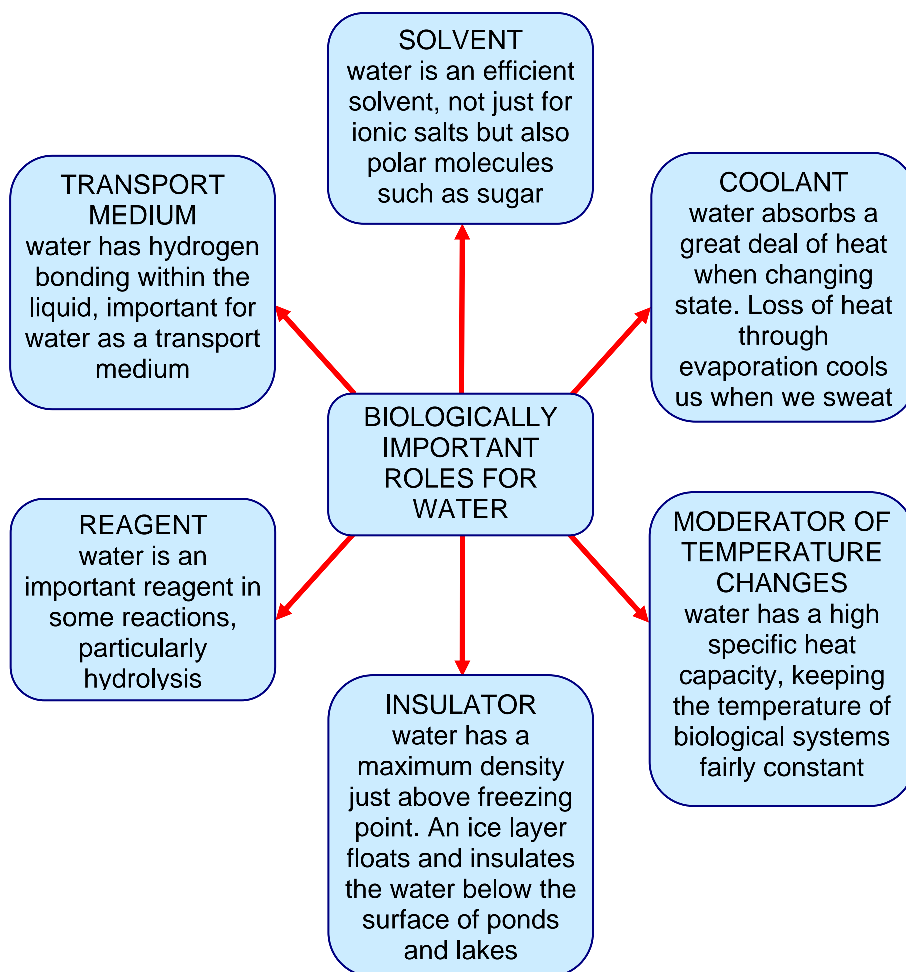


Figure 1.2 – biologically important roles of water

Possibly the most remarkable property of water is that it is a liquid at the normal temperatures found on Earth. Because of its small molecular size, water should be a gas – just like the similarly sized methane, CH_4, hydrogen sulphide, H_2S, and ammonia, NH_3, molecules. This property of water stems from the highly polar nature of the molecule. The electrons present in the covalent bonds between oxygen and hydrogen are shared unequally - oxygen being a highly electronegative atom. In each water molecule, the oxygen atom draws the bonding electrons towards itself. Thus the oxygen atom gains a partial negative charge, while the hydrogen atoms are left with a partial positive charge (Figure 1.3).

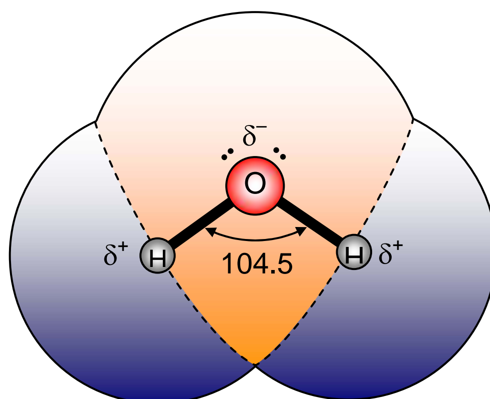


Figure 1.3 – the polarized nature of the water molecule

© University of Cambridge International Examinations 2006

www.theallpapers.com

As a result of these partial shifts of charge water is a polarized molecule. The partial positive and negative charges provide a force that attracts water molecules together through ‘hydrogen bonds’. A hydrogen atom on one molecule is attracted to the oxygen atom of another (Figure 1.4). Hydrogen bonds are much weaker than normal covalent bonds but these interactions provide an additional force between water molecules that are not present in methane, for instance. This results in water having a higher than expected boiling point.

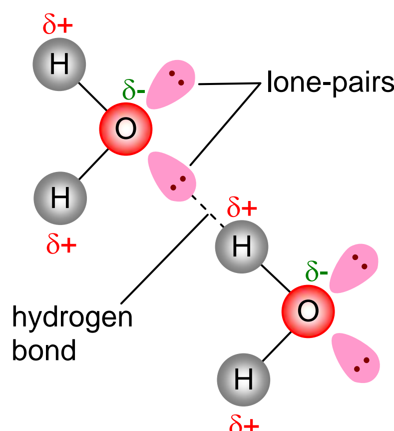


Figure 1.4 – hydrogen bonding between water molecules

The importance of hydrogen bonding to life

The polar nature of water means that ionic compounds can dissolve in water. But hydrogen bonding also means that polar covalent molecules, sugars and amino acids, for example, are also soluble in water.

Three properties of hydrogen bonds make them important for life, and these are:

- their transience – they are made and broken relatively easily,
- they have direction – the atoms involved become aligned,
- they have specificity – only certain groups can participate.

Water itself illustrates the first two of these properties. Liquid water consists of a network of hydrogen-bonded molecules, but an individual hydrogen bond lasts for no more than a trillionth (10^{-12}) of a second. Water molecules are constantly jostling with each other, moving past each other, breaking and re-making hydrogen bonds with different molecules. The relative weakness of the hydrogen bond is important in DNA replication, for instance, as the DNA molecule must ‘un-zip’ for the strands to be copied.

The expanded structure of ice illustrates that hydrogen bonds have direction. In ice, hydrogen bonds contribute to the ‘diamond-like’ tetrahedral arrangement of the atoms in the lattice. The three atoms involved in a hydrogen bond lie in a straight line. DNA again illustrates how important this alignment associated with hydrogen-bonding is. Hydrogen-bonding between the bases lies at the core of the double helical structure.

The specificity of the interaction between the bases involves hydrogen bonding. Hydrogen bonds can only form between certain groups, and this plays a part in the complementary base-pairing which is essential to the function of DNA.

Condensation polymerisation

Given the importance of a water environment to the emergence of life on Earth it is not surprising that condensation polymerisation is ‘the method of choice’ in the natural world when it comes to making the macromolecules important to life. It is within an aqueous environment that small biological precursor molecules can associate and polymerise through condensation reactions (Figure 1.5). The reverse process of hydrolysis can also break down defunct macromolecules.

© University of Cambridge International Examinations 2006

www.theallpapers.com

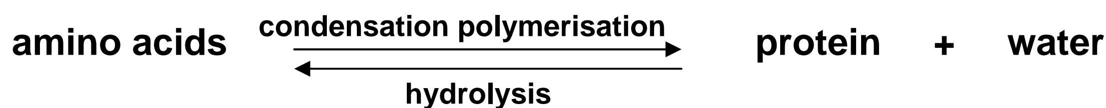


Figure 1.5 – the processes of condensation and hydrolysis

These twin processes allow the recycling of the components of life's molecules. All the main classes of macromolecules involved in life's processes are condensation polymers:

- nucleic acids;
- proteins;
- carbohydrates;
- lipids.

Shaping up

Hydrogen bonding mentioned earlier is one of the forces that play a significant role in the 3-dimensional folding of proteins. Whether an enzyme, antibody or structural protein, the correct functioning of all proteins depends upon their shape. Life depends on molecules recognizing and interacting with each other in specific ways. Such interaction is most efficient if the molecules involved have complementary structures – they fit together as would a lock and key. Thus an enzyme recognizes its substrate, a DNA strand associates with its complementary strand, and a hormone recognizes its receptor. Interaction of biological molecules begins with recognition of complementary molecules, proceeds by short-lived binding through forces such as hydrogen bonding, and ends in those activities that support and propagate life. The processes of recognition and self-assembly are key to the functioning of structures such as chromosomes, ribosomes, and cell membranes where the specific aggregation of macromolecules of different types gives rise to the functional component of the cell.

© University of Cambridge International Examinations 2006

www.theallpapers.com

1.2 – Protein chemistry

By the end of this section you should be able to:

- *recall that proteins are condensation polymers formed from amino acids,*
- *recognize and describe the generalized structure of amino acids,*
- *distinguish between the primary, secondary and tertiary structure of proteins,*
- *explain the stabilization of secondary and tertiary structure,*
- *describe and explain the characteristics of enzyme catalysis, including*
  (i) specificity [using a simple lock and key model] and the idea of competitive inhibition, and
  (ii) structural integrity in relation to denaturation and non-competitive inhibition,
- *given information, use core chemistry to explain how small molecules interact with proteins and how they modify the structure and function of biological systems.*

Proteins – the workhorses of life

Proteins are complex biological polymers. For instance, the haemoglobin in your red blood cells, responsible for transporting oxygen around your body, has a formula of $C_{2952}H_{4664}O_{832}N_{812}S_8Fe_4$ and a molecular mass of about 65000 (Figure 1.6). Collagen, the major structural protein in our bodies, is made of three chains, each around 1000 amino acids long, coiled round each other in a triple helix.

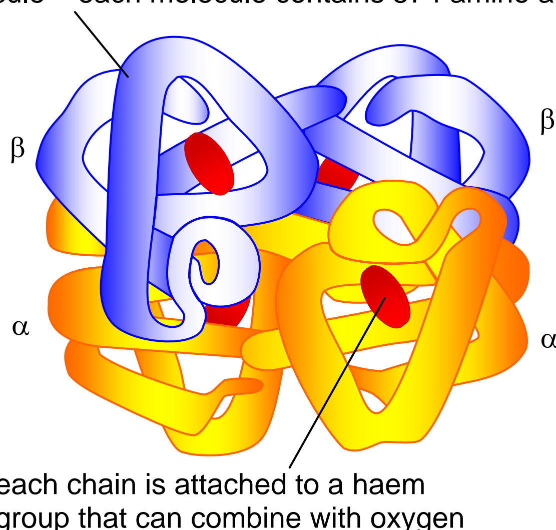


Figure 1.6 – diagram of haemoglobin

Until the 1930's proteins were thought to be random aggregates consisting of chains of polymerised amino acids. Later, with developments in amino acid sequencing and X-ray crystallography (for which Sanger and Pauling, with their respective colleagues, won Nobel Prizes), proteins were shown to have a much more ordered structure than originally believed. Progressively we also came to realize that proteins are involved in virtually every biological process.

Many proteins, such as antibodies, enzymes and haemoglobin, are water-soluble molecules. Others, such as collagen and keratin, are insoluble and aggregate to form very tough and resistant structures. Proteins make up 18% of the mass of the average person (Figure 1.7) and some of their functions are listed in Table 1.1.

© University of Cambridge International Examinations 2006

www.theallpapers.com

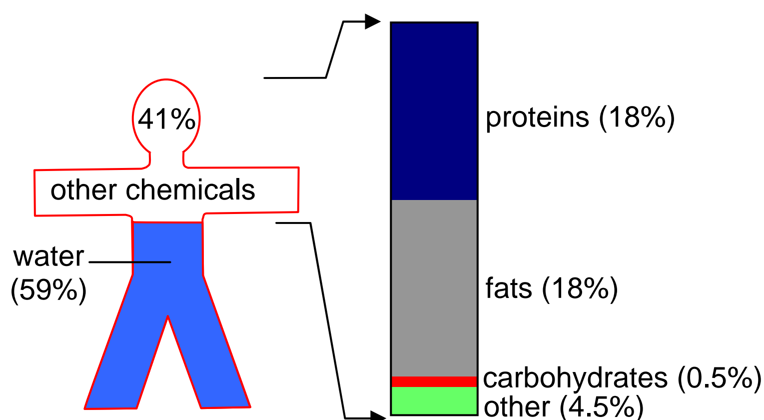


Figure 1.7 – A summary of the chemical components of the average person

Table 1.1: Some proteins and their functions

Protein(s)	Function	Location
myosin actin	muscle contraction	muscle tissue
chymotrypsin pepsin	digestive enzymes	small intestine stomach
insulin	hormone	blood
immunoglobulins	antibodies	blood
collagen keratin	structural proteins	skin, tendon hair
haemoglobin	transport	blood
ferritin	storage	bone marrow, liver, spleen

Proteins are unbranched polymer chains made by linking together large numbers (from hundreds to several thousand) of amino acid monomer units by peptide bonds. Such chains are often referred to as polypeptide chains.

Amino acids – the building blocks of proteins

Common features of amino acids

Amino acids are important organic molecules because living organisms use them as the building blocks of proteins. Protein chains are synthesized from twenty different amino acids. Nineteen of these molecules contain two functional groups: a carboxylic acid group (-COOH) and a primary amino group ($-NH_2$). The amino acid proline is the exception in that it is a cyclic compound and contains a secondary amino group rather than a primary amino group. However, the twenty molecules all have one common feature: the two functional groups are both attached to the same carbon atom (Figure 1.8).

© University of Cambridge International Examinations 2006

www.theallpapers.com

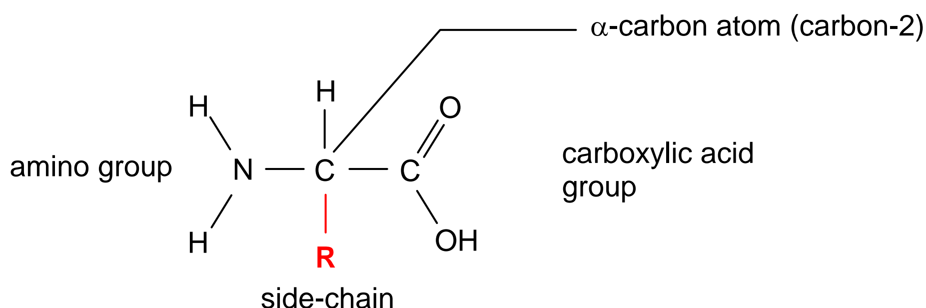


Figure 1.8 – generalised structure of an amino acid, highlighting the key features

When naming organic compounds systematically, the carbon atom of an acid group is always counted as the first in the structure. This means that in these molecules the amino group is always attached to the second carbon atom (C-2): the one immediately adjacent to the carboxyl group. This carbon atom is also sometimes known as the α-carbon atom. These important molecules are therefore all 2-amino acids (or α-amino acids).

The 20 different amino acids that cells use to build proteins differ in the nature of the R-group (the side-chain) (Figure 1.8). These side-chains vary considerably in their complexity. In the simplest case the R-group is just a hydrogen atom, resulting in the simplest of the amino acids. This is known systematically as 2-aminoethanoic acid, but is more commonly referred to as glycine.

The 20 different amino acids can usefully be categorized into separate sub-groups according to the nature of the R-group. There are three broad categories depending on whether the side-chain group is non-polar, polar, or can be ionised (charged) under appropriate conditions (see Figure 1.9).

sub-group (based on type of R-group)	example	structure
non-polar	alanine (ala)	$NH_2-C(H)(CH_3)-COOH$
	valine (val)	$NH_2-C(H)(CH(CH_3)_2)-COOH$
polar	serine (ser)	$NH_2-C(H)(CH_2OH)-COOH$

© University of Cambridge International Examinations 2006

www.theallpapers.com

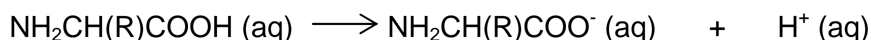
sub-group (based on type of R-group)	example	structure
electrically-charged (acidic or basic side-chains)	**aspartic acid (asp)**	$NH_2-C(H)(CH_2COOH)-COOH$ (R-group: **CH_2COOH**)
	lysine (lys)	$NH_2-C(H)((CH_2)_4NH_2)-COOH$ (R-group: **$(CH_2)_4NH_2$**)

Figure 1.9 – examples of the 20 different amino acids found in proteins

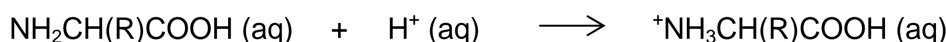
The nature of the R-groups in these amino acids is of crucial importance. Once the amino acids have condensed together to form a polypeptide chain, the R-group is the remaining feature of a particular amino acid which is still distinctive. Interactions between the different R-groups profoundly influence the folding of the polypeptide chain, and hence the shape of the final protein.

The ionisation of amino acids

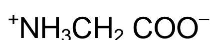
As we have seen, each amino acid molecule contains an acidic group (-COOH) and a basic group, ($-NH_2$). As a consequence, these molecules show both the properties of an acid and those of a base. Such compounds are said to be amphoteric. The acid grouping can lose a proton (H^+ ion).



While, under different conditions, the basic amino group can accept a proton.



In this latter case the proton attaches to the lone pair of electrons on the N atom. In aqueous solution there are no conditions under which amino acids are uncharged. At physiological pH, which is just above neutrality, both the acid and amino groups are ionised. Thus at pH 7 the structure of glycine is as shown:



Thus molecules of amino acids frequently carry charged groups of opposite polarity. Under such circumstances they exist as dipolar ions, or zwitterions. X-ray crystallography of crystalline amino acids has shown that amino acids exist in the zwitterionic form in the solid. Because of this they show the physical properties characteristic of ionic compounds: they are white solids that are soluble in water.

The ionisation of some amino acids is further complicated by the presence of another carboxyl or amino group in the side-chain. This increases the number of possible charged groups present in such molecules. These additional groups (see *Figure 1.9*) will also be charged at around pH 7.

Thus amino acids such as glycine or alanine, with non-polar R-groups, will have no net charge at pH 7 (overall charge = 0). However, at this pH, arginine will have a net charge of +1 because of its additional amino group. Similarly, aspartic acid will have an overall charge of –1 because of the additional acid group in its side-chain. These and related differences in charge can be used to separate amino acids by electrophoresis or ion exchange chromatography.

© University of Cambridge International Examinations 2006

www.theallpapers.com

Structure – the key to protein function

Condensation polymerisation of amino acids

All amino acid molecules contain two functional groups – an amino group and a carboxylic acid group – and can react as bi-functional monomers to form a long chain polymer. Two amino acid molecules can react to form an amide. This is known as a condensation reaction because water is formed in the process. When many amino acids react to form a polymer they produce a condensation polymer – a protein (or polypeptide chain).

For example, glycine and alanine can react to form a dipeptide (Figure 1.10). This dipeptide is an amide made up of two amino acids joined by a peptide bond (or amide link).

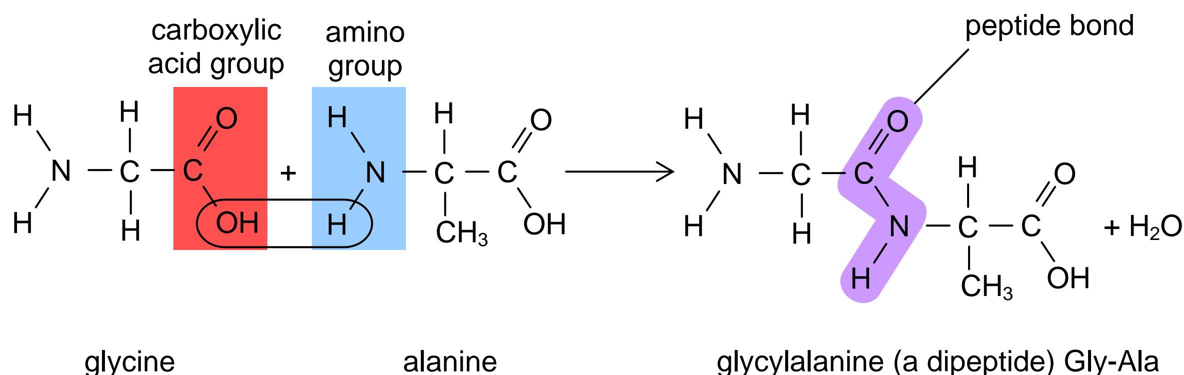


Figure 1.10 – Diagram showing the formation of a gly-ala dipeptide

Additional amino acids can react with the dipeptide to form first a tripeptide and then eventually a polypeptide. In this way a protein can be put together.
The peptide bond consists of the group –CONH- in which the four atoms lie in one plane, with all bond angles being about 120°.

All proteins are made of one or more unbranched polymer chains formed from many amino acid monomer units. Each protein chain is a linear polymer built from its own unique selection from the amino acid pool. It is not only the mix that is unique but also the sequence in which the amino acids are joined together along the chain. The sequence is genetically determined and characterizes that particular protein. A particular protein has a specific biological function because of its particular sequence of amino acids and consequently the particular three-dimensional shape which the sequence generates.

Levels of protein structure

The structure of a single protein chain in its functional form can be considered on three levels:

- primary structure – the sequence of amino acids in a polypeptide chain – the direct product of the coding sequence in the gene.;
- secondary structure – regular structural arrangements of the polypeptide chain that result from hydrogen-bonding between peptide bond regions of the chain;
- tertiary structure – the overall folding of a polypeptide chain that arises from interactions between the amino acid side-chains.

The primary structure of proteins

Each polypeptide chain is a linear polymer of amino acids and as such has an amino- (or N-)terminal end and a carboxy- (or C-)terminal end (Figure 1.11). Each polypeptide has direction and the sequence of amino acid residues in a chain is known as the primary structure of the polypeptide.

© University of Cambridge International Examinations 2006

www.theallpapers.com

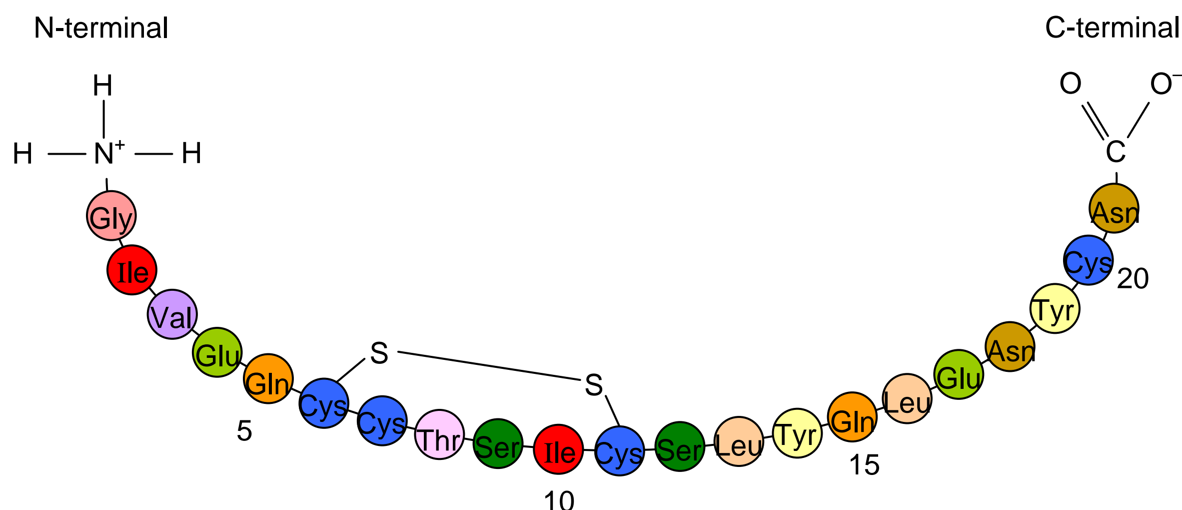


Figure 1.11 – The primary sequence of the insulin A chain, a short polypeptide of 21 amino acids

The primary structure of the insulin A chain is shown here. In the cell, a polypeptide chain is always synthesized from the N-terminal end to the C-terminal end. Thus, when writing out the primary sequence of a polypeptide chain the amino acids are numbered from the N-terminal end. Of particular significance in the primary sequence are the positions of any cysteine residues, as these will determine the possible formation of disulphide bridges to stabilise the 3D-tertiary structure of the protein.

The primary structure of a polypeptide chain is genetically controlled and is crucial in determining the other levels of structure that the protein can adopt.

The secondary structure of proteins

Each polypeptide has a 'backbone' that runs the length of the chain. As the only difference between the different amino acids lies in their R-groups, this backbone is essentially the same for all protein chains [-C-C-N-C-C-N- etc]. This polypeptide backbone is flexible and in certain regions of the protein can fold in a regular manner, known as secondary structure. These structures are stabilized by hydrogen bonding between the peptide bond regions of the chain's backbone (Figure 1.12). The N-H of one peptide link hydrogen bonding to the C=O of another.

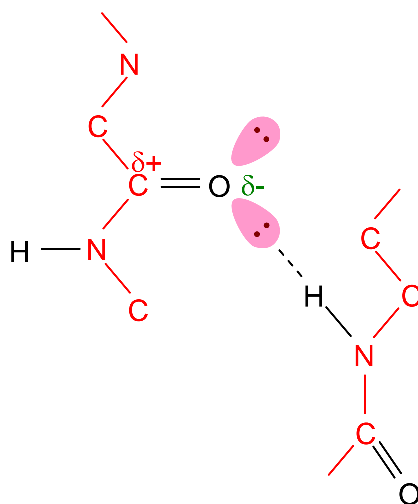


Figure 1.12 – Illustration of hydrogen bond between two polypeptide link regions

This type of folding, stabilized by the intramolecular hydrogen bonding, gives rise to certain structural features which are found in many different types of protein. Collectively this level of structural organisation is known as secondary structure. Two of the most stable arrangements at this level of

© University of Cambridge International Examinations 2006

www.theallpapers.com

protein folding are the α-helix and the β-pleated sheet (Figure 1.13). In both these structures the chain folds on itself in a very stable arrangement because of the many hydrogen bonds formed between adjacent peptide bond regions.

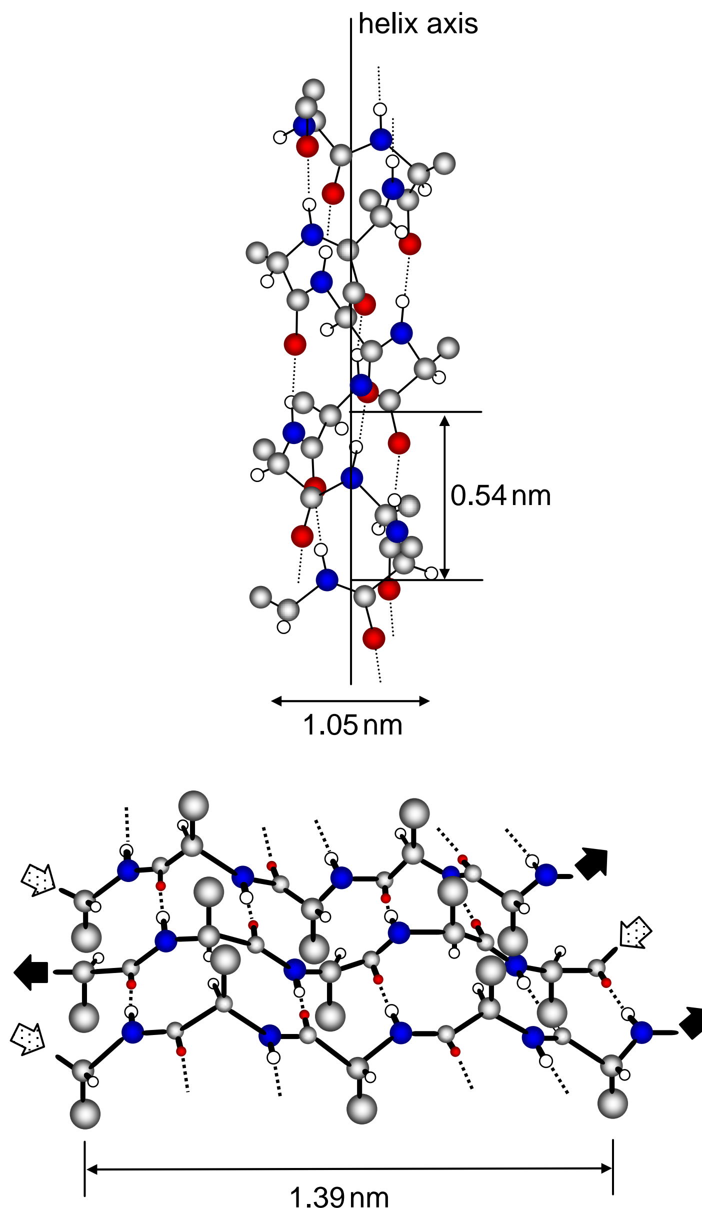


Figure 1.13 – the structure of (a) the α-helix [as found in keratin]; (b) the β-pleated sheet [as found in silk fibroin]

Regions of regular secondary structure occur in many proteins. Figure 1.14 shows a computer-generated graphic of the structure of pepsin – a protease present in our stomach which helps digest our food. The structure has distinct α-helical regions, represented by the 'cylindrical rods', and β-pleated sheet regions, represented by the 'arrows'.

© University of Cambridge International Examinations 2006

www.theallpapers.com

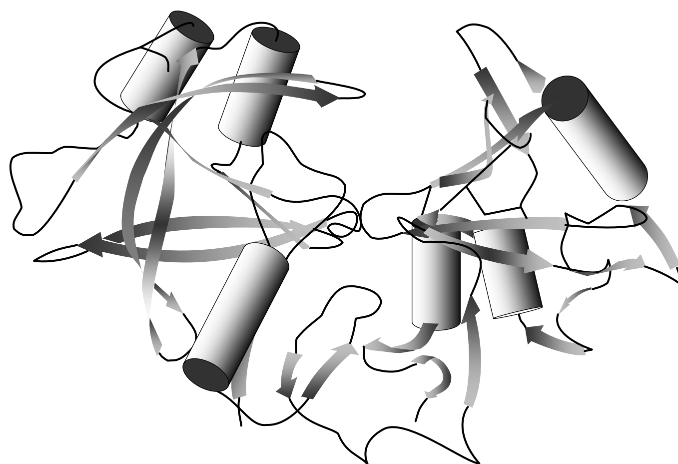


Fig 1.14 – A computer graphic model of the structure of pepsin

Protein tertiary structure

A series of possible interactions between the R-groups of the different amino acid residues in a protein chain produces a third level in the hierarchy of protein folding. This is known as tertiary structure and is crucially important to a protein's function. The three-dimensional shape or conformation of a protein chain is maintained by a series of mainly non-covalent, intramolecular interactions between the R-groups of the amino acids making the chain. At this level, the chemical nature of the different R-groups (see Figure 1.9 earlier) becomes particularly significant. Some of these interactions are relatively easily disrupted, others not so, and include:

- van der Waals' forces between non-polar side-chains,
- hydrogen bonding between polar R-groups,
- ionic bonds (salt bridges) between ionised R-groups, and
- covalent disulphide bridges formed between cysteine residues at different locations in the primary sequence (Figure 1.15)

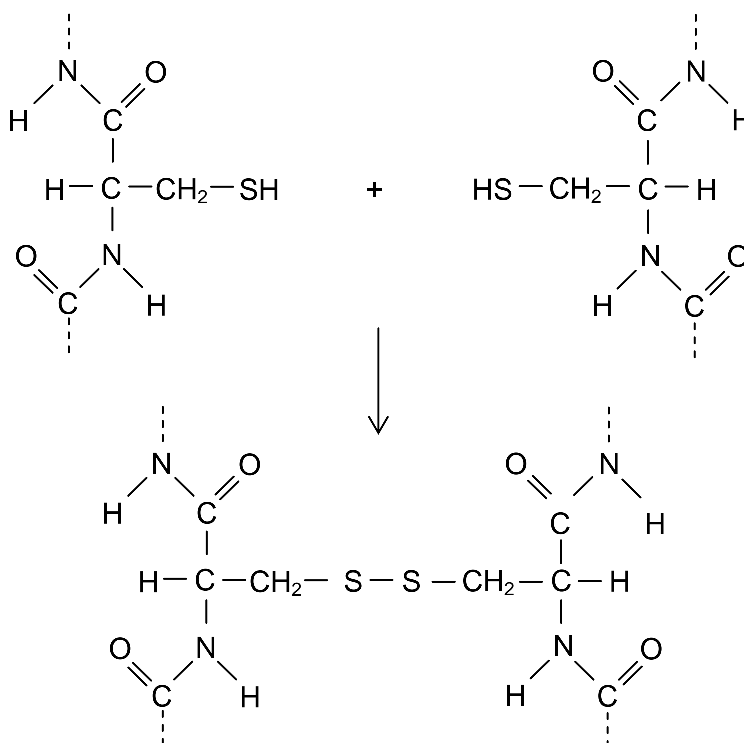


Figure 1.15 – diagram showing the formation of a disulphide bridge

© University of Cambridge International Examinations 2006

www.theallpapers.com

The formation of disulphide bridges is of particular significance. Because of their covalent nature disulphide bonds can have the effect of locking a particular tertiary structure in place. The different possible interactions responsible for maintaining the tertiary structure of a polypeptide chain are summarized in Figure 1.16.

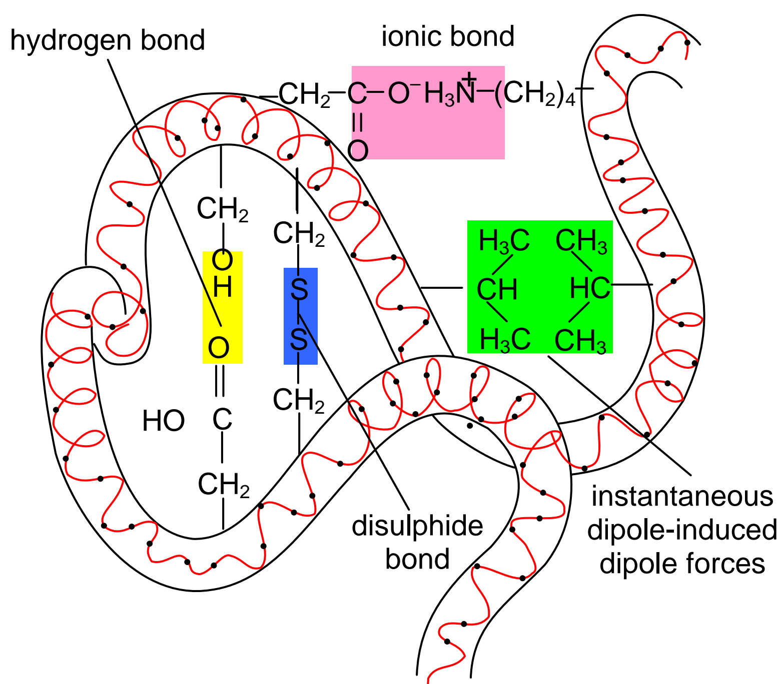


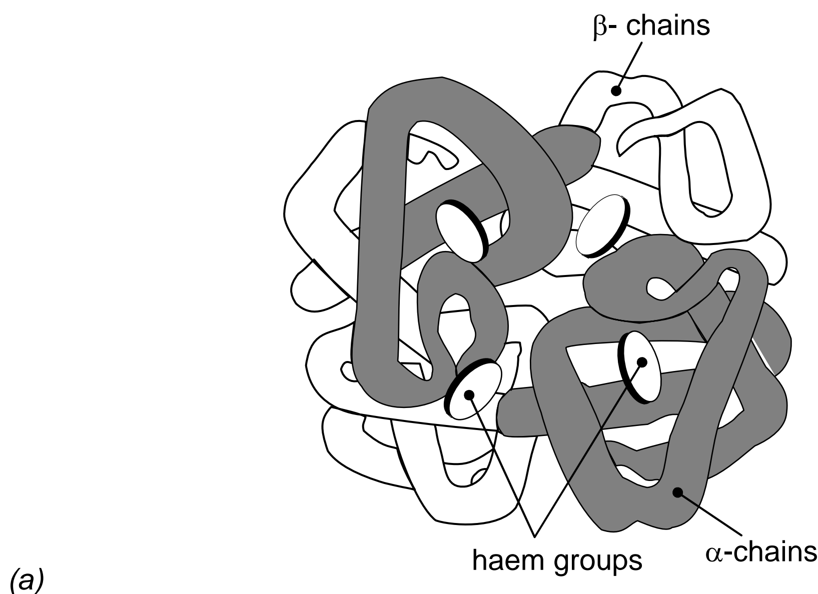
Figure 1.16 – diagram illustrating the nature of the interactions responsible for protein tertiary structure

Haemoglobin – the molecular breathing 'machine'

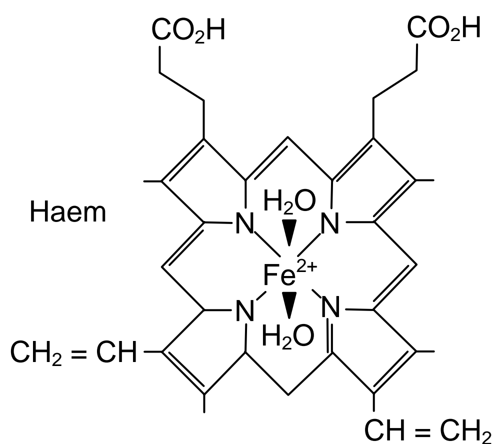
Human haemoglobin consists of two pairs of identical protein chains (subunits) known as α- and β-chains. These assemble together to form the functional, oxygen-carrying protein in our red blood cells (Figure 1.17). Although similar in size (141 and 146 amino acid residues respectively), the human globin α- and β-chains are different polypeptide chains coded for by different genes. When folded, each of the globin chains has a high degree (about 70%) of α-helical secondary structure. On the other hand there are very few, if any, β-pleated sheet regions to the structure. Each of the protein subunits in one molecule is also bound to a non-protein haem group containing an iron(II) ion (Fe^{2+}).

© University of Cambridge International Examinations 2006

www.theallpapers.com



(a)



(b)

Figure 1.17 – (a) the structure of haemoglobin showing the α and β chains; (b) the structure of the haem group.

In haemoglobin, the iron(II) ions in the haem groups bind oxygen and enable the molecules to perform their function of transporting oxygen around the body. The non-covalent interactions similar to those involved in tertiary structure bind these haem groups in place in the structure. The inter-connected nature of the whole structure is emphasised by the fact that the overall structure of the whole molecule changes as oxygen binds to the haem groups. Remarkably the protein structure actually co-operates in the attachment of O_2, subtly changing its shape as the oxygen molecules attach themselves. The molecule undergoes a 'breathing motion'. The positive co-operative effect achieved by these changes in structure means that there is a favoured sequence for the attachment and detachment of oxygen from the molecule. Either all four subunits have oxygen attached to their haem groups, or none of them do. This leads to the most efficient transfer of oxygen from the lungs to the other tissues of the body. The haemoglobin molecule is not simply an inert carrier, but a functioning piece of machinery of molecular dimensions. You will come across the idea of molecular machinery again in the section on nano technology.

SAQ 1. ***(a)*** *Name and give the formulas of the two functional groups in α-amino acids.*

(b) ***(i)*** *Which one of the 20 amino acids found in proteins has an unusual structure involving one of these functional groups?*

(ii) *What is distinctive about its structure?*

© University of Cambridge International Examinations 2006

www.theallpapers.com

(c) *Name an example of the following types of amino acid:*

(i) *a non-polar amino acid,*

(ii) *a polar amino acid with an –COOH group in the side-chain,*

(iii) *a polar amino acid with an $–NH_2$ group in the side-chain.*

(d) *Name the amino acid which gives rise to disulphide-bridging between different regions of a polypeptide chain.*

SAQ 2. **(a)** *What type of polymerisation takes place when a protein chain is formed from α-amino acids?*

(b) *A polypeptide chain is said to have direction. How are the two ends of the chain referred to?*

(c) *What type of chemical bonding is responsible for maintaining the primary structure of a protein chain?*

SAQ 3. **(a)** *Draw a diagram to show how hydrogen bonds may be formed between two peptide bond regions of a polypeptide chain.*

(b) *List the different types of interaction responsible for stabilizing the tertiary structure of a protein.*

Summary

- Proteins are condensation polymers. The amino acid monomers are linked by peptide bonds to form the polypeptide chain.
- α-amino acids are the biological monomers from which protein molecules are built.
- α-amino acids contain both a carboxylic acid group and amino group. These groups are both attached to the same carbon atom. They have the general formula $NH_2CH(R)COOH$.
- α-amino acids exist in a zwitterionic form in the solid state and when in aqueous solution around pH 6 to 7. In this state both the amino group and the acid group are ionised
- α-amino acids have different R groups, which are categorized according to their polar or non-polar nature. There are 20 different amino acids used by cells to build proteins.
- The primary structure of a protein chain is the sequence of amino acids in the chain.
- Polypeptide chains have direction. They have an amino- (or N-) terminal end and a carboxy- (or C-) terminal end.
- The secondary structure of a protein involves the folding of the polypeptide backbone and is stabilised by hydrogen-bonding between peptide bond regions of the chains.
- The two most stable types of secondary structure are the α-helix and the β-pleated sheet.
- The tertiary structure of a protein consists of the folding of the polypeptide chain arising from interactions between the R groups (side chains) of the amino acids in the chain. These interactions include van der Waals' forces, ionic bonds, hydrogen-bonding, and covalent disulphide bridges.
- The functioning of proteins in their biological role is very much linked to their three-dimensional shape.
- Proteins can be hydrolysed back to their constituent amino acids by treatment with hot concentrated hydrochloric acid.

Enzymes – Nature's catalysts

Nature is the most prolific of chemical industries. It turns out billions of tonnes of a vast range of products every year using the simplest of starting materials. The catalysts that make all this possible are enzymes – all large protein molecules. As with inorganic catalysts, enzymes speed up chemical reactions without themselves being used up in the course of the reaction. Enzymes are able to catalyse reactions in aqueous solutions under exceptionally mild conditions of temperature and pH.

© University of Cambridge International Examinations 2006

www.theallpapers.com

Catalytic efficiency

In common with other catalysts enzymes provide an alternative reaction pathway that has a lower activation energy barrier than the uncatalysed reaction.

Most inorganic catalysts catalyse a range of reactions involving similar substances. They are able to function over quite a wide range of temperatures and pressures. Although they are able to produce significant effects in speeding up reactions, the price of this relatively broad specificity is that the increases in reaction rate achieved are not as spectacular as those produced by enzymes. Dramatically restricted as to the conditions under which they can function, enzymes are very effective catalysts. Table 2 compares the efficiency (turnover number) of several well-characterized enzymes.

Table 2: Comparison of the catalytic efficiency of certain enzymes

enzyme	turnover number (molecules reacted per enzyme molecule per minute)
carbonic anhydrase	36 000 000
catalase	5 600 000
β-amylase	1 100 000
β-galactosidase	12 500
phosphoglucose isomerase	1 240
succinate dehydrogenase	1 150

Enzymes are very specific: generally catalysing only one particular reaction. Carbonic anhydrase, for instance, is an enzyme in red blood cells that catalyses the reaction:

$$CO_2 + H_2O \rightarrow H_2CO_3$$

This enzyme increases the rate of this reaction up to a million fold, increasing the efficiency of removal of carbon dioxide from our bloodstream. In a cellular environment this specificity is absolutely essential – an enzyme must, for instance, be able to distinguish one amino acid from another, one nucleotide from another, or a particular sequence of such residues in a macromolecule from all other sequences. Each enzyme has a specific substrate, the substrate being the target molecule acted upon during the enzyme-catalysed reaction.

Thus, although enzymes are functioning within the rules that define catalytic activity, they differ from ordinary chemical catalysts in several important respects.

- Higher reaction rates – the rates of enzyme catalysed reactions are typically increased by factors of 10^6 to 10^{12} times compared to the uncatalysed reaction and are several orders of magnitude greater than those of the corresponding chemically catalysed reaction.
- Milder conditions – enzyme catalysed reactions occur under relatively mild conditions: temperatures below 100°C, atmospheric pressure, and at pH's around neutrality.
- Greater reaction specificity – enzymes have a vastly more defined specificity with regard to their substrate and products: enzyme catalysed reactions are 'clean' and do not produce side products.
- Capacity for regulation – the catalytic activities of many enzymes can be varied by the concentrations of substances other than the substrate: the mechanism of these regulatory processes can be complex.

Shapely molecules

The vast majority of enzymes are water-soluble globular (i.e. curled up in a ball) proteins. The complicated folding of the protein chain to form the tertiary structure gives rise to 'clefts' or 'crevices' of precise geometric shape on the surface of the enzyme. The precise shape of these clefts is designed to 'recognise' and hold in place a particular substrate molecule while it reacts. This substrate-binding site has a shape which matches the shape of the substrate. As this region is where the enzyme-catalysed reaction takes place it is known as the active site of the enzyme.

© University of Cambridge International Examinations 2006

www.theallpapers.com

The catalytic properties and specificity of an enzyme are determined by the chemical nature of the amino acid R-groups located at the active site. The active site usually occupies less than 5% of an enzyme's surface area and involves only a small number (3 to 12) of amino acids. The rest of the enzyme structure functions as the scaffolding that maintains and protects the shape of the active site.

The 'lock and key' model

The precise specificity shown by enzymes led to Fischer (in 1894) proposing a model of enzyme activity often referred to as the 'lock and key' mechanism. He suggested that enzymes catalysed reactions by binding to substrates in a manner similar to how a key (the substrate) fits into a lock (the enzyme) (Figure 1.18). Locks and keys are complementary structures and this would also explain enzyme specificity. Only one substrate will fit into the active site, just as only one key fits a lock.

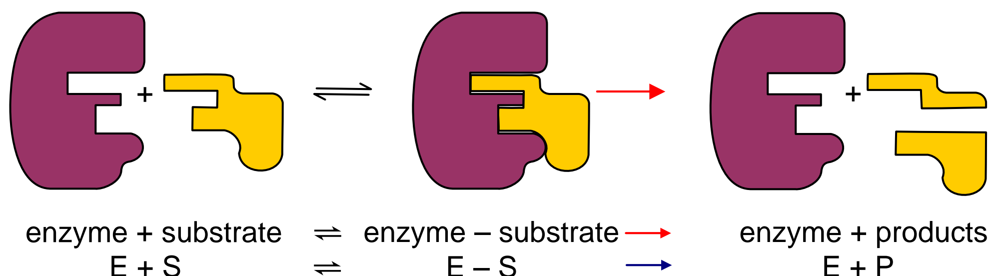


Figure 1.18 – the 'lock and key' mechanism

Enzymes, as other catalysts, function by providing an alternative reaction pathway that requires a lower activation energy (E_a). Thus more molecular interactions possess sufficient energy to produce products. The energy profile shown in Figure 1.19 shows how the formation of the enzyme-substrate complex reduces the energy requirement for the reaction to proceed.

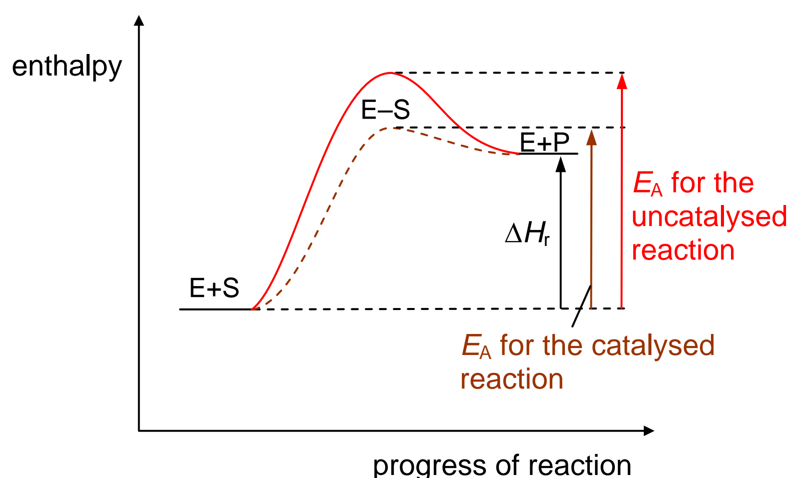
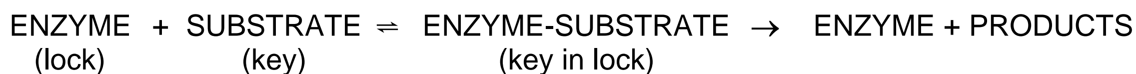


Figure 1.19 – energy profile for enzyme-catalysed and non-catalysed reaction

The overall reaction between enzyme and its substrate can be represented by the following equation:



ENZYME + SUBSTRATE ⇌ ENZYME-SUBSTRATE → ENZYME + PRODUCTS
(lock) (key) (key in lock)

The first stage of the reaction is reversible since if the available energy is not greater than E_a the complex may dissociate without product being formed. In some cases the second stage is also reversible, making the whole enzyme catalysed process capable of proceeding in either direction depending on the cells metabolic requirements. Once the products have been formed, they leave the active site of the enzyme. The enzyme is then free to combine with a new substrate molecule. Enzymes, like inorganic catalysts, are not used up in the reaction they catalyse so they can be used again and again.

© University of Cambridge International Examinations 2006

www.theallpapers.com

Lysozyme, the first protein whose 3-dimensional structure was determined by X-ray crystallography (Figure 1.20), illustrates the 'fit' between enzyme and substrate very well. It is a water-soluble enzyme present in tears, egg-white, and nasal mucus. It has an important role in protecting us from bacteria because it breaks down the carbohydrates present in the bacterial cell wall.

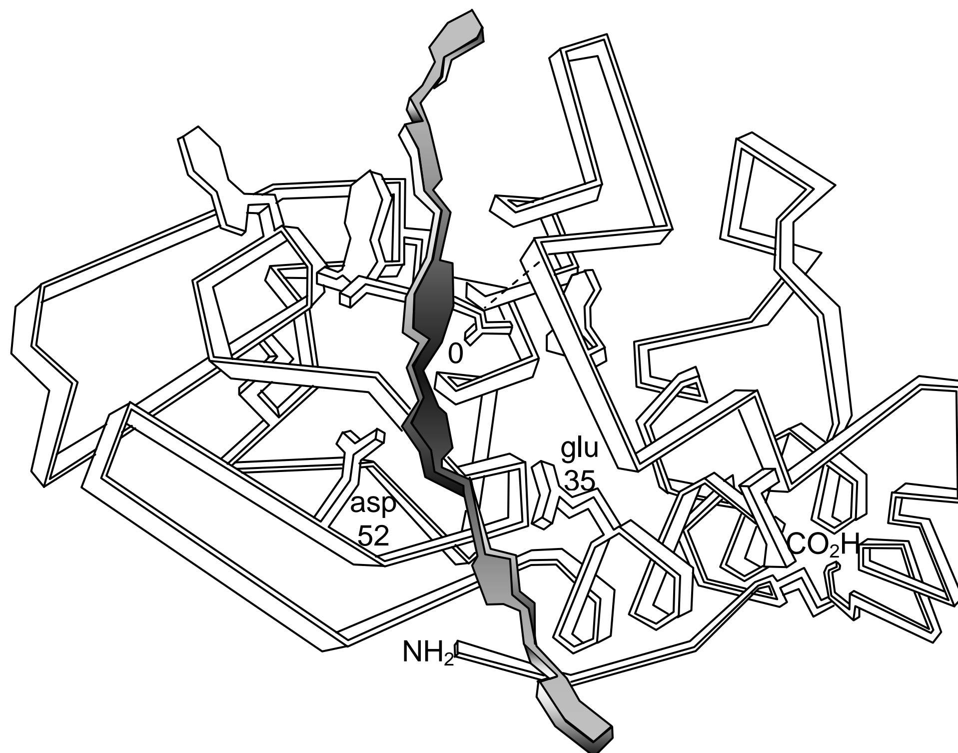


Figure 1.20 – the structure of the enzyme, lysozyme, showing the carbohydrate substrate lying in the active site.

Competitive inhibition of enzyme activity

Further evidence for the 'lock and key' model of enzyme activity is given by studies on enzyme inhibitors, and, in particular by a type of inhibitor which imitates the substrate in the way it binds to the enzyme. This type of inhibitor is known as a competitive inhibitor. Competitive inhibitors of a particular enzyme are molecules that have a similar shape to the substrate molecule. Such molecules can bind to the active site but cannot participate in the catalysed reaction. When they are present in the active site no reaction is taking place and the correct substrate cannot attach to the enzyme.

When such an inhibitor is added to an enzyme/substrate mixture there is indeed a competition between the substrate and the inhibitor to occupy the active sites on the enzyme molecules. The result of this competition depends on the relative concentrations of the substrate and inhibitor. The functionality of the enzyme molecules is not interfered with – the active sites are merely blocked (Figure 1.21). This type of inhibition is reversible by an increase in substrate concentration.

© University of Cambridge International Examinations 2006

www.theallpapers.com

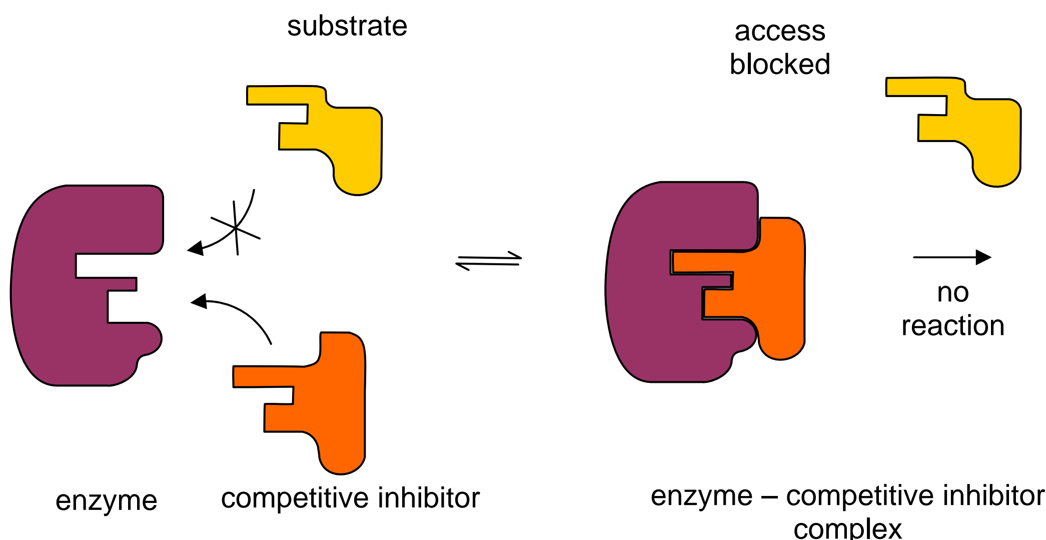


Figure 1.21 – model of action of a competitive inhibitor

The classical example of competitive inhibition is the inhibition of succinate dehydrogenase by various ionic species that structurally resemble the substrate, succinate (Figure 1.22).

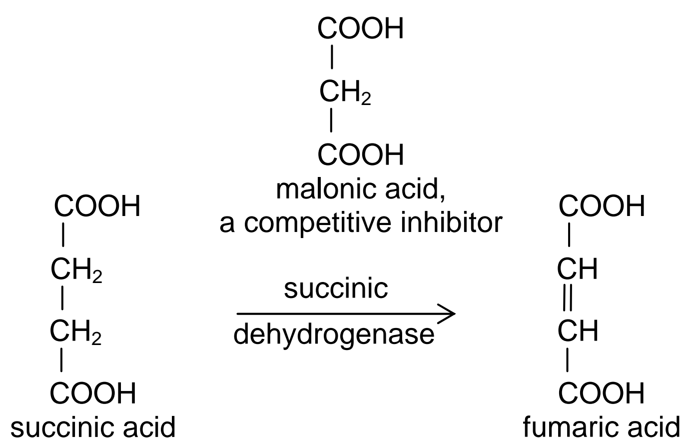


Figure 1.22 – succinate, the reaction, and some competitive inhibitors

Sulphanilamide acts as an anti-bacterial agent because it is a competitive inhibitor for an enzyme producing a compound essential to the living bacteria. This area is an immense field of research centred on computer-aided molecular 'design'. Many medicinal drugs, toxins, pesticides and herbicides act as enzyme inhibitors by virtue of their molecular shape and so it is important to know how they work. Knowledge of the shapes of active sites and inhibitors enables more effective drugs and pesticides to be 'designed'.

Non-competitive inhibition of enzymes

Enzyme structure is even more subtle than suggested above, and molecules can bind on to regions of the enzyme other than the active site and affect enzyme activity. This is known as non-competitive inhibition.

In non-competitive inhibition the inhibitor again binds to the enzyme, preventing the catalysed reaction from occurring. However, in this case, the inhibitor does not bind to the active site. Instead it binds to another position on the enzyme. This binding is thought to cause one of the following:

- the active site to change shape so that the substrate cannot bind (Figure 1.23).
- the enzyme-substrate complex to change shape so that the reaction cannot take place.

© University of Cambridge International Examinations 2006

www.theallpapers.com

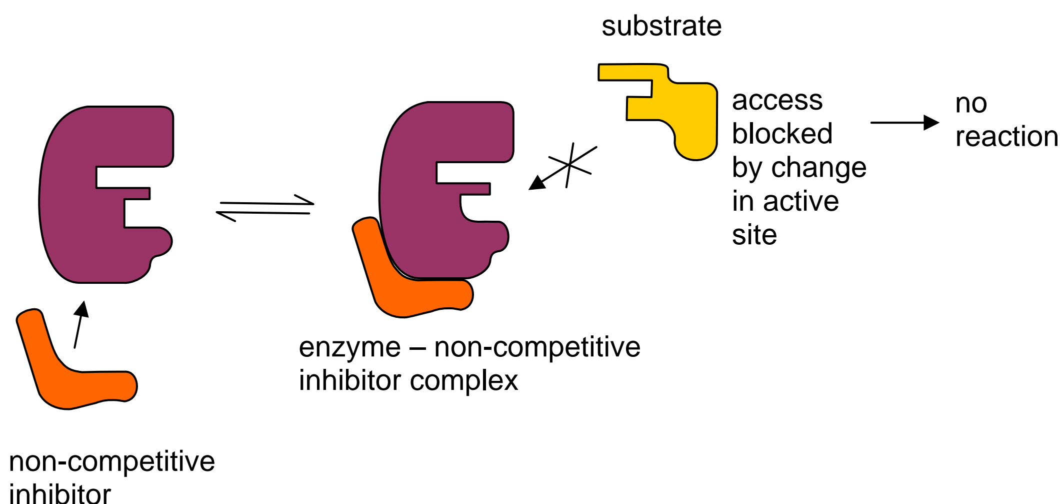


Figure 1.23 – scheme for non-competitive inhibition

In non-competitive inhibition, the inhibitor is not shaped like the substrate and there is no competition between the substrate and the inhibitor. The inhibition cannot be overcome simply by adding more substrate. The effect is to reduce the number of active enzyme molecules available.

As with competitive inhibition, most non-competitive inhibitors only bind weakly to the enzyme. If the concentration of inhibitor falls, the enzyme-inhibitor complex falls apart and the functional shape of the enzyme is restored. This type of inhibition is reversible and can provide an important mechanism for feedback control of a metabolic pathway in cells.

One example of non-competitive inhibition involves the effect of heavy metal ions, such as silver or mercury, on a range of enzymes. Such enzymes contain amino acid side-chains that contain –SH groups. The heavy metal ions react reversibly with one or more –SH groups, replacing the hydrogen atom with a heavy metal atom (Figure 1.24).

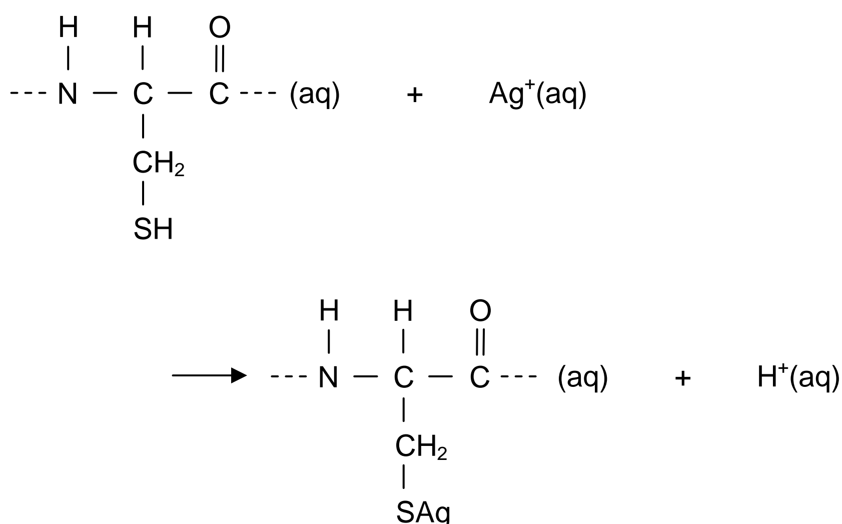


Figure 1.24 – the reaction between heavy metal ions and an –SH group

The resulting modification changes the shape of the enzyme sufficiently enough to prevent the catalysed reaction taking place. The enzyme is temporarily denatured.

Factors affecting enzyme activity

The catalytic activity of any enzyme is highly dependent on those relatively weak, but highly significant, interactions that give rise to the three-dimensional tertiary structure of the protein. The

© University of Cambridge International Examinations 2006

www.theallpapers.com

same non-covalent interactions are also involved in the interaction between the substrate and the enzyme. Recognition and molecular 'fit' are the key ideas behind enzyme function. Even subtle changes in pH or temperature can modify the interactions involved in molecular shape and recognition, resulting in an enzyme working at less the maximum efficiency. Interactions with other chemical substances that cause irreversible changes in structure can also result in loss of the correct structure and the destruction of enzyme activity known as denaturation.

The effect of temperature

The effect of temperature on enzyme activity is complex because it is the outcome of several different factors:

- the speed of the molecules;
- the activation energy of the catalysed reaction;
- the thermal stability of the enzyme and the substrate.

At relatively low temperatures (around 0°C) the rate of most enzyme-catalysed reactions is very low. The molecules involved in the reaction have low kinetic energy. They do not collide frequently, and even when they do the molecules do not possess the minimum energy (E_a) required for reaction to occur. The enzyme is said to be deactivated by low temperature. Increasing temperature increases the rate of enzyme activity since the molecules involved have greater kinetic energy as the temperature rises. Between 0°C and approximately 40°C the rate of enzyme activity increases almost linearly (Figure 1.25) for reasons similar to any other chemical reaction:

- the molecules are moving more quickly, increasing the frequency of collision, and
- a greater proportion of the collisions involve molecules with energy greater than the activation energy for the catalysed reaction.

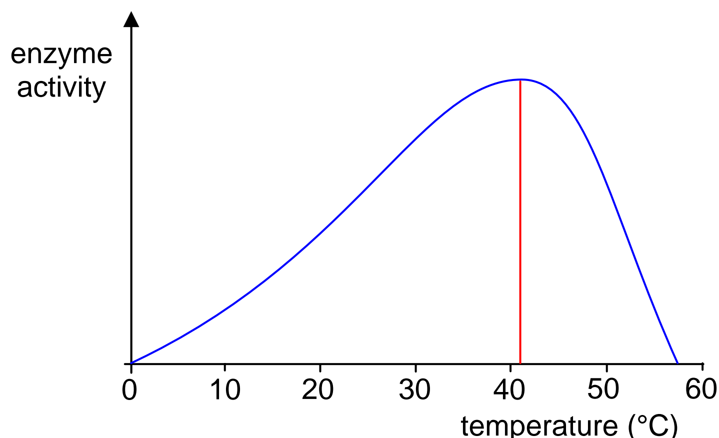


Figure 1.25 – profile of enzyme activity with temperature

For most enzymes the rate of reaction starts to decrease above 40°C. Above this temperature a different effect comes into play. Increased thermal motion of the polypeptide chain is causing disruption of the forces maintaining the shape of the enzyme molecules. The enzyme molecules are progressively denatured, causing the shape of the active site to change. Above 65°C the enzymes from most organisms are completely heat denatured.

However, there are thermophilic organisms that show adaptation to the high temperatures of hot springs and deep-sea thermal vents. Enzymes from such 'extremophiles' retain activity at 80°C or higher. They provide fascinating models for studying the modifications of protein structure necessary to maintain enzyme function at such temperatures.

© University of Cambridge International Examinations 2006

www.theallpapers.com

The effect of pH changes

Enzyme activity is also dependent upon pH.

- Extremes of pH (high acidity or alkalinity) will denature proteins by disrupting the precise three-dimensional arrangement of the protein chains.
- Even small changes around neutral pH can affect the ionization of amino acid side-chains in the active site and/or the substrate itself (Figure 1.26). If enzyme activity depends on particular residues in the active site being charged or not, then a shift of just one pH unit (remember that this represents a ten-fold change in H^+ ion concentration) can change the enzyme activity significantly.

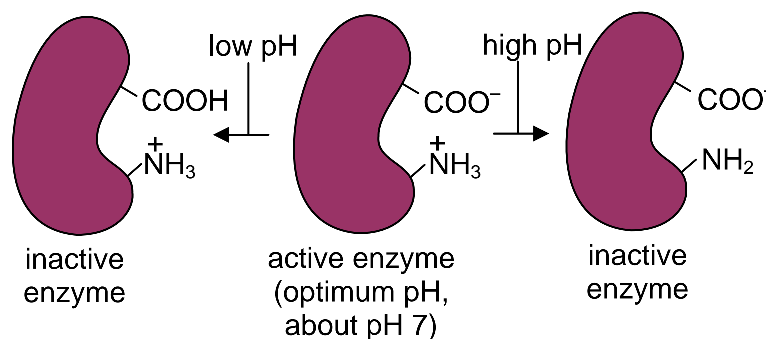


Figure 1.26 – pH changes can affect the ionization of the active site

Most enzymes are active over a fairly narrow range of pH. Each enzyme has its own distinct optimum pH (Figure 1.27).

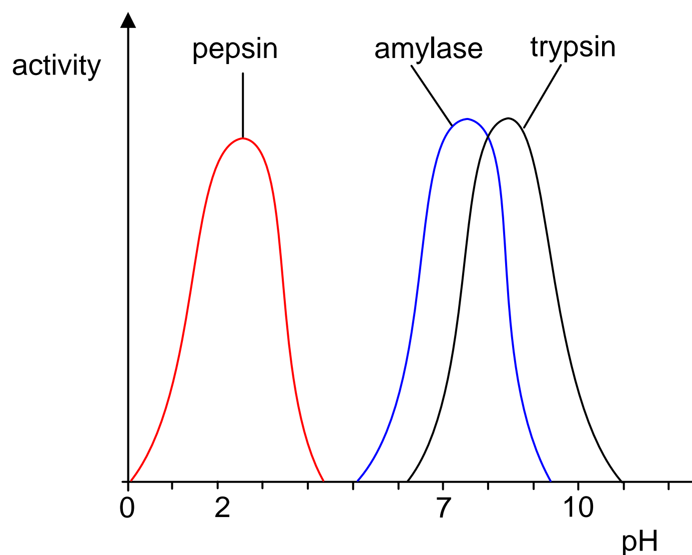


Figure 1.27 - curves showing pH optima for several enzymes

Not all enzymes have the same pH optimum. Digestive enzymes show this clearly.

- Pepsin hydrolyses proteins to peptides in the very acidic conditions of the stomach.
- Amylase, found in saliva, hydrolyses starch to a mixture of glucose and maltose. The pH of saliva is approximately neutral.
- Trypsin hydrolyses peptides to amino acids in the mildly alkaline conditions of the small intestine.

© University of Cambridge International Examinations 2006

www.theallpapers.com

Chemical denaturation

Isolated enzymes extracted from tissues can be denatured by changes in chemical conditions. Thus,

- high salt concentration changes the ionic environment of an enzyme, disrupting ionic interactions between different regions of the chain, while
- urea denatures proteins by disrupting the hydrogen bonds that maintain the secondary and tertiary structure of proteins.

Certain chemical inhibitors totally inactivate enzymes; their effects are irreversible. Some of these inhibitors have been useful in establishing the key amino acid residues responsible for the catalytic function of the active site of a particular enzyme. DFP (systematic name: di-(1-methylethyl)fluorophosphate) is a reagent that binds to serine residues in enzymes. It has been used to show the importance of serine in the active sites of chymotrypsin and acetylcholinesterase. Its effect on this latter enzyme explains its function as a nerve gas. The horrific attack on the Tokyo underground in March 1995 used the nerve gas, sarin (Figure 1.28). Sarin is very similar in structure to DFP.

MASSACRE BY POISON GAS

Many die, 1,200 collapse in Tokyo Tube attack

Figure 1.28 – report of the nerve gas attack in Tokyo, 1995

Prosthetic groups and coenzymes – 'little helpers'

Many enzymes require a non-protein group, or cofactor, that is necessary for them to function as catalysts. Without this cofactor the protein molecule, or apoenzyme, has no enzyme activity.

APOENZYME + COFACTOR → HOLOENZYME
(protein molecule) (functional enzyme)

Some cofactors are actually integrated into the enzyme itself, remaining part of the enzyme under all circumstances. For example, carbonic anhydrase, one of a group of metalloenzymes, contains a metal ion at the heart of the active site. In carbonic anhydrase the metal ion is a Zn^{2+} ion, and its presence is absolutely essential for enzymic activity. Other enzymes have the haem group as a cofactor. These include cytochrome oxidase, an enzyme involved in the crucial process of respiration. Cofactors such as these, which are an integral part of the enzyme structure, are known as prosthetic groups.

Other cofactors, known as coenzymes, are not integral parts of the enzyme structure. They are complex organic molecules, often derived from a vitamin. These coenzymes, such as NAD^+ (nicotinamide-adenine dinucleotide), work together with the enzyme to bring about the required reaction. They bind temporarily to the active site of the enzyme and effectively function as a co-

© University of Cambridge International Examinations 2006

www.theallpapers.com

substrate, providing for the transfer of groups or electrons not readily available from the side-chains of the enzyme protein (Figure 1.29). They are released from the enzyme at the end of the catalysed reaction.

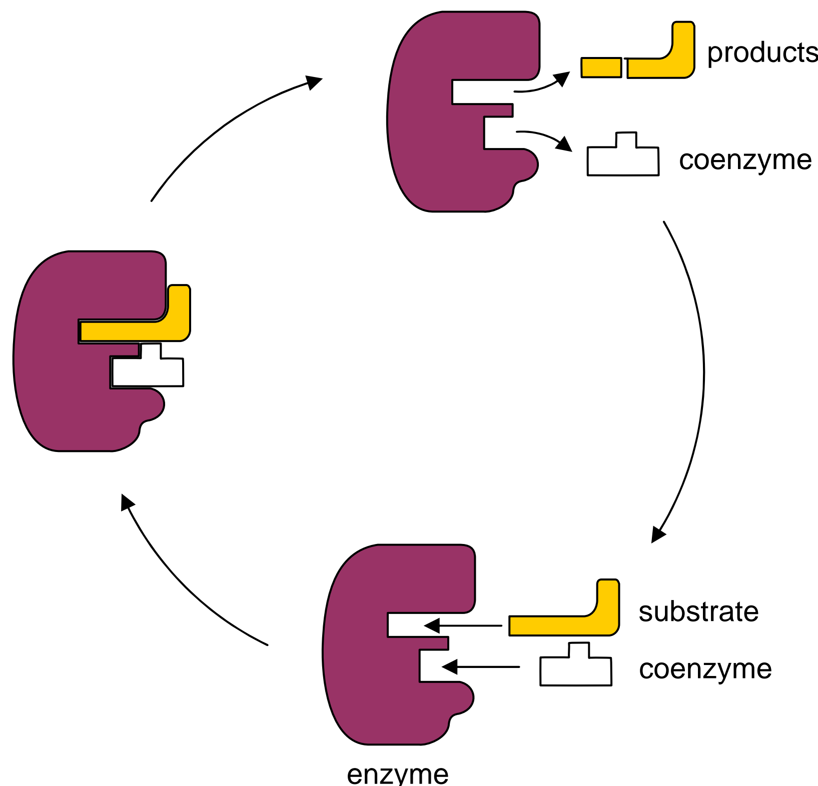


Figure 1.29 – binding of substrate and coenzyme to active site – the catalytic cycle

NAD^+, $NADP^+$, and FAD are important coenzymes because of their ability to accept H^+ ions and electrons and therefore take part in redox reactions. They are sometimes referred to as 'hydrogen carriers' because of their ability to transfer hydrogen atoms between reactions. Coenzyme A acts as a carrier of CH_3COO- groups (acyl groups) and is therefore important in the metabolism of fatty acids (the long chain carboxylic acids in food). The vitamin precursors of many of these coenzymes are essential components of the diet of an organism. For instance, both NAD^+ and FAD are derived from the water-soluble B group vitamins. In humans, an inadequate supply of these precursors gives rise to deficiency diseases.

Ion channels in biological membranes

The forces that contribute to protein folding in enzymes and their interactions with small molecules to bring about the catalysed reaction are also important in maintaining the function of other protein structures.

One recent focus of crucial importance has been the proteins that form the water and ion channels in cell membranes. These channels are found to consist of protein subunits which sit across the plasma membrane – they are trans-membrane proteins. They have enzyme-like function and, by their interaction with ions such as Na^+ and K^+ and relatively small molecules such as ATP (adenosine triphosphate), these proteins are able to selectively control the transport of ions into, and out of, the cell. In a manner similar to enzyme inhibitors certain plant alkaloids, oubain, for instance, can interfere with ion transport by interacting with the protein units of the ion channels.

SAQ 4. (a) *Sketch the energy profile of an uncatalysed exothermic reaction, showing :*

(i) *the activation energy (E_a), and*

(ii) *the enthalpy change of reaction (ΔH_r).*

© University of Cambridge International Examinations 2006

www.theallpapers.com

(b) *Sketch a similar energy profile for the above reaction when it is catalysed by an enzyme.*

SAQ 5. (a) *Explain what you understand by the term 'active site' of an enzyme.*

(b) *Explain how the 'lock-and-key' model describes the mechanism of enzyme action.*

SAQ 6. (a) *Sketch a graph to show how the activity of an enzyme varies with temperature.*

(b) *Explain the shape of the graph in terms of kinetic theory and the effect of temperature on the integrity of the enzyme's structure.*

SAQ 7. (a) *Sketch a graph to show how pH affects the activity of a neutral protease.*

(b) *Explain the shape of the graph in terms of changes taking place in the region of the enzyme's active site.*

Summary

- Enzymes are biological catalysts – they increase the rate of a chemical reaction without being altered themselves by the overall reaction.
- Enzymes achieve their effect by providing an alternative reaction pathway that has a lower activation energy than the uncatalysed reaction.
- Enzymes are proteins and are able to catalyse reactions in aqueous solutions under mild conditions of temperature and pH. They show a high degree of specificity and are remarkably efficient.
- Each enzyme has a specific substrate – the target molecule, or class of molecules, acted upon in the catalysed reaction.
- The function of an enzyme depends on its three-dimensional shape – in particular the precise shape of the active site. This region of the enzyme's surface is arranged to recognise the particular substrate.
- The recognition and binding of a substrate molecule by the active site of an enzyme is often referred to as a 'lock-and-key' mechanism. The active site not only recognises the substrate, it is also the site at which the catalysed reaction takes place.
- Enzyme-catalysed reactions are particularly sensitive to conditions of temperature and pH. They show characteristic temperature and pH optima.
- There are two common types of reversible inhibition of enzymes, competitive and non-competitive inhibition. In competitive inhibition, the inhibitor has a similar shape to the substrate and competes with it to bind to the active site. In non-competitive inhibition the inhibitor does not bind to the active site but elsewhere on the enzyme. This binding alters the shape of the enzyme sufficiently to prevent the catalysed reaction taking place.
- Ion channels in cell membranes are found to be made of proteins with enzyme-like properties that enable the transport of ions across cell membranes to be controlled in a selective way.

© University of Cambridge International Examinations 2006

www.theallpapers.com

1.3 – Genetic information

By the end of this section you should be able to:

- *describe the double helical structure of DNA in terms of a sugar-phosphate backbone and attached bases,*
- *explain the significance of hydrogen-bonding in the pairing of bases in DNA in relation to the replication of genetic information,*
- *explain in outline how DNA encodes for the amino acid sequence of proteins with reference to mRNA, tRNA and the ribosome in translation and transcription,*
- *explain the chemistry of DNA mutation from provided data,*
- *discuss the genetic basis of disease (for example, sickle cell anaemia) in terms of altered protein structure and function,*
- *explain how modification to protein/enzyme primary structure can result in new structure and/or function.*

DNA – the source of heredity

'We wish to suggest a structure for the salt of deoxyribonucleic acid (D.N.A.). This structure has features which are of considerable scientific interest ……… . It has not escaped our notice that the specific base pairing (inherent in the proposed structure) *suggests a possible copying mechanism for the genetic material.'*

Watson & Crick, Nature, 1953.

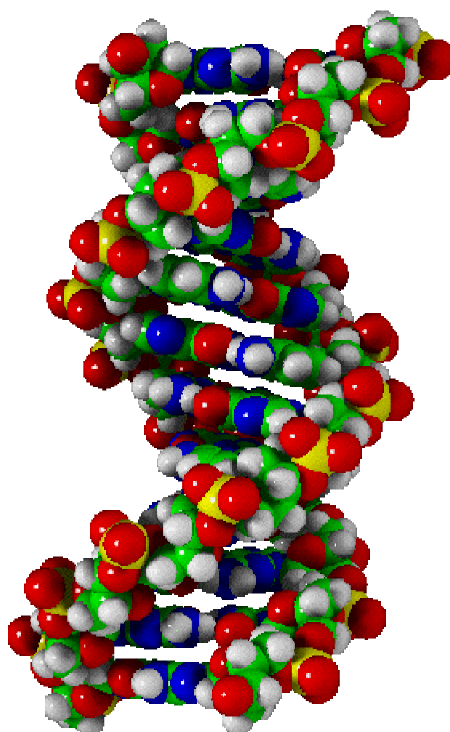


Figure 1.30 – Computer-generated picture of the double helix

DNA was discovered in 1869 by the Swiss biochemist Friedrich Meischer, ten years after the publication of Darwin's 'Origin of the Species'. At this stage there was no suspicion of the immense significance of the molecule isolated from white blood cells in pus sticking to discarded bandages. Not until 1944 did Avery demonstrate that DNA was the material that transferred genetic information from one cell to another.

The recent dramatic spiralling of information about the origins of life, evolutionary development, and the transfer of genetic information from one generation to another, starts on a one-page letter by James Watson and Francis Crick. This letter was published on the 25th April 1953 in the scientific

© University of Cambridge International Examinations 2006

www.theallpapers.com

journal *Nature*. The advent of molecular biology and genetic engineering stems from this event – arguably the most important scientific development of the twentieth century.

The development of science often builds on previous results. Elucidating the structure of DNA would have been impossible without the discovery of X-rays in 1895. In 1925, Max von Laue showed that the diffraction of X-rays could be used to find the arrangement of atoms in crystals. The method was successfully applied to determine the structure of proteins, including myoglobin and insulin, for example. Then, in 1952, Rosalind Franklin, working with Maurice Wilkins, shone X-rays onto crystalline forms of DNA and produced diffraction patterns that were both beautiful and complex (Figure 1.31).

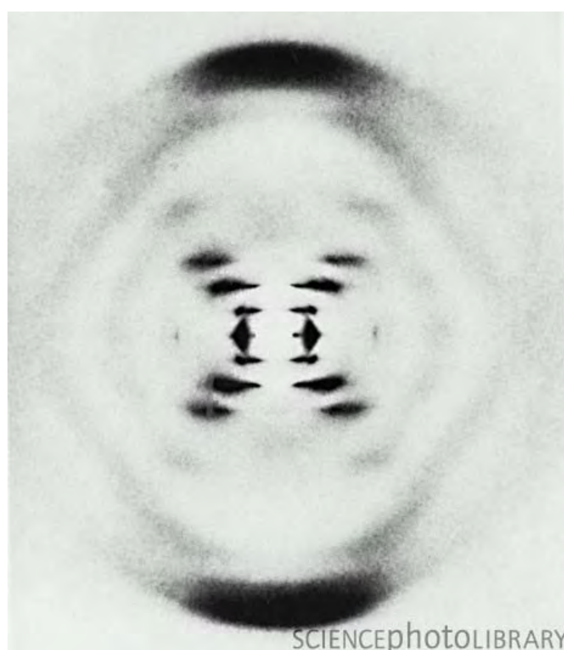


Figure 1.31 – X-ray diffraction pattern of crystalline DNA

The ordered X-ray patterns produced reflected the regularity of a double helical structure. Two DNA strands, running in opposite directions, are linked together in a ladder-like molecule – but a twisted ladder – a right-handed helix (Figure 1.30). Each DNA strand is a condensation polymer of sugar molecules and phosphate groups. Attached to this sugar-phosphate backbone is a sequence of organic bases (Figure 1.32) constructed from four alternatives – often referred to simply by the first letter of their names – A, C, G and T. Hereditary information is stored as the sequence of these bases along the chain. The genetic message is written in a language of only four letters.

© University of Cambridge International Examinations 2006

www.theallpapers.com

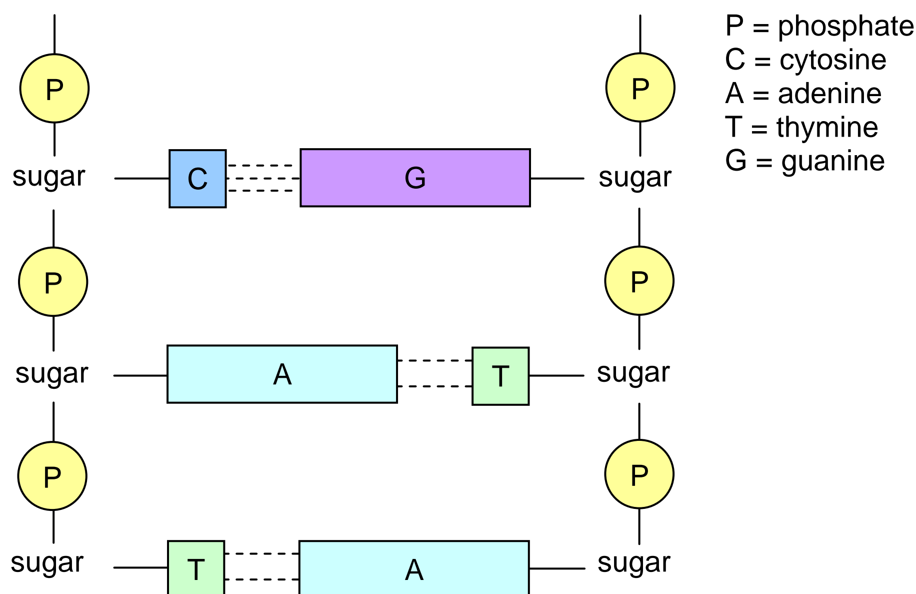


Figure 1.32 – The sugar-phosphate-base structure of DNA

The structure of DNA

Deoxyribonucleic acid (DNA) controls heredity on a molecular level:

- it is a self-replicating molecule capable of passing genetic information from one generation to the next;
- it contains in its base sequence the genetic code used to synthesise proteins.

A strand of DNA is a macromolecule made by the condensation polymerisation of units called nucleotides. Nucleotides (Figure 1.33) are themselves made from three components:

- a sugar,
- a phosphate group,
- a nitrogen-containing organic base.

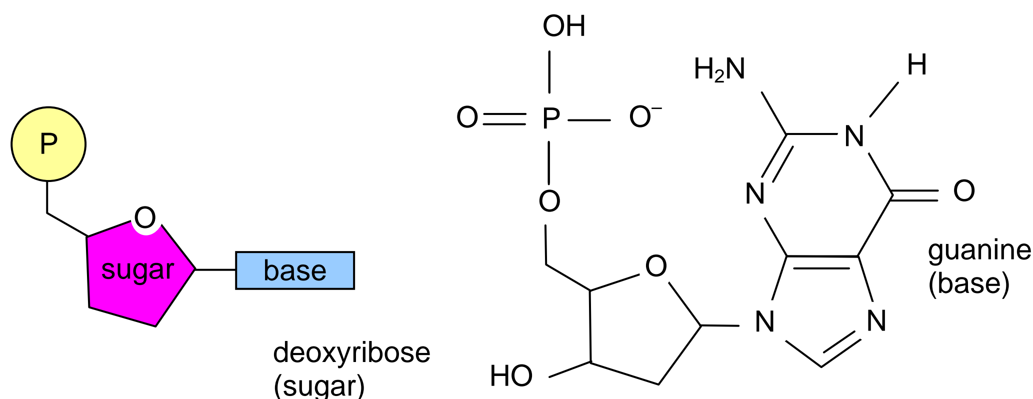


Figure 1.33 – the three components that make up a nucleotide

The sugar molecule in the nucleotides that make up DNA is deoxyribose (a pentose sugar with a five-membered ring). The phosphate group is attached by an ester link to the deoxyribose. The final components of the nucleotides in DNA are the four different bases (all of which are cyclic compounds formed from carbon, nitrogen and hydrogen):

- adenine (A),
- guanine (G),
- thymine (T),
- cytosine (C).

© University of Cambridge International Examinations 2006

www.theallpapers.com

Two of the bases, adenine(A) and guanine(G), have planar two-ring structures [they are purines]. The other two bases, thymine(T) and cytosine(C), are planar single-ring molecules [they are pyrimidines].

Each strand of DNA has a sugar-phosphate backbone with the bases hanging off the side.

The DNA double helix

A DNA molecule consists of two strands. The two strands sit next to each other running in opposite directions, linked together by hydrogen-bonding between the bases.

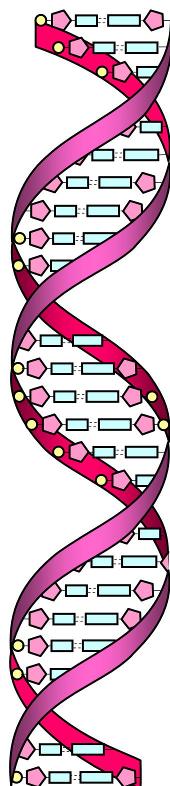


Figure 1.34 – the DNA double helix

These two anti-parallel strands are twisted together in a double helix with the bases on the inside and the sugar-phosphate backbones on the outside (Figure 1.34). The bases positioned between the two chains lie at right angles to the backbone, filling the space between the strands. The bases between the two chains are paired to fill the available space. The bases in each pair interact with each other through hydrogen bonding (Figure 1.35). Two hydrogen bonds form between each adenine-thymine pair (A = T). Three hydrogen bonds are formed between a guanine-cytosine pair (G ≡ C). This difference in the hydrogen bonding between the pairs, together with the size considerations [each pair of bases consists of a two-ring structure and a single-ring structure in order to fit in the space between the backbones], gives rise to the specificity of the pairing of the bases in DNA.

The bases always pair up as follows:

- adenine is always paired with thymine;
- guanine is always paired with cytosine.

This is known as complementary base pairing and is key to the transfer of the information stored in the sequence of the bases along the DNA chains (Figure 1.35). Hydrogen bonding and van der Waals' forces between the stacks of bases are responsible for holding the chains together. The precise sequence of the bases carries the genetic information.

© University of Cambridge International Examinations 2006

www.theallpapers.com

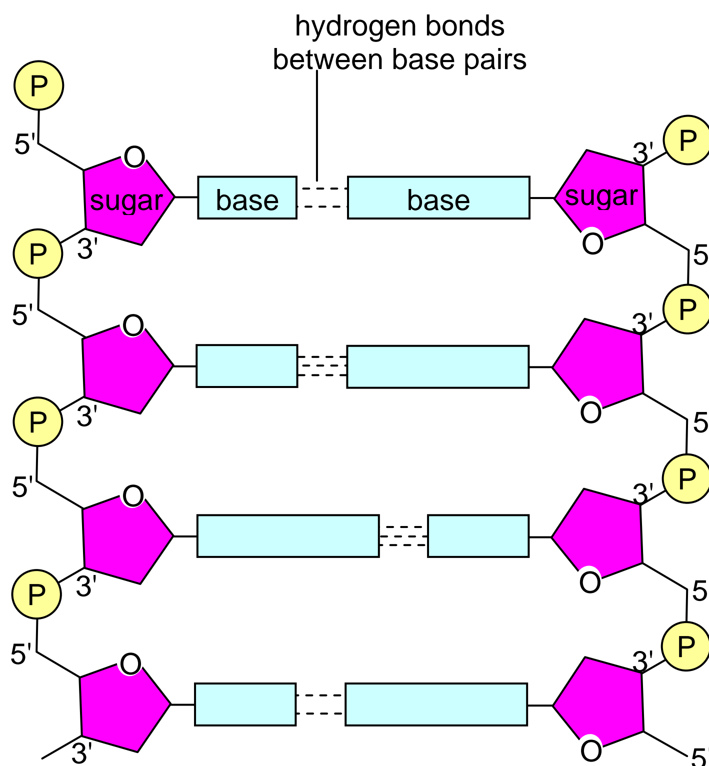


Figure 1.35 – the outline structure of DNA

The structure of RNA

DNA is not the only nucleic acid present in a cell. In order to express the genetic message Nature has devised a system of 'information transfer' that involves various forms of ribonucleic acid (RNA) as well. The major different forms of RNA in gene expression are:

- messenger RNA (mRNA);
- ribosomal RNA (rRNA);
- transfer RNA (tRNA).

The different forms of RNA are also polynucleotides like DNA but there are significant differences (Table 1.3).

Table 1.3: Comparison of the structures of DNA and RNA

factor	deoxyribonucleic acid, DNA	ribonucleic acid, RNA
sugar	pentose sugar present is deoxyribose	pentose sugar present is ribose
bases	• adenine • cytosine • guanine • <u>thymine</u>	• adenine • cytosine • guanine • <u>uracil</u>
structure	a double helix made of two anti-parallel strands	single stranded, though the chain can fold on itself to form helical loops

The first major difference is that the sugar component in RNA is ribose rather than deoxyribose. The second major difference is that the base, uracil (U), replaces thymine in the set of four bases used to build the polymer. Like thymine, uracil is a single-ring structure and can form a complementary base pair with adenine.

© University of Cambridge International Examinations 2006

www.theallpapers.com

The third major difference is that RNA molecules are single-stranded rather than double stranded. However, although an RNA molecule is a single long chain it can bend back on itself to form hairpin loops. These loops are hydrogen-bonded and are important features of the structure of ribosomal RNA (rRNA) and transfer RNA (tRNA) molecules. For example the cloverleaf folding of tRNA molecules (Figure 1.36) enables them to carry out their important function in protein synthesis.

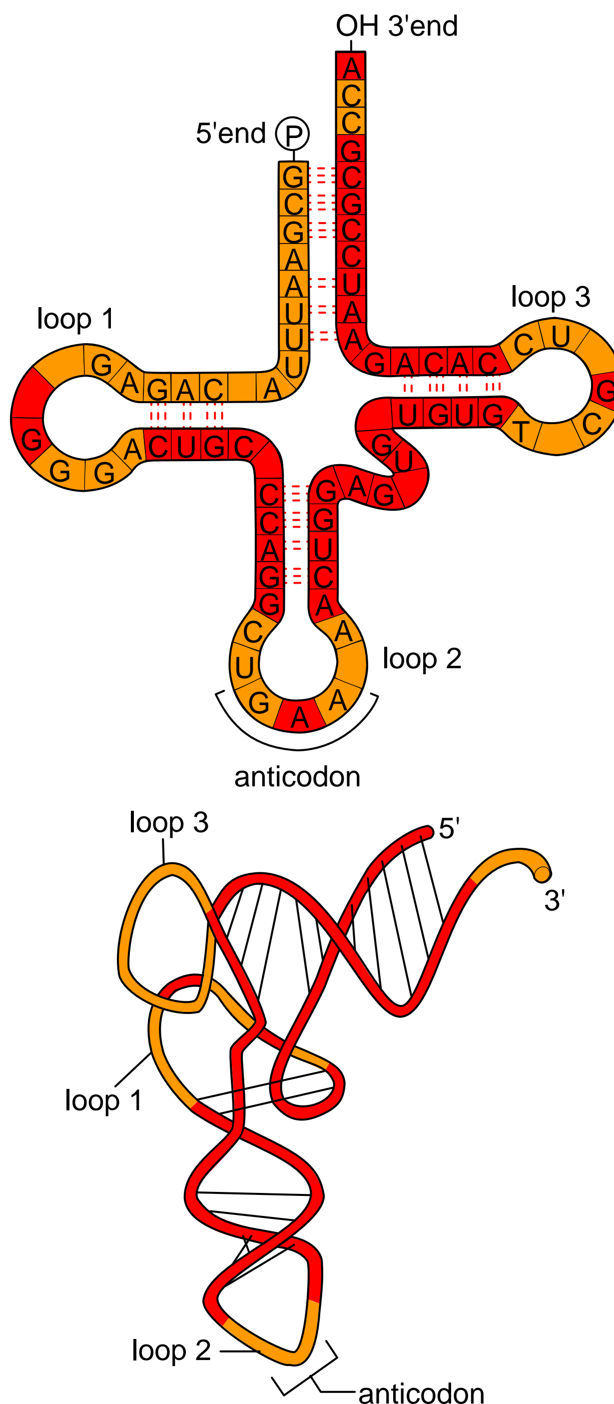


Figure 1.36 – the cloverleaf structure of tRNA molecules

Gene expression

The role of DNA - chains of information

The astounding and compelling ‘neatness’ of the discovered structure of DNA was that it contained a built-in mechanism for information transfer. DNA strands are capable of self-recognition and self-

© University of Cambridge International Examinations 2006

www.theallpapers.com

replication. Duplication of the genetic information can take place every time a cell divides. This means that information encoded in the structure can be passed from one generation to another. Parent DNA molecules can produce identical daughter copies in a process known as replication.

The major interest of DNA, however, is not as the material in itself, but as the blueprint for the synthesis of proteins – the enzymes, antibodies, and structural proteins which determine the nature and function of an organism. Broadly speaking the amino acid sequence of each polypeptide chain is encoded in a specific stretch of DNA, or gene. The 'message' coded in a gene generates copies of a particular polypeptide chain through a two-stage process.

- **Transcription:** The DNA template is first copied, or transcribed, into an intermediary nucleic acid molecule, messenger ribonucleic acid (mRNA).
- **Translation:** mRNA molecules copied from the gene programme the assembly of the polypeptide chain. The translation process involves ribosomes attaching to, and moving along, the mRNA as the chain is synthesised.

Thus DNA, by these two processes of transcription and translation is ultimately responsible for the nature of all the proteins synthesised by cells.

The double helix of DNA controls heredity on the molecular level. DNA both preserves the genetic information (replication), and uses it to direct the synthesis of proteins (transcription and translation). The overall flow of genetic information between generations, and its expression by the cells of an organism, are controlled by the processes summarised in Figure 1.37.

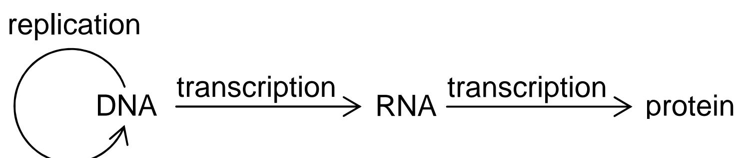


Figure 1.37 – summary diagram of replication, transcription, & translation

This scheme is sometimes referred to as the 'central dogma of molecular biology'. It indicates the predominant direction for the expression of genetic information. The scheme applies across the evolutionary spectrum of organisms from simple bacterial cells to complex animal cells.

Replication – the biological assembly of new DNA

The process of formation of new DNA strands is catalysed by the enzyme DNA polymerase. The monomer units are fed into the reaction process in the triphosphate form (Figure 1.38). The breakdown of the triphosphates into the monophosphate form is exothermic. The energy released by this breakdown 'drives' the addition of the next nucleotide unit to the growing DNA copy.

© University of Cambridge International Examinations 2006

www.theallpapers.com

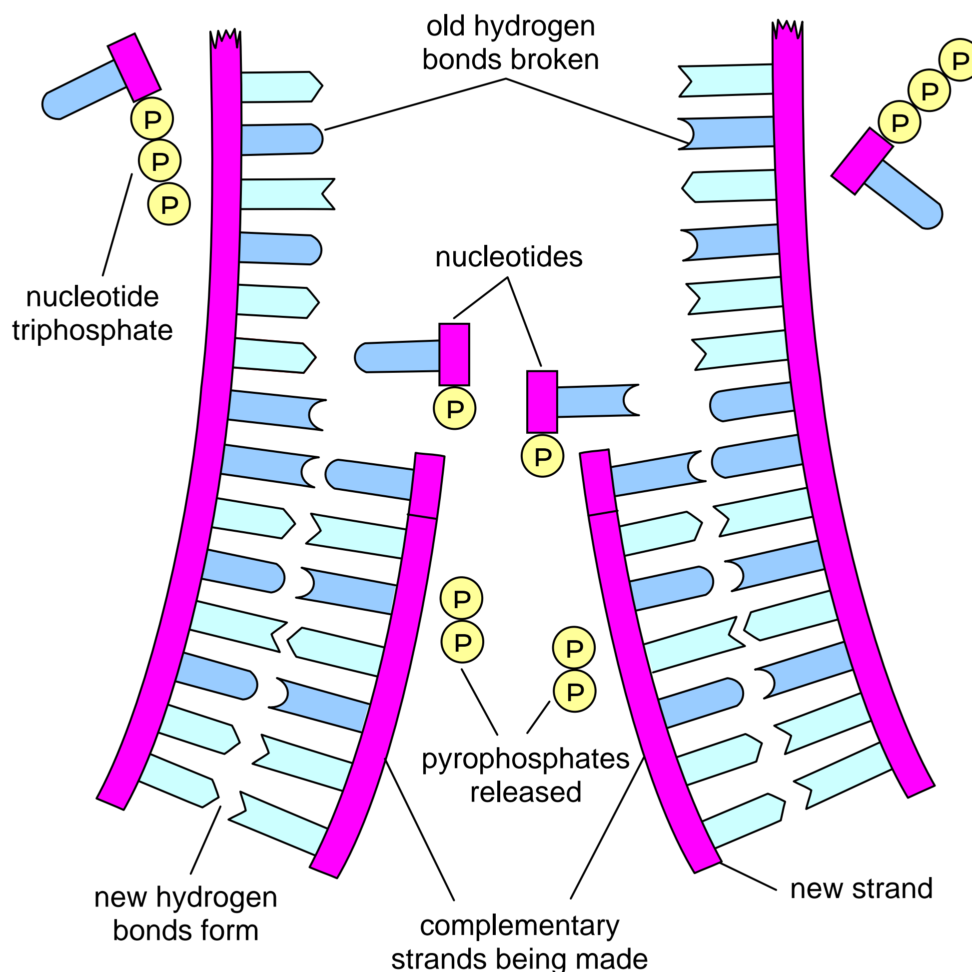


Figure 1.38 – strand of DNA acting as template for replication

Semi-conservative replication

During replication, the hydrogen bonds and van der Waals' forces between the base pairs in the double helix are broken. Two new strands are formed using the original strands as templates for the synthesis. Each new strand contains a complementary sequence of bases as dictated by the order of the bases in the original strand. Hydrogen bonds and Van der Waals' interactions then form between the original and new strands creating a stable helical structure. Thus two daughter molecules are formed from the parent double helix (Figure 1.38). This form of replication is known as semi-conservative replication because each daughter molecule contains one new strand and one original strand.

© University of Cambridge International Examinations 2006

www.theallpapers.com

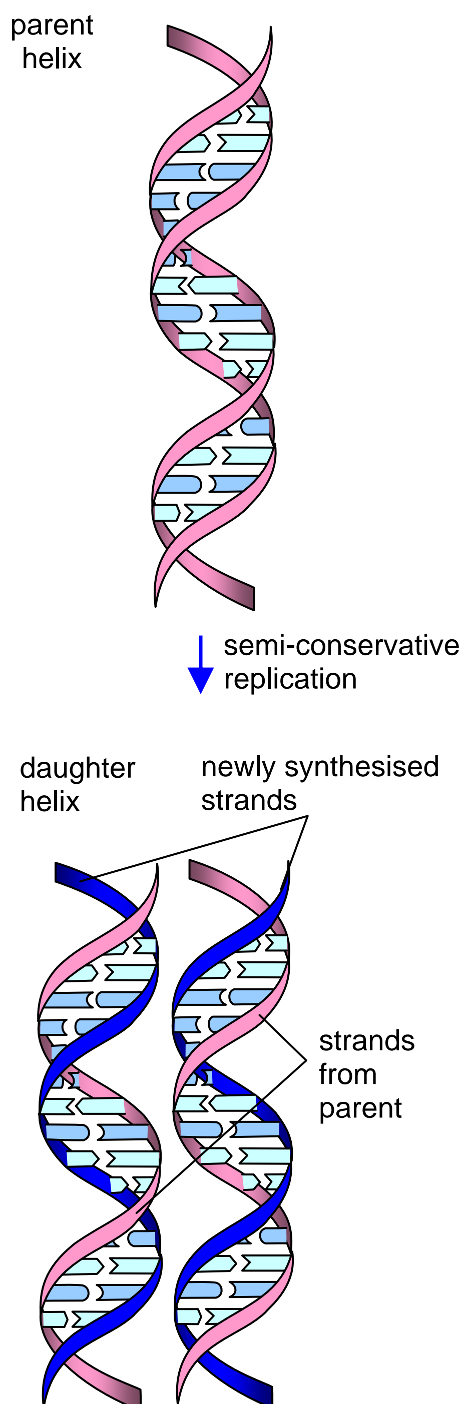


Figure 1.39 – semi-conservative replication

Since the nucleotides in DNA differ only in the bases they carry, the sequence of the monomers in a DNA strand can be represented simply as the base sequence. Each nucleotide can be thought of as a single letter in an alphabet that has only four letters, A, T, G, and C. Different genes have different sequences of these four nucleotides and so code for different polypeptide chains. Gene sequences are always written in the 5'→ 3'direction, for example,

5'-ATGCCGTTAGACCGT _ _ _ _ _ _ _ _ GT-3'.

The DNA in almost every cell in our bodies (the exceptions are certain white blood cells and sex cells) should be an identical copy of the DNA in the fertilised egg.

© University of Cambridge International Examinations 2006

www.theallpapers.com

Expressing the message – the role of RNA

The genetic message encoded in the DNA of cells is used to form protein molecules through the processes of transcription and translation. Gene expression through these two processes involves several different ribonucleic acid (RNA) molecules. The most significant of these is the intermediary molecule, messenger RNA (mRNA) – the genetic messenger.

Each of the three main types of RNA has a different role within the complex mechanism of translation:

- Ribosomal RNA (rRNA) – there are a number of different rRNA's that form part of the structure of the ribosomes. Ribosomes are the small organelles where protein synthesis takes place within the cell. Ribosomal RNA makes up to 80% of the RNA within a cell, and the larger molecules contain over 3500 nucleotides.
- Transfer RNA (tRNA) – there are a group of small RNA molecules, each one specific for a particular amino acid. Their role is to 'carry' the amino acids to the ribosomes for protein synthesis. Each tRNA recognises the coding sequence for a particular amino acid in the messenger RNA. tRNA molecules are about 75 nucleotides long and represent up 15% of cellular RNA.
- Messenger RNA (mRNA) – the RNA copied from the DNA gene sequence for a particular polypeptide chain. The 'message' encoded in the mRNA molecule is translated into the primary sequence of a polypeptide chain.

Delivering the message – transcription

Each gene contains a unique sequence of the four nucleotide bases and codes for a particular protein chain. The gene sequence is always written in the 5'→ 3'direction. However, the gene for a particular polypeptide chain is not copied directly into an amino acid sequence. First the code is transcribed into the mRNA for the protein chain by the enzyme RNA polymerase. Part of the DNA double helix unravels and an RNA copy of the gene is synthesised using the appropriate nucleotides (Figure 1.40). The mRNA molecule is synthesised from the 5'end to the 3'end. This is also the direction in which the 'message' will subsequently be translated on the ribosomes.

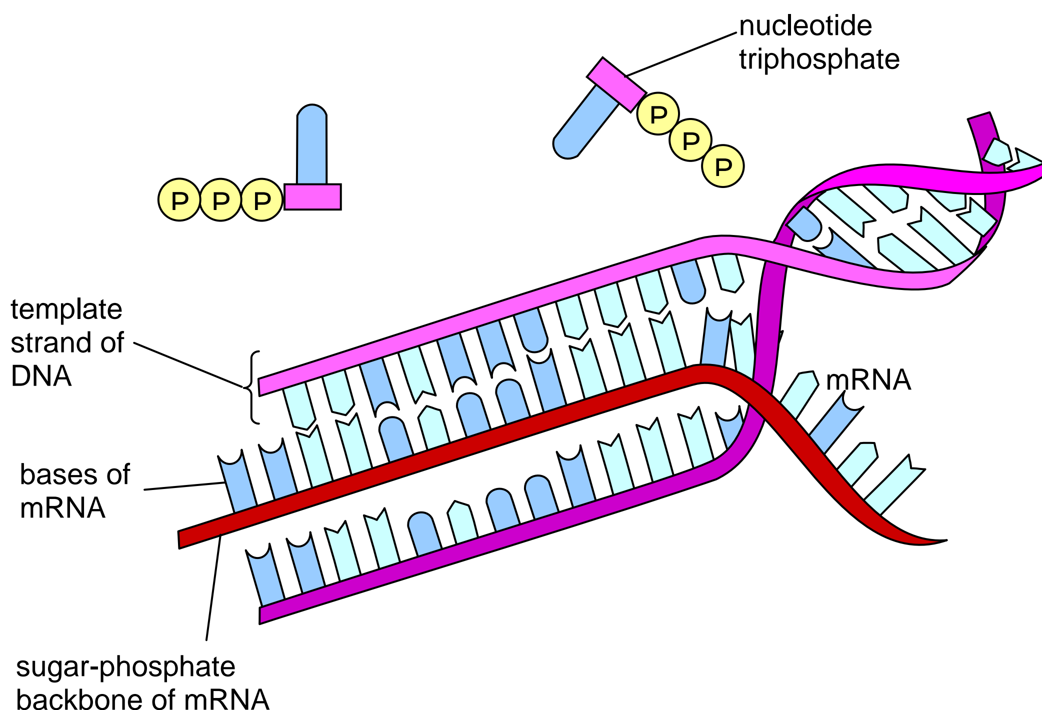


Figure 1.40 – diagram of the transcription process

© University of Cambridge International Examinations 2006

www.theallpapers.com

Protein synthesis – translating the message

The ribosomes are the cellular 'machines' that synthesise protein chains. They are involved in the final stage of the fascinating sequence of events by which the genetically encoded message is expressed in the range of proteins made in a cell. During translation several ribosomes can attach to a particular mRNA molecule at any one time. As the ribosomes move along the mRNA the sequence of bases directs the bringing together of amino acids in the correct order to produce proteins. The language of the transcribed RNA, the order of the bases along the mRNA, is translated into the language of proteins, an order of amino acids along a polypeptide chain.

The genetic code

DNA and mRNA molecules each contain just four nitrogen-containing bases, but there are 20 amino acids used in making proteins. These numbers suggest that the unique sequence of amino acids found in a specific polypeptide chain must be encoded by groups of bases.

If the bases were taken two at a time – AA, AC, AG, AT, CA, CC, etc.- only 16 combinations (4^2) would be possible. A three-base (or triplet) code would provide 64 possible combinations (4^3). This would allow coding for 64 different amino acids if all the codes were unique – each triplet coding for just one amino acid. Biochemical and genetic evidence established that the coded information in mRNA was in the form of a comma-less, non-overlapping triplet code. The direction of readout of the message was found to be from the 5'end of the mRNA to the 3'end. Each triplet of bases is known as a codon.

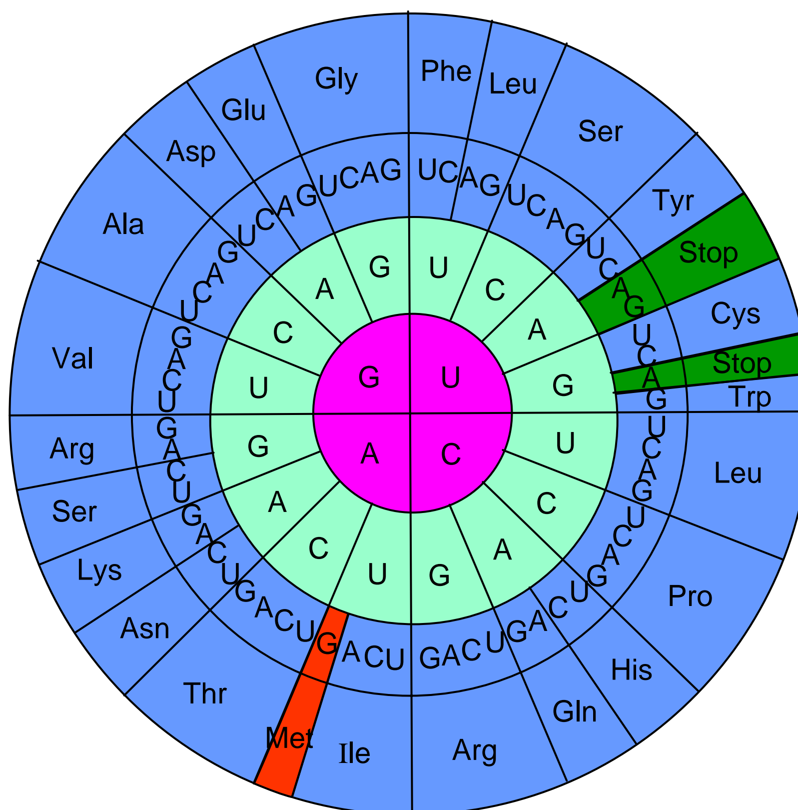


Figure 1.41 – the genetic code

The complete genetic code is shown in Figure 1.41. Most amino acids are coded for by more than one triplet codon. Indeed, some amino acids (e.g. arginine) have up to six possible codons. For all amino acids except methionine and tryptophan, more than one base is allowed in the third position of the combination. This arrangement offers some protection from mutations since a base change in the third position of a codon will often still mean that it still codes for the same amino acid.

All polypeptide chains have a defined length and sequence and so there must be a codon for the first amino acid in the chain (the amino-terminal end). This 'START' signal is 5´-AUG-3´, which codes for

© University of Cambridge International Examinations 2006

www.theallpapers.com

methionine. The 'start' signal ensures that the series of triplet codons is read in the correct groups of three. Consequently the first amino acid in any newly-synthesised protein chain is always methionine, though in many cases it is removed after translation is completed. There are also three codons that do not code for any amino acid. These codons act as 'STOP' signals to end the assembly of a polypeptide chain.

Translating the message

Amino acids on their own cannot bind to mRNA. Transfer RNA molecules (tRNA's) act as the vehicles for these interactions. Each tRNA binds a specific amino acid at one end of the molecule. At the other end it has a specific triplet of bases (the anticodon) which can bind to the codon triplet on the mRNA. Each tRNA, carrying its specific amino acid, can interact with the ribosome and the correct codon on the messenger RNA to continue the process of translation.

The translation process is a complex one involving three steps – initiation, elongation, and termination (Figure 1.42). The correct amino acids are incorporated into the chain as dictated by the sequence of codons in the mRNA. When a 'STOP' codon is reached synthesis is complete and the protein chain is released.

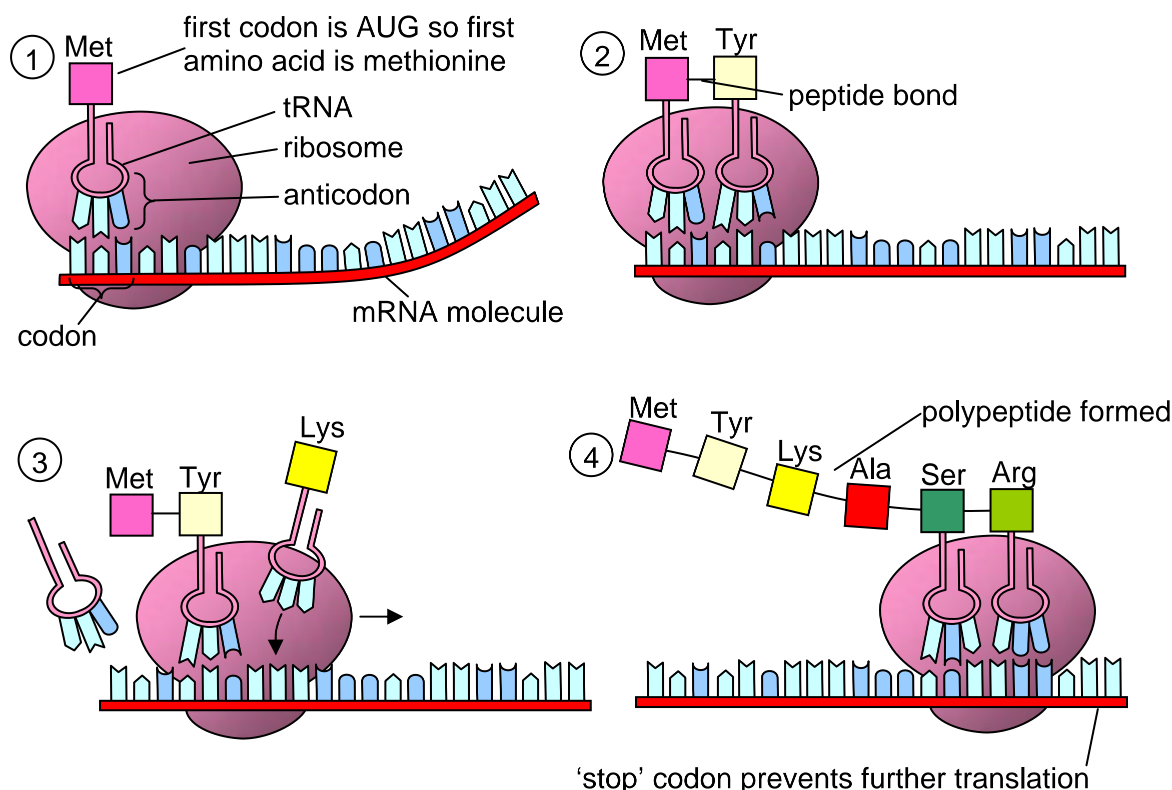


Figure 1.42 – the process of translation

Mutations

Errors can happen during DNA replication. However, it is only on very rare occasions that an error is not corrected by the cell's own mechanisms. Changes from the original DNA are known as mutations. Apart from errors in the replication process, mutations can also be caused by any process that damages DNA. UV light, cigarette smoke and many other chemical compounds can cause mutations.

In some cases the change in the DNA may be very small. For instance, a single base may be miscopied, and a single base pair may then be altered in the DNA molecule in future generations. Such mutations are not uncommon. Since changes in the sequence of base pairs alter the amino acid coding, the end result may be a change in the structure and functioning of a protein.

In many cases a single change in the base sequence has no effect on the protein that is being produced. This is because most amino acids have several codons, and a change in the DNA from a CAA sequence to a CAG, for instance, will still produce a protein containing valine in the correct place.

© University of Cambridge International Examinations 2006

www.theallpapers.com

On the other hand mutations which result in the deletion of a base would alter the way the message is read and produce a different sequence of amino acids in the protein chain. Such mutations together with those that remove a start or stop codon may have serious consequences. A crucial protein may not be produced or may be so changed that it is unable to function properly. Such a situation may result in a genetically based condition such as sickle cell anaemia or cystic fibrosis.

Sickle cell anaemia

Sickle cell anaemia is a condition that affects the red blood cells. The red blood cells of these patients do not have the normal disc shape, but have a crescent moon (or sickle) shape. People with sickle cell anaemia have sickle haemoglobin (HbS) which is different from the normal haemoglobin (HbA).

The disease arises from a single mutation in the DNA of the gene for one of the haemoglobin chains. The result of the mutation is an abnormal amino acid sequence in one of the protein chains in haemoglobin (the β-chain). The abnormality alters a single amino acid at the sixth position of the 146 amino acid chain:

Normal β-chain	Val	His	Leu	Thr	Pro	**Glu**	Glu
Sickle cell β-chain	Val	His	Leu	Thr	Pro	**Val**	Glu

When sickle haemoglobin gives up its oxygen to the tissues, it sticks together to form long rods inside the red blood cells, making these cells rigid and sickle-shaped. Normal red blood cells can bend and flex easily. However, because of their shape, sickle-shaped red blood cells cannot squeeze through small blood vessels as easily as the almost doughnut-shaped normal cells. This can lead to these small blood vessels getting blocked, stopping oxygen from getting through to where it is needed. This can then lead to severe pain and damage to organs in the body.

The different kinds of sickle cell anaemia and the different traits are found mainly in people whose families come from Africa, the Caribbean, the Eastern Mediterranean, Middle East and Asia.

Cystic fibrosis

Cystic fibrosis is a relatively common genetic disorder. It occurs in 1 in 2000 live births, and 1 in 22 Caucasians are carriers of the gene. The condition affects the lungs, pancreas, gut and sweat glands. Instead of the normal fluid secretions a thick sticky mucous forms. This viscous mucous blocks and damages the intestines and lungs. Because the supply of digestive enzymes from the pancreas is blocked, nutrients cannot be absorbed and babies fail to thrive. These babies have repeated chest infections and in particular can get intestinal obstruction. The malfunctioning of the sweat glands results in abnormally salty sweat, which is used in the diagnosis of the condition.

Thanks to better understanding of the disease and its treatment, people with cystic fibrosis are living longer than ever before. Until the 1930s, the life expectancy of a baby with cystic fibrosis was only a few months. Today the average life expectancy for someone with cystic fibrosis is around 31 years. There is no cure for cystic fibrosis, but the faulty gene has been identified and doctors and scientists are working to find ways of repairing or replacing it.

Cystic fibrosis affects the cells that line the cavities and tubes inside organs such as the lungs. The membranes of these cells have a mechanism for pumping chloride ions into the cells from the blood supply. In lungs the chloride ions normally diffuse out of the cells through channels in the cell membrane lining the airways. This is part of the process for keeping a runny layer of watery mucus on the surface of the cells.

The chloride ions diffuse out of the cell through a channel created by a protein. The protein channel is only open in the presence of ATP. The name of this membrane protein is CFTR protein (short for cystic fibrosis transmembrane regulatory protein). In a person with cystic fibrosis the CFTR protein may be missing or, if present, it does not work properly. It does not allow chloride ions which are being pumped into the cell to leave (Figure 1.43). The chloride ion concentration in the cell builds up. The high solute concentration in the cell causes water to move into the cell instead of out of it by osmosis. As a result the mucus covering the cells lining the airways becomes thick and sticky.

© University of Cambridge International Examinations 2006

www.theallpapers.com

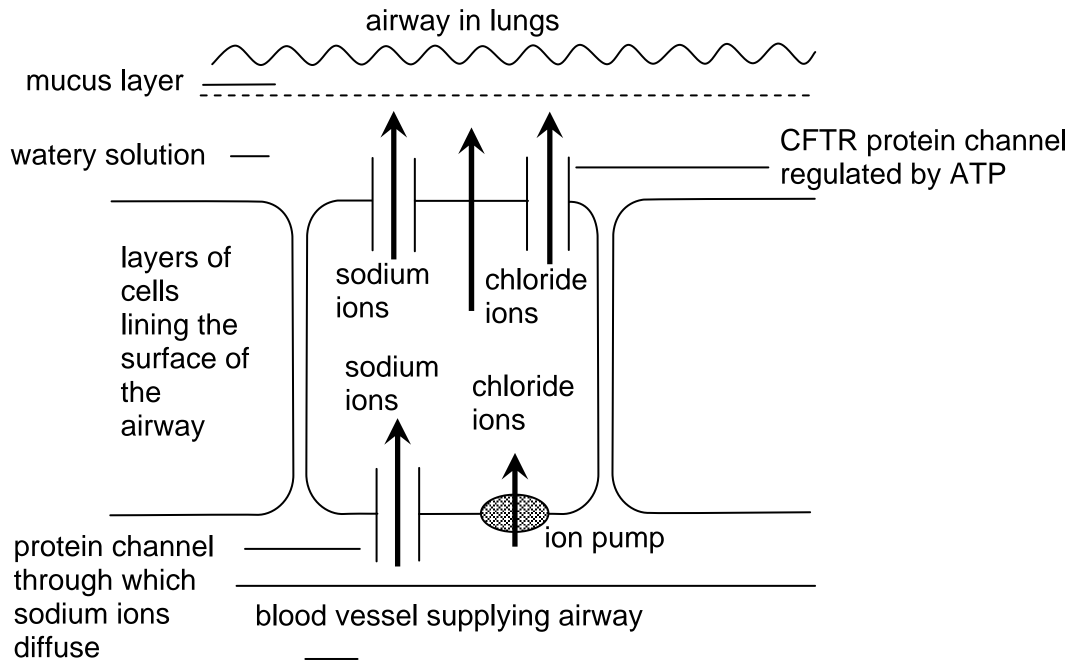


Figure 1.43 – the movement of ions across cell membranes in the lungs

The genetics of cystic fibrosis is not as simple as that of sickle cell anaemia. Hundreds of different mutations have been identified that can give rise to the disease. The various mutations affect the CFTR protein in different ways. In some cases ATP is unable to bind to it so the channel cannot open. In other cases the channel opens but in a way that does not let the chloride ions escape. The gene involved is on chromosome 7. The commonest mutation is the deletion of three nucleotides which result in the loss of phenylalanine, the 508th amino acid in the structure of the protein (Figure 1.44).

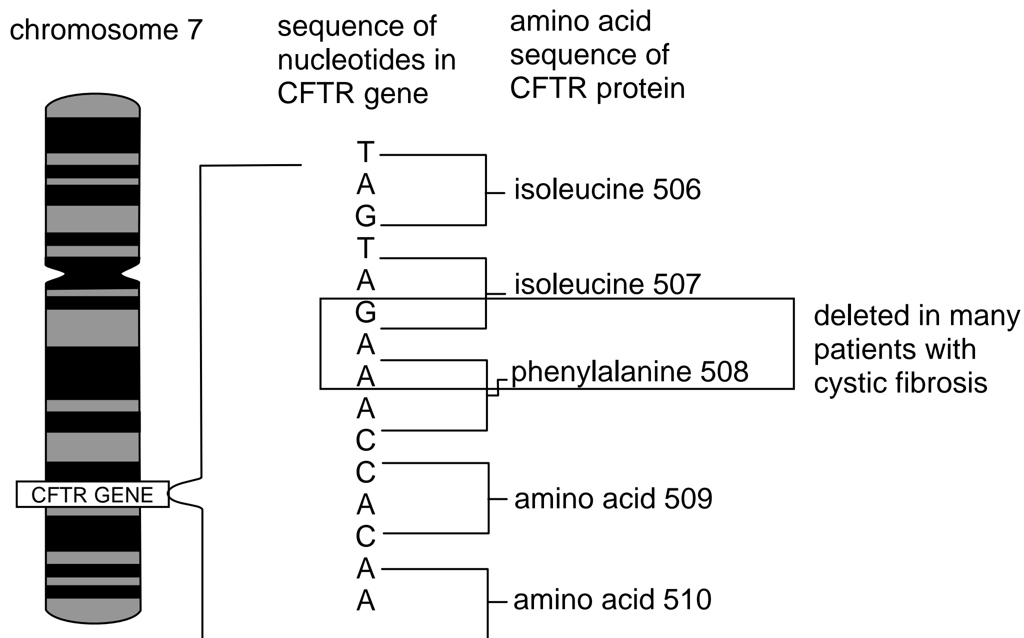


Figure 1.44 - the site of the commonest mutation that causes cystic fibrosis

Improving nature – re-designing enzymes

Despite the staggering complexity of biological enzyme systems it is intriguing to note that evolution has not necessarily developed particular enzymes to perfection. Indeed there was no need for that.

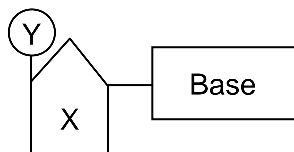
© University of Cambridge International Examinations 2006

www.theallpapers.com

Having reached a certain level of efficiency in the natural world, there was no reason to develop the structure of a particular enzyme any further. A research team in Cambridge led by Sir Alan Fersht has been able to synthesise an 'improved' enzyme by modifying its gene.

The scientists experimented with key regions of the gene for the enzyme and so synthesised modified versions of the enzyme. By this genetic manipulation they were able to optimise key areas of the 3-dimensional shape of the enzyme; significantly improving the efficiency of the enzyme. Such studies help to develop our understanding of the mechanisms involved in the folding of proteins, a key area in the progress of computer-aided drug design.

SAQ 8. *The diagram below represents the basic chemical unit from which the nucleic acid DNA is formed.*



(a) *State the name of:*

(i) *the whole chemical unit shown;*

(ii) *the component labelled X;*

(iii) *the component labelled Y.*

(b) *Name the four nitrogen-containing base present in DNA.*

SAQ 9. ***(a)*** *Representing the nitrogen-containing bases by* ***B****, sugars by* ***S*** *and the phosphate groups by* ***P****, and using no other symbols, draw a diagram to show how these are linked in a short length of double-stranded DNA.*

Use full lines (_________) for normal covalent bonds
and dotted lines (- - - - - -) for hydrogen bonds.

(b) *Your sketch makes the two strands look identical. Ignoring the difference between the bases explain:*

(i) *how the two strands differ;*

(ii) *give the technical term which describes this difference;*

(iii) *state how it is indicated on diagrams of DNA.*

(c) *An analysis of the bases in a sample of double-stranded DNA gave the partial result below.*

Adenine 23 mole %

Guanine 27 mole %

What would you expect the rest of the analysis to show? Explain your answer.

SAQ 10.(a) *What role do hydrogen bonds play in the accurate replication of DNA?*

(b) *DNA is replicated* ***semi-conservatively****. What is meant by the term in* ***bold*** *type?*

SAQ 11.(a) *State three ways in which the structure of DNA differs from that of RNA.*

(b) *Outline the role of the several kinds of RNA in the biosynthesis of protein.*

(c) ***(i)*** *The peptide fragment*

-Tyr-Ser-Ala-Ala-Glu-Gly-Ala-Val-

is known to be coded somewhere inside the fragment of mRNA below. ***The start of this fragment may not coincide with the start of a codon.***

5'-G U U A C U C U G C U G C U G A A G G A G C U G U A C-3'

© University of Cambridge International Examinations 2006

www.theallpapers.com

Use the above information to work out the codons for alanine and tyrosine.

(ii) *Give the base sequence matching the codon for tyrosine in the DNA from which the RNA was transcribed, indicating the direction of the bases in the DNA strand.*

Summary

- There are two forms of nucleic acid: one is DNA (deoxyribonucleic acid), in which the sugar is deoxyribose; the other is RNA (ribonucleic acid), where the sugar is ribose.
- Nucleotides are the monomers from which the nucleic acids are built. The nucleotides themselves are made from a sugar, a nitrogen-containing base and a phosphate group.
- Both forms of nucleic acid are linear condensation polymers made up of a sugar-phosphate backbone to the chain. The nitrogen-containing bases are attached to the sugars in this chain.
- DNA has a double-stranded structure. The strands are arranged so that the sugar-phosphate backbones are on the outside of the structure, with the bases pointing inwards towards each other.
- The two chains run in opposite directions to each other - they are anti-parallel - and interact with each other through hydrogen bonds between the bases. The double-stranded DNA molecule is twisted on itself into a right-handed double helix. Van der Waals' forces also contribute to the forces holding the two strands together.
- There are four bases in DNA: adenine (A), guanine (G), thymine (T) and cytosine (C). The hydrogen bonding between these bases is quite specific and means that an adenine in one chain always pairs with a thymine in the other strand, while guanine always pairs with cytosine. This type of pairing is known as complementary base pairing.
- RNA differs from DNA in three major ways. As well as the difference in the sugar unit, RNA molecules contain the base uracil rather than thymine and they are single-stranded polymers.
- Complementary base pairing is the molecular basis for the process of replication - the production of identical copies of the genetic code from one generation of cells to the next. The process of replication is semi-conservative: after cell division the new DNA molecules consist of one parent strand and one daughter strand.
- DNA carries the genetic code for the production of proteins.
- Using the genetic code to synthesise proteins involves several types of RNA: messenger RNA (mRNA), transfer RNA (tRNA) and ribosomal RNA (rRNA).
- The processes involved in synthesising proteins based on the genetic code in DNA are known as transcription and translation. Each gene codes for a single polypeptide chain. Transcription involves the copying of the gene into an mRNA molecule. This molecule is then translated into the polypeptide chain in a complex process involving the ribosomes of a cell.
- mRNA molecules contain a triplet code in which three bases in the RNA sequence (a codon) code for one amino acid in the polypeptide chain.
- tRNA brings specific amino acids to the ribosomes and binds to specific codons in the mRNA molecule. Protein synthesis then takes place in three stages: initiation, elongation and termination.

© University of Cambridge International Examinations 2006

www.theallpapers.com

1.4 – ATP, Life's energy currency

By the end of this section you should be able to:

- *outline in terms of the hydrolysis of ATP to ADP + P_i, the provision of energy for the cell.*

The reactions that sustain life and its functions in cells form complex sequences of reactions that are linked in pathways. The sequences of reactions together make up the metabolism of the cells. Each reaction in a pathway is catalysed by a particular enzyme, with the product of one reaction becoming the substrate for the next. The steps in these metabolic pathways tend to require relatively small, manageable amounts of energy which need to be readily accessible to continuously 'drive' the activities of the cell.

The nucleotide, adenosine triphosphate (ATP), has a crucial role to play in making energy available for metabolic reactions in all living organisms. ATP is the short term energy source for cellular activity. In animal cells this nucleotide is synthesised in the mitochondria of the cell. The structure of this nucleotide is shown in Figure 1.45.

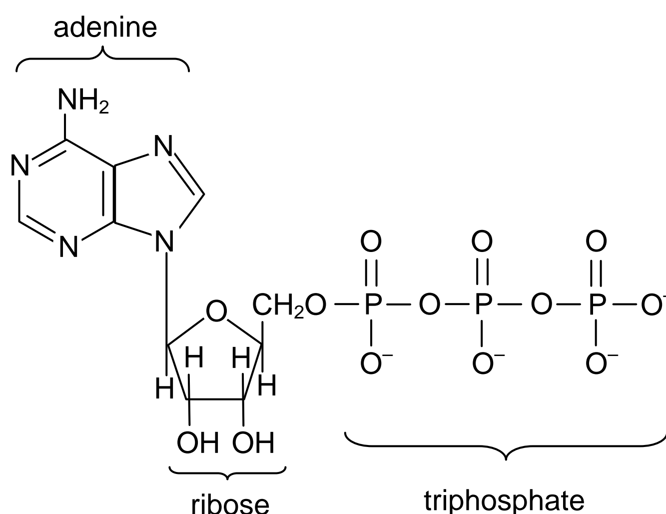
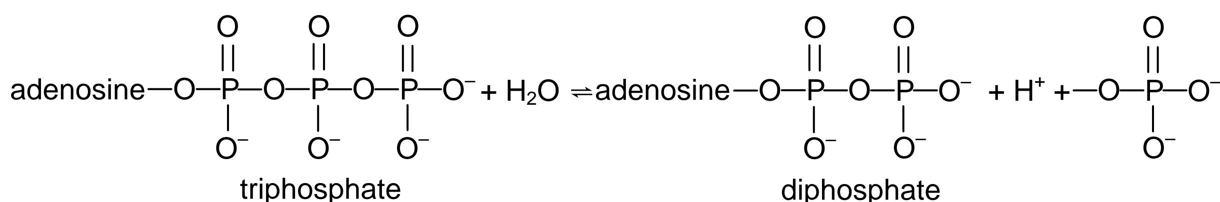


Figure 1.45 – the structure of ATP

The molecule consists of three phosphate groups linked in a short chain and covalently bonded to the hydroxyl group of a sugar, ribose. The last part of the molecule is adenine, an organic base.

The breakdown of ATP is an exothermic reaction and this released energy is used by enzymes to power the catalysed reactions. ATP is hydrolysed to ADP (adenosine diphosphate) and an inorganic phosphate ion (often represented as P_i) in a reaction that is energetically favourable. Energy is required to break bonds between phosphate groups, and those in water, but there is a net gain of energy when the products are formed. The release of the end phosphate group is favoured by the repulsion between the negatively charged O atoms on the adjacent phosphate groups.

$$\text{ATP} + H_2O \rightleftharpoons \text{ADP} + P_i \qquad \Delta H = -30\ \text{kJ mol}^{-1}$$



In many chemical transformations and synthetic reactions involving enzymes, ATP plays an essential role. The proteins involved in ion channels across membranes also require ATP for their function.

Metabolic pathways use large amounts of ATP all the time, so its synthesis in the mitochondria of cells is an important process. Ultimately each ATP-dependent process produces ADP, and a key step in metabolism is

© University of Cambridge International Examinations 2006

www.theallpapers.com

the regeneration of ATP from this. This reaction is of course energetically endothermic and this is where the energy available from the oxidation of food is needed.

Plants can make ATP from ADP and phosphate ions using energy from sunlight (photosynthesis); but for animals the energy require must come from the oxidation of food. Energy-rich molecules in our diet, such as carbohydrates or fats, are metabolised by a series of oxidation reactions ultimately producing carbon dioxide and water. Respiration is a metabolic process in cells that oxidizes glucose and produces ATP to drive varies cell and tissue activities (Figure 1.46).

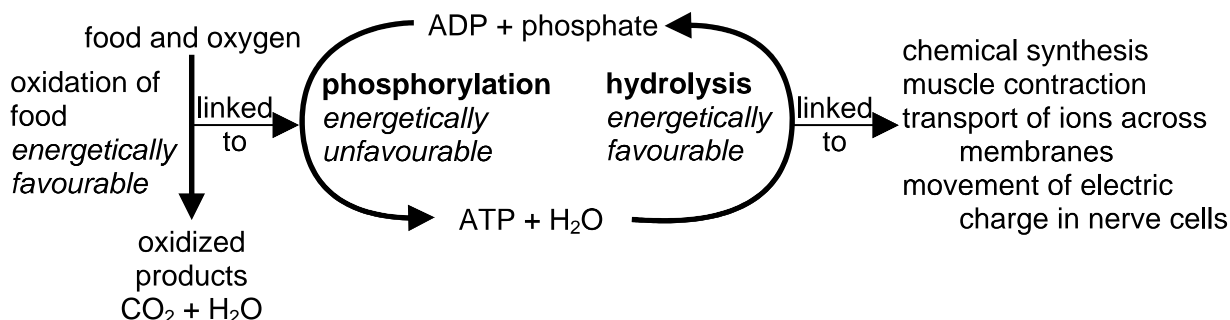


Figure 1.46 – the role of ATP in metabolism

The oxidation of one mole of glucose produces 38 moles of ATP. The breakdown products of proteins (amino acids) and fats (long-chain carboxylic or fatty acids) in our diet can also 'feed' into the respiration pathway at different stages. In this way all the three major components of our food can produce energy for our cells and tissues.

If cellular metabolic activity is to be controlled then spontaneous reactions must not occur in the absence of enzymes. Although the hydrolysis of ATP to ADP is energetically favourable, the activation energy for the reaction is high. This means that spontaneous hydrolysis of ATP without the presence of an enzyme does not occur. Consequently all the ATP produced during the oxidation of food is available for controlled cellular processes.

SAQ 12.(a) *What type of compound is ATP?*

(b) *List the different components of the ATP molecule.*

(c) *What feature of the structure of the molecule favours its hydrolysis to ADP?*

(d) *Draw the structure of ADP.*

SAQ 13.(a) *Where in the cell is ATP synthesised?*

(b) *What is the name of the metabolic process in which the oxidation of glucose produces ATP?*

Summary

- The chemical reactions involved in cell and tissue functions are linked in metabolic pathways of enzyme catalysed reactions.
- The nucleotide, adenosine triphosphate (ATP), is the short term energy source for these metabolic reactions.
- The energy required by metabolic reactions is released in the hydrolysis of ATP to ADP and inorganic phosphate (P_i).
- ATP is synthesised in the mitochondria of cells.
- The synthesis of ATP from ADP requires energy and in animals this is produced by the oxidation of food.
- The metabolic process by which ATP is produced from glucose is known as respiration.

© University of Cambridge International Examinations 2006

www.theallpapers.com

1.5 – Metals in biological systems

By the end of this section you should be able to:

- *understand why some metals are essential to life and, given information and using previous syllabus content, be able to explain the chemistry involved.*
- *recognise that some metals are toxic and discuss, in chemical terms, the problems associated with heavy metals in the environment entering the food chain – for example, mercury.*

The key elements involved in the 'chemistry of life' are predominantly non-metals. Carbon, hydrogen, oxygen, nitrogen and phosphorus are the main elements from which molecules such as proteins, nucleic acids, fats and carbohydrates are built. However, some metals do have vital roles to play in the effective functioning of biological structures and processes. Some metals are naturally found in the body and are essential to health. Iron, for example, is present as an essential component of a range of haem proteins, while zinc is a co-factor in over 100 enzyme reactions. Because these metals normally occur at low concentrations in the body they are known as trace metals. Indeed, at high levels these metals may be toxic or produce deficiencies in other trace metals.

Iron and the haem proteins

We have already referred to the most well-known of the iron-containing proteins, namely, haemoglobin. This protein is the oxygen-carrying protein present in red blood cells. Each of the four protein chains in human haemoglobin is bound to a non-protein haem group that contains an iron(II) ion, Fe^{2+}. It is the Fe^{2+} ions that bind oxygen to the haemoglobin. Each haem group can bind one oxygen molecule, and each of the four haem groups binds oxygen simultaneously, so the overall reaction is:

$$\underset{\text{haemoglobin}}{Hb} + \underset{\text{oxygen}}{4O_2} \rightleftharpoons \underset{\text{oxyhaemoglobin}}{HbO_8}$$

In each case the Fe^{2+} ions act as the centres of complex ions; the ligands being the haem group, the protein chain, and the attached oxygen molecule (see Figure 1.47). The haem group binds the Fe^{2+} ion using four N atoms, and the protein chain also binds through a nitrogen atom. The binding of the oxygen is reversible so that it can be 'delivered' to the tissues of the body where it is required.

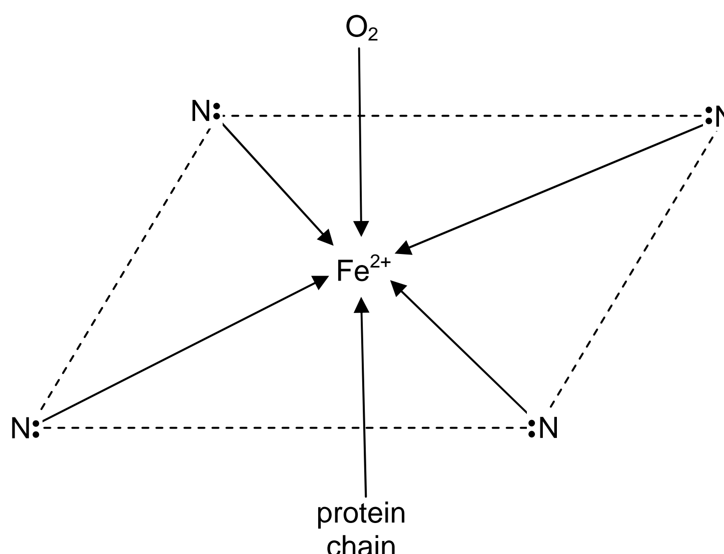


Figure 1.47 – the complex ion in haemoglobin

As in other complexes the oxygen ligand can be replaced by another ligand that binds more strongly. Thus, in carbon monoxide poisoning the CO molecules replace the oxygen in oxyhaemoglobin. This binding is 200 times stronger and irreversible, resulting in the haemoglobin molecule losing its function.

© University of Cambridge International Examinations 2006

www.theallpapers.com

Haem is also involved in the functioning of other proteins such as the cytochromes present in the mitochondria. These proteins are responsible for the production of ATP in the final stage of the respiration process. Redox processes involving electron transfer are crucial at this stage and the ability of the Fe^{2+} ions to form Fe^{3+} ions relatively easily, and vice versa, is important here.

Zinc as an enzyme cofactor

Carbonic anhydrase, as we have seen earlier, is one of the most efficient enzymes in the human body. Present in our red blood cells, it is responsible for the removal of carbon dioxide from the blood, producing hydrogen carbonate ions. Key to the activity of this enzyme is the zinc ion (Zn^{2+}) present in the active site of the enzyme. The zinc is bound to the enzyme as part of a complex using nitrogen atoms on the protein chain.

Water is also bound to the zinc ion. Since the zinc ion has a high charge density it assists the breakdown of this water molecule into an H^+ and an OH^- ion. The hydroxide ion is then in a position to attack the carbon dioxide molecule. The product of this nucleophilic attack is the hydrogen carbonate ion which is released from the active site.

$$CO_2 \quad + \quad OH^- \quad \rightarrow \quad HCO_3^-$$

Following release of the hydrogen carbonate ion a further water molecule binds to the zinc and the catalytic cycle begins again.

Sodium and potassium ion transfer across cell membranes

The ionic composition within living cells is different from that of their surroundings. Within cells the Na^+ ion concentration is lower, and the K^+ ion concentration higher, than the surrounding liquid outside. How this difference is achieved is important for all cells, but particularly for nerve cells.

When a nerve is stimulated sodium ions pour into the nerve cell. When this 'signal' has passed the Na^+ and K^+ ion concentrations have to be restored to normal by the sodium being transported out of the cell once again. The energy to 'drive' this transport has been shown to come from the hydrolysis of ATP. Research into this phenomenon led to the discovery of the ion-transporting enzyme, Na^+, K^+-ATPase (sodium, potassium – adenosine triphosphatase). Because of its function and its dependence on energy from ATP, this enzyme is often referred to as the 'sodium-potassium pump'.

These enzyme molecules are located in the cell membrane. They sit across the membrane with parts of the protein exposed on the outer and inner surfaces (they are trans-membrane proteins). Studies on the mechanism of transport have shown that initially three Na^+ ions and an ATP molecule bind to the inner protein surface of the enzyme. The ATP is then hydrolysed, with the Pi binding to the protein. The enzyme changes shape so that the Na^+ ions move to the outside surface. Here they are released and two K^+ ions attach to the protein instead. The release of the phosphate group from the enzyme results in the K^+ ions moving into the cell. When a new ATP molecule attaches to the enzyme, the K^+ ions are released inside the cell and the cycle of transport can begin again. Thus the ATP-driven sodium, potassium pump restores the concentrations of K^+ and Na^+ to their normal levels following a nerve impulse.

The maintenance of ion balance in cells, and the generation and transmission of electrical impulses, does not solely depend on ATP-dependent ion pumps. There are also specific water and ion channels that have been identified in cell membranes. These are also protein structures but the energy required and their selectivity is dependent on the hydration and size of the ions concerned. The potassium specific channel has been worked on in detail – and the explanation found as to why K^+ ions, and not the smaller Na^+ ions are allowed through the channel. The key lies in the fact that the aqueous K^+ ions ($K^+(aq)$) must lose their hydration shells before they can pass through the channel. The $K^+(aq)$ ions are stripped of the associated water molecules as they enter the channel, linking instead to oxygen atoms in certain R-groups of the protein.

The enthalpy required to lose the hydration shell around the ions is compensated for by that given out when the new association is formed with the protein. The K^+ ions pass through the channel and then re-associate with water on the other side. The hydration shell is re-formed around the ion and energy is released. The selectivity of the channel depends on the distances between the oxygen atoms in the

© University of Cambridge International Examinations 2006

www.theallpapers.com

protein side-chains and the K^+ ions. The smaller Na^+ ions will not 'fit' the channels as the distances are too great for the complex to form.

Knowledge of how membrane channels and 'pumps' work is of crucial importance in understanding how cells function both in health and disease. Diabetes and other serious diseases of the nervous system, muscles, and heart can be attributed to malfunctioning cellular water and ion channels.

Toxic trace metals in the environment

Certain metals, particularly certain heavy metals, can have serious health effects even though they may occur only in small amounts in the environment. Part of the problem lies in the fact that they can accumulate in the food chain and so build up to toxic levels over a period of time. Lead and mercury are two particular metals for which this has proved a significant problem.

We have seen earlier that such metals can interfere with enzyme function by disrupting the disulphide bridges involved in protein tertiary structure.

Van der Waals' forces between non-polar side-chains also contribute to protein tertiary structure. The presence of certain salts can have a disruptive effect on this type of force too. Ions such as Li^+, Mg^{2+}, and Ca^{2+}, together with heavy metal ions such as lead or mercury, have all been shown to interfere with Van der Waals' interactions. Shape is so dramatically important to protein function that any disruption of the forces holding such molecules in their functional structure can have serious consequences for effective activity or function.

The mechanisms by which heavy metals can accumulate vary from metal to metal but lead and mercury illustrate some aspects of the problems involved. Lead was once used for pipes in many water supply systems and so the metal can be ingested over time in this way. Lead compounds form car exhausts can settle on fruit and vegetables grown near roadsides and so can get into the food chain in this way. Levels from this form of pollution are falling as the use of lead-free petrol increases. Lead can cause mental health problems, particularly in young children.

Mercury contamination is perhaps the most notorious case of heavy metal poisoning and its effects – with the effects of mercury leaking into the environment at Minamata in Japan still remembered. Mercury can enter the food chain by a number of routes:

- in waste water discharged into rivers from factories that use mercury compounds in their processes,
- mercury compounds have been used as fungicides and these can be washed off crops into the soil,
- mercury compounds have been used to treat timber and again they can be washed into rivers and streams,

a mercury cathode cell is one which is used in the large scale production of sodium hydroxide – again any leakage of mercury is dangerous as micro-organisms can convert mercury salts into organomercury compounds e.g. methylmercury salts, and these can be ingested by water-borne organisms. Here they accumulate and are passed through the food chain, via fish, for instance, and finish up in man.

The effects of mercury toxicity are a loss of muscle co-ordination and mental function.

SAQ 14. *Haem is an important prosthetic group which contains iron at its centre.*

(a) *How many haem groups are there associated with a molecule of human haemoglobin?*

(b) *Give two important features of the nature of the attachment of oxygen to each haemoglobin molecule.*

(c) *What forces hold the haem groups in place in the protein?*

(d) *Name another important group of proteins that contain the haem group. Where are these proteins located in the cell?*

© University of Cambridge International Examinations 2006

www.theallpapers.com

SAQ 15. *The enzyme carbonic anhydrase is present in red blood cells and is a metalloenzyme.*

(a) *Which metal is present in this enzyme?*

(b) *Write the equation for the reaction catalysed by the enzyme.*

(c) *The metal ion present is thought to induce the ionisation of a water molecule to produce an hydroxide ion, (OH^- ion). This ion then attacks a carbon dioxide molecule to produce the hydrogen carbonate ion, (HCO_3^-).*

(i) *What type of attack is the OH^- ion taking part in on the CO_2 molecule?*

(ii) *Using appropriate 'curly arrow notation', suggest the mechanism for the production of the HCO_3^- ion.*

SAQ 16.(a) *Which ions are involved in the transmission of the signal that stimulates nerve cells?*

(b) *Name the enzyme involved in restoring the ionic balance of the interior of the cell after stimulation.*

(c) *How is this enzyme orientated with the membrane of the cell?*

(d) *Where is the energy obtained from to drive this transport of ions across the membrane?*

SAQ 17. *Ion-specific channels are important in maintaining the correct ion balance of cells. The K^+-specific channel depends on the hydrated potassium ions losing their hydration shell in order for transport to take place.*

(a) *Is the loss of the 'shells' of water molecules round the ions an exothermic or endothermic process?*

(b) *What interactions of the K^+ ions replace those with water?*

(c) *Why are the smaller Na^+ ions not able to use the same channel?*

SAQ 18.(a) *How do heavy metal ions affect the structural integrity of proteins and impair their function?*

(b) *Outline how traces of a heavy metal such as mercury can accumulate progressively up the food chain.*

Summary

- There are several important non-metal elements that make up the molecules central to the biochemical processes of life. These include carbon, hydrogen, oxygen, nitrogen and phosphorus.
- However, some metals do have key roles in the effective functioning of biological systems. These include the following:
  - iron is important in proteins that function with the help of the haem group; haemoglobin and the cytochromes, for example.
  - zinc is important in the role of certain key enzymes; carbonic anhydrase, for example.
  - the balance of the concentration of sodium and potassium ions within cells is important in maintaining cell structure and the transmission of nerve impulses.
- Certain metals – in particular, heavy metals such as lead and mercury are toxic. They interfere with the tertiary structure of proteins and hence their effective function.
- The effects of these heavy metals are evident at even relatively low concentrations. However, the problems associated with them are made worse by the fact that these metals can be concentrated within the food chain.

Conclusion

The tremendous increase in our understanding of the biochemical basis of life has been developed over the past 60 years or so. The beauty of the DNA story and its universality across the different levels of the evolutionary spectrum of life is of great intellectual interest and excitement. The overall mechanisms involved in the generation of proteins, and the manner in which they function to support the different

© University of Cambridge International Examinations 2006

www.theallpapers.com

processes involved in sustaining life, are intriguing. Although the molecules and the processes are complex, they can be seen to rely on certain basic chemical principles and forces of interaction. The factors that determine and maintain the three-dimensional structures of these complex polymers are absolutely crucial to the correct functioning of life's biochemical reactions.

The continued development of our understanding of these processes will aid our understanding of both health and disease; and will surely throw up significant moral, ethical and intellectual challenges for the future.

1.6 – Revision self-assessment questions

1. *The following table shows the types of interaction and bonds that can stabilize protein tertiary and quaternary structure. Complete the table by filling in spaces* ***(a)*** *to* ***(e).***

amino acid	*side-chain*	*side-chain*	*amino acid*	*bond/interaction between side-chains*
alanine	*__ CH_3*	*$(CH_3)_2$ CH __*	*valine*	***(a)***
serine	*__ CH_2OH*	*NH_2 $COCH_2$ __*	*asparagine*	***(b)***
aspartic acid	*__ $CH_2CO_2^-$*	*$^+NH_3(CH_2)_4$ __*	*lysine*	***(c)***
cysteine	*__ CH_2SH*	*$HSCH_2$ __*	***(e)***	***(d)***

2. *A Bad Hair Day! – changing the style*

Hair is made of an insoluble fibrous protein. The shape of hair can be changed in a way that lasts for some time by changing the disulphide (–S-S-) bridges present. Small sulphur-containing molecules such as thioglycollate can bring this about.

Hairdressers first use rollers to create a new style for the hair. They then apply the thioglycollate solution to break apart the disulphide bonds, producing –SH groups. This allows the protein chains to re-arrange themselves to the new shape of hair. The thioglycollate is thoroughly washed away. The hair is fixed in its new shape ('permed') using a dilute hydrogen peroxide solution which reforms new disulphide bridges.

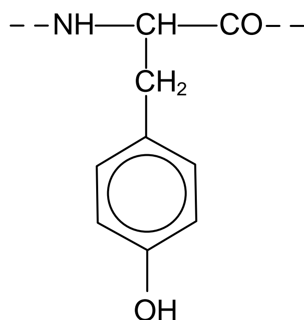
(a) *Which insoluble fibrous protein is hair made of?*

(b) *Are the disulphide bonds mainly responsible for the secondary, tertiary, or quaternary structure of proteins?*

(c) *What types of reaction are* ***(i)*** *the thioglycollate solution, and* ***(ii)*** *the hydrogen peroxide solution carrying out?*

(d) *What causes the shaping of the hair to eventually be lost?*

3. *The structure of a tyrosine residue in a polypeptide, at pH 7.0, is shown:*

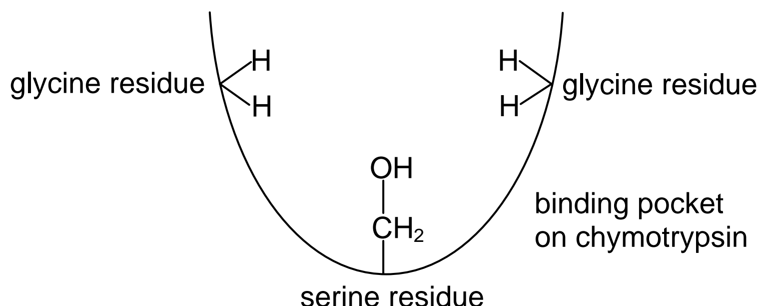


The tyrosine residue binds the polypeptide to the enzyme chymotrypsin.

© University of Cambridge International Examinations 2006

www.theallpapers.com

(a) *Re-draw the tyrosine residue and the binding pocket of chymotrypsin with the tyrosine in position in the pocket. State the two kinds of intermolecular attraction involved.*



(b) *By what technique may the three-dimensional structure of an enzyme like chymotrypsin be determined?*

(c) ***(i)*** *What is the function of a protease?*

(ii) *Why are proteases often included in biological washing powders?*

4. *The following sequence of bases is part of a nucleic acid.*

. . . . .AGAAGAGAAGCU . . . .

(a) ***(i)*** *What information is missing from the above which enables the full structural formula to be drawn unambiguously?*

(ii) *Is the nucleic acid DNA or RNA? Explain your answer.*

(iii) *Give the names of two of the bases represented in the above sequence.*

(b) *When the* ***complete*** *sequence (of which the above 12 bases are a small part) is expressed, the hormone insulin is synthesised.*

(i) *What kind of compound is insulin?*

(ii) *Apart from the possibility of translating sequences of bases into molecules of the wrong* ***length****, there is another important reason why there must be a START codon somewhere in the sequence. What is this reason?*

5. *Mutation of the mRNA of a T4 bacteriophage leads to the omission of one base at the beginning of a sequence of 15 bases so that the rest are displaced by one position in the 5' direction, as shown below.*

Normal 5'- AGUCCAUCACUUAAU - 3'

Mutant 5'- GUCCAUCACUUAAUG - 3'

(a) *Use the genetic code in Figure 6.24 to translate each of these base sequences into amino acid sequences in the normal and mutant protein.*

(b) *Write down the sequence of bases in the piece of DNA which would produce the normal mRNA sequence after transcription, identifying the 3' and 5' ends.*

(c) *How are the base, phosphate and sugar parts of each nucleotide linked in a single strand of DNA?*

(d) *The normal amino acid sequence is part of the enzyme lysozyme. Explain how the mutation might affect the activity of the enzyme; in your answer, refer to the DNA and RNA involved in its biosynthesis.*

© University of Cambridge International Examinations 2006

www.theallpapers.com

1.7 – Key definitions

amino acids: the monomers from which proteins are built; molecules with two functional groups - an amino group and a carboxylic acid group - attached to the same carbon atom

active site (of an enzyme): the most important region of a functional enzyme; the active site has two functions - it has a structure that recognises and binds the substrate, and a catalytic region that helps bring about the reaction catalysed by the enzyme

amylases: enzymes that bring about the breakdown (hydrolysis) of carbohydrates such as starch

carbohydrates: compounds containing carbon, hydrogen and oxygen. Carbohydrates include simple sugars such as glucose, disaccharides, and complex polysaccharides such as starch and cellulose

chromosome: a coiled thread of DNA and protein, found in the nucleus of cells

competitive inhibition: a form of enzyme inhibition by molecules that bind to the active site of the enzyme but do not take part in a reaction

complementary base pairing: the basis of how the two helical strands of DNA bond to each other; adenine (A) in one strand is always paired with thymine (T) in the other and cytosine (C) is always paired with guanine (G)

condensation polymerisation: a type of polymerisation in which a molecule of water is eliminated each time a monomer molecule is added to the chain; it is used in the building of proteins, nucleic acids and polysaccharides

cytoplasm: the fluid within cells surrounding the nucleus and organelles

denaturation: processes by which the complex three-dimensional structure of functional biological molecules such as proteins is destroyed, leading to temporary or permanent loss of activity

deoxyribonucleic acid (DNA): a double helical polymer which carries the genetic message; each molecule is made up of two anti-parallel polynucleotide chains consisting of a sugar-phosphate backbone with nitrogenous bases attached to them

disulphide bonding: a type of covalent bond important in many proteins for maintaining their tertiary and quaternary structure; the bond is formed between the -SH groups of two cysteine residues

endoplasmic reticulum: a network of membranes in the cytoplasm of cells, where large molecules are built up from small ones

enzymes: protein molecules that function as biological catalysts; they are generally more efficient than inorganic catalysts and have a high degree of specificity

extracellular: outside cells

fatty acids: long chain carboxylic acids consisting of hydrocarbon chains with a terminal acid group (-COOH); such fatty acids form one of the components of triglycerides and phosphoglycerides

gene: a length of DNA that codes for the making of a particular protein

genetic engineering (modification): the manipulation of genetic material to produce new types of organisms

haemoglobin: the iron-containing protein found in red blood cells which is responsible for transporting oxygen around the body; it is made up of two α-chains and two β-chains

hydrogen bonding: a type of attraction between molecules which is stronger than other types of inter-molecular force; a hydrogen bond involves a hydrogen atom attached to an electronegative atom an oxygen or nitrogen atom, for example)

hydrolysis: a reaction important in the breakdown of condensation polymers, such as proteins or carbohydrates, in which the elements of water (H and OH) are added to the molecular fragments

intracellular: inside cells

© University of Cambridge International Examinations 2006

www.theallpapers.com

lipids: compounds grouped together because of their non-polar nature; they tend to be insoluble in water, but soluble in organic solvents such as hexane; biochemically important lipids include triglycerides, phosphoglycerides and steroids

lock-and-key mechanism: a model of enzyme activity (first put forward by Fischer) that stresses the importance of molecular shape in explaining the high degree of specificity in enzyme activity

membrane: the boundary surrounding all cells and also surrounding organelles within eukaryotic cells. Average thickness 7 nm

metabolism: the chemical reactions taking place in a living organism

mitochondrion: the organelle in cells in which aerobic respiration takes place – where ATP is produced

monasaccharides (simple sugars): molecules with the general formula $(CH_2O)_n$, where n ranges from 3 to 9; glucose, deoxyribose and ribose are biologically important examples

mutation: unpredictable change in the structure of DNA, or in the structure of a number of chromosomes

nitrogenous bases (in DNA and RNA): nitrogen-containing bases involved in the structure of DNA and RNA; in DNA they are adenine (A), guanine (G), thymine (T) and cytosine (C); in RNA uracil (U) replaces thymine

non-competitive inhibition: a form of enzyme inhibition in which the inhibitor molecule binds to a region of the enzyme surface other than the active site and thus distorts the shape of the enzyme so that the active site no longer functions

nucleotides: the basic structural units of DNA and RNA; each nucleotide is made from a sugar, deoxyribose or ribose, a phosphate group, and a nitrogen-containing base

nucleus: the large membrane-bound organelle in a cell containing DNA for the majority of the cell cycle

organelle: a functionally and structurally distinct part of a cell, for example a ribosome or mitochondrion

peptide bond: the link present between amino acids in a polypeptide (protein) chain; the link is formed by a condensation reaction between the amino group ($-NH_z$) of one amino acid and the carboxylic acid group (-COOH) of another amino acid

phosphodiester link: a link present in the sugar-phosphate backbone of DNA and RNA strands formed between the -OH groups on the sugar molecules and the intervening phosphate groups

primary structure (of proteins): the first of several levels of protein structure; the sequence of amino acids in a polypeptide chain as determined by the gene for that chain

proteases: enzymes which catalyse the breakdown (hydrolysis) of proteins into peptides and amino acids.

proteins: condensation polymers of amino acids joined together by peptide bonds; proteins have a range of important functions ranging from structural proteins to enzymes, hormones and antibodies

replication: the process by which new DNA molecules are generated when cells divide, during the process the double helix unwinds and each strand is copied

ribonucleic acid (RNA): a single-stranded polynucleotide molecule; there are several different types of RNA serving different functions in the mechanism of gene expression

ribosomes: very small organelles (diameter 18-22 nm) found in all cells, where protein molecules are assembled from amino acids

secondary structure (of proteins): the second level of protein structure; a-helix and P-pleated sheet, for example, are structures stabilised by hydrogen bonding between peptide bond regions of the polypeptide

substrate: the molecule upon which an enzyme acts to bring about a reaction

© University of Cambridge International Examinations 2006

www.theallpapers.com

tertiary structure (of proteins): the third level of protein structure involving the overall folding of a polypeptide chain; the chain is stabilised by ionic interactions, van der Waals' forces, hydrogen bonding and covalent disulphide bond formation

transcription: the process in which the genetic message encoded on the template strand of DNA is copied into a messenger RNA (mRNA) molecule

translation: the process by which the message encoded in mRNA is translated into a polypeptide chain by a process involving ribosomes and transfer RNA (tRNA) molecules

triglycerides (triglyceryl esters): lipids that occur in animal fats and vegetable oils; they are formed by the addition of three long-chain fatty acid molecules to a molecule of glycerol (propan-1,2,3-triol) via ester links

turnover number: a measure of the efficiency of an enzyme - it is the number of substrate molecules reacted per enzyme molecule per minute

van der Waals' forces: weak intermolecular forces that occur between covalent molecules; they occur where weak forces of attraction between dipoles in adjacent molecules result in an interaction

© University of Cambridge International Examinations 2006

www.theallpapers.com

1.8 – Resources

Texts & booklets:

Biochemistry, by Richard Harwood; published by Cambridge University Press.
ISBN: 0-521-79751-9
Chemistry for Biologists at Advanced Level, by B. Rockett & R. Sutton; published by John Murray.
ISBN: 0-7195-7146-4
Biochemistry and Food Science, by E. Ramsden; published by Nelson Thornes Ltd.
ISBN: 0-7487-1806-0
Biochemistry and Molecular Biology, by M.Sheehan; published by Nelson Thornes Ltd;
ISBN: 0-17-448207-8
Molecules and Cells, by J. Adds, E. Larkcom, and R. Miller; published by Nelson Thornes Ltd;
ISBN: 0-17-448293-0
Biochemistry for Advanced Biology, by S.Aldridge; published by Cambridge University Press;
ISBN: 0-521-43781-4
Chemistry and the Human Genome, by T. Lister; published by The Royal Society of Chemistry;
ISBN: 0-85404369
Medicinal Chemistry, published by The Royal Society of Chemistry; ISBN: 1-870343-42-5, see
http://www.chemsoc.org/networks/learnnet/medicinal.htm

Resources available for practical and project work on biochemical topics

'In search of more solutions' published by the RSC has a method of isolating the amino acids that make up the sweetener 'aspartame' and separating them by chromatography – this could be extended to a project on the chromatography of amino acids in general, including 2-D thin layer chromatography.

The **National Centre for Biotechnology Education** (NCBE) at the University of Reading, UK [http://www.ncbe.reading.ac.uk/] has a range of equipment and kits for carrying out experimental work in schools. Their website also has a range of practical protocols which can be downloaded from the site e.g. 'Illuminating DNA'.

The **Science and Plants for Schools** (SAPS) website also has kits and resources for practical work in this area [http://www-saps.plantsci.cam.ac.uk/].

The following companies market kits and equipment for practical work on DNA, proteins, and enzymes. These can be used directly as kits or adapted to form the basis of project work.

Bio-Rad [http://www.biorad.com] – go to the Life Science Education section,
Carolina Biological [http://www.carolina.com/],
Edvotek [www.edvotek.com/]
Philip Harris Education market a useful set of enzyme project kits for studies on a range of different enzymes, including urease, various amylases, lipase, and various proteases.
Sigma-Aldrich market a 'DNA Spooling Educational Kit', D8666, which is useful to demonstrate the fibrous nature of DNA.

© University of Cambridge International Examinations 2006

www.theallpapers.com