13 Option B: Biochemistry

Essential ideas

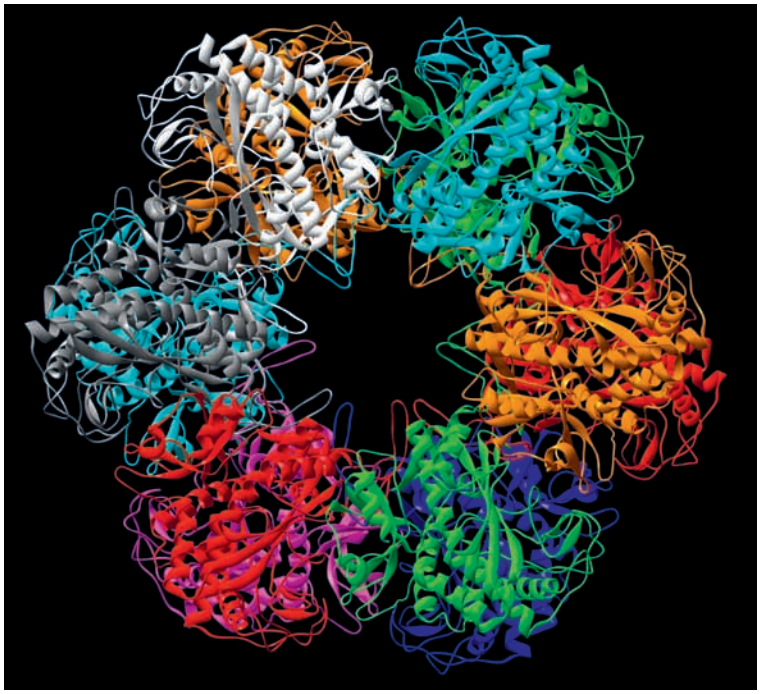
B.1	Metabolic reactions involve a complex interplay between many different components in highly controlled environments.
B.2 and B.7	Proteins are the most diverse of the biopolymers responsible for metabolism and structural integrity of living organisms. Analyses of protein activity and concentration are key areas of biochemical research.
B.3	Lipids are a broad group of biomolecules that are largely non-polar and are therefore insoluble in water.
B.4	Carbohydrates are oxygen-rich biomolecules that play a central role in metabolic reactions of energy transfer.
B.5	Vitamins are organic micronutrients with diverse functions and which must be obtained from the diet.
B.8	DNA is the genetic material that expresses itself by controlling the synthesis of proteins by the cell.
B.9	Biological pigments include a variety of chemical structures with diverse functions and which absorb specific wavelengths of light.
B.10	Most biochemical processes are stereospecific and involve only molecules with a certain configuration of chiral carbon atoms.
B.6	Our increasing knowledge of biochemistry has led to several environmental problems, while also helping to solve others.

Calcium carbonate ($CaCO_3$) plates on the surface of a marine organism, shown in blue in this coloured scanning electron micrograph (SEM). These plates are shed and sink to the ocean floor, forming an important part of deep-sea sediments and rocks such as chalk and limestone. Metabolic processes, controlled by enzymes, synthesize the structural components of organisms, which are later returned to the environment.

Living things show unique properties. They extract and transform energy from their environment to build complex structures, and have the ability to self-regulate and self-replicate. Yet all living organisms are made of molecules that individually are no different from any other molecules found in non-living matter. Somehow the interactions between component molecules in living things must explain the complex processes that we associate with life. Biochemistry is the subject that seeks to find such molecular level explanations for biological phenomena.

Although biochemistry is a relatively young discipline, it is a rapidly advancing area of knowledge, and applications of research are found in the fields of health, diet, and medicine as well as in new technologies. In recent years many of the Nobel Prizes in chemistry – generally considered to be the most prestigious recognition of leading research in chemistry – have been given to scientists working in this field. Prizes have been awarded for major advances made in our understanding of proteins, of chemical interactions at receptors, and of the detailed action of DNA.

Biochemical molecules are an excellent example of the relationship between molecular structure and function, a theme that will be carried through this chapter. As biomolecules are all organic in nature, it is recommended that you are first familiar with the work in Chapter 10, Organic Chemistry. After studying the features of the major groups of biomolecules, this chapter ends with a consideration of some of the impacts and possible solutions to environmental problems related to biochemistry.



Computer model of the enzyme glutamine synthetase. Enzymes are proteins that catalyse all biochemical reactions. Their activity depends on their specific three-dimensional structure, which is determined by their chemical composition. In turn this is controlled by the exact molecular sequence in the nucleic acid DNA. Glutamine synthetase is a complex protein, having a quaternary structure made of six different polypeptide subunits – shown in different colours here. It catalyses the reaction for the synthesis of the amino acid glutamine.

B.1 Introduction to biochemistry

Understandings:

- The diverse functions of biological molecules depend on their structures and shapes.
- Metabolic reactions take place in highly controlled aqueous environments.
- Reactions of breakdown are called catabolism and reactions of synthesis are called anabolism.
- Biopolymers form by condensation reactions and are broken down by hydrolysis reactions.
- Photosynthesis is the synthesis of energy-rich molecules from carbon dioxide and water using light energy.
- Respiration is a complex set of metabolic processes providing energy for cells.

Guidance

Intermediates of aerobic respiration and photosynthesis are not required.

Applications:

- Explanation of the difference between condensation and hydrolysis reactions.
- The use of summary equations of photosynthesis and respiration to explain the potential balancing of oxygen and carbon dioxide in the atmosphere.

Biochemical reactions are organized in metabolic pathways

The scientific study of the origin of life has been described as one of the last frontiers of the life sciences, as so many questions remain unanswered. There is though general agreement that life began in an aqueous environment, and that a key event was the development of a membrane that enclosed and defined a volume of space, known as a **cell**. Within this membrane-bound environment, sequestered from the outside, specific conditions such as temperature, pH, and chemical concentrations can be

generated and controlled. Cells are the basic unit of structure and function in all living things. Complex organisms contain vast numbers of cells, which together carry out all the life processes.

At any one time within the microscopic volume of a living cell, thousands of chemical reactions occur (Figure 13.1). The sum of all these reactions taking place in an organism is known as **metabolism**. This complex chemistry is dependent on a high level of order where every compound has a distinct function. Some features of metabolism are:

- Reactions are controlled in sequences and cycles known as **metabolic pathways**. The product of each step is the reactant for the next. Compounds taking part in metabolism are known as **metabolites**.
- Every reaction is controlled by a specific catalyst called an **enzyme**.
- Similar pathways and enzymes exist in a wide range of different organisms.
- Reactions can be coupled so that energy from one reaction is used to drive another.

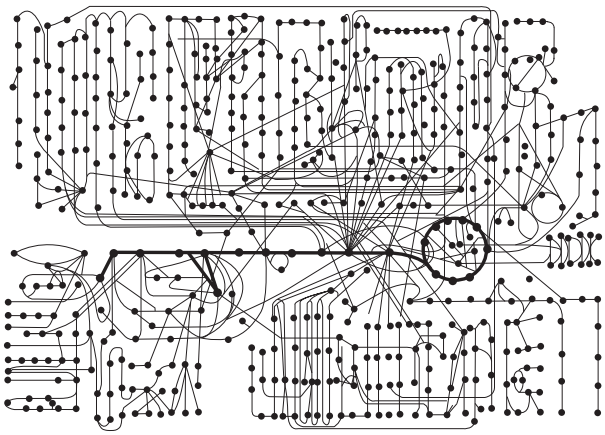


Figure 13.1 A simplified diagram of some metabolic pathways. Each dot represents a metabolite and every line represents a chemical reaction. Pathways, cycles, end-products, and coupled reactions can be identified.

Given the complexity of metabolism, it is useful to classify pathways according to their broad purpose. Later in this chapter, specific examples of metabolic processes will be discussed in more detail.

Anabolism: pathways of synthesis

The part of metabolism concerned with building up or synthesis is known as **anabolism**. The reactants are small molecules called **precursors**, and the products are larger, more complex molecules of higher energy. Anabolic pathways therefore require energy. Examples include the synthesis of proteins from amino acids, nucleic acids from nucleotides, and carbohydrates from the process of photosynthesis.

Catabolism: pathways of breakdown

Metabolic reactions of breakdown or degradation are known as **catabolism**. Catabolic reactions release energy and produce energy-poor end-products, such as carbon dioxide and water. Examples include the breakdown of glucose in respiration or the oxidation of fatty acids.

The energy from catabolic reactions is used to drive anabolic reactions. This is known as **energy coupling** and involves an intermediary energy carrier called adenosine triphosphate, which is known as **ATP**.

The term **emergent properties** is used in biology to describe the properties that result from interactions between components in cells and between cells in multicellular organisms. It is a manifestation of 'the whole is greater than the sum of the parts'.

Metabolism is the sum of the chemical reactions occurring in a living organism.

Metabolism is dependent on a regular supply of diverse nutrients in the diet. Many differences in human health across the world are the result of differences in the supply of nutritious food.

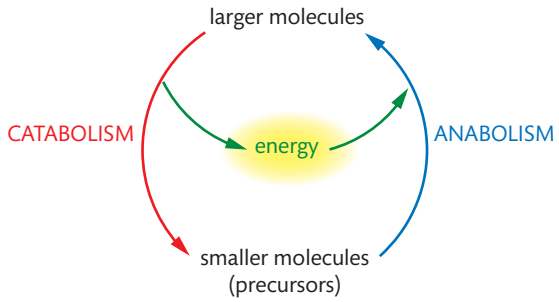


Figure 13.2 The relationship between anabolic and catabolic pathways of metabolism.

Anabolism refers to metabolic pathways of synthesis; catabolism refers to metabolic pathways of breakdown.

Anabolic and catabolic processes take place simultaneously in a cell, but they are separately controlled. This is possible because the metabolic pathway for the biosynthesis of a molecule is usually different from the pathway for its breakdown, and involves different enzymes. Otherwise, metabolism would involve so-called **futile cycles** where stable complex structures would not exist in cells, but would be broken down as they form.

NATURE OF SCIENCE

The complexity of the chemical environment in biochemical systems makes it necessary to use advanced techniques for the isolation of different compounds. Elucidation of metabolic pathways has often been the result of analyses carried out *in vitro*, that is where components are isolated from their usual biological environment. As much as possible, biochemists need to check that data obtained in this indirect way are fully transferable to *in vivo* conditions, that is those within living cells. Similar considerations arise when data obtained from within one living system are used as a basis of knowledge for a different living system, such as carrying out drug trials on animals.

Biomolecules are diverse organic molecules

Of the approximately 100 elements found on Earth, only 27 have been found to be essential components of living things, and the majority of these are present in very small amounts – known as **trace elements**. The molecules found in cells are mostly organic compounds containing about 96% by mass hydrogen, oxygen, carbon, and nitrogen. Sulfur, phosphorus, calcium, and iron are also present in significant amounts.

Biomolecules, the molecules present in living things, represent an immense diversity of structures. Many of them, such as proteins, nucleic acids, and polysaccharides, are **macromolecules**, having relative molecular masses of several thousand. Despite their size and complexity, macromolecules can be described relatively easily in terms of the small units from which they are made, commonly referred to as building blocks, which are linked together by covalent bonds. Most macromolecules are polymers and so the building blocks are monomers – molecules such as glucose and amino acids. Many of these monomers have functions in their own right, for example as energy carriers or molecular messengers. The reactions of build-up and breakdown of all these molecules follow a similar pattern, described below.

Condensation and hydrolysis reactions

Biopolymers are **condensation polymers**, because their synthesis involves the loss of a molecule of water for each covalent bond that forms between two monomers. An example of this, esterification, was described in Chapter 10. But in order to form a polymer, the monomers must each have *two functional groups*, which can be considered as two active ends of the molecule, so they can link together in a similar way to people forming a chain by linking hands on both sides. The reaction is catalysed by enzymes, which are often known as **polymerases**.

Note that condensation polymerization differs from addition polymerization, also described in Chapter 10. Addition reactions lead to only one product, with no elimination of a small molecule.

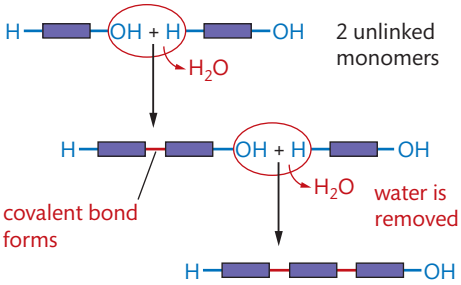


Figure 13.3 Condensation reactions involve loss of a molecule of water for every covalent bond formed.

The breakdown of these molecules reverses the above reaction, adding a molecule of water for each covalent bond broken. The water is split with –H and –OH attaching separately to the product molecules. These are **hydrolysis** reactions and occur during chemical **digestion**. They are catalysed by enzymes, and can also be favoured by heat and acidic or alkaline conditions.

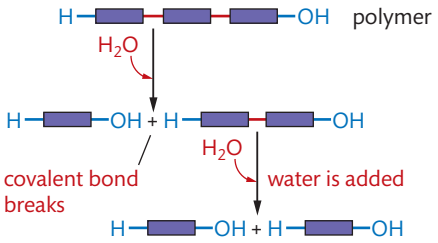
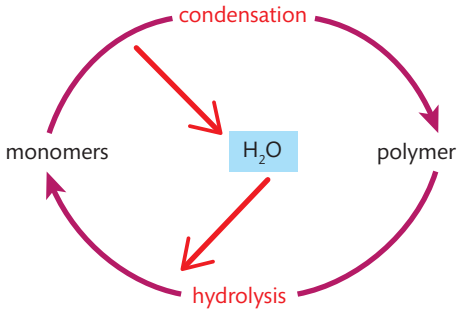


Figure 13.4 Hydrolysis reactions involve addition of a molecule of water for every covalent bond broken.

Examples of condensation and hydrolysis reactions include the synthesis and breakdown of proteins from amino acids and polysaccharides from sugars. Lipids, although they do not form polymers, also involve condensation and hydrolysis reactions between their sub-units. These are all discussed in more detail later in this chapter. These reactions are summarized below.



The activity of biological molecules is dependent on their structures and shapes

A recurring theme in biochemistry is the relationship between molecular structure and function. In other words, molecules seem to be 'chosen' for particular tasks in cells on the basis of their chemical nature and often their three-dimensional shape. Structural molecules like collagen and cellulose are tough and insoluble, whereas a chemical messenger such as the hormone insulin must interact specifically with receptor molecules. The catalytic action of enzymes depends entirely on their ability to form a temporary binding to their substrate, while nucleic acids are able to store and transmit genetic information using chemical sequences.

NATURE OF SCIENCE

As more and more data have been obtained on biochemical pathways, similar reaction patterns are seen in metabolic processes in species that may not be closely related. For this reason biochemistry is now a major part of the study of evolutionary biology. This is an example of how an interdisciplinary approach can contribute to deeper knowledge and understanding.

Living cells transform energy

Living things absorb energy from their environment and use it to synthesize their own complex structures from simple starting molecules, and to carry out functions such as movement and reproduction. As energy is used in life processes, most of it is returned to the environment as heat.

CHALLENGE YOURSELF

1 From a thermodynamic standpoint, what must occur while living things create molecular order using energy from their environment?

Photosynthesis converts light energy into chemical energy

The primary source of energy for life is the Sun. Green plants (and some other organisms) are able to capture solar energy and use it to synthesize energy-rich biomolecules. This process is known as **photosynthesis**, and all organisms, directly or indirectly, depend on it for their supply of food. Simply put, without photosynthesis there would be no life on Earth.

Photosynthesis is the synthesis of energy-rich molecules from carbon dioxide and water using light energy.

The key to photosynthesis is the absorption of light by **photosynthetic pigment** molecules. The primary pigment is **chlorophyll**, whose structure is described on page 740. The light energy trapped drives a series of redox reactions, in which – remarkably – water is split into hydrogen and oxygen. The oxygen is released as a waste product and the hydrogen ultimately reduces carbon dioxide to simple sugar molecules. This summary does not attempt to do justice to the complexity of the process, which involves many intermediates, electron carriers, and enzymes. But, in essence, photosynthesis transforms the energy-poor molecules carbon dioxide and water into energy-rich sugar with the release of oxygen. The overall reaction of photosynthesis can be summarized as follows.

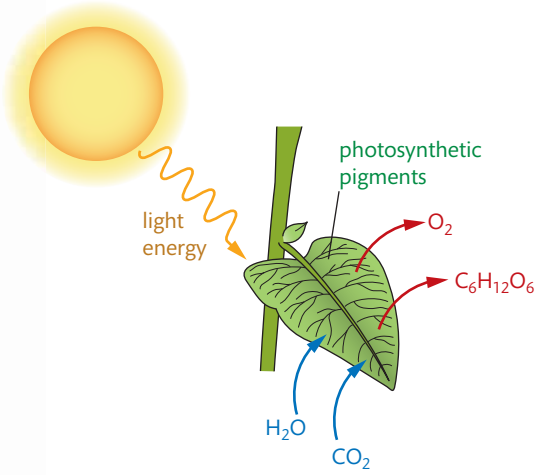


Figure 13.5 Summary of the reactants and products of photosynthesis.

$$6CO_2 + 6H_2O \rightarrow \rightarrow \rightarrow C_6H_{12}O_6 + 6O_2$$

(The multiple arrows indicate that this is only a summary of the process, not a representation of a single reaction.)

Note that *respiration*, which is a biochemical process, is completely different from *breathing*, which is a physical process used to bring about gas exchange. Confusion between the two is generated by many sources, including the fact that the word for breathing is related to 'respire' in several languages, including French, Italian, and Spanish.

Respiration makes energy available for life processes

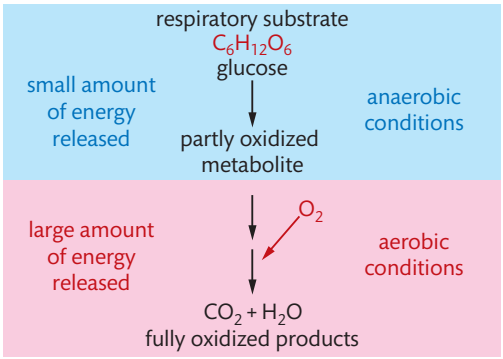
Living things obtain energy-rich molecules such as glucose either through their own process of photosynthesis, or by the intake of energy-rich molecules in food. Broadly speaking, green plants do the former and animals the latter. The release of this energy in a controlled way inside cells is known as **respiration**. It is essential to life, taking place in every living cell continuously.

Respiration is often likened to burning a fuel in oxygen, though in reality it is a much more complex and highly controlled process. But, like burning, it does involve reactions of oxidation where the amount of energy released depends on the extent of oxidation achieved. The details of respiration are complex, involving up to 50

different chemical reactions, each controlled by a specific enzyme. Although different metabolites can be used as respiratory substrates, they are usually first converted into glucose, $C_6H_{12}O_6$.

The first stage of respiration, known as **glycolysis**, is common to all cells and does not use oxygen. Only a small proportion of the energy in glucose is released, as most is trapped in the products of this stage. In the absence of oxygen, known as **anaerobic conditions**, this is the only energy released, and it is enough to keep some cells alive, temporarily in the case of muscle cells and permanently in the case of some bacteria. Products of anaerobic respiration such as lactate and ethanol are therefore energy-rich molecules (as we know from the burning of ethanol described on page 491). In some ways this is like the incomplete combustion of fuels discussed in Chapter 10.

In the presence of oxygen, known as **aerobic conditions**, the oxidation of glucose is complete, and much more energy is released. This is why most cells are dependent on a continuous supply of oxygen. The end products of aerobic respiration are the energy-poor molecules carbon dioxide and water.



Aerobic respiration involves a series of coupled redox reactions, where reactants known as **cytochromes** are successively reduced and then re-oxidized. Ultimately, oxygen acts as the **terminal electron acceptor** when it is reduced to water. The structures of cytochromes are described on page 744.

The overall equation for the aerobic respiration of glucose can be represented as follows:

$$C_6H_{12}O_6 + 6O_2 \rightarrow \rightarrow \rightarrow 6CO_2 + 6H_2O$$

Multiple arrows are again used to show that this is a multi-step reaction, and the representation is a summary only. Note that the reactants and products are the same as for the combustion of glucose.

Aerobic respiration uses O_2 and completely oxidizes glucose to CO_2 and H_2O. It releases much more energy than anaerobic respiration, which occurs in the absence of oxygen.

Respiration is the process of chemical breakdown of energy-rich molecules in cells with the release of energy. It takes place in all living cells all the time.

The products of anaerobic respiration of the single-celled fungus yeast have been used by human cultures for millennia. The release of CO_2 is used to make dough rise in baking, and the production of ethanol is the basis of all brewing.

Figure 13.6 Summary of aerobic and anaerobic respiration.

Respiration releases heat. Here water vapour is rising from decaying organic matter on an active compost heap. Microorganisms living in the compost are releasing significant heat as a result of their high rate of respiration.

CHALLENGE YOURSELF

2 Compare the oxidation states of carbon in $C_6H_{12}O_6$ and in CO_2 to examine the redox processes involved.

Summary of photosynthesis and respiration

Photosynthesis and respiration are metabolic redox processes. Photosynthesis is an anabolic, energy-storing process which reduces carbon dioxide to sugar. Respiration is a catabolic, energy-yielding process which oxidizes sugar to carbon dioxide.

The overall chemical changes in photosynthesis and respiration are the reverse of each other – but in no way do these processes involve a reversal of their chemical steps. Yet together they provide a balance of chemicals, as the waste products of respiration are the raw materials for photosynthesis.

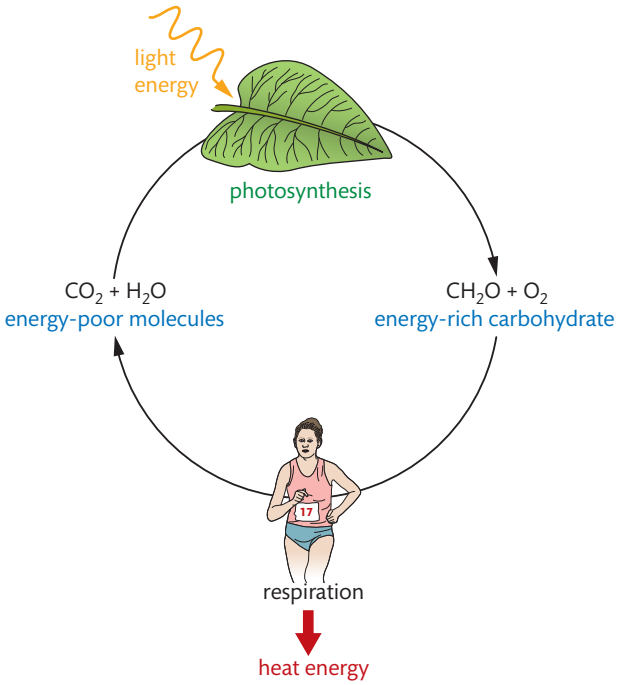


Figure 13.7 Chemical cycling and energy flow through photosynthesis and respiration.

Matter cycles and energy flows through biochemical processes.

Photosynthesis acts as a **carbon sink**, as it removes carbon dioxide from the atmosphere, while respiration and combustion, which release carbon dioxide, are **carbon sources**. Human activities such as deforestation have decreased the carbon sink, and increased the carbon source by burning fossil fuels. Consequently, there is a change in the balance of gases in the atmosphere, and the increasing level of carbon dioxide linked to climate change is a source of widespread concern.

Data from the Mauna Loa research station in Hawaii provide a record of atmospheric carbon dioxide concentration since 1958. Taking into account seasonal fluctuations and variances, the overall trend has been an increase in carbon dioxide levels from about 315 ppm to about 390 ppm in the last 50 years. Carbon dioxide is a greenhouse gas that absorbs and re-radiates infrared radiation, and so this dramatic increase in its concentration is widely accepted as a major factor influencing climate change.

Note in Figure 13.7 that while chemicals are cycled and reused through living things, energy flows and is ultimately lost as heat. This means that life is dependent on a continuing supply of solar energy.

Exercises

1 Write an equation for the hydrolysis of the molecule $C_{18}H_{32}O_{16}$ (a trisaccharide) into its monomers, which are all isomeric with glucose.

2 State the features of molecules that are able to form condensation polymers, and describe what happens during the polymerization reaction.

3 State whether the following metabolic processes are anabolic or catabolic.

(a) protein synthesis
(b) oxidation of fatty acids
(c) anaerobic respiration
(d) DNA replication

4 Outline the conditions required for photosynthesis to take place, and explain why it can be described as a redox process.

5 State and explain the difference in energy yield between aerobic and anaerobic respiration.

B.2 & B.7 Proteins and enzymes

B.2 Proteins and enzymes

Understandings:

- Proteins are polymers of 2-amino acids, joined by amide links (also known as peptide bonds).
- Amino acids are amphoteric and can exist as zwitterions, cations, and anions.
- Protein structures are diverse and are described at the primary, secondary, tertiary, and quaternary levels.
- A protein's three-dimensional shape determines its role in structural components or in metabolic processes.
- Most enzymes are proteins that act as catalysts by binding specifically to a substrate at the active site.
- As enzyme activity depends on the conformation, it is sensitive to changes in temperature, pH, and the presence of heavy metal ions.
- Chromatographic separation is based on different physical and chemical principles.

Guidance

- *The names and structural formulas of the amino acids are given in the IB data booklet in section 33.*
- *Reference should be made to alpha helix and beta pleated sheet, and to fibrous and globular proteins with examples of each.*

Applications and skills:

- Deduction of the structural formulas of reactants and products in condensation reactions of amino acids, and hydrolysis reactions of peptides.
- Explanation of the solubilities and melting points of amino acids in terms of zwitterions.
- Application of the relationships between charge, pH, and isoelectric point for amino acids and proteins.
- Description of the four levels of protein structure, including the origin and types of bonds and interactions involved.
- Deduction and interpretation of graphs of enzyme activity involving changes in substrate concentration, pH, and temperature.
- Explanation of the processes of paper chromatography and gel electrophoresis in amino acid and protein separation and identification.

Guidance

In paper chromatography the use of R_f values and locating agents should be covered.

B.7 Proteins and enzymes

Understandings:

- Inhibitors play an important role in regulating the activities of enzymes.
- Amino acids and proteins can act as buffers in solution.
- Protein assays commonly use UV-vis spectroscopy and a calibration curve based on known standards.

Applications and skills:

- Determination of V_{max} and the value of the Michaelis constant K_m for an enzyme by graphical means, and explanation of their significance.
- Comparison of competitive and non-competitive inhibition of enzymes with reference to protein structure, the active site, and allosteric site.
- Explanation of the concept of product inhibition in metabolic pathways.

Coloured scanning electron micrograph (SEM) of hair shafts growing from the surface of human skin. Hair is made of the fibrous protein keratin, anchored in hair follicles. The outer layer of skin consists of dead keratinized cells that detach giving this flaky appearance.

- Calculation of the pH of buffer solutions, such as those used in protein analysis and in reactions involving amino acids in solution.
- Determination of the concentration of a protein in solution from a calibration curve using the Beer–Lambert law.

Guidance

- *The effects of competitive and non-competitive inhibitors on K_m and V_{max} values should be covered.*
- *The Henderson–Hasselbalch equation is given in the IB data booklet in section 1.*
- *For UV-vis spectroscopy, knowledge of particular reagents and wavelengths is not required.*

The functions of proteins

Proteins are in many ways the most remarkable of the biomolecules. It is estimated that as many as ten million different protein molecules may exist in nature, each with a unique structure and function. From this almost overwhelming diversity, we can classify two main types of proteins.

Fibrous proteins

- structural components
- elongated molecules with dominant secondary structure
- insoluble in water

Globular proteins

- tools that operate at the molecular level – as enzymes, carriers, receptors
- compact spherical molecules with dominant tertiary structure
- soluble in water

Fibrous proteins are responsible for structure, support, and movement, whereas globular proteins drive the reactions of metabolism. Some specific examples of each type are given below.

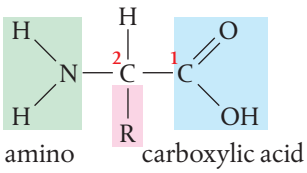
Role of protein	Named example	Type of protein	Specific function
structure	keratin	fibrous	protective covering in hair, wool, claws
structure	collagen	fibrous	connective tissue in skin and tendons
enzyme (catalyst)	polymerase	globular	catalyzes anabolic reactions such as DNA synthesis
communication	insulin	globular	controls and maintains the concentration of glucose in the blood
transport	hemoglobin	globular	carries oxygen

How can it be that this same type of molecule, the protein, is used for both the walls of the reactor and the reactions within? The answer lies in the fact that proteins are as diverse and unique in their structures as they are in their functions, and that this variety is rooted in their molecular building blocks.

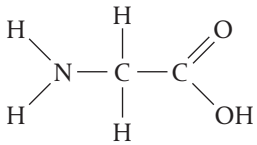
The structure of proteins

Amino acids are the building blocks of proteins

Proteins are polymers – long-chain molecules – of monomer units called **amino acids**. Each amino acid contains an amino group (—NH_2) and an acid group (—COOH) bonded to the same carbon atom.



They are called 2-amino acids. Because the chain is numbered starting with the carboxylic acid group, the amino group is attached to carbon 2. This carbon atom is also bonded to a hydrogen atom and to a group usually known as 'R'. The R group differs from one amino acid to the next, and is therefore the feature that defines the amino acid. About 20 different amino acids are found in naturally occurring proteins. Each is given a standard three-letter abbreviation, for example the smallest amino acid glycine, where R = H, is known as Gly.



A complete list of all the amino acids used in proteins is given in section 33 of the IB data booklet. Amino acids can be classified according to the chemical nature of their R group, usually on the basis of their different polarities, as shown in the examples below.

In addition to the 20 common amino acids of proteins, over 150 other amino acids occur in biology in different forms other than in proteins. Fungi and higher plants have a particularly large variety of non-protein amino acids, some of which are toxic to other forms of life.

Type of amino acid	R group contains	Named example	Structure
non-polar / hydrophobic	hydrocarbon	alanine, Ala	$H_2N-CH(CH_3)-COOH$
polar but uncharged	hydroxyl, —OH, sulfhydryl, —SH, or amide, —$CONH_2$	serine, Ser	$H_2N-CH(CH_2-OH)-COOH$
basic (positively charged at pH 6.0–8.0)	amino, —NH_2	lysine, Lys	$H_2N-CH(CH_2-CH_2-CH_2-CH_2-NH_2)-COOH$
acidic (negatively charged at pH 6.0–8.0)	carboxylic acid, —COOH	aspartic acid, Asp	$H_2N-CH(CH_2-COOH)-COOH$

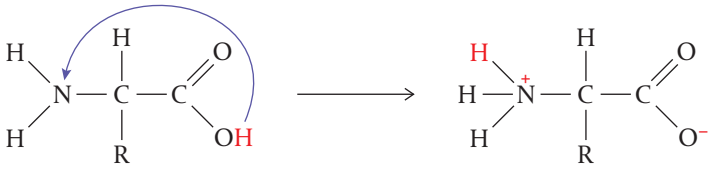
Amino acids are crystalline compounds with high melting points, usually above 200 °C, and they have much greater solubility in water than in non-polar solvents. In addition, amino acids usually move in an electric field. These properties are all typical of ionic compounds, and so suggest that amino acids contain charged groups. The charges are a result of acid–base behaviour, which is best explained in terms of Brønsted–Lowry theory.

In aqueous solution and in crystalline form, amino acids commonly exist with both positive and negative charges within the molecule, known as **zwitterions**. They are sometimes referred to as **internal salts**, as the charges result from an internal acid–base reaction, with the transfer of a proton (H^+) from the acid —COOH group to the basic –NH_2 group in the same amino acid.

CHALLENGE YOURSELF

3 All amino acids except glycine are chiral, and exist in the L form in nature. By reference to section 33 of the IB data booklet, determine which two amino acids found in proteins have *two* chiral carbon atoms.

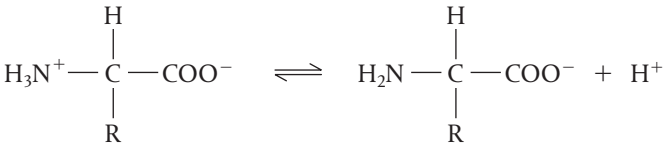
Zwitterions are neutral molecules that have unit electrical charges of opposite sign.



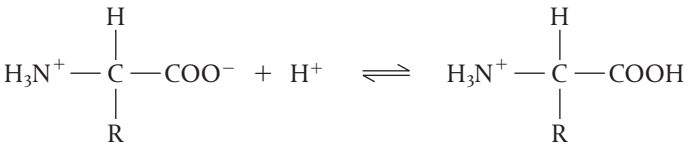
undissociated form zwitterion

As amino acids contain both an acidic group and a basic group, they are **amphoteric** or **amphiprotic**, as described in Chapter 8. In aqueous solution they will accept and donate H^+ according to changes in the pH of the medium as shown below.

1 As an acid, donating H^+:



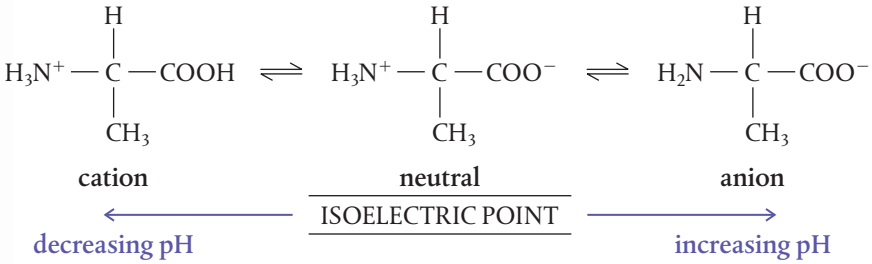
2 As a base, accepting H^+:



Note that in the zwitterion it is the conjugates of the acid and the base that are responsible for this property. As the pH of the medium affects the equilibrium position of these reactions, it influences the charge of an amino acid as follows:

- at high pH (low $[H^+]$), reaction 1 above is favoured as the $-NH_3^+$ group loses its H^+ and forms an anion
- at low pH (high $[H^+]$), reaction 2 above is favoured as the $-COO^-$ group gains H^+ and forms a cation.

These effects of pH on charge are summarized below. This example assumes that the R group is an uncharged group, for example the amino acid alanine.



The charge carried by an amino acid depends on the pH of the medium. The isoelectric point is the pH at which it is electrically neutral.

So amino acids tend to be positively charged at low pH and negatively charged at high pH. The intermediate pH at which the amino acid is electrically neutral is known as its **isoelectric point**. With no net charge at this pH, amino acids will not move in an electric field. Also at this point, the molecules will have minimum mutual repulsion and so be the least soluble. Section 33 in the IB data booklet gives the pH of the isoelectric point of each amino acid alongside its structure, and some examples are shown in the table on page 683.

Common name	Symbol	Structural formula	pH of isoelectric point
glycine	Gly	$H_2N—CH_2—COOH$	6.0
alanine	Ala	$H_2N—\underset{\displaystyle CH_3}{\underset{\vert}{CH}}—COOH$	6.0
lysine	Lys	$H_2N—\underset{\displaystyle CH_2—CH_2—CH_2—CH_2—NH_2}{\underset{\vert}{CH}}—COOH$	9.7
aspartic acid	Asp	$H_2N—\underset{\displaystyle CH_2—COOH}{\underset{\vert}{CH}}—COOH$	2.8

This shows that amino acids such as alanine and glycine, which have uncharged R groups, have the same isoelectric point of pH 6.0. But where the R group contains an acidic or a basic group, then the pK_a and pK_b values of these groups will also influence the charge as pH changes. This is why aspartic acid and lysine, for example, have very different isoelectric points. This difference is exploited in techniques for separating amino acids, to be discussed later in this chapter.

The equations above also show that amino acids act as **pH buffers**. By reacting with both H^+ ions and OH^- ions, amino acids cause the pH to be resistant to change on addition of small amounts of acid or alkali, as explained in Chapter 8. The buffering role of amino acids is important in helping to maintain a constant pH in cells, a crucial need for biological solutions. Many of the protein components, especially enzymes, are extremely sensitive to changes in pH and can be made inactive by significant fluctuations. For example, human blood has a pH of 7.4, and an increase or a decrease of more than 0.5 pH units can be fatal. Clearly, effective buffering is a must. There are several different buffer systems at work in the human body, including those that use amino acids and proteins.

Like amino acids, proteins also have isoelectric points. Fresh milk has a pH of about 6.7 and at this pH the protein casein carries a negative charge and is dispersed in solution. As it sours, bacteria growing in the milk produce acids that lower the pH and so reduce the charges on the molecules. When it reaches pH 4.6 this is the isoelectric point of casein, so now the protein becomes less soluble, precipitating from solution as the familiar 'curdled milk'. This is the first step in cheese making.

Amino acids are amphoteric and can act as buffers in solution.

A trough of curdled milk in a cheese-making factory. The curd is solidified milk formed by lowering the pH and so precipitating the protein at its isoelectric point.

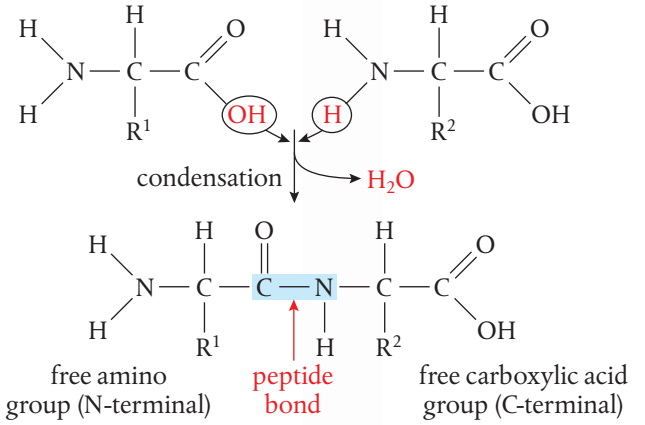


Figure 13.8 Formation of a dipeptide by condensation of two amino acids.

Amino acids link together through condensation reactions

Amino acids are able to react together in a condensation reaction in which a molecule of water is eliminated and a new bond is formed between the acid group of one amino acid and the amino group of the other. This bond is a substituted **amide link** known as a **peptide bond**, and two amino acids linked in this way are known as a **dipeptide**. By convention the free $-NH_2$ group (known as the N-terminal) is put on the left of the sequence and the free $-COOH$ group (C-terminal) on the right.

We can see that the dipeptide still has a functional group at each end of the molecule, with $-NH_2$ at one end and $-COOH$ at the other. So it can react again by another condensation reaction, forming a **tripeptide** and eventually a chain of linked amino acids known as a **polypeptide**.

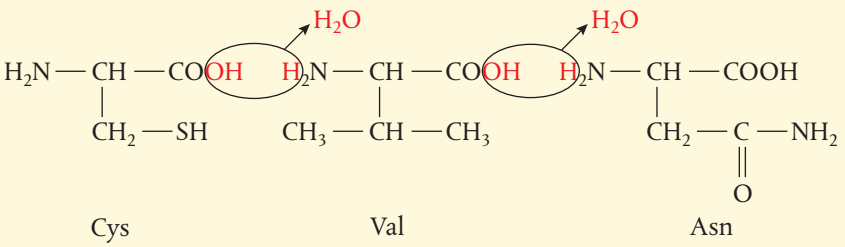
Worked example

Draw a tripeptide with the following sequence:

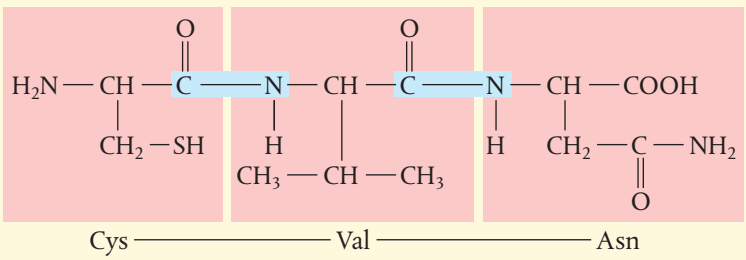
Cys — Val — Asn

Solution

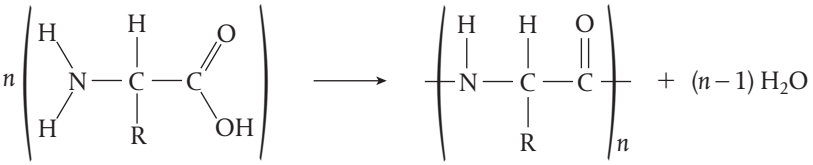
Look up the structures (the R groups) of the amino acids in section 33 of the IB data booklet and draw them out in the same order as given in the question.



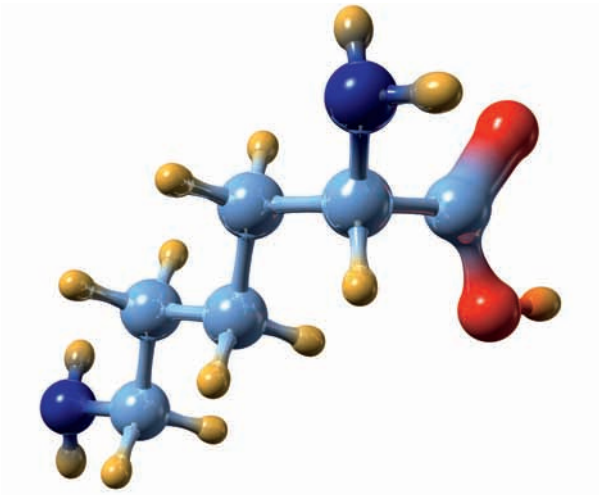
Now draw peptide bonds between the carbon of the $-COOH$ group and the nitrogen of the $-NH_2$ group, ensuring that H_2O is released, and that each atom has the correct number of bonds in the final structure.



We can write a general equation for the synthesis of a polypeptide from its amino acids as follows



repeat unit



Molecular model of the amino acid lysine $NH_2(CH_2)_4CH(NH_2)COOH$. Lysine is known as an **essential amino acid** because it cannot be synthesized by the body and so must be obtained in the diet. Carbon atoms are shown in blue, hydrogen in gold, oxygen in red, and nitrogen in dark blue.

In just the same way as letters linked together in different orders make different words (e.g. eat, ate, tea), so amino acids linked together in different sequences make different peptides. The three amino acids above could have been linked for example as: Asn—Cys—Val or Cys—Asn—Val, all of which would be different tripeptides with different properties.

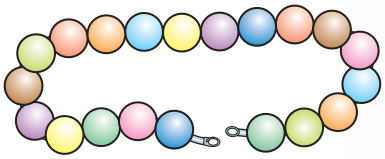


Figure 13.9 Different coloured beads can be strung together in different combinations to make a wide variety of different necklaces. In a similar way, different sequences of amino acids link together to give rise to an almost infinite variety of proteins.

The sequence in which the amino acids are linked to form the chain is of great significance as it will determine the exact nature of the polypeptide. This is where the extraordinary variety of protein structures comes from.

Consider making a necklace by joining together 20 different colours of beads, with each colour being used as many times as you like. In every position you would have a choice of 20 different possibilities, so just imagine how many different combinations of beads you could have. Building polypeptides from amino acids presents a similar situation – at any point in the chain there are 20 different possibilities. Even for a tripeptide there are $20 \times 20 \times 20 = 8000$ different possible combinations. But proteins are typically made from polypeptide chains with 50 or more amino acids, so the number of possible structures becomes enormous. A polypeptide with 50 amino acids has 20^{50} different possible sequences: 1.1×10^{65} variations.

We will now move on to look at how the sequence of amino acids in the polypeptide uniquely determines the structure and therefore the function of the protein. Understanding protein structure can seem quite complex, so for convenience it is divided into four levels of organization, known as primary through to quaternary.

Many synthetic compounds are polyamides, formed by condensation of a carboxylic acid and an amine group. Examples include nylon and Kevlar.

CHALLENGE YOURSELF

4 By reference to section 33 of the IB data booklet, determine which amino acid contains a secondary amine group. How might this affect its properties in a peptide chain?

The primary structure of proteins is the amino acid sequence

The **primary structure** of a protein refers to the number and sequence of amino acids in its polypeptide chain. Held together by peptide bonds, this forms the covalent backbone of the molecule. Interestingly, once the primary structure has been determined, all the other levels of protein structure follow – so it really does dictate the entire structure and function of the protein.

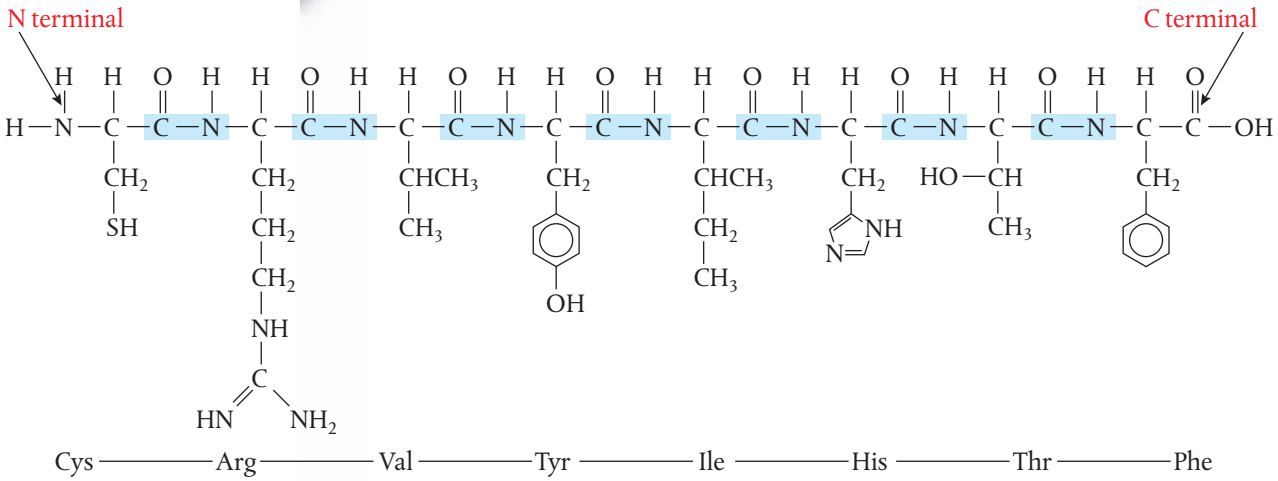


Figure 13.10 Primary structure of a small polypeptide. Peptide bonds are highlighted in blue.

Sickle-cell anaemia is a condition that results from a single amino acid change in one chain of 146 amino acids in the protein haemoglobin. The resulting change in the conformation of haemoglobin means that it is not able to carry oxygen efficiently. It is a common condition in many African countries, as the altered red blood cells provide some immunity to malaria.

Often considered the father of modern molecular biology, Fred Sanger of Cambridge, England, is one of only four people to have been awarded two Nobel Prizes. In 1958 he was awarded the Chemistry Prize for establishing the sequence of the 51 amino acids in insulin chain B. This was the first protein to have its primary structure elucidated in this way and it was the culmination of 12 years of work. In 1980 Sanger shared the Nobel Prize in Chemistry for similar work on the base sequencing of nucleic acids.

So what determines the primary structure of a protein? The answer to this question when it came in the early 1960s was considered to be one of the most important discoveries of molecular biology. DNA (deoxyribose nucleic acid), which determines genetic information, acts by dictating to cells the primary structure of their proteins. In other words, the expression of genes is through the proteins that an organism synthesizes. So it is the primary structure of our proteins that gives each of us our unique genetic characteristics.

Protein sequencing is now a routine operation in biochemical research, with machines able to deduce the entire sequence of a large protein very quickly. This is a major part of **proteomics**, the study that explores the relationship between structure and function of proteins. The synthesis of new so-called **designer proteins** through protein engineering has many applications in pharmaceutical and environmental research. Protein sequences are also the aspect of protein structure used in studies of biochemical evolution to determine the relationships between organisms.

The secondary structure of proteins is regular hydrogen bonding

The **secondary structure** refers to folding of the polypeptide chain as a result of hydrogen bonding between peptide bonds along its length. Hydrogen bonds can form between the —C=O group of one peptide bond and the —N—H group of another peptide bond further along the chain which will cause the chain to fold. The exact

configuration of this will be influenced by the R groups along the chain and so differs in different proteins and even in different sections of the same protein.

We can distinguish two main types of secondary structure, the α-**helix** and the β-**pleated sheet**.

The α-helix is a regular coiled configuration of the polypeptide chain resulting from hydrogen bonds forming between two peptide bonds four amino acid units apart. This twists the chain into a tightly coiled helix, much like a spiral staircase, with 3.6 amino acids per turn, as shown in Figure 13.11.

The α-helix is flexible and elastic as the intra-chain hydrogen bonds easily break and re-form as the molecule is stretched. A good example of the α-helix is found in keratins, structural proteins found in hair, skin, and claws.

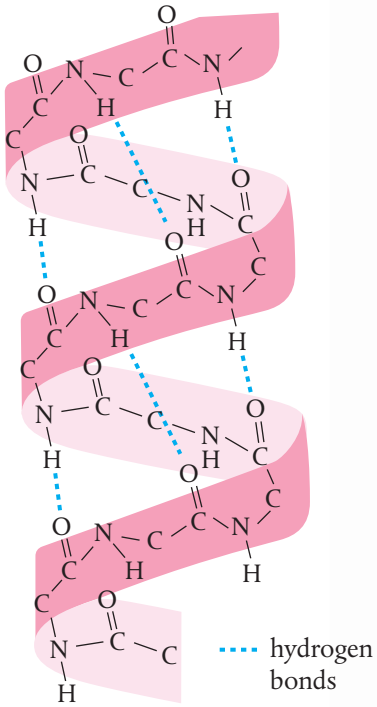


Figure 13.11 α-helical secondary structure of keratin. The amino acid backbone winds in a spiral, held by hydrogen bonds shown in blue.

Human hair (made of the protein keratin) grows approximately 15 cm in one year, which means that 9.5 turns of the α-helix must be produced every second.

Hair stretches to almost double its length when exposed to moist heat but contracts to its normal length on cooling, because the hydrogen bonds in the secondary structure easily break and re-form. This is why hair is often much curlier in humid conditions.

The β-pleated sheet is a structure composed of 'side by side' polypeptides which are in extended form, not tightly coiled as in the α-helix. They are arranged in pleated sheets that are cross-linked by inter-chain hydrogen bonds, as shown in Figure 13.12.

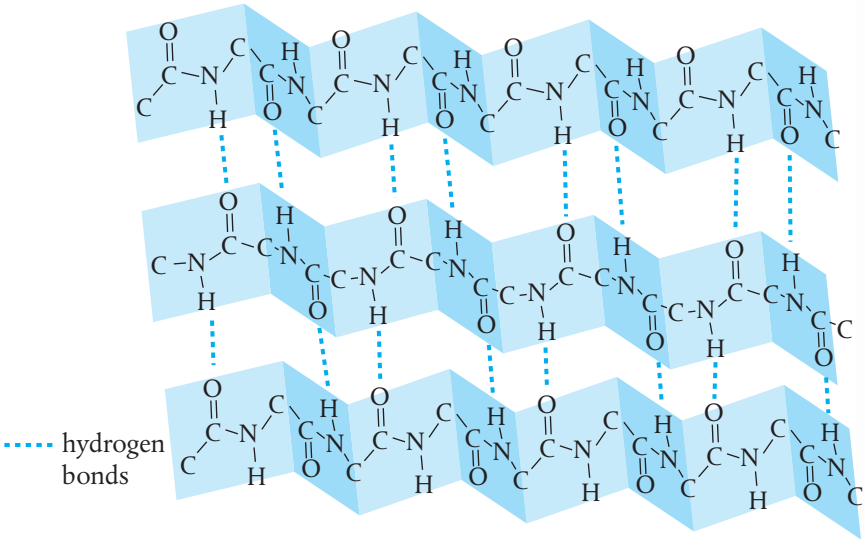
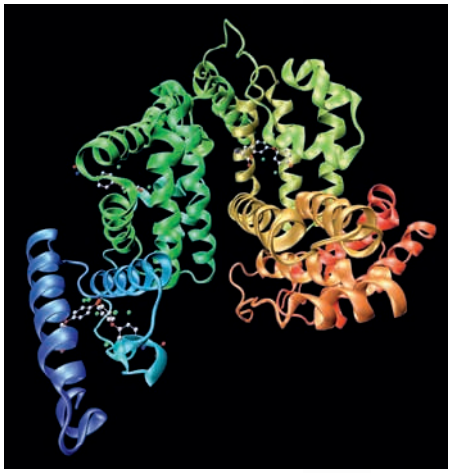


Figure 13.12 β-pleated sheet secondary structure of silk fibroin. The polypeptides run parallel to each other, held by hydrogen bonds shown in blue.

A spider's web outlined with drops of morning dew. The web's fibres are made of the protein fibroin containing a β-pleated sheet secondary structure and spun from special secreting glands. A typical web may contain 20 metres of fibroin.

The β-pleated sheet is flexible but inelastic. A good example is found in fibroin, the protein in the fibres spun by spiders and silkworms and also found in the beaks and claws of birds.

Keratin and fibroin with their well-defined secondary structure are fibrous proteins. Their toughness and insolubility in water make them ideal for their structural roles. They are simpler structures with more repetitive amino acid sequences than globular proteins, in much the same way as the bricks in a building are mostly alike.



Computer artwork showing the structure of the protein albumin from human blood. The spiral regions of α-helical structure can be clearly seen as well as the overall three-dimensional conformation.

The tertiary structure of proteins is the result of interactions between the R groups

The **tertiary structure** refers to the further twisting, folding, and coiling of the polypeptide chain as a result of interactions between the R groups, known as **side chains**. The structure that results is a very specific compact three-dimensional structure, known as the protein's **conformation**. It is the most stable arrangement of the protein, taking into account all the possible interactions along the entire length of the polypeptide. Note that the interactions between the side chains are all *intra*-molecular forces, as they occur within the one polypeptide chain.

This conformation is particularly important in the globular proteins, which include all the enzymes and protein hormones. They are water soluble because their structure positions nearly all of the polar (or hydrophilic) side chains on the outer surface of the molecules where they can interact with water, and most of the non-polar (or hydrophobic) side chains in the interior out of contact with water. The interactions that stabilize this conformation are of the following types, summarized in Figure 13.13.

(a) *Hydrophobic interactions* – between non-polar side chains.
For example, between two alkyl side chains in valine; these weak interactions, based on London (dispersion) forces between induced dipoles, produce non-polar regions in the interior of the protein.

(b) *Hydrogen bonding* – between polar side chains.
For example, between the $-CH_2OH$ group in serine and the $-CH_2COOH$ group in aspartic acid.

(c) *Ionic bonding* – between side chains carrying a charge.
For example, between the $-(CH_2)_4NH_3^+$ group in lysine and the $-CH_2COO^-$ group in aspartic acid.

(d) *Disulfide bridges* – between the sulfur-containing amino acid cysteine.
These are covalent bonds, and hence the strongest of these interactions.

Make sure you are clear on the difference between hydrogen bonds in the secondary structure, which form between groups in the peptide bonds, and hydrogen bonds in the tertiary structure, which form between the side chains.

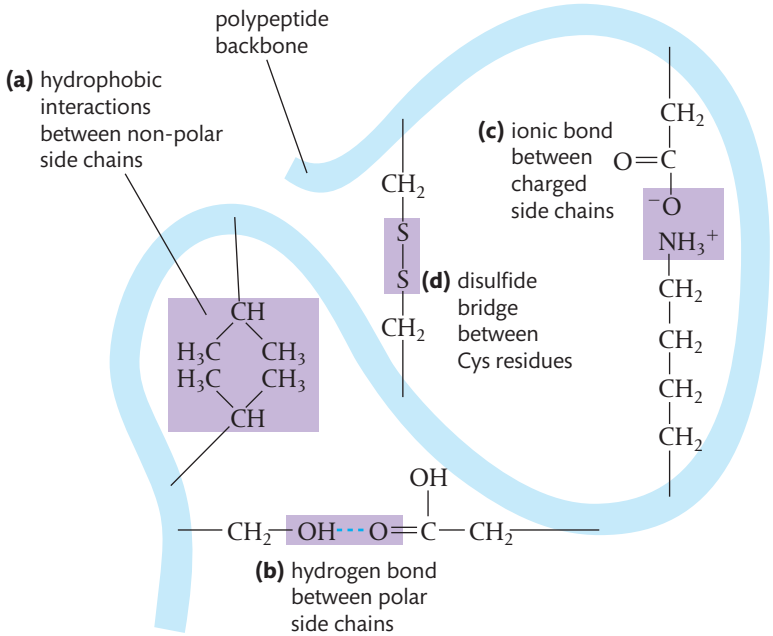


Figure 13.13 Summary of the interactions that contribute to the tertiary structure of a protein.

These interactions can all be upset by changes in the medium such as to temperature, pH, or the presence of metal ions. When a protein loses its specific tertiary structure as a result of such disruptions, it is said to be **denatured**. The familiar sight of the white of an egg solidifying on heating is an example of this. Denaturation of enzymes renders them biologically inactive, which is one of the reasons why intracellular conditions must be tightly controlled. We will come back to this on page 695.

The process of 'perming' that introduces more curls into hair involves first breaking the disulfide bridges between the cysteine residues in the keratin using a reducing agent. New disulfide bridges are formed when the hair is chemically re-oxidized while it is twisted around rollers. The size of the roller determines the position of the new disulfide bridges that form. As these are covalent bonds they do not break on normal treatments like washing and combing, and so are said to be 'permanent'. Similar processes are used in straightening curly hair.

Quaternary structure of a protein is the association between different polypeptides

Some proteins comprise more than one polypeptide chain, and in these cases the association between these chains is known as the **quaternary structure**. This association involves similar forces and bonds to those found in the tertiary structure – hydrophobic interactions, hydrogen bonds, ionic bonds, and disulfide bridges.

For example, the protein collagen, which is found in skin and tendons and is actually the most abundant protein in the human body, is a triple helix of three polypeptide chains, with inter-chain hydrogen bonds between them. This helps to give it a stable rope-like structure that is resistant to stretching.

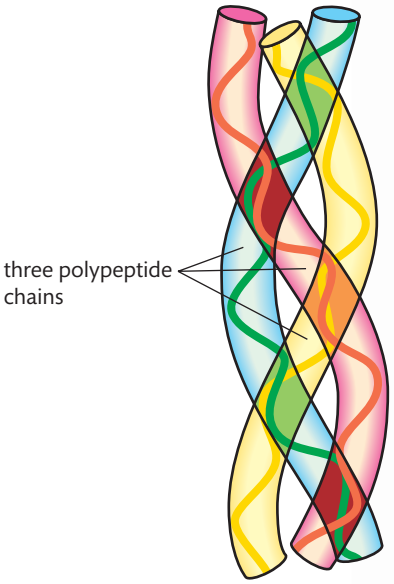


Figure 13.14 Triple-helical quaternary structure of collagen.

Another example is the protein hemoglobin, responsible for carrying oxygen in the blood, which is made up of four polypeptide chains that fit together tightly in the protein assembly. The structure–function relationship of this pigment molecule is discussed further on page 741.

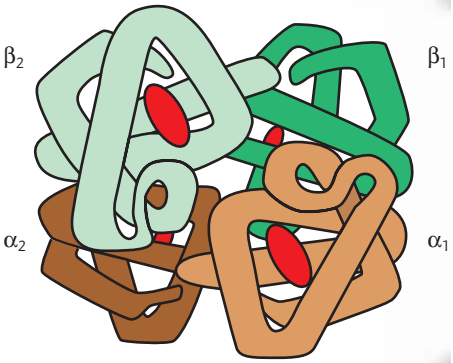


Figure 13.15 Quaternary structure of hemoglobin, an assembly of four polypeptides.

Many proteins consist of only one polypeptide chain, and so do not have a quaternary structure.

Summary of the bonds and forces involved in protein structure

Protein structure	Nature of interactions
primary	covalent (amide / peptide) bonds between amino acids
secondary	hydrogen bonds within the polypeptide chain involving groups in the peptide bonds
tertiary	interactions between side chains including hydrophobic interactions, hydrogen bonds, ionic bonds, disulfide bridges
quaternary	interactions between polypeptide chains, including the same interactions as in the tertiary structure

NATURE OF SCIENCE

Because of the almost infinite variety of protein structures, for a long time they were considered to be more likely than DNA to be the store of hereditary information. Biochemists were surprised to find from experimental results that in fact it is DNA and not protein which carries the genetic code. The evidence came from separate experiments on the ability of DNA to achieve transformation of bacteria, and on the use of radioactive isotopes to demonstrate that viruses inject DNA and not protein into their host cells. Data from both experiments were necessary for broad conclusions to be reached. Subsequent evidence proved that although proteins are not the carriers of hereditary information, they are the molecules through which it is expressed.

Exercises

6 **(a)** Using the three-letter word symbols for amino acids, show all the possible tripeptides that can form from the three amino acids tyrosine, valine, and histidine.
(b) Deduce the number of different peptides that could form from the four amino acids tyrosine, valine, histidine, and proline.

7 Consult section 33 of the IB data booklet to consider the following amino acids:
leucine, glutamic acid, threonine, lysine, serine.

Which of these amino acids:

(a) is most likely to be found in the interior of a globular protein?
(b) contains a secondary alcohol group?
(c) will be negatively charged at pH 5.0?
(d) will be positively charged at pH 7.0?

8 Explain the differences in the structure and properties of fibrous and globular proteins.

9 Hydrogen bonds occur in the secondary and the tertiary structures of proteins. Describe the different origins of these.

Enzymes are globular proteins

Enzyme action is the basis of the brewing and cheese-making industries, possibly the oldest forms of biotechnology. Cheeses and fermented drinks are often associated with specific place names, according to the different flavours produced by enzymes in local microorganisms.

Enzymes are biological catalysts that control all biochemical reactions. They are protein molecules.

The complexity of metabolism demands a highly sensitive control system which can respond to the changing needs of the cell. This is achieved through the action of **enzymes**, the biological catalysts that control every reaction in biochemistry. Because enzymes are specific for each reaction and can be individually controlled, they determine the cell's reactivity at the molecular level.

Enzymes are globular proteins, and exist in compact spherical shapes when in aqueous solution in cells. Their well-defined tertiary structure gives them a specific three-dimensional shape, which is essential for enzyme activity. Enzymes are typically relatively large molecules, containing several hundred amino acids, and some also have a quaternary structure. For example, many of the enzymes involved in the first stage of respiration are dimeric proteins, meaning they contain two polypeptide chains.

In addition, some enzymes require non-protein molecules to be bound for activity. These are known as **co-factors** and may be organic, when they are known as **coenzymes**, or inorganic, such as metal ions. Common examples include vitamins, many of which act as precursors for coenzymes.

Many serious or fatal illnesses are the results of the failure of a single enzyme. For example, the condition phenylketonuria (PKU) that can lead to mental retardation is the consequence of a malfunction in the enzyme responsible for the breakdown of the amino acid phenylalanine in the liver. This condition is the reason why many food and drinks that contain aspartame are labelled 'contains a source of phenylalanine'.

Enzymes form a complex with the substrate

Enzymes are catalysts and so increase the rate of a chemical reaction without themselves undergoing permanent chemical change. The reactant in the reaction catalysed by the enzyme is known as the **substrate**. The presence of the enzyme provides a reaction route of lower activation energy and so enables the reaction to occur more quickly at the same temperature.

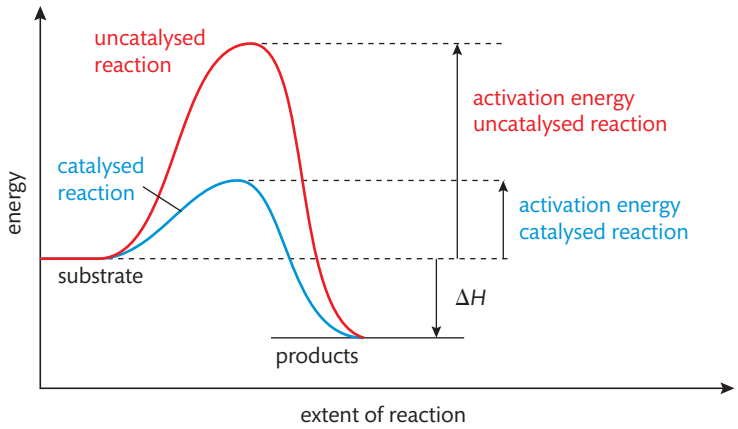
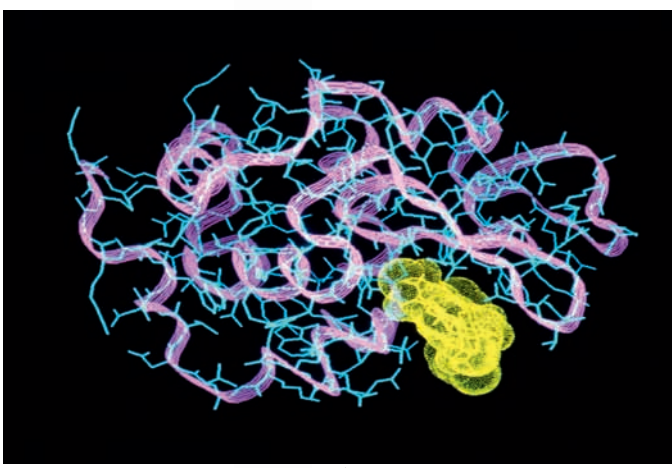


Figure 13.16 Graph showing reaction route of lower activation energy in the presence of an enzyme. This means that at a specified temperature a higher proportion of particles will have sufficient energy to react and so the overall rate of reaction is increased.

The action of the enzyme is due to its ability to form a temporary binding to the substrate where it is held by relatively weak forces of attraction, forming an **enzyme–substrate complex**. This binding occurs at a small region of the enzyme known as the **active site**, which is typically a pocket or groove on the surface of the protein. The substrate is usually a much smaller molecule than the enzyme, and therefore fits within it. Formation of the complex depends on a 'chemical fit' or compatibility between the substrate and the side chains of the amino acids at the active site of the enzyme. This involves non-covalent interactions such as hydrophobic interactions, dipole–dipole attractions, hydrogen bonds, and ionic attractions. The binding in the complex puts a strain on the substrate molecule, and this facilitates the breaking and the forming of bonds. Once the substrate has reacted, the product formed no longer fits in the active site and so it detaches. The enzyme is then released unchanged and is able to catalyse further reactions.



Computer graphics representation of the enzyme lysozyme, which breaks down polysaccharides. The protein is shown in blue with its backbone traced out as the magenta ribbon. The substrate is shown in yellow bound to the active site. Note that the substrate is much smaller than the enzyme.

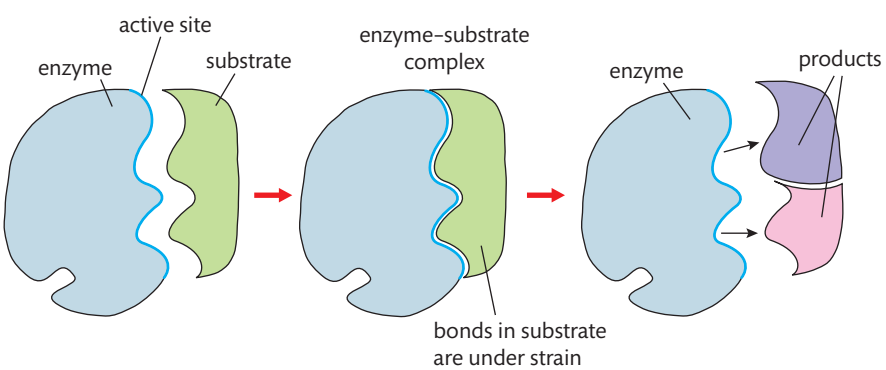


Figure 13.17 Enzymes operate as catalysts by forming a complex with their substrate in which the reaction occurs.

Using **E** for enzyme, **S** for substrate, and **P** for product, we can summarize the action of enzymes as follows:

$$E + S \rightleftharpoons E\text{-}S \rightleftharpoons E\text{-}P \rightleftharpoons E + P$$

Note that all the reactions are shown as equilibrium reactions, and so are reversible depending on the conditions.

Enzymes are specific for the reaction they catalyse

Many enzymes are known to be very specific for their substrate. For example, different proteases, enzymes that hydrolyse proteins, target only specific peptide bonds between certain amino acids. The specificity of an enzyme for its substrate results from its conformation, as this determines the arrangement of the side chains of the amino acids at its active site and therefore its ability to bind precisely with the substrate. The exact nature of the interactions leading to the enzyme–substrate complex is a subject of ongoing research.

NATURE OF SCIENCE

Development of enzyme theory provides a case study of how a scientific model may change over time. In 1890 the German chemist Emil Fischer proposed a model known as the **lock and key** mechanism to describe the fit between an enzyme and its substrate, and to explain the specificity of enzymes. This is illustrated in Figure 13.18. But this model cannot readily explain all aspects of enzyme inhibition, and does little to explain the catalytic action, because a lock does not change a key the way an enzyme changes a substrate. Later work on proteomics recognized that enzymes are less rigid structures than the lock and key model suggests, and in 1958, Daniel Koshland of Rockefeller University, New York, proposed a modification known as the **induced-fit** mechanism of enzyme action. This suggests that in the presence of the substrate, the active site undergoes some conformational changes, shaping itself to allow a better fit. So instead of the substrate fitting into a rigid active site, this is a more dynamic relationship where the side chains at the active site change into the precise positions that allow the binding to occur. It has been compared to putting on a rubber glove that is shaped to the specific shape of the hand as it is pulled over the fingers. It is thought that the distortion that occurs as enzyme and substrate approach may be critical to catalysis.

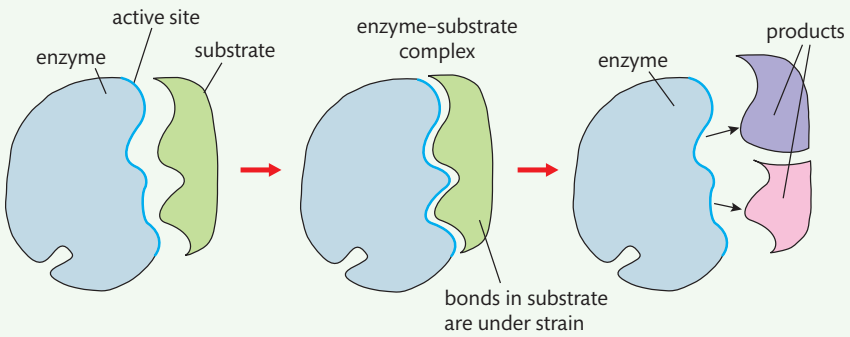


Figure 13.18 Induced fit model of enzyme action. The approach of the substrate induces a conformational change at the active site. This enables the substrate to bind and undergo reaction. As the products are released, the enzyme reverts to its original state.

More recent research suggests that enzymes may be even more flexible structures in which the active site continually reshapes by its interactions with the substrate until the time the substrate is completely bound to it. This is known as the **shifting specificity model** and recognizes the involvement of the entire enzyme molecule rather than just the active site in the catalytic event. Evidence for these developments in enzyme theory has come largely from X-ray protein crystallography, and from advances in computer modelling that enable accurate three-dimensional predictions of enzyme structure to be made.

Enzyme kinetics indicate that saturation occurs in the formation of the complex

The rate of enzyme-catalysed reactions can be followed using the same principles as those for other reactions discussed in Chapter 6. When graphs of substrate concentration against rate of reaction are plotted, the curves show the distinctive shape of **saturation**. Interpretation of these data lends support to the mechanism of enzyme action proceeding via an enzyme–substrate complex.

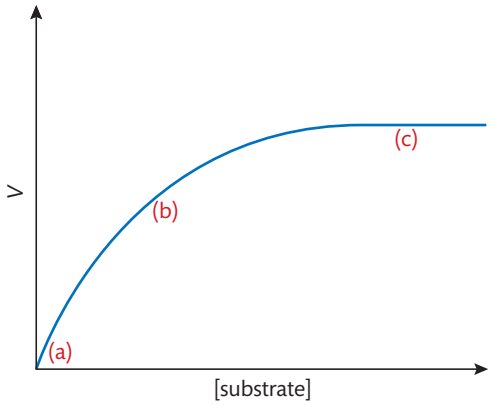


Figure 13.19 The relationship between substrate concentration and rate for an enzyme-catalysed reaction. Note the symbol *V* (velocity) is used to describe the rate of enzyme reactions.

The following points can be deduced from the graph.

(a) At low substrate concentration, the rate of the reaction is proportional to the substrate concentration. Enzyme is available to bind to the substrate.

(b) As the substrate concentration is increased the rate decreases and is no longer proportional to the substrate concentration. Some of the enzyme has its active sites occupied by substrate and is not available.

(c) At high substrate concentration the rate is constant and independent of substrate concentration. At this point the enzyme is saturated with substrate.

All enzymes show this saturation effect, but they vary widely with respect to the substrate concentration required to produce saturation. The mathematical expression describing the kinetics is known as the **Michaelis–Menten equation** after the German biochemist Leonor Michaelis and the Canadian medical scientist Maud Menten.

There are two features of Michaelis–Menten kinetics to note, which are shown in Figure 13.20 and explained below.

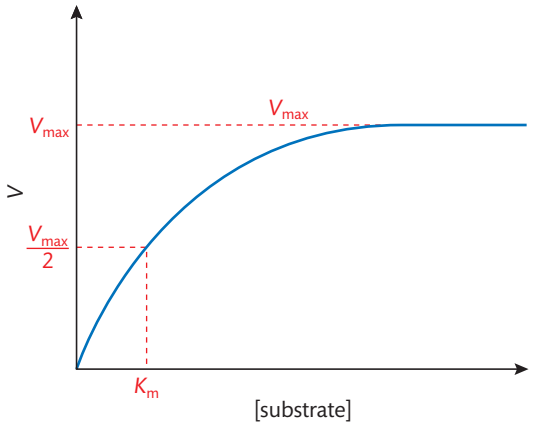


Figure 13.20 The derivation of V_{max} and K_m from the rate–concentration graph.

A small group of enzymes, known as **ribozymes**, have catalytic activity in cells but are made not of protein, but of DNA. The discovery of the catalytic properties of RNA in the 1980s led to a new theory regarding the origin of life, known as the 'RNA world hypothesis'. This suggests that in early life the cell used RNA as both the genetic material and the structural and catalytic molecule, instead of dividing these functions between DNA and proteins. In this case the ribozymes that have persisted could be considered as 'living fossils' of a life based only on nucleic acids.

CHALLENGE YOURSELF

5 What do you think is the order of the reaction with respect to the substrate in Figure 13.19 for each of the regions (a) to (c)?

1 The maximum velocity V_{max}

This is the maximum velocity of the enzyme under the conditions of the experiment. V_{max} has the units of rate. It varies greatly from one enzyme to another, and with pH and temperature. The rate of enzyme reactions is sometimes expressed as the **turnover number**, defined as the number of molecules of substrate that can be processed into products per enzyme molecule per unit of time. For example, the enzyme catalase is a very fast enzyme with a turnover rate of up to 100 000 molecules of its substrate H_2O_2 per second.

2 The Michaelis constant K_m

This is the substrate concentration at which the reaction rate is equal to one half its maximum value. In other words, [S] = K_m when the rate is $V_{max}/2$.

K_m has the units of concentration and its value gives information about the affinity of the enzyme for its substrate. It is an inverse relationship – a low value of K_m means that the reaction is going quickly even at low substrate concentrations. A higher value means that the enzyme has a lower affinity for its substrate. These relationships can be seen in Figure 13.21. The value of K_m also varies with pH and with temperature.

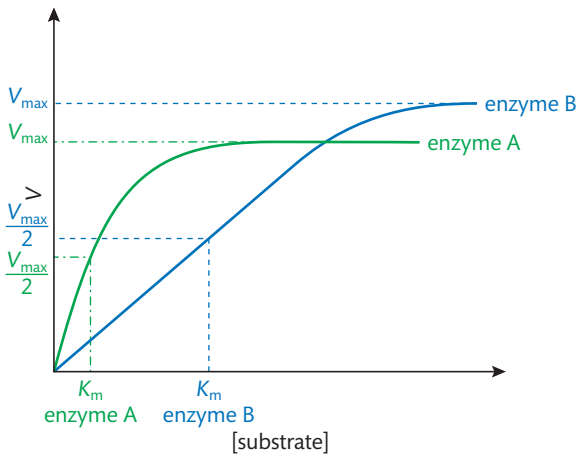


Figure 13.21 The rate–concentration graphs for two different enzymes, A and B. It can be seen that the rate is higher at low substrate concentration for A, and so the value for its K_m is lower.

The table below gives the K_m values for some enzymes.

Enzyme	Substrate	K_m / mmol dm^{-3}
catalase	hydrogen peroxide	25.0
hexokinase	glucose	0.15
carbonic anhydrase	hydrogencarbonate ion HCO_3^-	9.0

These differences also determine how responsive an enzyme will be to changes in substrate concentration. An enzyme with a low K_m, such as hexokinase, is saturated with substrate under most cell conditions and so will act at a more or less constant rate regardless of variations in substrate concentration. An enzyme with a high K_m, such as catalase, is not normally saturated with substrate so its activity will be more sensitive to changes in the concentration of substrate.

Enzyme names are often quite long and appear complex, but in most cases they are logically derived. The name of the substrate is followed by the type of reaction followed by the suffix –ase. So, for example, 'pyruvate dehydrogenase' acts on pyruvate to remove hydrogen; 'glucose isomerase' converts glucose into its isomer fructose.

Enzyme activity is influenced by its physical and chemical environment

We have seen that the action of an enzyme depends on how its specific three-dimensional shape enables it to bind to the substrate. So any conditions that alter its conformation will affect its catalytic action. We will describe three of these factors here.

TOK

To what extent do the use of metaphors and models, such as 'lock and key' and 'induced fit', enhance our knowledge of enzyme action?

1 Temperature

We know from Chapter 6 that the rate of a reaction is increased by a rise in temperature due to the increase in the average kinetic energy of the particles. For enzymic reactions, this means there is an increase in the frequency of collisions between molecules of enzyme and substrate that have greater than the activation energy, leading to a higher rate of reaction. But this is only true up to a certain temperature. Beyond this, the effect of the increase in kinetic energy is to change the conformation of the protein by disrupting the bonds and forces responsible for holding it in its tertiary structure. Consequently, the enzyme is no longer able to bind the substrate at the active site and its catalytic activity is diminished. This explains the shape of the curve shown in Figure 13.22.

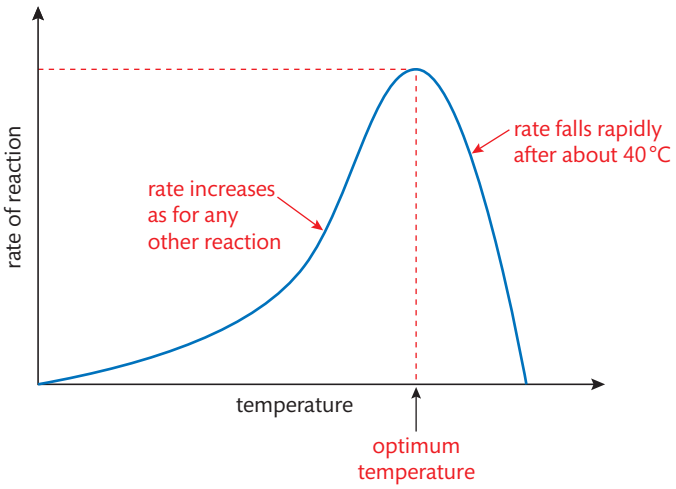
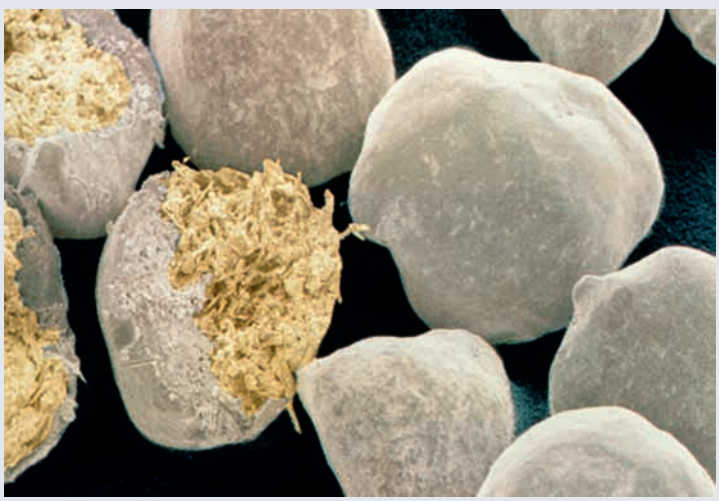


Figure 13.22 The effect of temperature on the activity of an enzyme.

The temperature corresponding to the maximum rate of reaction for a particular enzyme is known as its **optimum temperature**. Most enzymes in the human body have an optimum value close to 37 °C, which is body temperature. Organisms that are adapted to very different environments, such as bacteria in hot springs or algae on glaciers, produce enzymes with optimum values closer to their ambient temperatures.



Coloured scanning electron micrograph of granules of biological detergents showing the enzymes encapsulated in their structure. The enzymes help break down biomolecules such as those in blood, which may stain clothing. These enzymes are often the result of protein engineering with different optimum temperatures. Synthetic enzymes are also used in biodegradable plastics and in the textiles and food industries.

The effect of temperature on enzyme activity
Full details of how to carry out this experiment with a worksheet are available online.

Note that loss of the tertiary structure, which is known as **denaturation**, does not mean loss of the covalent backbone of the protein molecules, which is known as **digestion**. Nonetheless, denaturation is usually an irreversible process, as you might know from the impossibility of uncooking an egg! Lowering the temperature usually causes what is called **deactivation** of an enzyme rather than denaturation. This prevents the enzyme from working but as it does not change the tertiary structure, it is usually reversible. This is why, for example, food that has been preserved by freezing soon spoils on thawing due to resumption of microbial activities as the temperature rises.

The effect of temperature on enzymes and other proteins is one of the main reasons why controlling the temperature is so important to many organisms, and why a change in the human core temperature of a couple of degrees Celsius is usually fatal.

2 pH

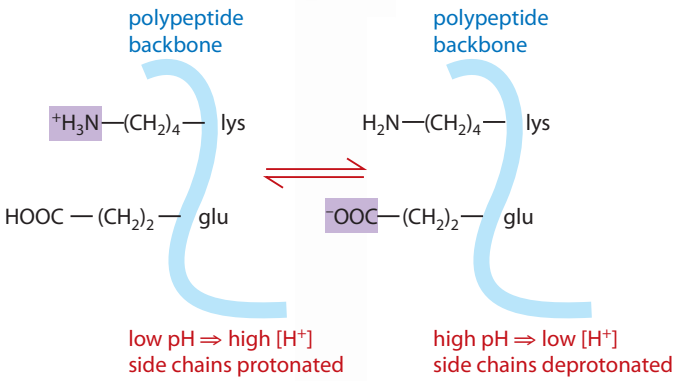


Figure 13.23 The influence of pH on the state of ionization of acidic and basic groups in the side chains of a protein. For simplicity the diagram shows side chains of only two amino acids, lysine (lys) and glutamic acid (glu). In lysine R = $-(CH_2)_4-NH_2$ and in glutamic acid R = $-(CH_2)_2-COOH$.

The pH of a solution has several effects on the activity of an enzyme. It directly influences the state of ionization of acidic or basic groups in the side chains of the protein. From Le Chatelier's principle we can predict that at low pH (high $[H^+]$), acidic and basic groups become protonated, and at high pH (low $[H^+]$) they become deprotonated. The precise influence of pH depends on the pK_a and pK_b values of the acidic and basic groups, and so varies in different enzymes. These changes in ionization within the protein structure alter the attractive forces stabilizing the enzyme, and crucially affect its shape and its ability to recognize the substrate.

More specifically, changes in the ionization of side chains of amino acids at the active site will affect the ability of the enzyme to form a complex with the substrate. In most cases there is a clear optimum value for pH at which the enzyme has maximum activity. The optimum pH value for an enzyme gives a clue to the likely amino acids at its active site – typically, enzymes show maximum activity when the pH is close to the pK_a value of the groups at its active site. Unlike for temperature, different enzymes in the same organism can have a wide variation in the optimum values of their pH, as shown in Figure 13.24.

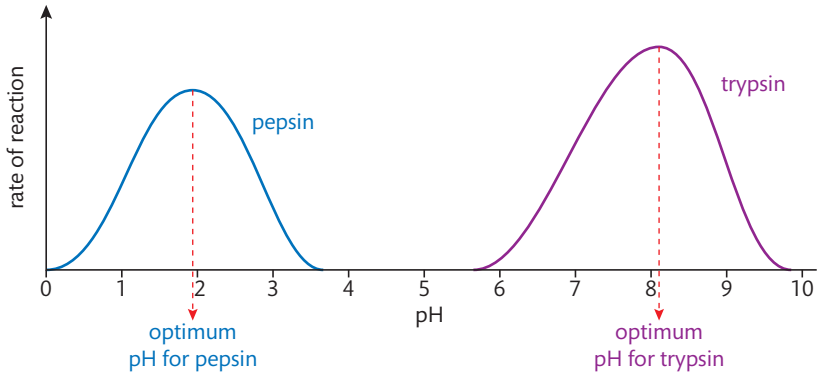


Figure 13.24 The effect of pH on two different digestive enzymes in the human body.

These differences in the pH optima of different enzymes can be an important way of controlling enzyme activity. In the example shown in Figure 13.24, the enzyme pepsin

is most active in the stomach, where the pH is very low. However, it becomes inactive once it is moved with the digested food into the more alkaline environment of the intestine, where trypsin is active. Their optimum values correspond to the pH in the different regions of the digestive tract where they are active. Extremes of pH denature an enzyme in much the same way as described for high temperature; an egg which has been dropped into strong acid looks 'cooked' in the same way as if it had been heated, due to changes in the tertiary structure of its protein.

The optimum values of temperature and pH for an enzyme refer to the conditions where the enzyme and substrate are maximally able to bind and form a complex.

3 Heavy metal ions

Heavy metals such as lead, copper, mercury, and silver are poisonous, primarily due to their effects on enzymes. When these metals are present as positive ions, they react with sulfhydryl groups, –SH, in the side chains of cysteine residues in the protein, forming a covalent bond with the sulfur atom and displacing a hydrogen ion. This disrupts the folding of the protein, and may change the shape of the active site and its ability to bind substrate. This is a form of inhibition, discussed below.

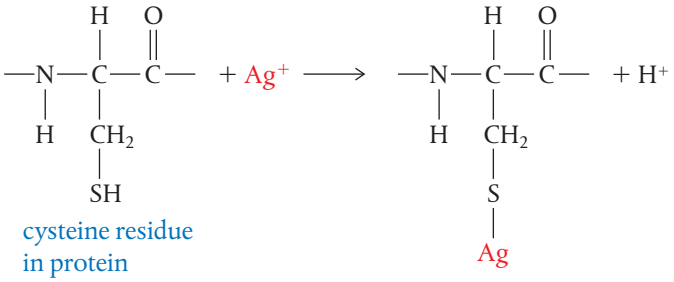


Figure 13.25 Enzyme inhibition by Ag^+ ions. Binding of the metal to a sulfhydryl group in the amino acid cysteine may change the conformation of the enzyme at the active site.

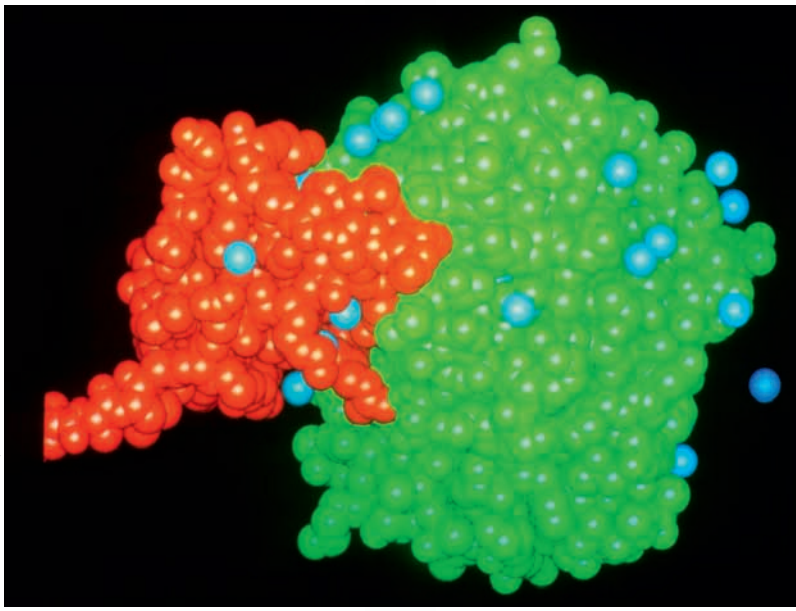
Inorganic dyes used to darken hair, such as 'Grecian Formula', contain a soluble lead salt which interacts with the sulfur atoms in the hair protein keratin to produce lead sulfide, PbS, which is black. The amount of darkening depends on the proportion of cysteine residues in the hair. Ancient Greeks used to line their waterways with lead and it was noticed that people who bathed often in this water had darker hair. Lead poisoning from the drinking water is believed to have contributed to the decline of both the Greek and the Roman empires. Concerns over the toxicity of lead salts have caused these hair dyes to be banned in many countries.

Chemical inhibitors modify enzyme activity

Inhibitors are chemicals that bind to enzymes and decrease their catalytic activity. Enzyme inhibition is an important aspect of the regulation of enzyme activity. Biochemical research using inhibitors has helped to provide a lot of information on details of enzyme action and metabolic pathways, and often plays a central role in drug design.

There are two main types of inhibitor, distinguished by where they bind to the enzyme:

- competitive inhibitors that bind at the active site
- non-competitive inhibitors that bind at a different location on the enzyme.



Computer graphics model of the digestive enzyme chymotrypsin, shown in green, with an inhibitor, shown in red. The small blue objects are water molecules. Chymotrypsin's action is to break down protein molecules, but the inhibitor prevents this from happening by binding at the active site.

1 Competitive inhibitors

Chemicals that bind at the active site of the enzyme are known as **competitive inhibitors** because they compete with the substrate for the binding position. They usually have a chemical structure similar to that of the substrate, so in a sense they mimic its ability to bind. But once they are bound they do not react to form products, instead blocking the active site and making it unavailable to the substrate.

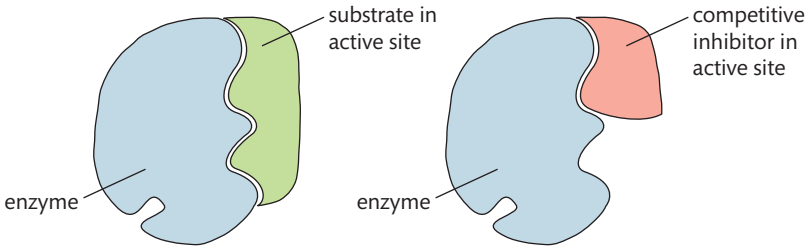


Figure 13.26 Binding of a competitive inhibitor at the active site of an enzyme.

Increasing the concentration of substrate reduces the extent of inhibition as relatively fewer of the inhibitor molecules are able to bind. In this type of inhibition V_{max} is not altered as there is still a substrate concentration where full activity of the enzyme can be achieved. But as it takes a higher substrate concentration to reach this rate, K_m is increased. These relationships are shown in Figure 13.27.

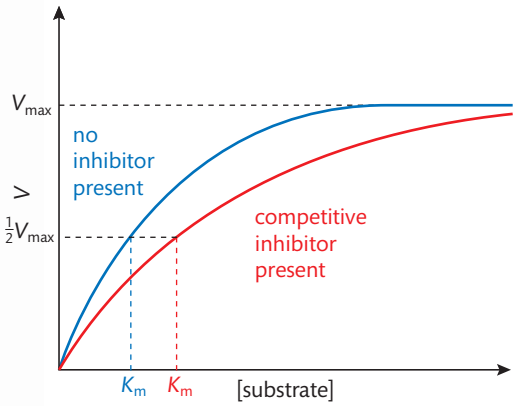


Figure 13.27 Effect of a competitive inhibitor on the rate of an enzyme-catalysed reaction. V_{max} is not altered by the inhibitor but K_m is increased.

A good example of a competitive inhibitor is malonate inhibiting the enzyme succinate dehydrogenase. As its name suggests, the enzyme acts to remove hydrogen from succinate, which converts it into fumarate. This reaction occurs during aerobic respiration. The structures of the substrate succinate and the inhibitor malonate are sufficiently similar for them both to be able to bind at the same active site.

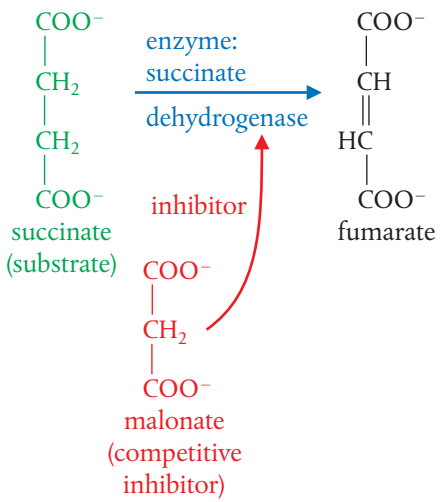


Figure 13.28 The action of malonate as a competitive inhibitor for the enzyme succinate dehydrogenase. It possesses the correct chemical groups to be able to compete with succinate for binding at the active site.

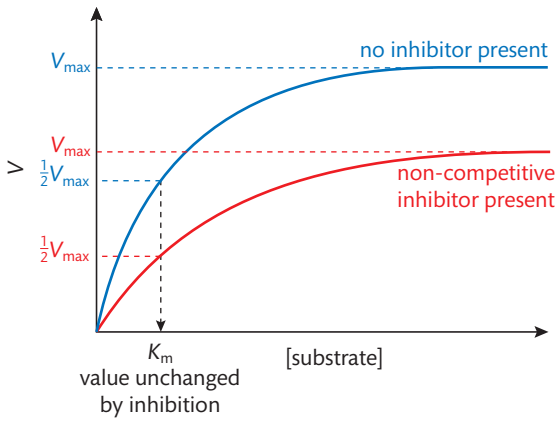
CHALLENGE YOURSELF

6 What are the IUPAC names for succinate, fumarate and malonate?

2 Non-competitive inhibitors

Chemicals that bind to the enzyme at a position other than the active site are known as **non-competitive inhibitors**. Their binding causes a conformational change in the protein structure that alters the active site, inhibiting its ability to bind to the substrate. The term **allosteric site** refers to the binding site of the non-competitive inhibitor on the enzyme.

Increasing the concentration of substrate will not reduce the extent of this type of inhibition as the enzymes have effectively been decommissioned by the inhibitor and are unavailable. V_{max} is decreased and cannot be restored no matter how high the substrate concentration. But, as is shown in Figure 13.30, the value of K_m is unchanged because the uninhibited enzymes are perfectly functional.



As non-competitive inhibitors do not bind at the active site, they do not necessarily possess similar chemical features to the substrate.

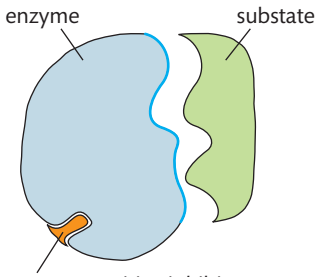


Figure 13.29 Binding of a non-competitive inhibitor to an enzyme changes the active site so that the substrate cannot bind.

Figure 13.30 Effect of a non-competitive inhibitor on the rate of an enzyme-catalysed reaction. V_{max} is reduced by the inhibitor but K_m is unchanged.

Allosteric inhibition is not covered here. This is a broader term which can include both competitive and non-competitive inhibition, as well as negative and positive modulators. Here we refer to the allosteric site only in the context of the binding site of non-competitive inhibitors.

Competitive inhibitors bind *at* the active site of the enzyme, non-competitive inhibitors bind *away* from the active site.

Summary

The table below summarizes some of the main differences between competitive and non-competitive inhibitors.

	Competitive inhibition	Non-competitive inhibition
binding site on enzyme	binds at active site	binds at allosteric site
effect on V_{max}	not affected	decreased
effect on K_m	increased	not affected

We have seen that metal ions such as Ag^+, Hg^{2+}, Pb^{2+}, and Cu^{2+} can act as non-competitive inhibitors of enzyme activity, and many poisons such as DDT and cyanide behave similarly. Antibiotics, such as penicillin (described in Chapter 15), kill bacteria by inhibiting one of their key enzymes. Many anticancer drugs also work in this way, in order to block cell division in the tumour.

In healthy cells, enzyme inhibition is often an important means of controlling metabolic activity. One example is that the product of a reaction sometimes acts as an inhibitor of the enzyme for its synthesis, thereby setting up a feedback loop regulating its own concentration. This is known as **product inhibition**.

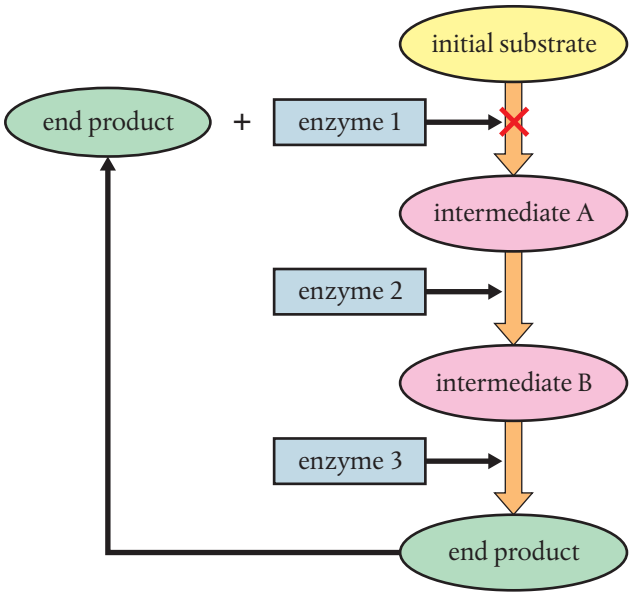


Figure 13.31 Product inhibition in a biochemical pathway. The end product of the reaction acts as an inhibitor for the first enzyme leading to its production, so as its concentration builds up it effectively switches off its own synthesis. The synthesis of many amino acids is regulated in this way.

Inhibitors are described as **irreversible** when their binding to the enzyme is permanent, usually because it involves the formation of a covalent bond. Cyanide is an example of an irreversible inhibitor of the enzyme cytochrome oxidase, which catalyses the reduction of oxygen to water in the last step of aerobic respiration.

Exercises

10 **(a)** State four characteristics of enzymes.
(b) Sketch a graph of the rate of an enzyme reaction against temperature and explain its shape.

11 **(a)** Draw a graph of the rate of an enzyme reaction against substrate concentration.
(b) On the same axes sketch how this graph would change in the presence of a competitive inhibitor.
(c) State the effect of a competitive inhibitor on the values of:
(i) V_{max}
(ii) K_m

12 Compare enzymes and inorganic catalysts, with reference to two similarities and three differences.

Analysis of proteins

Analysis of protein content is a major aspect of research in biochemistry. For example, clinical analysis of protein levels in body fluids can be used to confirm pregnancy or can be used to help diagnose a variety of diseases. Food analysts are interested in knowing the total concentration and type of proteins in foods, and the pharmaceutical industry often uses altered protein levels to test the efficacy of a drug.

Many different approaches and techniques are used in protein analysis. We will consider here two main aspects of this work:

- analysis of the amino acid composition of an isolated protein
- analysis of the protein concentration of a sample.

Analysis of the amino acid composition of a protein

The first step in the analysis is to break the peptide bonds between the amino acids in the protein structure through **hydrolysis** reactions, usually using acid. These reactions reverse the condensation reactions discussed on page 675, and occur commonly in

cells during enzyme-catalysed protein digestion. (Note that determination of the amino acid composition of a protein is not the same as knowing its primary structure, as the *sequence* of the amino acids will not be known.)

Separation of the resulting amino acid mixture into its components can then be achieved in two ways.

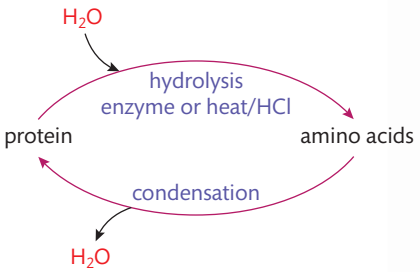


Figure 13.32 Hydrolysis of protein produces amino acids; condensation of amino acids produces proteins.

1 Chromatography

Chromatography is a useful technique for separating and identifying the components of a mixture. The basic principle is that the components have different affinities for two phases, a **stationary phase** and a **mobile phase**, and so are separated as the mobile phase moves through the stationary phase. As amino acids are colourless in solution they are usually treated with a **locating reagent** at the end of the process which makes them take on colour to aid their identification.

Paper chromatography is an example of partition chromatography, in which the components are separated on the basis of their different solubilities in the two phases. It is used mainly for qualitative analysis. The paper contains about 10% water, and this is the stationary phase. Water is adsorbed by forming hydrogen bonds with the −OH groups in the cellulose of the paper. The solvent is the mobile phase as it rises up the paper by capillary action. As it does so it dissolves the components of the mixture to different extents, so carrying them at different rates.

When measuring the distances to calculate R_f values, be careful to measure from the origin and not from the bottom of the paper. Similarly, the solvent moves as far as the solvent front, not to the top of the paper.

The procedure is simple to run. A small sample of the amino acid mixture is spotted near the bottom of the chromatographic paper, and this position, known as the **origin**, needs to be clearly marked (in pencil so as not to interfere with the experiment). The paper is then suspended in a chromatographic tank containing a small volume of solvent, ensuring that the spot is above the level of the solvent.

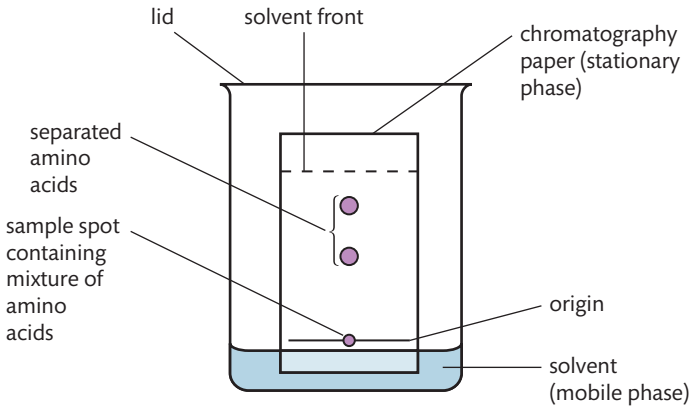


Figure 13.33 Apparatus used to separate amino acids by paper chromatography.

As the solvent rises up the paper it will pass over the spot. Amino acids in the spot will distribute themselves between the two phases and so move up the paper at different speeds. As a result they become spread out according to their different solubilities. When the solvent reaches almost to the top of the paper, its final position is marked and is known as the **solvent front**. The paper is removed from the tank and developed by spraying it with the locating reagent **ninhydrin**. Most amino acids will now appear

Separation of amino acids by chromatography
Full details of how to carry out this experiment with a worksheet are available online.

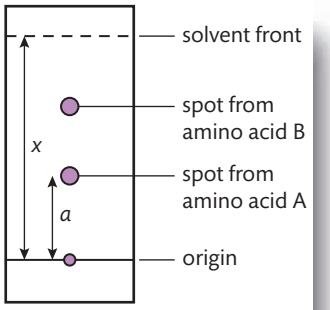


Figure 13.34 Calculation of R_f values in chromatography.

purple and can be distinguished as separate isolated spots up the length of the paper. The final result is known as a **chromatogram**.

The position of each amino acid on the chromatogram can be represented as an R_f value (retention factor), which is calculated as shown below.

$$R_f = \frac{\text{distance moved by amino acid}}{\text{distance moved by solvent}}$$

So, for amino acid A, $R_f = \frac{a}{x}$

Specific amino acids have characteristic R_f values when measured under the same conditions, so can be identified by comparing the values obtained with data tables. It is helpful to spot known amino acids alongside the mixture to act as markers for the experiment.

In chromatography, the R_f value can be used to identify the components of a mixture.

$$R_f = \frac{\text{distance moved by amino acid}}{\text{distance moved by solvent}}$$

2 Electrophoresis

Electrophoresis is a technique for the analysis and separation of a mixture based on the movement of charged particles in an electric field. As we saw on page 682, amino acids carry different charges depending on the pH, and these differences can be exploited to separate them when placed in a buffered solution at a particular pH.

- When the pH is equal to their isoelectric point, amino acids will not move as they carry no net charge.
- When the pH > their isoelectric point, the amino acids exist as anions and move to the anode.
- When the pH < their isoelectric point, the amino acids exist as cations and move to the cathode.

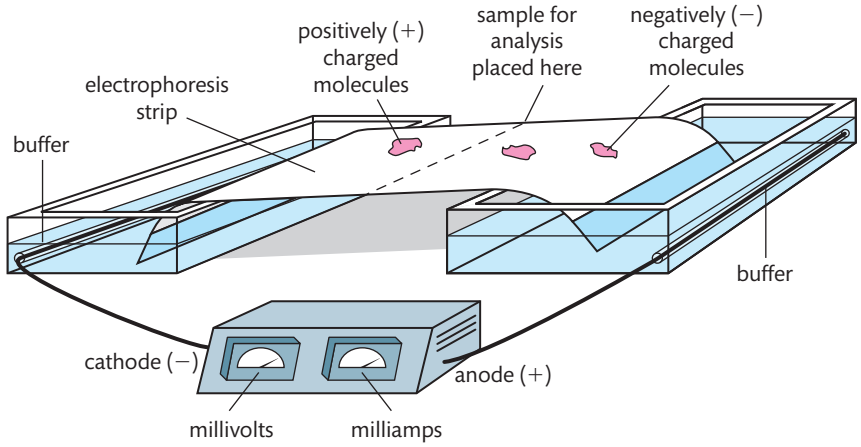


Figure 13.35 Electrophoresis apparatus for amino acid or protein separation.

The rate of movement of the ions will depend on the number of charges on the ion and the molecular mass – smaller, more highly charged ions migrate faster. The voltage used and the temperature also affect the rate of movement.

In gel electrophoresis, the medium is a gel, typically made of polyacrylamide. The amino acid mixture is placed in wells in the centre of the gel and a voltage is applied. Depending on the pH of the buffer used, different amino acids will move at different rates and in different directions towards the oppositely charged electrodes. When

separation is complete they can be detected with a stain such as ninhydrin or made to fluoresce under UV light, and identified by comparison with known samples or from data tables.

Electrophoresis can also be used to separate and identify intact proteins according to their different rates of migration towards the poles.

Worked example

Gel electrophoresis was carried out using a buffer at pH 6.0 on a mixture of glycine, alanine, lysine, glutamic acid, and aspartic acid. Use section 33 of the IB data booklet to identify which amino acids

(a) will remain close to the origin
(b) will move fastest towards the anode
(c) will move fastest towards the cathode.

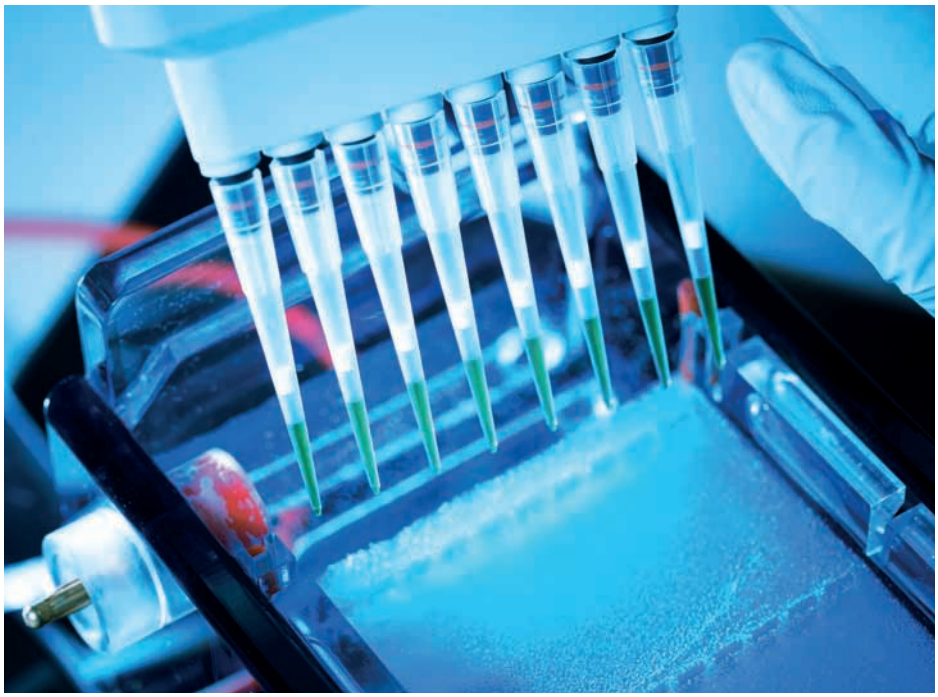
Solution

From their isoelectric points, at pH 6.0 the amino acids will exist as follows:

glycine: neutral; alanine: neutral; lysine: cation; glutamic acid: anion; aspartic acid: anion

Therefore:

(a) glycine and alanine will not move
(b) glutamic acid and aspartic acid will move towards the anode, and aspartic acid will move more quickly as it is a smaller molecule
(c) lysine will move towards the cathode.



Multi-pipette used to place protein samples into agarose gel in a buffer solution for electrophoresis. The protein molecules will separate according to their different attractions to opposite electrical poles across the polyacrylamide gel (polyacryl amide gel electrophoresis or PAGE). This technique is used to detect whether proteins associated with particular diseases are found in the blood.

Determination of the pH of buffer solutions

Electrophoresis and other aspects of experimental work used in protein studies and analysis require the use of buffer solutions. Buffers resist the change in the pH of a solution on the addition of small amounts of acid or base, and can be prepared to operate at a wide range of pH. The composition and mode of action of buffers was discussed in Chapter 8, and you might find it useful to review this first before going on to this section on the calculation of the pH of a buffer solution.

The pH of a buffer solution, that is its H^+ concentration, will depend on the interactions among its components. We will consider here an acidic buffer made of the generic weak acid HA and its salt MA.

The equilibria that exist in the buffer will be:

$$HA(aq) \rightleftharpoons H^+(aq) + A^-(aq)$$

$$MA(aq) \rightarrow M^+(aq) + A^-(aq)$$

We can make two approximations, based on some assumptions about these reactions, which will help to make the calculations easier.

1 The dissociation of the weak acid is so small that it can be considered to be negligible. So we can make the approximation

$$[HA]_{initial} \approx [HA]_{equilibrium}$$

2 The salt is considered to be fully dissociated into its ions. So we can approximate

$$[MA]_{initial} \approx [A^-]_{equilibrium}$$

The equilibrium expression for the acid is

$$K_a = \frac{[H^+][A^-]}{[HA]}$$

Therefore $[H^+] = K_a \frac{[HA]}{[A^-]}$

remembering that all values in this expression must be *equilibrium* concentrations. But from the approximations justified above, we know that $[HA]_{equilibrium} \approx [HA]_{initial}$ and $[A^-]_{equilibrium} \approx [MA]_{initial}$, so we can substitute these values as follows:

$$[H^+] = K_a \frac{[HA]_{initial}}{[MA]_{initial}}$$

which is usually given as

$$[H^+] = K_a \frac{[acid]}{[salt]}$$

By taking the negative logarithms of both sides of the equation, we can derive:

$$pH = pK_a + \log_{10} \frac{[salt]}{[acid]}$$

For basic buffer solutions the equivalent equations are:

$$[OH^-] = K_b \frac{[base]}{[salt]} \text{ and } pOH = pK_b + \log_{10} \frac{[salt]}{[base]}$$

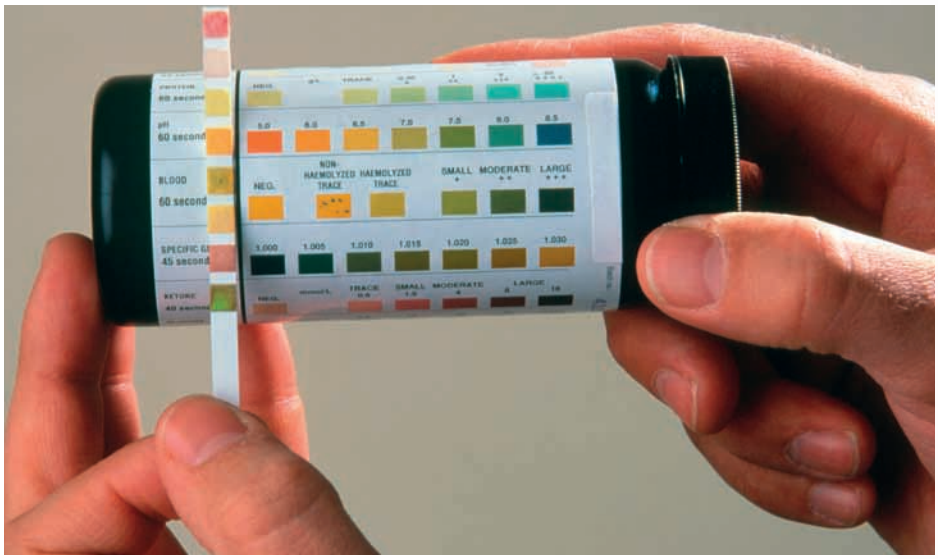
Joseph Henderson (1878–1942) was an American biochemist who developed equations showing that acid–base balance in the blood is regulated by buffers. Karl Hasselbalch (1874–1962), a Danish chemist and a pioneer in the use of pH measurement in medicine, converted the equations to logarithmic form in his work on studying acidosis in the blood. We now know that different buffers in the blood work together to keep the pH tightly controlled at 7.4. Any fluctuation in this value is so crucial that pH levels below 7.0 (acidosis) and above 7.8 (alkalosis) are, in the words of the medical profession, 'incompatible with life'.

These equations, known as the Henderson–Hasselbalch equations, are given in section 1 of the IB data booklet. The beauty of these expressions is that they enable us to know the pH of a buffer solution directly from:

- the K_a or K_b values of its component acid or base and
- the ratio of initial concentrations of acid and salt used to prepare the buffer.

$$pH = pK_a + \log_{10}\frac{[salt]}{[acid]}$$

$$pOH = pK_b + \log_{10}\frac{[salt]}{[base]}$$



Urine analysis using a multiple test stick against a reference chart. The pads test for pH as well as for various components of the urine. Body fluids such as urine are effectively buffered for pH, so alterations in the urine pH can provide insights into medical conditions and disease.

Worked example

Calculate the pH of a buffer solution at 298 K, prepared by mixing 25 cm^3 of 0.10 mol dm^{-3} ethanoic acid, CH_3COOH, with 25 cm^3 of 0.10 mol dm^{-3} sodium ethanoate, $Na^+CH_3COO^-$. K_a of $CH_3COOH = 1.8 \times 10^{-5}$ at 298 K.

Solution

pK_a of $CH_3COOH = -\log_{10}(1.8 \times 10^{-5}) = 4.74$

As there are equal volumes and concentrations of CH_3COOH and $NaCH_3COO$, then [acid] = [salt].

$$pH = pK_a + \log_{10}\frac{[salt]}{[acid]} = 4.74 + \log_{10}(1) = 4.74 + 0 = 4.74$$

Note that $\log_{10}(1) = 0$

This example shows that when a buffer solution contains equal amounts in moles of acid and salt (or base and salt), the last term in the Henderson–Hasselbalch expression becomes zero, so $pH = pK_a$ (or $pOH = pK_b$). This relationship is extremely useful when it comes to preparing buffers of a specified pH. All we have to do is choose an acid with a pK_a value close to the required pH and then, if necessary, adjust the concentrations of acid and salt accordingly.

In a buffer solution:
when [acid] = [salt], $pH = pK_a$;
when [base] = [salt], $pOH = pK_b$

The buffer solution can be prepared by reacting the acid with enough strong alkali to convert one half of it into salt, as described on page 381.

Worked example

How would you prepare a buffer solution of pH 3.75 starting with methanoic acid (HCOOH) and NaOH?

Solution

From the IB data booklet we have pK_a (HCOOH) = 3.75, so a buffer with equal amounts in moles of this acid and its salt NaHCOO will have pH = 3.75.

This equimolar solution is prepared by reacting the acid with enough NaOH, so that one half of it is converted into salt and therefore [HCOOH] = [$HCOO^-$].

Alternatively, the buffer can be prepared by mixing the acid directly with an appropriate amount of its salt.

Worked example

How much 0.10 mol dm^{-3} butanoic acid solution and solid potassium butanoate should be used to make 1.00 dm^3 of pH 5.00 buffer solution? State the assumptions made in the calculation.

Solution

From the IB data booklet butanoic acid has $pK_a = 4.83$.

$$pH = pK_a + \log_{10}\frac{[A^-]}{[HA]}$$

$$\therefore\ 5.00 - 4.83 = \log_{10}\frac{[\text{butanoate ion}]}{0.10\ \text{mol dm}^{-3}}$$

taking antilogs of both sides give

$$10^{0.17} = \frac{[\text{butanoate ion}]}{0.10\ \text{mol dm}^{-3}} = 1.5$$

$$\therefore\ [\text{butanoate ion}] = 0.15\ \text{mol dm}^{-3}$$

The molar mass of potassium butanoate is 126.12 g mol^{-1}

$$\therefore\ 1.00\ \text{dm}^3 \text{ of } 0.15\ \text{mol dm}^{-3} \text{ solution} = 0.15\ \text{mol} \times 126.12\ \text{g mol}^{-1} = 19\ \text{g}$$

So 19 g potassium butanoate should be added to 1.00 dm^3 of 0.10 mol dm^{-3} butanoic acid.

The following assumptions were made:

- $[\text{butanoate ion}]_{\text{equilibrium}} = [\text{potassium butanoate}]_{\text{initial}}$
- $[\text{butanoic acid}]_{\text{equilibrium}} = [\text{butanoic acid}]_{\text{initial}}$
- no volume change occurs on mixing the solution.

Analysis of the protein concentration of a sample

Protein assays refer to investigative procedures used to measure the concentration of protein in a sample. A technique commonly used in protein assays is **UV-visible spectroscopy**, often known as **UV-vis**.

As with other forms of spectroscopy, UV-vis depends on the fact that molecules interact with different parts of the electromagnetic spectrum according to their chemical composition. In this case we are interested in their interaction with visible and ultraviolet light. This radiation has the range approximately 180–750 nm, and has sufficient energy to excite the electrons in the occupied higher energy levels in complex molecules, such as proteins. When a full range of wavelengths of UV-vis radiation is passed through a sample, an **absorption spectrum** is obtained in which each characteristic absorption corresponds to an electronic transition. As seen in Figure 13.36, absorption spectra show wavelength (λ) on the horizontal axis and the intensity of absorption (*A*) on the vertical axis.

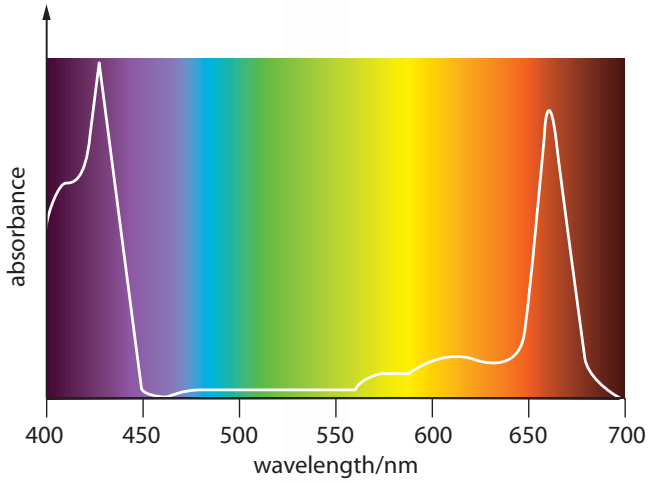


Figure 13.36 Absorbance spectrum of chlorophyll, showing absorption peaks at approximately 430 nm and 660 nm, the red and blue parts of the spectrum. Green light is barely absorbed and so is transmitted, giving the colour that we see. Chlorophyll is a pigment molecule, associated with proteins in plant cells. It is discussed further on page 740.

Substances that appear coloured absorb certain wavelengths of light in the visible region and transmit the remaining wavelengths. This is why chlorophyll appears green, as we see in Figure 13.36. When the energy needed to excite an electron is in the ultraviolet region of the spectrum and all visible light is transmitted, the substance appears colourless.

Absorption spectra are obtained from a **spectrophotometer**, which is commonly used as a data-logging device. The main parts of a spectrophotometer are shown in Figure 13.37.

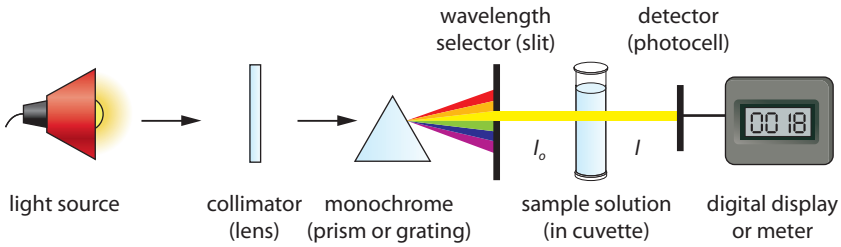


Figure 13.37 Simplified diagram of a spectrophotometer. The desired range of wavelength of light is selected and passed through the solution of a sample in the cuvette. The photometer detects the light that is absorbed and sends a signal to a digital display.

UV-vis absorption spectra are useful for quantitative measurements, such as protein assays. Generally the wavelength of maximum absorbance is selected for the analysis. The amount of light absorbed at this wavelength (the absorbance, *A*) is given by the relationship $\log_{10}(I_0/I)$ where:

I_0 = the intensity of light before passing through the sample

I = the intensity of light after passing through the sample

The absorbance depends on:

- the molar absorptivity, ε, defined as the absorbance of a 1.00 mol dm^{-3} solution in a 1.00 cm cell at a specified wavelength
- the concentration of the solution, *c*
- the path length, *l*.

This can be expressed as:

$$\log_{10}(I_0/I) = \varepsilon l c$$

This is known as the Beer–Lambert law and is given in section 1 of the IB data booklet.

The Beer–Lambert law expresses the linear relationship between the absorbance and concentration of a compound at a fixed wavelength.

Biuret reagent showing the presence of protein in the right-hand flask. The biuret reagent causes peptide bonds in the protein to form a complex with copper(II) ions in alkaline solution, giving the solution a deep violet colour. The intensity of the colour increases with the concentration of protein present, and can be analysed by UV-vis spectroscopy.

It follows from this relationship that the absorbance of a compound at a fixed wavelength is directly proportional to its concentration. This can be used as the basis of determining the concentration of protein in a sample.

Protein analysis by UV-vis typically involves first reacting the sample with a reagent that generates a colour change which is dependent on the amount of protein present. This will promote absorption in the UV-vis range. Several different reagents can be used which react with different groups within the protein molecules, such as peptide bonds, aromatic side-groups, or basic groups. A common process uses the biuret reagent, which generates a purple colour by reaction with peptide bonds.

A relatively easy way to convert absorbance data into concentration is by use of a **calibration curve**. This is based on **standard solutions**, those with a known concentration of protein, which are prepared to cover a range of concentrations on either side of the value being investigated. Typically, at least five dilutions of a standard of known concentration are prepared. These are then each treated with the appropriate reagent to generate colour. Their absorbance is measured at the selected wavelength and plotted on a graph to obtain the calibration curve, as shown in Figure 13.38.

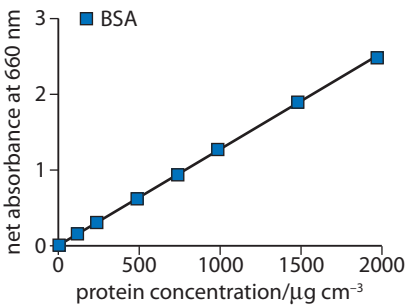


Figure 13.38 Calibration curve using the protein standard BSA, bovine serum albumin.

The absorbance of the solution to be analysed is then measured at the same wavelength as the standard solutions, and its protein concentration determined from the calibration curve, as shown in Figure 13.39. Where calibration curves have been generated by data-logging, appropriate software can often be used for this calculation. It is important that the sample and standard solutions are treated in the same way with respect to the type and amount of reagent added, buffer solutions used, and the temperature.

Although we refer to calibration *curves*, the best results are obtained by interpolation on the *linear* part of the graph. Ideally, several calibration standards should be used and a regression function fitted to the data.

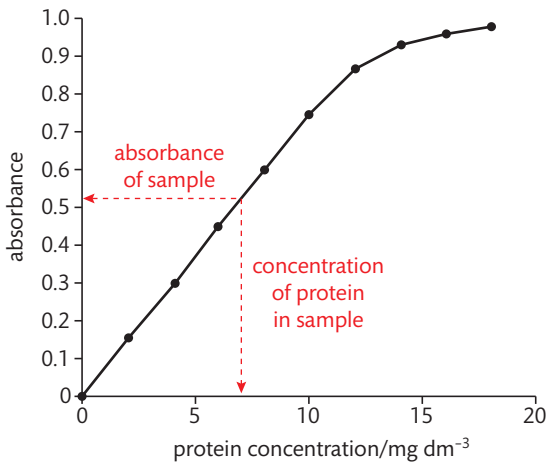


Figure 13.39 A calibration curve used to find the concentration of a sample. The graph is linear at low concentrations, in accordance with the Beer–Lambert law.

Worked example

The following tubes were set up using pipettes to measure volume. Tubes 1–6 contained dilutions of the protein standard solution bovine serum albumin (BSA) with an initial concentration of 5.00 mg protein cm^{-3}. The last row in the table refers to a tube of protein solution of unknown concentration. The absorbance of all tubes was measured at 540 nm.

Tube number	Volume BSA / cm^3	Volume distilled H_2O / cm^3	Volume reagent / cm^3	Conc. protein / mg cm^{-3}	Absorption at 540 nm
1	0.00	2.00	2.00		0.00
2	0.10	1.90	2.00		0.34
3	0.30	1.70	2.00		0.67
4	0.50	1.50	2.00		1.02
5	0.70	1.30	2.00		1.35
6	1.00	1.00	2.00		1.65
2.00 cm^3 protein sample		0.00	2.00		1.11

(a) Complete the column of conc. protein (mg cm^{-3}) for tubes 1–6 in the table.

(b) Plot a calibration curve of absorbance versus concentration for tubes 1–6.

(c) Use the curve to calculate the protein concentration in the unknown sample.

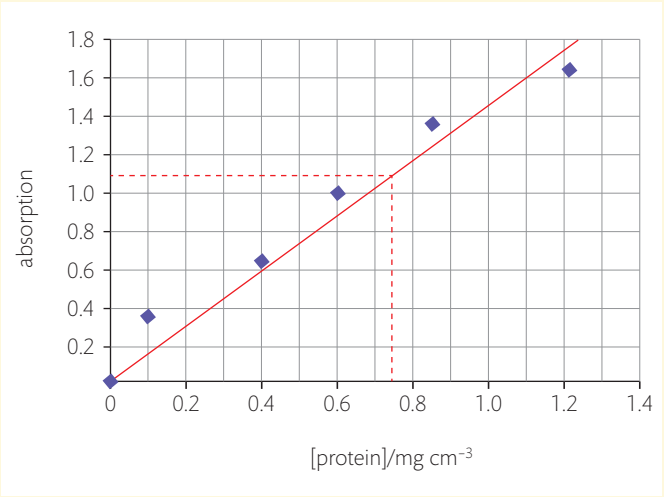
Solution

(a) $$\text{conc. protein (mg protein cm}^{-3}) = \frac{\text{conc. BSA (mg protein cm}^{-3}) \times \text{vol BSA (cm}^3)}{\text{total volume in tube (cm}^3)}$$

$$\therefore \text{conc. protein} = \frac{5.00 \text{ (mg protein cm}^{-3}) \times \text{vol BSA added (cm}^3)}{4.00 \text{ (cm}^3)}$$

Tube number	Conc. protein / mg cm^{-3}
1	0.000
2	0.125
3	0.375
4	0.625
5	0.875
6	1.250

(b)



(c) Absorbance of 1.11 corresponds to protein concentration 0.750 mg cm^{-3} in the sample tested.

Exercises

13 (a) When carrying out paper chromatography on a sample of hydrolysed protein, explain why it is important to avoid handling the paper.
Make reference to section 33 in the IB data booklet to answer parts (b) and (c).

(b) Explain why in gel electrophoresis the amino acid isoleucine migrates towards the anode at high pH and towards the cathode at low pH.

(c) You are attempting to separate a mixture of glutamic acid and histidine by gel electrophoresis. Give a suggested pH for an appropriate buffer solution to use, and say in which direction each amino acid would migrate.

14 Given 100 cm^3 of 0.05 mol dm^{-3} methanoic acid, HCOOH, what volume of 0.05 mol dm^{-3} NaOH would you need to add to prepare a buffer of pH 4.23? (Assume K_a of HCOOH = 1.77×10^{-4}.)

15 Explain the use of protein standards in preparing a calibration curve for the calculation of protein concentration.

B.3 Lipids

Understandings:

- Fats are more reduced than carbohydrates and so yield more energy when oxidized.
- Triglycerides are produced by condensation of glycerol with three fatty acids and contain ester links. Fatty acids can be saturated, monounsaturated, or polyunsaturated.
- Phospholipids are derivatives of triglycerides.
- Hydrolysis of triglycerides and phospholipids can occur using enzymes or in alkaline or acidic conditions.
- Steroids have a characteristic fused ring structure, known as a steroidal backbone.
- Lipids act as structural components of cell membranes, in energy storage, thermal and electrical insulation, transport of lipid soluble vitamins, and as hormones.

Guidance

- *The structures of the fatty acids are given in the data booklet in section 34.*
- *Specific named examples of fats and oils do not have to be learned.*

Applications and skills:

- Deduction of the structural formulas of reactants and products in condensation and hydrolysis reactions between glycerol and fatty acids and/or phosphate.
- Prediction of the relative melting points of fats and oils from their structures.
- Comparison of the processes of hydrolytic and oxidative rancidity in fats with respect to the site of reactivity in the molecules and the conditions that favour the reaction.
- Application of the concept of iodine number to determine the unsaturation of a fat.
- Comparison of carbohydrates and lipids as energy storage molecules with respect to their solubility and energy density.
- Discussion of the impact of lipids on health, including the roles of dietary high-density lipoprotein (HDL) and low-density lipoprotein (LDL), cholesterol, saturated, unsaturated, and *trans* fat and the use and abuse of steroids.

Guidance

The structural differences between cis and trans fats are not required.

CHALLENGE YOURSELF

7 Even though the concept of oxidation number is of limited use in organic chemistry, predict how you would expect the oxidation state of carbon to compare in a simple carbohydrate and in a simple lipid molecule. See if you can check this with some simple examples from each group.

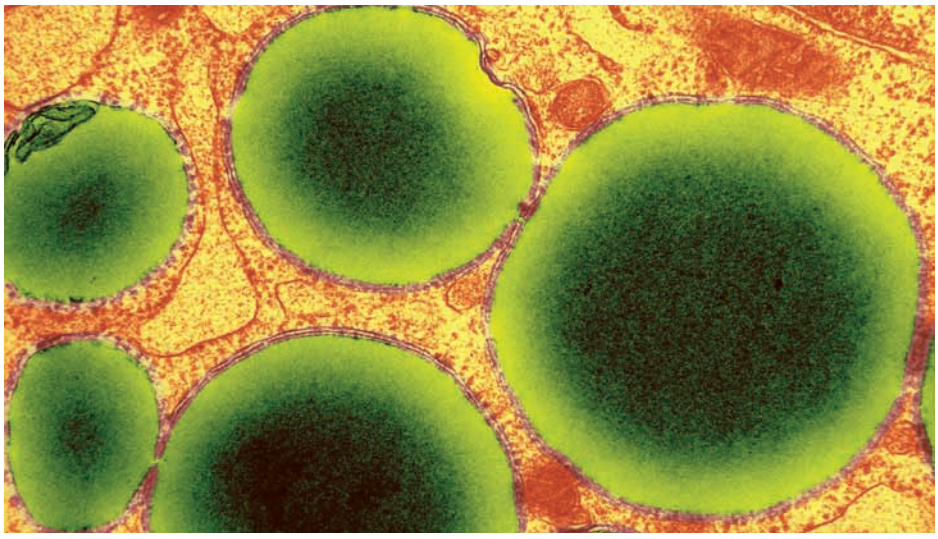
The term **lipid** is used as an umbrella term for a range of biomolecules such as fats and oils, steroids, and phospholipids. They are characterized by being hydrophobic or insoluble in water. They are, though, soluble in non-polar solvents, and this property is often used in extracting them from cells. Lipids contain the elements carbon, hydrogen, and oxygen, but the ratio of hydrogen to oxygen is greater than in carbohydrates – in other words they are more reduced molecules.

Functions of lipids

Lipids are essential molecules in a variety of roles in cells

Lipids contain stored energy that is released when they are broken down in the reactions of respiration in cells. Because lipids are more reduced than carbohydrates, they can effectively undergo *more* oxidation and so release *more* energy per unit mass when used as a respiratory substrate. The difference is significant: a gram of lipid releases almost twice as much energy as a gram of carbohydrate. However, partly due to their insolubility, the energy in lipids is not so readily available as it is in carbohydrates, as more reactions are involved in the breakdown. So the energy is released more slowly. The fat stores in animals, known as **adipose tissue** or **blubber**, serve as reservoirs of energy, swelling and shrinking as fat is deposited and withdrawn. Plants also sometimes store lipids for energy, for example as oils in seeds.

You are more likely to take a glucose tablet than suck a lump of cheese, for example, when you are running a marathon. But if you were going on an expedition to the Arctic, you would take lots of lipids like cheese and butter, as they make ideal storage molecules. These differences are because the energy in carbohydrates is more readily available, but lipids are a denser energy store.

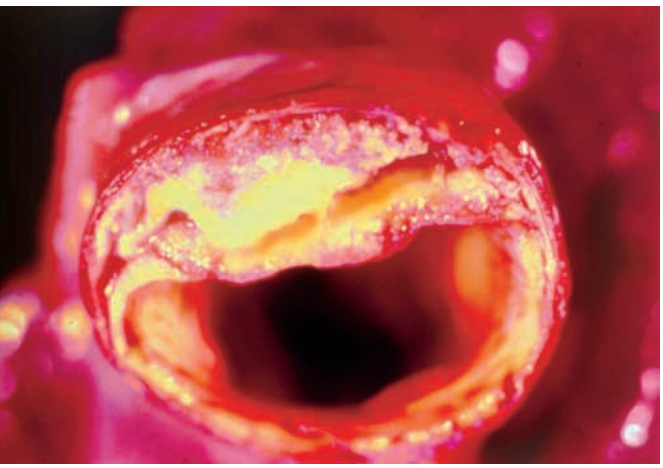


Electron micrograph of lipid droplets shown in green in a fat cell. These cells form adipose tissue, which stores energy as an insulating layer of fat.

Stored fat helps to protect some body organs, such as the kidneys, and a layer of fat under the skin acts as a thermal insulator. This is why animals that live in cold climates, such as seals and polar bears, have significant fat stores. Lipids also act as electrical insulators. In nerve cells a special layer of phospholipids called the myelin sheath gives electrical insulation to the nerves, and speeds up nervous transmission.

Some hormones, such as the sex hormones testosterone and estrogen, are made of lipids in the form of steroids. Bile acids, which aid digestion of fat in the intestine, are also steroid based. In addition, lipids help to absorb fat-soluble vitamins such as A, D, E, and K.

Lipids also play an important structural role. The phospholipids are a major component of membranes that enclose cells. Here they help to determine the selective transport of metabolites across cell boundaries. A different lipid molecule, cholesterol, is also important in plasma membrane structure, where it influences the fluidity and so the permeability of the membrane.



Magnified view of a slice through an artery showing a thick deposit caused by the disease atherosclerosis. The deposit is composed of a mixture of fats, cholesterol, and dead muscle cells. It disrupts blood flow and can break off in fragments blocking smaller blood vessels, leading to strokes and heart disease.

Excess lipids in the diet can have negative effects on health

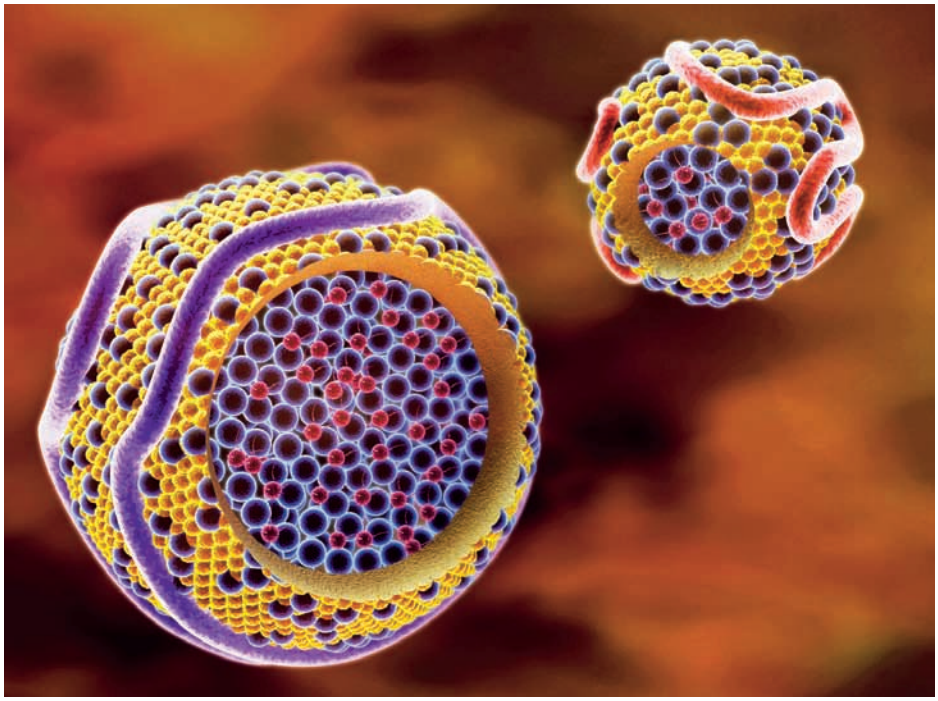
Excess lipids in the diet are increasingly linked to negative effects on health. These arise largely due to their low solubility that causes some lipids to be deposited in the walls of the main blood vessels. This can restrict blood flow, a condition known as **atherosclerosis**. It is usually associated with high blood pressure and can lead to heart disease.

In addition, because of the body's ability to convert excess fats into adipose tissue, a diet too rich in lipids can lead to the excess accumulation of body fat known as **obesity**. This is linked to many other health issues including diabetes and a variety of cancers. The molecule that is often the main culprit in circulatory diseases is **cholesterol**. It is present in the human diet, particularly from animal fat, and is also synthesized in the body. Because cholesterol is insoluble in blood, it is transported bound in different lipoproteins, the most well known of which are **LDL** (low-density lipoprotein) and **HDL** (high-density lipoprotein). These have gained the somewhat simplistic terms 'bad cholesterol' and 'good cholesterol' respectively. The names reflect the fact that high levels of LDL cholesterol are associated with increased deposition in the walls of the arteries, while high levels of HDL cholesterol seem to protect against heart attack. It is believed that HDL tends to carry cholesterol away from the arteries, so slowing its build-up. The main sources of LDL cholesterol are saturated fats and *trans* fats, the chemical nature of which is discussed in the next section.

Clearly the type of fat consumed is as important as the total amount. In general an intake of poly-unsaturated fats, such as those found in fish, many nuts, and corn oil, is considered beneficial in lowering levels of LDL cholesterol. Also, a type of fatty acid known as omega-3-poly-unsaturated fatty acid, found for example in fish oils and flax seeds, has been shown to be linked with reduced risk of cardiovascular disease as well as with optimum neurological development. These fatty acids cannot be manufactured by the body so are known as **essential fatty acids** and must be taken in the diet.

TOK

Debate exists as to whether people who need two airline seats because they are obese should have to pay for this themselves or have the second seat paid for as part of a disability coverage. To what extent should health problems perceived as resulting from life-style choices be distinguished from other forms of disability?



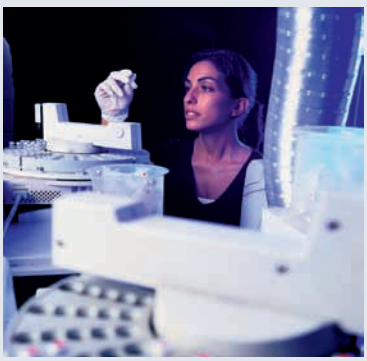
Computer artwork of LDL (right) and HDL (left) cholesterol, the major carriers of cholesterol in the blood. The purple spheres represent cholesterol molecules which are bonded to phospholipids, shown in yellow, and protein shown as purple and pink strands. LDL cholesterol is associated with increased risk of heart disease.

Uses and abuses of steroids

As noted above, lipids in the form of steroids are found in some hormones. Female steroid hormones are used in contraceptive pill formulations and in HRT (hormone replacement therapy) sometimes prescribed during menopause. Possible side-effects of these steroid treatments must be monitored and ongoing research is essential to provide data about long-term usage.

Male steroid hormones are collectively called androgens, of which **testosterone** is the most important. Medical uses of testosterone include treatment of disorders of the testes and breast cancer. These hormones are also known as **anabolic steroids** due to their role in promoting tissue growth, especially of muscles. Synthetic forms of them are used medically to help gain weight after debilitating diseases. They have also been used as **performance-enhancing drugs** by athletes in sports such as weight-lifting and cycling, as they can increase strength and endurance. Their use is banned by most sporting authorities for medical and ethical reasons.

Scientist centrifuging urine samples sent for anti-doping testing at the laboratories of the Italian National Olympic Committee. The samples are analysed for performance-enhancing drugs such as anabolic steroids, using techniques such as liquid chromatography/mass spectrometry. Drug testing like this has become a focus of national and international sporting authorities and uses increasingly accurate technology. Anabolic steroids can cause changes in secondary sexual characteristics resulting from hormone imbalances, and can be associated with increased risk of liver cancer.



The essential fatty acids linoleic acid and linolenic acid are known as omega-6 and omega-3. The names refer to the position of the first double bond in the molecule relative to the terminal —CH_3 group. This is referred to as omega (the last letter in the Greek alphabet), to represent its distance from the —COOH group. These structures are given in section 34 of the IB data booklet.

The American Food and Drug Administration (FDA) recommends that pregnant women and nursing mothers should limit their intake of fish to two servings per week, due to concerns about mercury contamination. However, studies have also shown the importance of omega-3 fatty acids, which are found in fish, for fetal health and development of the brain. The obvious tension between these opposing recommendations leaves women with a difficult choice. One possible solution is to choose fish with low mercury levels, such as salmon and sardines; another is to use omega-3 fatty acid dietary supplements.

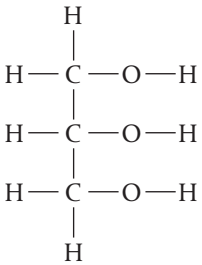
Structures of different lipids

The structures of the three main types of lipids – triglycerides, phospholipids, and steroids – will be considered here in turn.

Structure of triglycerides: fats and oils

Triglycerides are the major constituent of fats and oils. They are esters formed by condensation reactions between **glycerol** and **three fatty acids**.

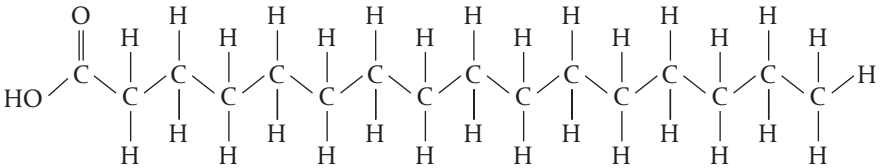
Glycerol is a molecule with three carbon atoms, each of which has an alcohol group, as shown below.



glycerol

Following the guidelines in Chapter 10, see if you can give glycerol its IUPAC name? The answer is propane-1,2,3-triol; however, we will continue to refer to it as glycerol as this name is widely used.

Fatty acids are long-chain carboxylic acids R-COOH. For example, the fatty acid palmitic acid, $C_{15}H_{31}COOH$, has the following structure:



An esterification reaction takes place between an acid —COOH group and each —OH group in glycerol, eliminating water as each ester link forms.

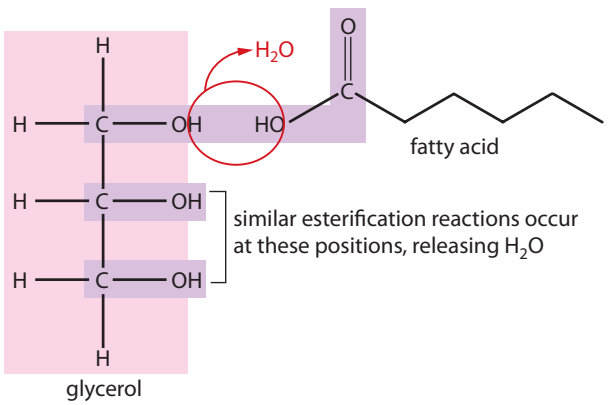
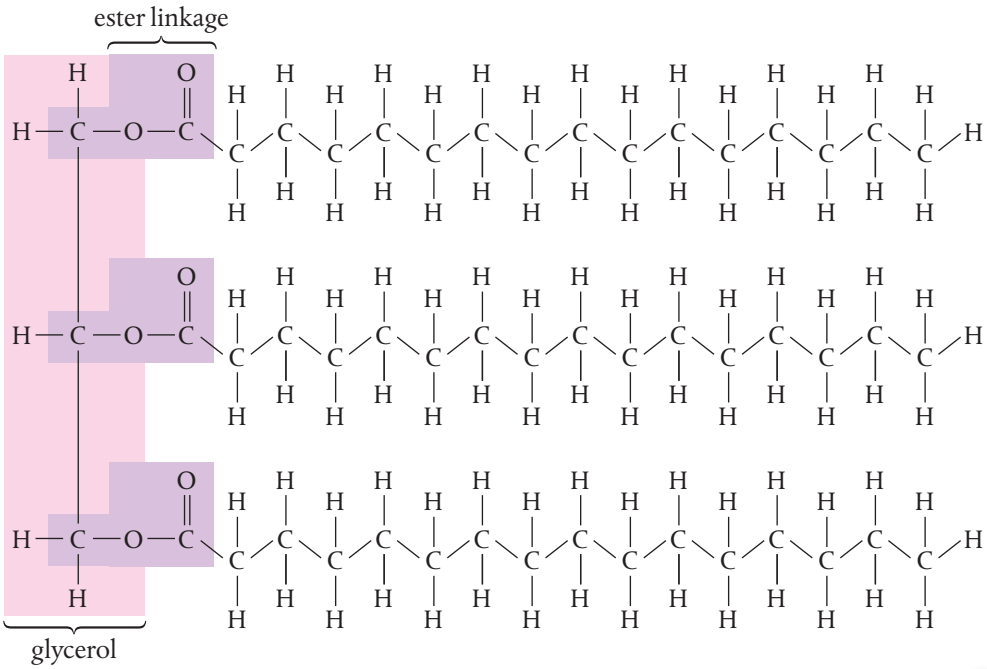


Figure 13.40 Esterification is a condensation reaction in which a molecule of water is released as each –OH group in glycerol condenses with the fatty acid –COOH group. The zig-zag lines represent the hydrocarbon chains of the fatty acid.

So one glycerol condenses with *three* fatty acids to form the *tri*glyceride.



a triglyceride

Figure 13.41 A triglyceride.

In most natural oils and fats the three fatty acids that form one triglyceride molecule are not all the same. They can be designated R^1, R^2, and R^3 as shown in Figure 13.42.

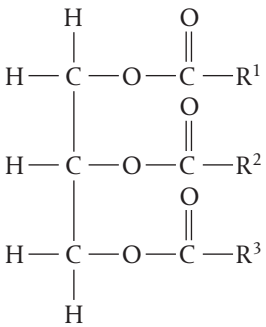


Figure 13.42 A triglyceride containing three different fatty acids.

The fatty acids differ from each other in the following two ways, which give rise to the specific properties of different fats and oils.

1. The length of their hydrocarbon chain: the most abundant fatty acids have an even number of carbon atoms with chains between 14 and 22 carbon atoms long.
2. The number and position of carbon–carbon double bonds in the hydrocarbon chain:
   - fatty acids with no double bonds are said to be **saturated**
   - fatty acids with a single double bond are described as **mono-unsaturated**
   - fatty acids with several double bonds are described as **poly-unsaturated**.

Because fats and oils usually contain a variety of fatty acids, they are classified according to the predominant types of unsaturation present.

Fatty acids with an odd number of carbon atoms are rarely found in land-based animals, but are very common in marine organisms.

Saturated fatty acids

Saturated fatty acids have the general formula $C_nH_{2n+1}COOH$. The carbon chain, made from only carbon–carbon single bonds, has tetrahedral bond angles (109.5°) which allow the molecules to pack relatively closely together. This leads to significant London (dispersion) forces between them. As a result they form saturated triglycerides with relatively high melting points that are solids at room temperature. They are known as **fats** and are derived mostly from animals. Common examples are butter and lard.

Saturated fatty acids contain all single carbon–carbon bonds; unsaturated fatty acids contain one or more double carbon–carbon bonds in the hydrocarbon chain.

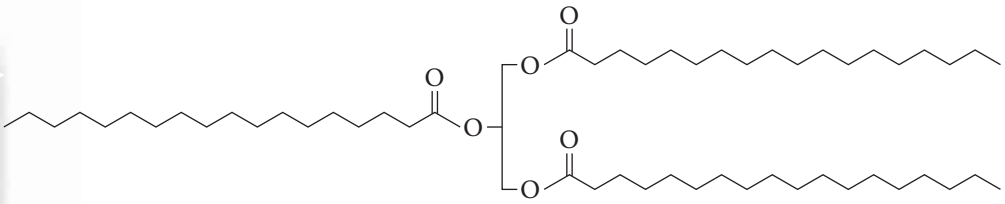


Figure 13.43 A saturated triglyceride.

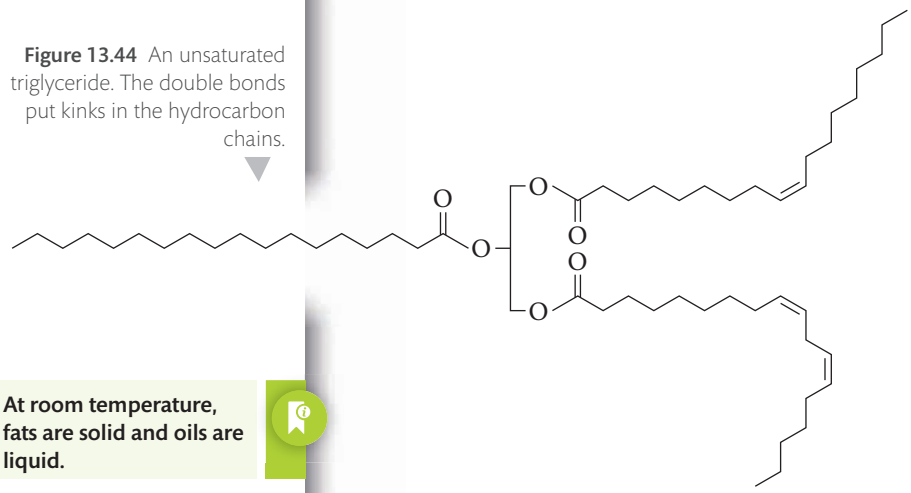


Figure 13.44 An unsaturated triglyceride. The double bonds put kinks in the hydrocarbon chains.

Unsaturated fatty acids

By contrast, the unsaturated fatty acids, containing one or more carbon–carbon double bonds with 120° bond angles, have kinks in the chains that make it more difficult for the molecules to pack closely together. They form unsaturated triglycerides which have weaker intermolecular forces and lower melting points, so are liquids at room temperature. They are known as **oils**, and are found mostly in plants and fish. Common examples are corn oil and cod liver oil.

At room temperature, fats are solid and oils are liquid.

The melting points of fatty acids increase with their molar mass and degree of saturation.

As commonly occurring fats and oils consist of a mixture of triglycerides, they usually do not have sharp melting points. Generally, the melting points increase with increasing molar mass (length of the hydrocarbon chains) and with increasing degree of saturation. The table below illustrates this by comparing the melting points of fatty acids with similar molar mass.

Fatty acid	Formula	M / g mol^{-1}	Number of double bonds	Melting point / °C	Common source	
linolenic	$C_{17}H_{29}COOH$	278	3	–11	soyabean oil	increasing saturation ↓
linoleic	$C_{17}H_{31}COOH$	280	2	–5	corn oil	
oleic	$C_{17}H_{33}COOH$	282	1	16	olive oil	
stearic	$C_{17}H_{35}COOH$	284	0	70	beef fat	

As noted earlier, a strong correlation has been shown between diets rich in saturated fats and elevated levels of LDL cholesterol, with associated increase in the incidence of heart disease.

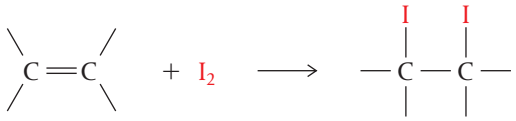
Assortment of dietary oils (liquids) and fats (solids). Oils have a lower melting point because they contain unsaturated fatty acids. These lower the level of cholesterol in the blood.

NATURE OF SCIENCE

Public knowledge concerning the health effects of excess lipids in the diet is the result of many published studies of long-term data. Scientists have shared their results through peer-reviewed publications and have helped to explain the links between diet and health. The widely accepted relationship between *trans* fats, excess saturated fat, and cholesterol with cardiovascular disease has helped to direct policies on food production and education.

Determination of the degree of unsaturation in a fat uses iodine

Unsaturated fatty acids are able to undergo **addition reactions** by breaking the carbon–carbon double bond(s) and adding incoming groups to the new bonding positions created on the carbon atoms. This is a characteristic reaction of alkenes, described in Chapter 10. Iodine (I_2) is able to react with unsaturated fats in this way.



The equation shows that one mole of iodine will react with each mole of double bonds in the fat. Therefore the higher the number of double bonds per molecule, the larger the amount of iodine that can react. This is expressed as the **iodine number**, defined as the number of grams of iodine which reacts with 100 grams of fat. It is therefore a measure of the amount of unsaturation in the fat.

Determination of the iodine number of a fat usually involves reacting a known amount of the fat with a known amount of iodine, and waiting for the reaction to be completed. The amount of excess iodine remaining can then be calculated by titration with $Na_2S_2O_3(aq)$, from which the amount of reacted iodine can be determined.

When linoleic and linolenic fatty acids were first discovered to be essential nutrients in 1923, they were originally designated as Vitamin F. Like vitamins, they give rise to deficiency disorders when absent from the diet. However, later work showed that they are better classified with the fats than with the vitamins.

The iodine number of a fat is defined as the number of grams of iodine that reacts with 100 g of fat.

Worked example

Linoleic acid has the formula $C_{18}H_{32}O_2$. Determine the iodine number of linoleic acid.

Solution

The formula for linoleic acid can be expressed as $C_{17}H_{31}COOH$, from which we can deduce that it has two carbon–carbon double bonds.

$\therefore$ 2 moles I_2 will react with 1 mole linoleic acid

$M(\text{linoleic acid}) = 280\ \text{g mol}^{-1}$

$M(I_2) = 254\ \text{g mol}^{-1}$

$\therefore$ 280 g linoleic acid reacts with $(2 \times 254) = 508\ \text{g } I_2$

100 g linoleic acid reacts with $\dfrac{508\ \text{g} \times 100\ \text{g}}{280\ \text{g}} = 181\ \text{g } I_2$

$\therefore$ iodine number = 181

The lower the iodine value, the higher the proportion of saturated fatty acids in the fat; the iodine number of saturated fatty acids will, of course, be zero. The table below gives some iodine values for some common fats and oils.

Fat or oil	Iodine number
soyabean oil	122–134
olive oil	80–90
bacon fat	47–67
beef fat	35–45

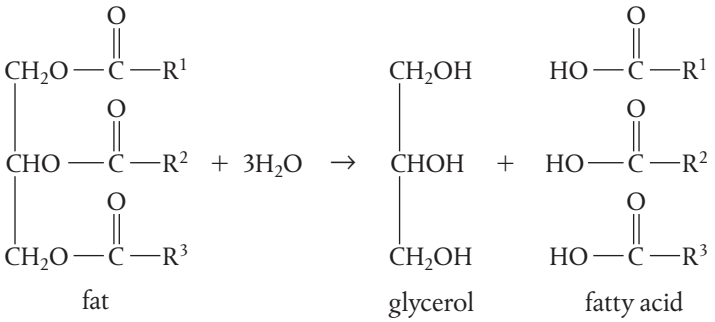
A similar reaction to the addition of iodine is addition of hydrogen, known as **hydrogenation**. This is sometimes carried out by the food industry to increase the saturation of oils. The process of partial hydrogenation produces *trans* fats, which are associated with some of the health problems discussed earlier. The structure of these fats is discussed in section 13.10, page 751.

Rancidity of fats

When fats used in the food industry are stored for long periods of time, they can undergo chemical change which causes them to become **rancid**. This is characterized by disagreeable smells, taste, texture, or appearance. There are two main causes of this.

Hydrolytic rancidity

This occurs when the fat breaks down by hydrolysis reactions, using the water present in food. The site of reactivity is the ester linkages in the triglycerides.



This is the reverse of the condensation reaction described on page 714. The reaction occurs more readily in the presence of heat, such as during deep-fat frying. It is also catalysed by the enzyme **lipase**, and can be favoured in the presence of certain bacteria. The rancid smell and flavour is due to the release of free fatty acids, such as butanoic and octanoic acids which are released from rancid milk. As hydrolytic rancidity is favoured by higher temperatures, it can be substantially reduced by refrigeration.

Oxidative rancidity

This occurs when unsaturated fats react with oxygen from the air. The site of reactivity is the carbon–carbon double bonds in unsaturated triglycerides. The products responsible for the rancidity are volatile aldehydes and ketones. The process, known as auto-oxidation, is often accelerated by light and enzymes or metal ions. It proceeds via a free-radical mechanism and so yields a mixture of products. Oxidative rancidity is characteristic of fats and oils that have a high proportion of carbon–carbon double bonds, such as oily fish like herring. It can be controlled, but not eliminated, by the addition of **antioxidants**.

As they cannot undergo auto-oxidation, saturated fats are more stable than unsaturated fats.

Different cultures use varied sources of lipids and so different methods to extend their shelf-life. Examples of preservation include processing, packaging, and chemical additives.

Structure of phospholipids

Phospholipids are similar to triglycerides in that they are also derived from fatty acids and glycerol, but have only two fatty acids condensed onto the glycerol molecule. The third –OH position of the glycerol has, instead, condensed with a phosphate group. Different phospholipids vary in their fatty acid chains and in the group attached to the phosphate. One of the most common phospholipids, **lecithin**, is shown in Figure 13.45.

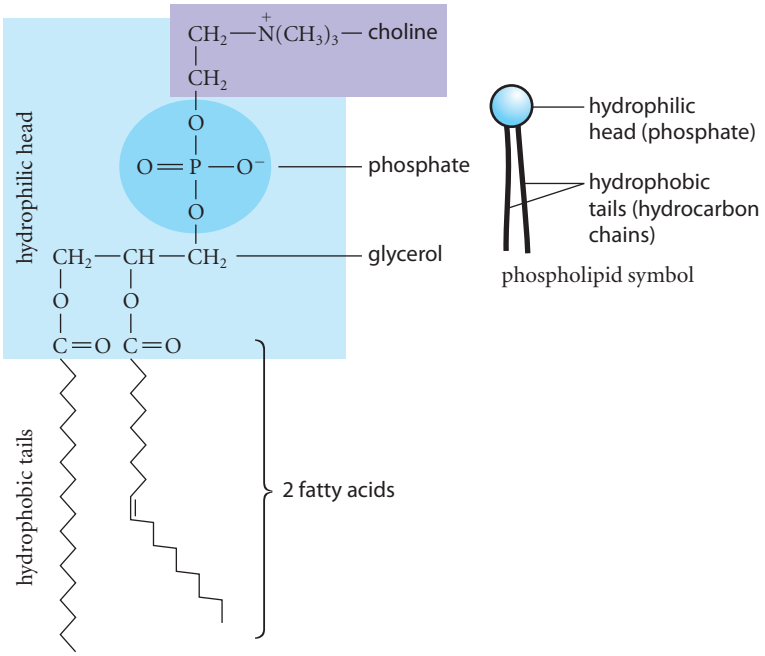


Figure 13.45 Representations of the structure of phospholipids.

Phospholipids are characterized by having a polar, or hydrophilic, 'head' (the phosphate group) and two non-polar, or hydrophobic, 'tails' (the hydrocarbon chains of the fatty acids). As a result they will spontaneously form a **phospholipid bilayer** which maximizes the interactions between the polar groups and water, while creating a non-polar, hydrophobic interior.

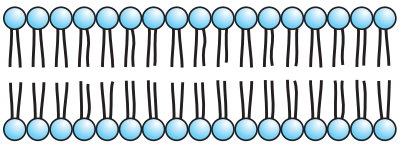


Figure 13.46 Phospholipid bilayer.

This phospholipid bilayer provides the basis of membrane structure.

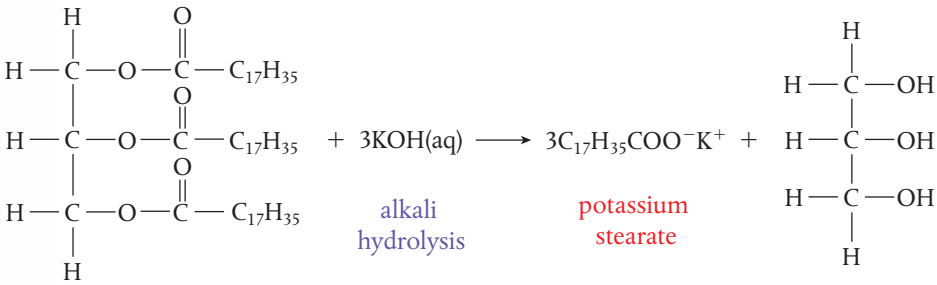
Hydrolysis of fats and phospholipids

Triglycerides and phospholipids are broken down in hydrolysis reactions to yield their component molecules. The reactions use water and can occur in acidic or in alkaline conditions, or catalysed by enzymes known as **lipases**. This occurs during the digestion of lipids in the gut, where the activity of the enzymes is controlled largely by local changes in pH. Lipid digestion is generally a slow process and may take several hours before soluble products are released. Hydrolysis reactions are also responsible for hydrolytic rancidity.

Alkaline hydrolysis produces the salt of the fatty acid and is used in the manufacture of soap. For this reason, this reaction of reverse esterification is sometimes called **saponification**.

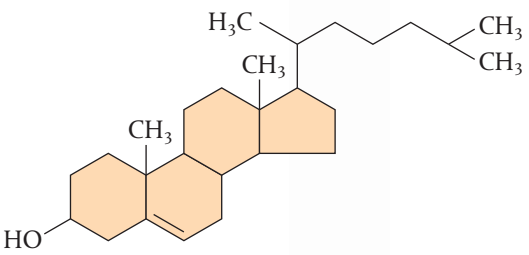
Whale oil was a traditional source of lipids for many countries, and was widely used as a food substance, a lubricant, and as a fuel for lighting. Over-harvesting of whales led to a depletion of stocks by the mid-1800s, and kerosene from crude oil became the major alternative fuel.

For example,



Structure of steroids

Steroids are lipids with a structure consisting of four fused rings, known as a **steroidal backbone**. One of the most important steroids is cholesterol, which has the following structure. It is given in the IB data booklet in section 34.



Cholesterol is used as a precursor in the synthesis of many biomolecules, including other steroids such as the sex hormones, bile acids, and Vitamin D. It is an essential component of cell membranes as it helps to provide fluidity and permeability to the structure. The hydroxyl group interacts with the polar head groups of phospholipids in the membrane, while the non-polar rings and hydrocarbon chain interact with the hydrophobic tails of the phospholipid bilayer. The uses, and problems, of cholesterol and other steroids were discussed earlier in this section.

Exercises

16 A sample of fat containing 0.02 moles of fatty acid was found to react with 10.16 g of iodine. Determine the number of carbon–carbon double bonds present in the fatty acid.

17 The following table shows the melting point for a number of common fatty acids found in dietary fats and oils.

Name of acid	Formula	Structural formula	Melting point / °C
lauric	$C_{11}H_{23}COOH$	$CH_3(CH_2)_{10}COOH$	44
myristic	$C_{13}H_{27}COOH$	$CH_3(CH_2)_{12}COOH$	58
palmitic	$C_{15}H_{31}COOH$	$CH_3(CH_2)_{14}COOH$	63
stearic	$C_{17}H_{35}COOH$	$CH_3(CH_2)_{16}COOH$	70
oleic	$C_{17}H_{33}COOH$	$CH_3(CH_2)_7CH{=}CH(CH_2)_7COOH$	16
linoleic	$C_{17}H_{31}COOH$	$CH_3(CH_2)_4CH{=}CHCH_2CH{=}CH(CH_2)_7COOH$	−5

(a) Which of the fatty acids are solids at a room temperature of 25 °C?
(b) Describe and explain the trend in the melting points in the first four fatty acids listed.
(c) Describe and explain the pattern in the melting points of the last three acids mentioned.

18 Explain the different chemical basis of the two types of rancidity which can occur in lipids.

B.4 Carbohydrates

Understandings:

- Carbohydrates have the general formula $C_x(H_2O)_y$.
- Haworth projections represent the cyclic structures of monosaccharides.
- Monosaccharides contain either an aldehyde group (aldose) or a ketone group (ketose) and several –OH groups.
- Straight-chain forms of sugars cyclize in solution to form ring structures containing an ether linkage.
- Glycosidic bonds form between monosaccharides forming disaccharides and polysaccharides.
- Carbohydrates are used as energy sources and energy reserves.

Guidance

- *The component monosaccharides of specific disaccharides and the linkage details of polysaccharides are not required.*
- *The distinction between α- and β-forms, and the structure of cellulose is not required.*

Applications and skills:

- Deduction of the structural formulas of disaccharides and polysaccharides from given monosaccharides.
- Relationship of the properties and functions of monosaccharides and polysaccharides to their chemical structures.

Guidance

The straight-chain and α-ring forms of glucose and fructose are given in the data booklet in section 34.

Carbohydrates (literally *hydrated carbon*) are composed of the three elements carbon, hydrogen, and oxygen, with the hydrogen and oxygen always in the same ratio as in water i.e. 2 : 1. They therefore can be expressed by the general formula $C_x(H_2O)_y$. Note that this higher ratio of oxygen represents a more oxidized state for the carbon atoms than in lipids. There are two main types of carbohydrates – the simple sugars or **monosaccharides** and the condensation polymers of these known as **polysaccharides**.

Functions of carbohydrates

The monosaccharides, for example **glucose** and **fructose**, are readily soluble in water and are mostly taken up by cells quite rapidly. They are used as the main substrate for respiration, releasing energy for all cell processes. They also act as precursors in a large number of metabolic reactions, leading to the synthesis of other molecules such as fats, nucleic acids, and amino acids.

Polysaccharides, being insoluble, are used as the storage form of carbohydrates. Animals mostly use **glycogen** for storage of carbohydrates in the liver and muscles, while plants store **starch** in their cells. These energy reserves can be broken down into monosaccharides, which are then oxidized in respiration to release energy for the cell's activities.

Animals generally make little use of carbohydrates for structural materials, but plant cells are very dependent on carbohydrates for their structure. The polysaccharide **cellulose**, which is found in the walls of all plant cells, is claimed to be the most abundant organic compound on Earth.

TOK

Aspartame is an artificial sweetener, widely used in some foods and drinks. However, its use is controversial, with claims and counter-claims made about its negative impact on health. What ways of knowing can be used to determine response when faced with contradictory information such as this?

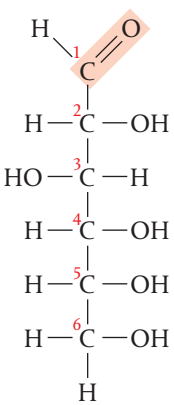
Structure of carbohydrates

Monosaccharides are simple sugars

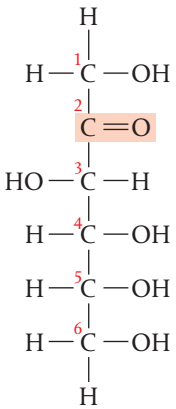
The monosaccharides are the simplest form of carbohydrates and are usually classified according to the number of carbon atoms that they contain. Some of the most common are the **triose** sugars (C3), the **pentose** sugars (C5), and the **hexose** sugars (C6). All monosaccharides can be represented by the empirical formula CH_2O. Hexose sugars, for example, all have the molecular formula $C_6H_{12}O_6$.

These sugar molecules all have two or more hydroxyl groups (—OH) and a carbonyl group (—C=O). Their large number of polar hydroxyl groups is responsible for their ready solubility in water.

Many isomers of sugars exist, representing different structural arrangements of the same number and type of atoms in different molecules. For example, $C_6H_{12}O_6$ exists in many forms including glucose and fructose. The straight-chain forms of these sugars are shown below.



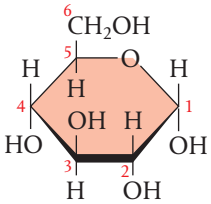
glucose: an aldose sugar



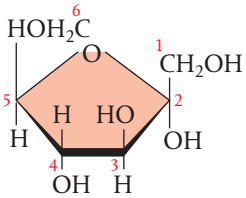
fructose: a ketose sugar

Monosaccharides contain a carbonyl group and at least two hydroxyl groups. They have the empirical formula CH_2O.

In aqueous solution these sugars undergo an internal reaction resulting in the more familiar ring structures shown below.



glucose



fructose

These representations of the ring forms of sugars are known as **Haworth projection formulas**. The edge of the ring nearest the reader is represented by bold lines, and the letter C for the carbons in the ring are usually omitted from the structure.

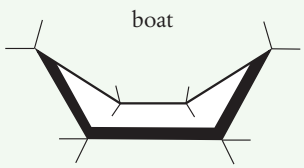
The formation of the ring in glucose and fructose makes possible another type of isomer, so-called alpha and beta forms which are discussed further in section B.10.

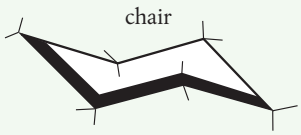
The straight-chain and α-ring forms of glucose and fructose are given in section 34 of the IB data booklet so they do not have to be learned.

NATURE OF SCIENCE

The stereochemistry of carbohydrates is essential to their structural and functional roles in cells. Different representations are used to symbolize these molecular structures. Fischer projections are two-dimensional representations of three-dimensional molecules, such as shown here for the straight-chain forms of the sugars. Haworth projection formulas, on the other hand, aim to give a simple three-dimensional perspective. They are often favoured for representing cyclic structures, as by making ring carbon (and often hydrogen) atoms implicit, they help to focus on the nature and position of the substituent groups.

Haworth formulas can, however, be misleading as they suggest that the five-membered ring of fructose and six-membered ring of glucose are planar, which is not the case. In reality these exist as puckered rings in different conformations which can interconvert by rotation around single bonds.



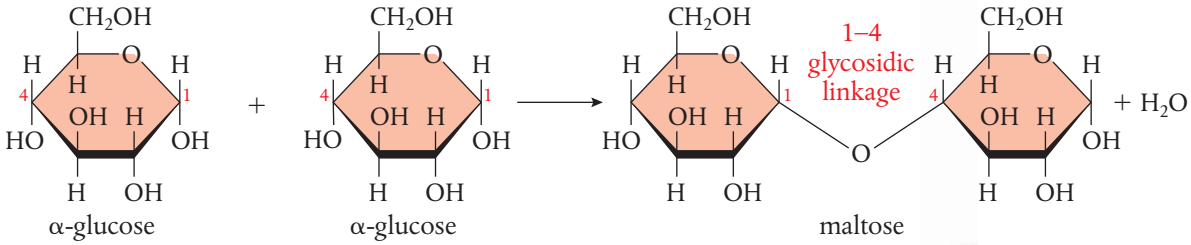


Scientists depend on models and visualizations to communicate knowledge, but it is often the case that a single representation is inadequate. Understanding can therefore be enhanced by the use of a variety of models, including computer simulations and mathematical models.

Disaccharides are two simple sugars linked together

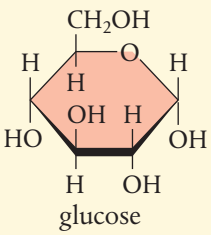
Disaccharides form by linking two monosaccharides together in a condensation reaction. A molecule of water is eliminated as an –OH group from each sugar molecule react together. The resulting bond between the monosaccharides is known as a **glycosidic link**. Disaccharides are all soluble molecules that can be hydrolysed into two monosaccharides by acid hydrolysis or by enzyme-catalysed reaction. Combining different monosaccharides will produce different disaccharides.

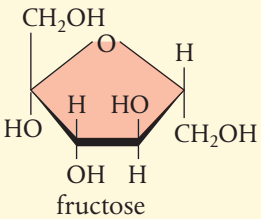
For example, two α-glucose molecules condense to form the disaccharide maltose, as shown here. The glycosidic link is known as 1–4 because C_1 in one molecule is bonded to C_4 in the other molecule. The molecular formula of maltose is $C_{12}H_{22}O_{11}$.



Worked example

Sucrose is a disaccharide of glucose and fructose, whose structures are shown below. The condensation reaction involves the –OH group at C_1 of glucose. Deduce the structural formula of sucrose.





You are not expected to learn the combinations of monosaccharides in different disaccharides, but will be given this information if it is needed in an examination question.

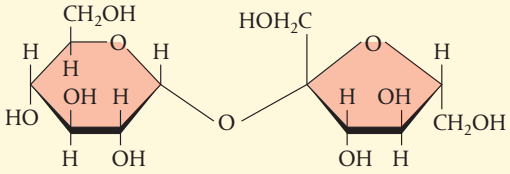
Sugar is an important trade product from many countries. Sugar cane is the world's largest crop, and is exported largely from tropical and sub-tropical regions of Brazil, India, China, and Thailand. Sugar beet grows only in the temperate zones of North America and Europe, though countries in these regions are usually net importers of sugar. Both crops yield sugar in the form of sucrose.

Lactose is a disaccharide found in milk. Many people suffer from 'lactose intolerance', which is usually a genetic condition, characterized by an inability to digest lactose owing to a lack of the enzyme lactase. The condition is more prevalent in many Asian and South African cultures where dairy products are less traditionally part of the adult diet. By contrast, people with ancestry in Europe, the Middle East, and parts of East Africa, where mammals are often milked for food, typically maintain lactase production throughout life.

The production of ethanol as a biofuel usually involves the fermentation of carbohydrates from crops such as corn and sugar cane. This is at the heart of the **food versus fuel** debate on land and resource use.

Solution

The monosaccharides are joined by a glycosidic link as follows:



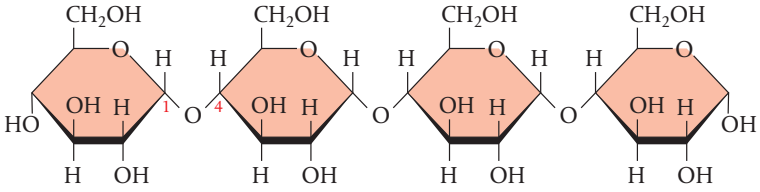
Polysaccharides are the polymers of sugars

Polysaccharides form by repetitions of the reaction shown above, leading to a long chain of monosaccharide units held together by glycosidic bonds. Due to their large size, polysaccharides are all insoluble molecules, and so make an ideal storage form of the energy-rich carbohydrates. Cellulose forms a rigid structure which provides support to plant cells.

The most common polysaccharides are:

- starch – carbohydrate store in plants
- glycogen – carbohydrate store in animals
- cellulose – structural material in plants.

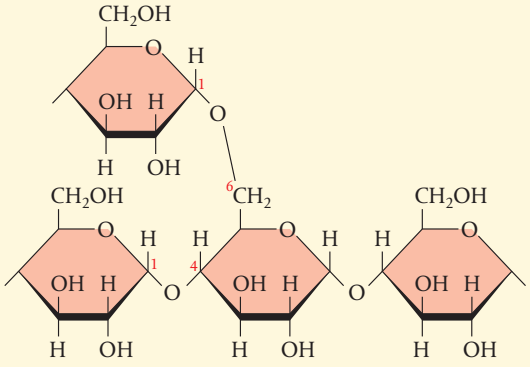
These are all polymers of glucose, but differ from each other in the isomer of glucose used and in the amount of cross-linking in the chain. For example, starch consists mostly of a straight-chain polymer of α-glucose with 1–4 linkages.



Worked example

Glycogen is a branched polysaccharide which contains α-glucose molecules linked with both 1–4 and 1–6 glycosidic bonds. Use the structure of α-glucose in your IB data booklet to draw a section of the molecule showing both types of linkages.

Solution

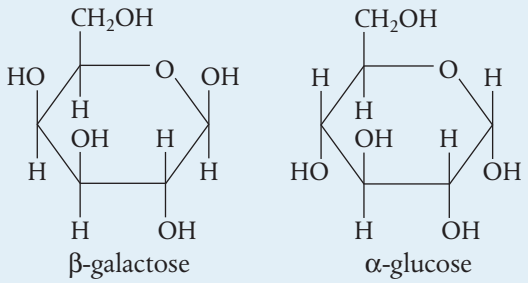


The structure of cellulose is described in section B.10.

Exercises

19 **(a)** State the empirical formula of all monosaccharides.
(b) State and explain the difference in solubility of monosaccharides and polysaccharides.

20 The disaccharide lactose is made from the monomers shown below.



Deduce the structural and molecular formula of lactose, and name the bond between the sugars.

B.5 Vitamins

Understandings:

- Vitamins are organic micronutrients which (mostly) cannot be synthesized by the body but must be obtained from suitable food sources.
- The solubility (in water or fat) of a vitamin can be predicted from its structure.
- Most vitamins are sensitive to heat.
- Vitamin deficiencies in the diet cause particular diseases and affect millions of people worldwide.

Applications and skills:

- Comparison of the structures of vitamins A, C, and D.
- Discussion of the causes and effects of vitamin deficiencies in different countries and suggestion of solutions.

Guidance

- *The structures of vitamins A, C, and D are provided in the IB data booklet section 35.*
- *Specific food sources of vitamins or names of deficiency diseases do not have to be learned.*

Given the complexity of metabolic processes, it is not surprising that organisms need an equally complex array of molecules to maintain health, and these are known as **nutrients**. In terms of the human diet, **micronutrients** are those needed in extremely small amounts, generally less than 0.005% of body mass, quantities so small that they are usually measured in mg or μg per day. These substances are needed to enable the body to produce enzymes, hormones, and other biomolecules. As tiny as the amounts are, however, the consequences of their absence are severe and are known as **deficiency diseases**. Micronutrients include the **vitamins** and many so-called **trace minerals** such as Fe, Cu, Zn, I, Se, Mn, Mo, Cr, Co, and B.

Vitamins are organic micronutrients

Vitamins are organic compounds, needed in small amounts for normal growth and metabolism, which (with the exception of vitamin D) are not synthesized in the body.

They are usually broken down by the reactions in which they are involved, so must be taken in from suitable food sources in the diet. They are often classified according to their relative solubility in water or in lipid.

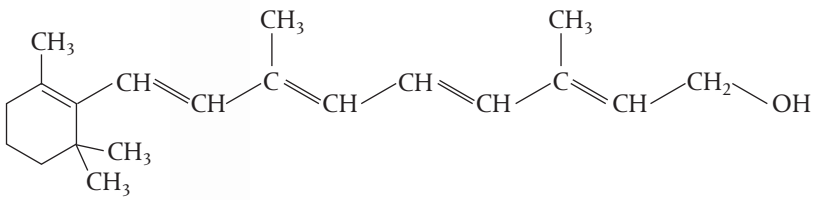
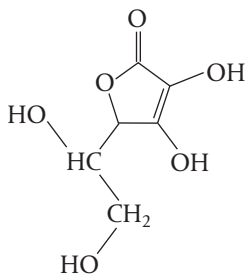
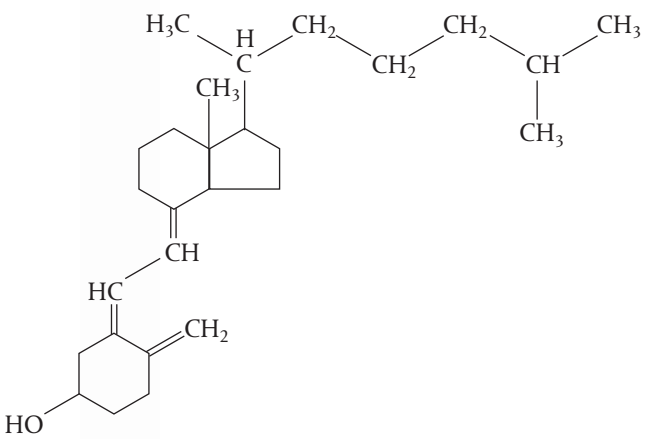
Water-soluble vitamins

Vitamins that are water soluble have polar bonds and the ability to form hydrogen bonds with water. They are transported directly in the blood, and excesses are filtered out by the kidneys and excreted. Vitamins B and C are water soluble.

The structures of some important vitamins are given in section 35 of the IB data booklet, so they do not have to be learned.

Lipid-soluble vitamins

Vitamins that are lipid soluble are mostly non-polar molecules with long hydrocarbon chains or rings. They are slower to be absorbed and excesses tend to be stored in fat tissues where they can produce serious side-effects. Vitamins A, D, E, and K are fat soluble.

Vitamin	Structure	Solubility and properties
A, retinol		• fat soluble • hydrocarbon chain and ring are non-polar and influence the solubility more than the one —OH group • involved in the visual cycle in the eye, and particularly important for vision in low light intensity
C, ascorbic acid		• water soluble • several —OH groups enable hydrogen bonds to form with water • acts as cofactor in some enzymic reactions, important in tissue regeneration following injury, and resistance to some diseases
D, calciferol		• fat soluble • predominantly a hydrocarbon molecule with four non-polar rings and only one —OH group • chemically similar to cholesterol • stimulates the uptake of calcium ions by cells and important in the health of bones and teeth

Vitamin C contains several functional groups (—OH and —C=C—) that are relatively easily oxidized. This is why the vitamin is easily destroyed by most methods of food processing and storage, and is therefore best obtained from *fresh* fruits and vegetables. In general, the water-soluble vitamins are the most sensitive to heat, but other vitamins also lose some activity after being heated.

British people are sometimes known as 'limies' in America, because of the earlier practice of the British Navy to supplement the diet of sailors with limes to prevent scurvy.

NATURE OF SCIENCE

The discovery of vitamins (*vital amines*) is an example of scientists seeking a cause for specific observations. In the 1700s, many sailors on long voyages suffered from a disease known as scurvy, with symptoms of bleeding gums, poor resistance to infection, and dark spots on the skin. Although the concept of a disease resulting from the *lack* of a dietary component was not understood, it was discovered that these symptoms could be prevented by providing a nutrient that is present in citrus fruits. The vital ingredient was later identified as vitamin C. Similarly, in 1905, it was noted that the disease beriberi could be prevented by eating an ingredient in unpolished rice, later identified as vitamin B.

These and similar correlations led to the **hypothesis of deficiency diseases** – the concept that a *lack* of something in the diet could cause disease. This was formulated in 1912 and now forms the basis of much healthcare practice.

Vitamin deficiencies are a form of malnutrition

The absence of a regular, balanced supply of the diverse nutrients needed in the diet is known as **malnutrition**. This describes a broad spectrum of conditions, including vitamin deficiency diseases, which are always associated with compromised health. The main focus of malnutrition has traditionally been the large variety of nutrient-deficiency diseases, and in particular the incidence of these in under-developed countries. But increasingly the world is seeing a dramatic increase in diseases caused by high consumption of processed, energy-dense but micronutrient-poor foods.

'Freedom from hunger and malnutrition is a basic human right and their alleviation is a fundamental prerequisite for human and national development.'

World Health Organization

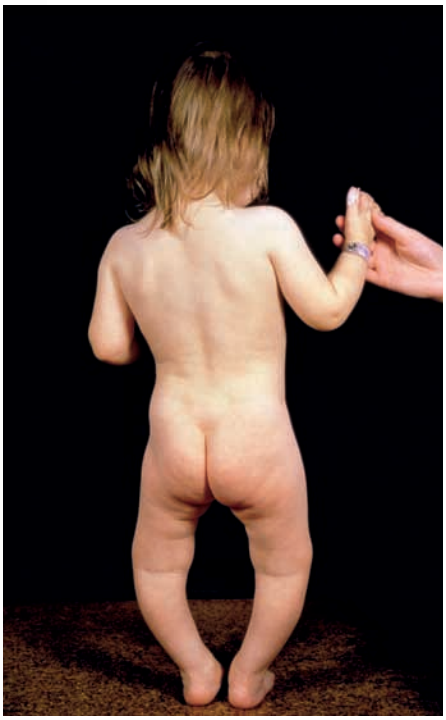
The causes of malnutrition arising from vitamin deficiency are varied and widespread. They include:

- lack of distribution of global resources
- depletion of nutrients in the soil and water
- lack of education about, or understanding of, the importance of a balanced diet
- over-processing of food for transport and storage
- the use of chemical treatments such as herbicides in food production.

Vitamin A is an antioxidant and is needed for healthy eyesight. It is related to carotene and found in orange and yellow fruit and vegetables. As it is fat soluble, it has been found that it can be effectively added to margarine in a process known as **vitamin fortification**. The potential of rice as a vehicle for vitamin A fortification is also being explored, given that rice is an important staple in many countries where the prevalence of vitamin A deficiency is high. Vitamin B is a term for a group of eight distinct water-soluble vitamins, which are most commonly found in unprocessed foods such as whole grains. Their deficiency causes a range of diseases including beriberi, forms of anaemia, and mental disorders. Cereals are commonly fortified with B vitamins. As vitamin B_{12} contains cobalt, which is available only from animal sources, vegetarians and vegans may be at particular risk of deficiency. Vitamin C deficiency is characterized by lower resistance to infection and can develop into scurvy. It is best prevented by a diet rich in fresh fruits and vegetables. Vitamin D is made by the action of sunlight on the skin and is important for healthy bones.

Food label showing fortification of foods with vitamins.

The World Health Organization has identified vitamin A as the most important vitamin deficiency in global health terms. It is estimated that more than half of all countries, especially in Africa and South-East Asia, have a significant incidence of vitamin A deficiency, which is the leading cause of preventable blindness in children and increases the risk of disease and death from severe infections.



Child suffering from rickets. Rickets is a disease in growing children in which the bones do not harden and are malformed. It is due to a deficiency in vitamin D, which is necessary for the uptake of calcium.

TOK

To what extent is mass medication through the food we eat an infringement of personal freedom? Are there ways of striking a balance between the claim of the scientists or government that they know what is best for health, and the right of an individual to choose what to eat?

In summary, there exist many possible solutions to the varying challenges of malnutrition. These include:

- the fortification of different staple foods with micronutrients
- the availability of vitamin supplements in many forms
- the possible improvements to nutrient content of food through genetic modification
- increased labelling of foods with content information
- education regarding the nature of a balanced diet and promotion of the importance of personal responsibility in dietary choices.

Exposure to sunlight is an important source of vitamin D, but it is also the case that ultraviolet light has a damaging effect on the skin and can be the cause of skin cancers. Protection from this necessitates the use of sunscreens, and those with a protection factor of 8 or greater will block UV rays that produce vitamin D. This makes it even more important to include good sources of vitamin D in the diet when sun exposure is limited in this way. Mandatory vitamin D fortification is increasing, usually applied to milk and margarine, as the vitamin is fat soluble.

NATURE OF SCIENCE

The manufacture and sale of vitamin supplements is a major industry in many countries. Debate is widespread on the extent to which a balanced diet can provide sufficient quantities of vitamins, or whether supplementation is necessary for optimum health. Part of the controversy is a result of claims made by Linus Pauling in 1970 that 'megadoses' of vitamin C could be effective in preventing the common cold, and possibly certain cancers and heart disease as well. As the only person ever to have won two unshared Nobel Prizes, Pauling was a scientist and humanitarian of significant authority, and so his claims led to a surge in the sales of vitamin C. This happened despite the fact that his megavitamin claims lacked the evidence needed for acceptance by the scientific community. The existence of conflicting information and advice, and the pressure of advertising, can make it difficult for individuals to make an informed choice. Scientists have a duty to share their findings with the public in ways that help people to understand the issues and make valid judgments.

Exercises

21 Make reference to section 35 in the IB data booklet.

(a) Identify two functional groups common to all three vitamins shown.

(b) Identify one vitamin that is water soluble and one vitamin that is fat soluble. Explain the differences in solubility in terms of their structures and intermolecular forces.

22 Suggest ways in which vitamin deficiency diseases can be alleviated.

B.8 Nucleic acids

Understandings:

- Nucleotides are the condensation products of a pentose sugar, phosphoric acid, and a nitrogenous base – adenine (A), guanine (G), cytosine (C), thymine (T), or uracil (U).
- Polynucleotides form by condensation reactions.
- DNA is a double helix of two polynucleotide strands held together by hydrogen bonds.
- RNA is usually a single polynucleotide chain that contains uracil in place of thymine, and a sugar ribose in place of deoxyribose.
- The sequence of bases in DNA determines the primary structure of proteins synthesized by the cell using a triplet code, known as the genetic code, which is universal.
- Genetically modified organisms have genetic material that has been altered by genetic engineering techniques, involving transferring DNA between species.

Guidance

- *Structures of the nitrogenous bases and ribose and deoxyribose sugars are given in the data booklet in section 34.*
- *Knowledge of the different forms of RNA is not required.*

Applications:

- Explanation of the stability of DNA in terms of the interactions between its hydrophilic and hydrophobic components.
- Explanation of the origin of the negative charge on DNA and its association with basic proteins (histones) in chromosomes.
- Deduction of the nucleotide sequence in a complementary strand of DNA or a molecule of RNA from a given polynucleotide sequence.
- Explanation of how the complementary pairing between bases enables DNA to replicate itself exactly.
- Discussion of the benefits and concerns of using genetically modified foods.

Guidance

- *Limit expression of DNA to the concept of a four-unit base code determining a twenty-unit amino acid sequence. Details of transcription and translation are not required.*
- *Details of the process of DNA replication are not required.*

The role of nucleic acids

Deoxyribonucleic acid (**DNA**) and **ribonucleic acid** (**RNA**) are collectively known as the nucleic acids. As is evident from their name they are acidic molecules found in the nucleus of cells (though RNA is also found elsewhere). DNA is responsible for storing the information that controls the genetic characteristics of an organism, and for passing it on to the next generation. RNA enables the information stored in DNA to be expressed by controlling the primary structures of proteins synthesized.

In order to carry out its functions, DNA needs to have the following features:

- it must be a very stable molecule, able to retain its precise chemical structure in cell conditions
- it must contain some 'code' that stores genetic information
- it must be able to replicate, in other words to produce an exact copy of itself.

The race to interpret all the known data about DNA and to come up with a model that could explain its structure and function was a major focus of biochemical research during the 1950s. When the double helical structure was suggested by Francis Crick and James Watson in their letter to *Nature* in 1953, it was immediately heralded as one

of the most significant discoveries of the time. In a remarkably simple and elegant way, this model explained the unique ability of DNA to store and copy information exactly, as we will see in this section.

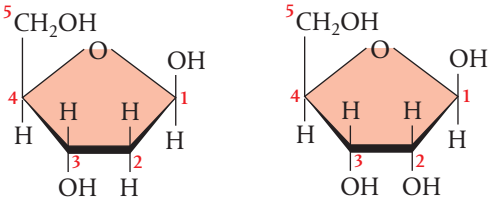
The structure of nucleic acids

Like other biological macromolecules, DNA and RNA are polymers. They are built from monomers known as **nucleotides** and so are described as **polynucleotides**. Nucleic acids typically contain thousands of nucleotides, and are among the largest macromolecules found in cells.

Nucleotides are the building blocks of nucleic acids

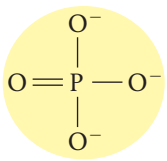
Nucleotides are made up from three components.

1 A pentose (C5) sugar: $C_5H_{10}O_5$
in DNA it is **deoxyribose**,
in RNA it is **ribose**.

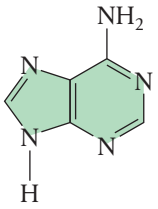
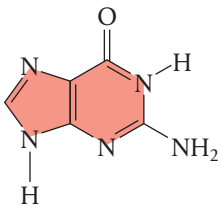
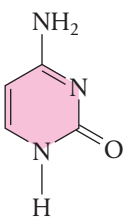
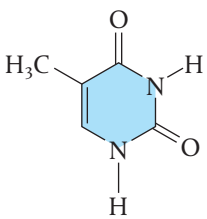
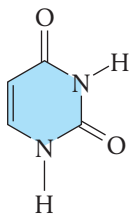


The difference between these sugars is in the groups attached to C^2 – *deoxy*ribose lacks an –OH.

2 A phosphate group PO_4^{3-}
(derived from phosphoric acid H_3PO_4)
It is often denoted as (P)



3 An organic nitrogenous base, of which there are two types. **Purines** are larger and contain two fused rings; **pyrimidines** are smaller and contain a single ring. There are two different purines and three different pyrimidines: each is described by the first letter of its name using a capital letter, as shown in the table below.

Purine base	adenine, A 	guanine, G 	
Pyrimidine base	cytosine, C 	thymine, T 	uracil, U 

Adenine, guanine, and cytosine are found in both DNA and RNA. Thymine is found exclusively in DNA and uracil is found exclusively in RNA.

The structures of the five bases are given in section 34 of the IB data booklet so they do not have to be learned. You should, however, be able to recognize which are purines, which are pyrimidines, and which pairs are able to link through hydrogen bonding.

The nucleotide forms as the pentose sugar, phosphate, and base join together by condensation reactions, releasing water. The base always condenses to C_1 of the sugar, and the phosphate to C_5. This is known as 'five prime' (5′).

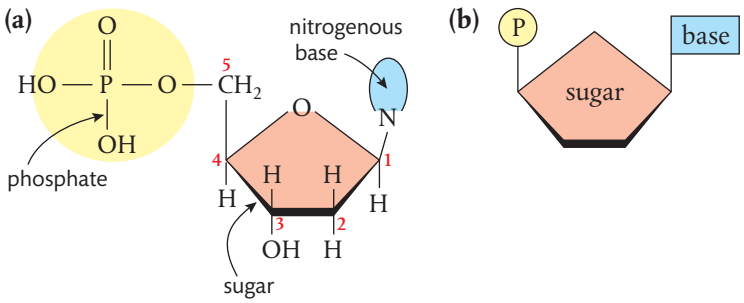


Figure 13.47 (a) The structure of a deoxyribonucleotide. (b) A convenient short-hand form for drawing a nucleotide.

Ribonucleotides are found in RNA: they contain ribose sugar and either A, G, C, or U. Deoxyribonucleotides are found in DNA: they contain dexoyribose sugar and either A, G, C, or T.

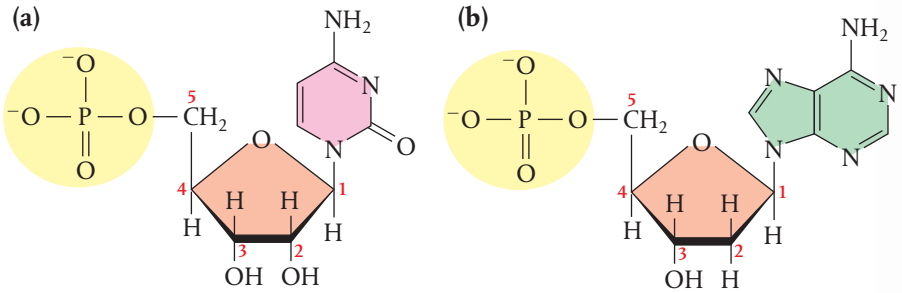


Figure 13.48 (a) A ribonucleotide; (b) a deoxyribonucleotide.

Nucleotides condense to form polynucleotides

Nucleotides link together in condensation reactions involving the phosphate at the 5′ end of one nucleotide and the –OH group at the 3′ (three prime) position of the other nucleotide. In this way they are able to build up a chain held together by covalent bonds between alternating sugar and phosphate residues. These bonds are phosphodiester links, as shown in Figure 13.49.

Note that the nitrogenous bases do not take part in the polymerization of the nucleotides but remain attached to the sugar at C_1.

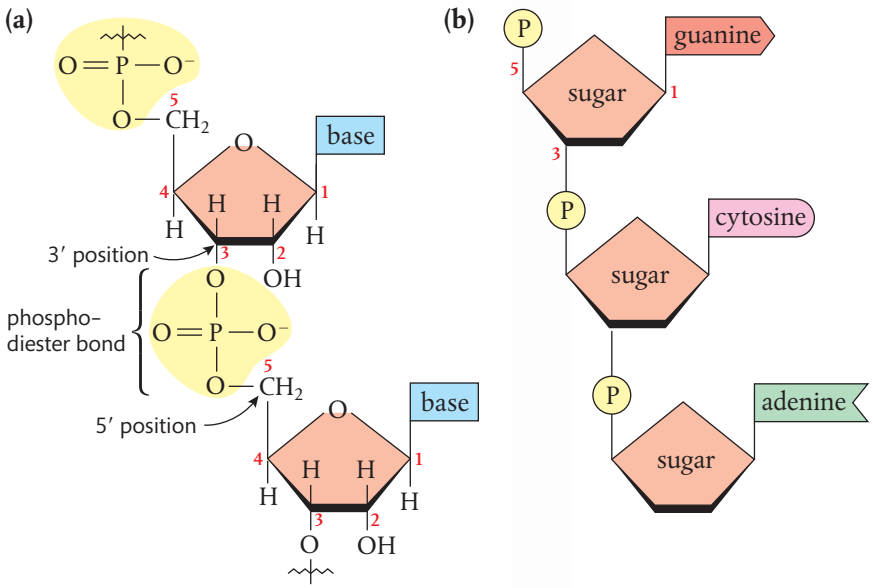
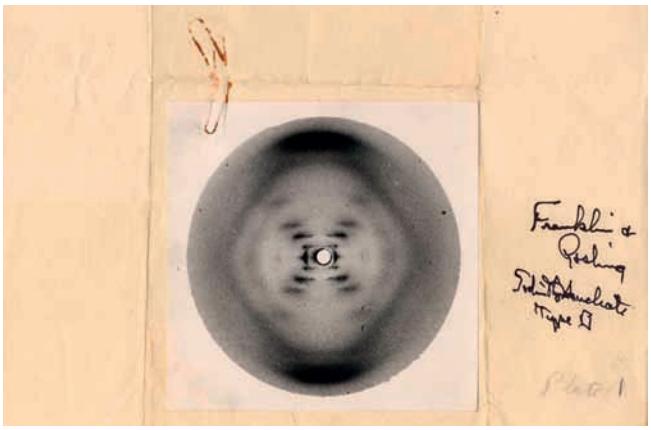


Figure 13.49 (a) Part of a polynucleotide showing the phosphodiester bond. Note that the phosphate group carries a negative charge. (b) A short-hand form of a polynucleotide.

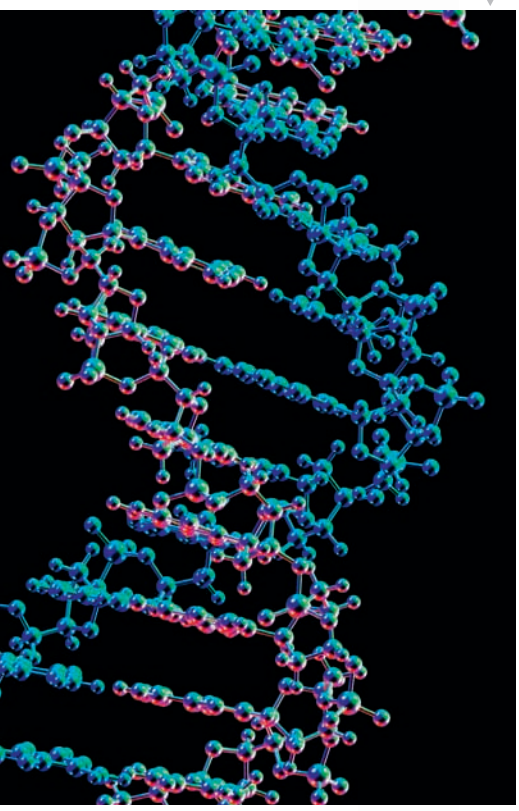
NATURE OF SCIENCE

The story of the discovery of the structure of DNA shows how different approaches to solving the same problem can lead to a consistent conclusion. James Watson and Francis Crick, working in Cambridge, England, tackled the question largely by building molecular models, while Maurice Wilkins and Rosalind Franklin, working in London, approached it through X-ray crystallography. It is, however, also a story interwoven with tales of personal ambition and human conflict, that reflects some failure of the different teams of scientists to collaborate and communicate effectively. It is now widely recognized that, although Nobel Prizes for the discovery of DNA were awarded to Crick, Watson, and Wilkins in 1962, the crystallography work done by Franklin was crucial to the discovery. Franklin died in 1958, aged 37, and Nobel Prizes are not awarded posthumously.

Rosalind Franklin's famous X-ray photograph of DNA, which involved over 100 hours of exposure. It has been described as 'the most beautiful X-ray photograph of any substance ever taken.' The cross of bands indicates the helical nature of DNA.



Computer artwork of the molecular structure of DNA. The two strands of the double helix, made of alternating sugar and phosphate groups, are shown in blue/green and pink/green. The diagonal lines show the stacked bases, which form complementary pairs held together with hydrogen bonds. The sequence of bases along the strand is the genetic code.



DNA is a double helix of two polynucleotides

In DNA two polynucleotide strands are coiled around the same axis forming a **double helix** with the sugar–phosphate backbone on the outside and the nitrogenous bases on the inside. The two strands are held together by hydrogen bonds that form between bases on each strand. Due to the chemistry of the bases and the conformation of the helix, only certain **base pairings** involving one purine with one pyrimidine are possible: adenine forms a pair with thymine, and guanine always forms a pair with cytosine. These are known as **complementary pairs**.

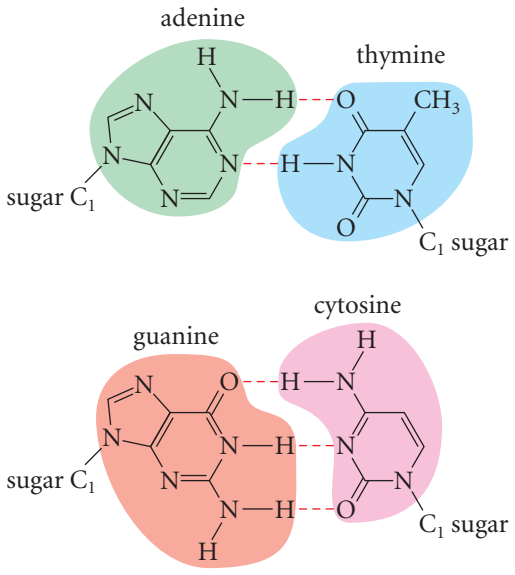


Figure 13.50 Base pairs form only between adenine and thymine, and between guanine and cytosine. Note there are two hydrogen bonds between A and T and three hydrogen bonds between G and C.

The double helical structure with its paired bases is often described as a twisted ladder where the sides are the sugar–phosphate backbones and the rungs are the base pair (Figure 13.51).

Ten nucleotide residues make up one complete turn of the helix and this has the length of 3.4 nm. The two polynucleotide strands in the helix are said to be **anti-parallel**, meaning they run in opposite directions (3′ → 5′ and 5′ → 3′) and so are effectively upside down relative to each other, as shown in Figure 13.52.

This model fulfils all the expectations of DNA outlined earlier.

- Its stability is achieved by the fact that it maximizes hydrophobic interactions between the non-polar stacked bases in the sequestered environment in the middle of the molecule, while allowing polar and charged groups in the sugar–phosphate backbone to interact with the aqueous solution.
- The sequence of bases in the polynucleotide strand is effectively a digital code of information, with infinite variety possible.

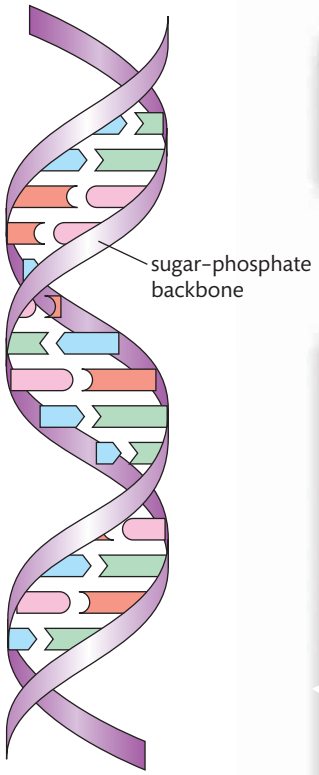


Figure 13.51 The double helical structure of DNA.

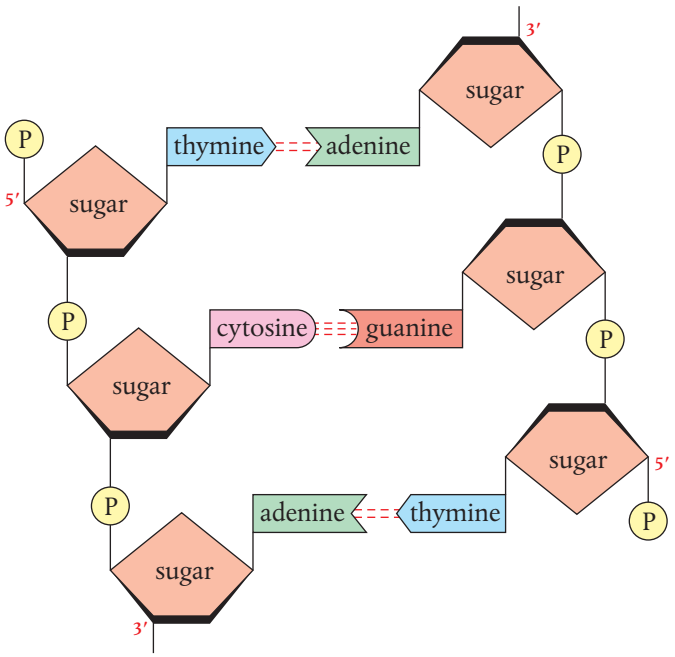


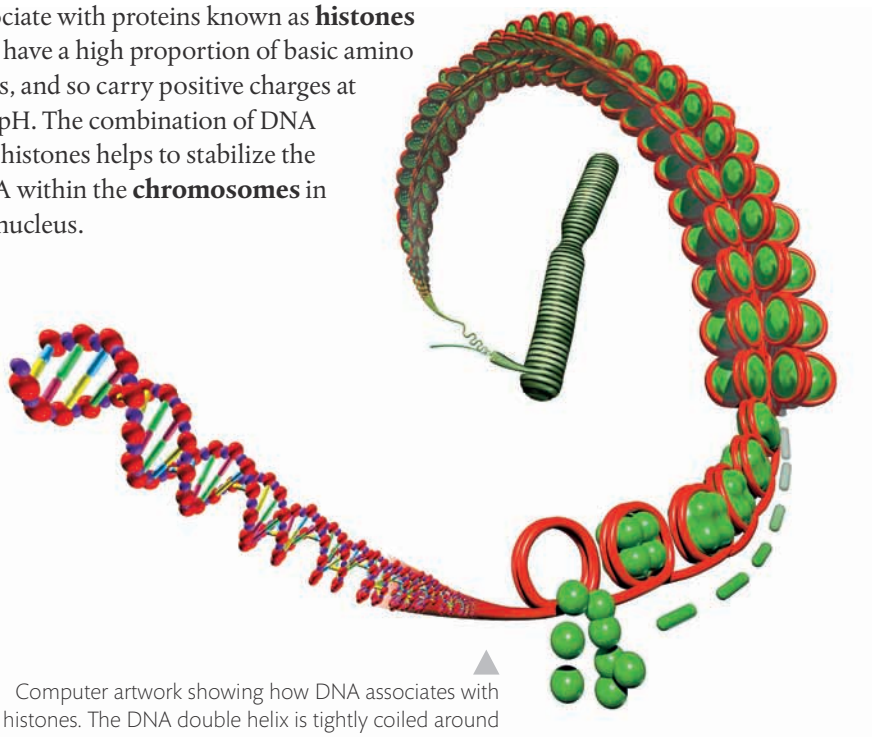
Figure 13.52 Part of a DNA molecule. The two anti-parallel polynucleotide strands are held together by hydrogen bonds between the complementary bases.

- The base pairing between complementary strands provides a means for replication of the code. In a famous piece of understatement, Watson and Crick's 1953 paper concludes: '*It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.*' We will look at this mechanism in more detail later.

In the structure of DNA, only the following base pairs exist: A═T, G≡C.

DNA fragments are observed to migrate towards the positive electrode in electrophoresis, indicating that they are negatively charged. The double helix structure of DNA shows that the origin of this charge is on the phosphate groups that link the sugars together in the backbone of the molecule. The negative charge causes DNA to associate with proteins known as **histones** that have a high proportion of basic amino acids, and so carry positive charges at cell pH. The combination of DNA and histones helps to stabilize the DNA within the **chromosomes** in the nucleus.

DNA carries a negative charge due to its phosphate groups.



Computer artwork showing how DNA associates with histones. The DNA double helix is tightly coiled around histone proteins, shown as green spheres, which are tightly packaged together in groups of eight. These are then coiled together and further packaged into chromosomes, shown in the centre, which are the structures visible within the nucleus.

RNA differs from DNA in that it has

- **ribose sugar instead of deoxyribose**
- **the base uracil instead of thymine**
- **a single-stranded structure.**

TOK

Watson laughed when asked if he and Crick would patent their discovery of DNA, saying that 'there was no use for it'. 20 years later when the process for making recombinant DNA was developed at Stanford University, USA, by Boyer and Cohen, the technique was patented – leading to a revenue of $255 million in the biotech industry before the patent expired in 1997. This has opened many controversies, including the question of right of ownership of biological knowledge. What limits do you think should apply to patenting, which implies sole rights to use information?

RNA is a single-stranded polynucleotide molecule

As we have noted, RNA differs from DNA in that it contains ribose sugar in place of deoxyribose and the base uracil in place of thymine. But in other ways its polynucleotide structure is constructed in the same manner as that of DNA, and it too carries information in its sequence of bases.

However, RNA exists as a single-stranded polynucleotide chain and does not generally form a double helix. It is a less stable molecule than DNA and is usually more short-lived in the cell. RNA is also able to cross the nuclear membrane and so can move between the nucleus and the cytoplasm.

DNA is expressed through protein synthesis

DNA is the genetic material containing all the information for the development of the individual coded in the base sequences along its length. This code, which is essentially built from four 'letters', directs the synthesis of proteins by determining the sequence of their amino acids – their primary structure. Essentially, this means that the 4-unit code of bases in DNA must be translated into a 20-unit code to account for all the amino acids found in proteins. This occurs in two main steps, known as **transcription** and **translation**.

DNA is confined to the nucleus, but protein synthesis occurs on **ribosomes** in the cytoplasm of the cell. So DNA must allow a copy to be made of the relevant part of its information in the form of RNA, which then takes it to the ribosome. We can think of it as similar to the way in which we might copy a recipe from a book that stays in the library, and then take it to the place where we will use the information. The synthesis of RNA from DNA, known as **transcription**, occurs when the two strands of DNA separate by breaking the hydrogen bonds between the paired bases, a process often referred to as **unzipping**. Each strand of DNA can then act as a **template** for the assembly of a complementary strand of RNA from ribonucleotides. The specific base pairing ensures that ribonucleotides complementary to the bases in DNA are aligned in sequence, as shown in Figure 13.53. Note that the base uracil (U) which is used in RNA is complementary to adenine (A) in DNA. In this way the code in DNA is copied exactly. The entire process is controlled by enzymes.

Once formed, the RNA detaches from its DNA template and leaves the nucleus for the ribosome while the DNA re-forms the double helix.

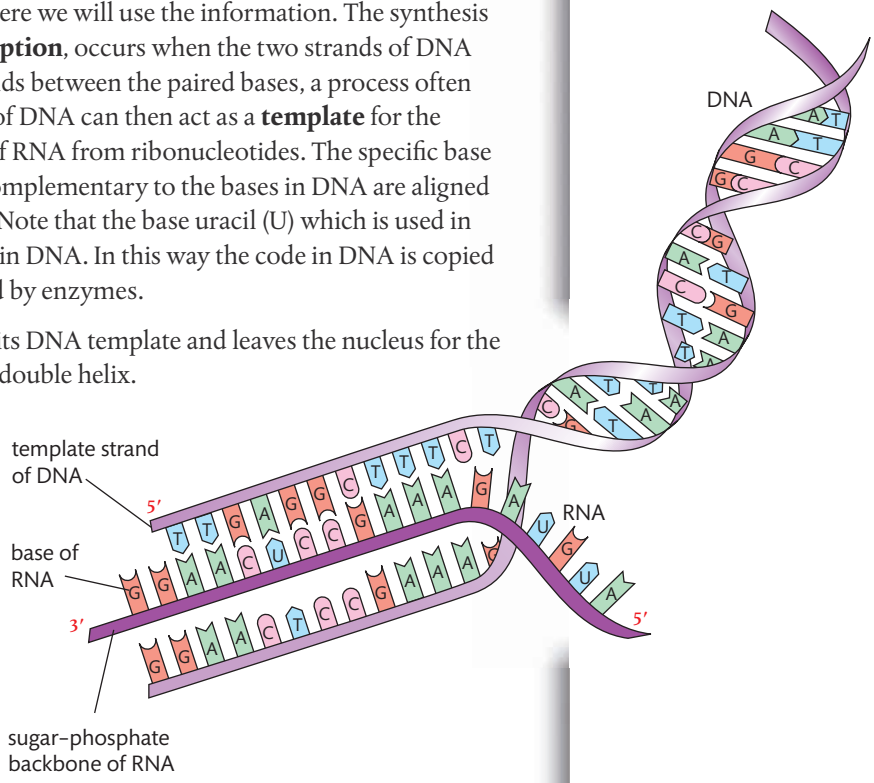
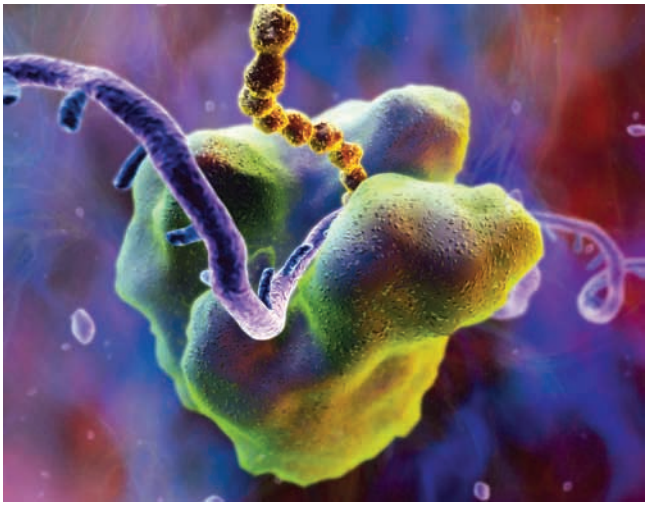


Figure 13.53 Production of RNA from DNA in transcription. Note that the base uracil which is present in RNA base pairs with adenine in DNA.



Computer artwork of a protein being synthesized at a ribosome. RNA, shown in purple, passes between the subunits of the ribosome and provides the instructions for the assembly of the protein. The sequence of bases in the RNA is a copy of the sequence in the DNA and determines the sequence of amino acids in the protein.

At the ribosome, the sequence of bases in RNA is used to determine the sequence of amino acids in a polypeptide. This is known as **translation** and involves the use of a different form of RNA that works like an adaptor. At one end this molecule recognizes a specific triplet of bases known as a **codon** in the transcribed RNA, and at the other end it recognizes a corresponding amino acid. As codons in the RNA are read sequentially, the adaptor molecule brings the appropriate amino acids into position, where they link together by peptide bonds to form a polypeptide.

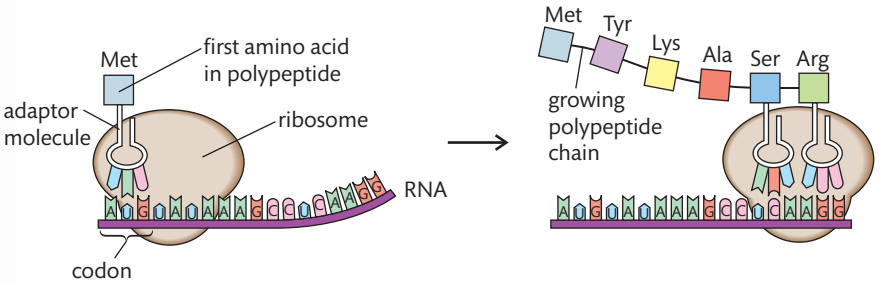


Figure 13.54 Assembly of a polypeptide from a strand of RNA at a ribosome. Codons are recognized by specific adaptor molecules which then insert the corresponding amino acid into the growing polypeptide chain.

Therefore, the sequence of bases in the DNA, via RNA, determines the sequence of amino acids in the protein. The specific relationship between the bases and amino acids is known as the **genetic code**. As we saw above, it is a **triplet code** which means that each sequence of three bases in RNA specifies one amino acid to be inserted into the protein. For example, the sequence GGA specifies glycine while GCA specifies alanine. The universal nature of the code, the fact that the same codon specifies the same amino acid in all organisms, makes possible many of the developments in biotechnology and genetic engineering which have occurred over the last 50 years.

The genetic code is a triplet code. A sequence of three bases in RNA codes for one amino acid and is known as a codon.

The Human Genome Project was the international, collaborative research programme whose goal was the complete mapping of the entire human DNA, three billion base pairs. The project began in 1990 and the complete sequence was published in 2003.

NATURE OF SCIENCE

That the genetic code uses a triplet of bases was postulated initially from logic, but needed experimental data to be confirmed:

- a single base code would have 4 letters, and so could specify only 4 amino acids
- a double base code would have 4 × 4 combinations and so could specify 16 amino acids – still not enough
- the triplet code gives 4 × 4 × 4 = 64 different combinations - more than enough to code for the 20 amino acids that exist.

The hypothesis led to the development of experiments that confirmed the existence of the triplet code, and then determined the amino acid specification of each codon. Elucidation of the genetic code was the subject of the Nobel Prize awarded in Physiology or Medicine in 1968. The 64 variants of the triplet code means that most amino acids are coded for by more than one codon.

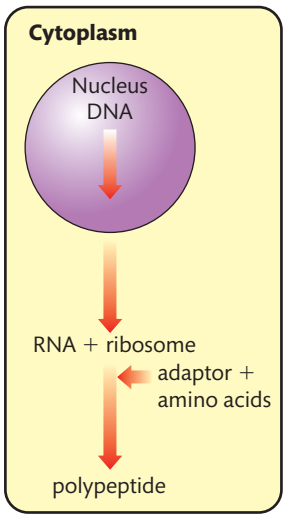


Figure 13.55 The central dogma of molecular biology. The genetic code in DNA is copied to RNA and used to direct the synthesis of a polypeptide at the ribosome.

The term **central dogma**, which was coined by Crick, is sometimes used to summarize these ideas that genetic information flows in one direction in cells, from DNA to RNA to protein. As the primary structure of a protein determines its higher levels of structure and therefore its function, DNA is expressed through protein activity.

DNA replication makes a copy of the genetic information

The process by which a molecule of DNA makes an exact copy of itself is known as **DNA replication**. It always occurs during cell division and ensures that every cell in an organism (with the exception of the sex cells used in reproduction) contains an identical set of genetic information. The process involves separation of the strands in the double helix by breaking the hydrogen bonds between the base pairs, and then using each strand as a template for the synthesis of a new strand. Again the specific base pairing ensures that only a base sequence complementary to that of the template strand will be produced. As DNA replication results in new molecules that contain one strand from the parent molecule and one newly synthesized strand, it is sometimes referred to as **semi-conservative replication**.

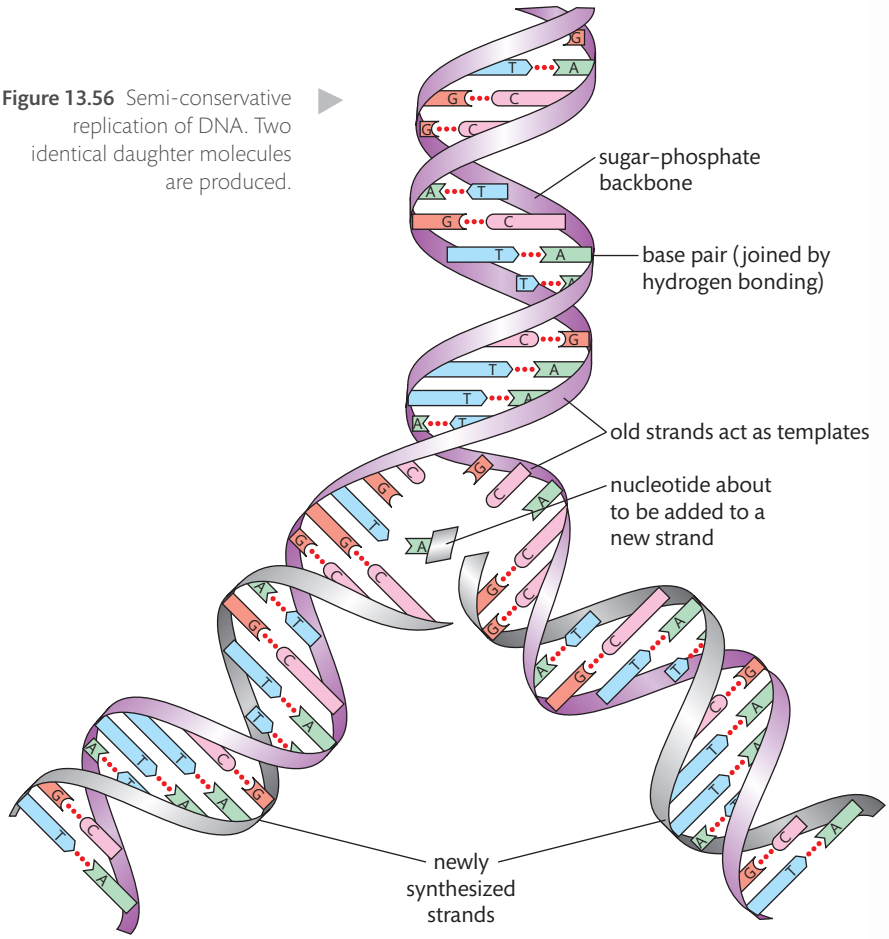


Figure 13.56 Semi-conservative replication of DNA. Two identical daughter molecules are produced.

Make sure you do not confuse DNA replication with transcription. Although they both involve the separation of DNA strands and the use of a single-stranded template, they are very different processes, controlled by different enzymes and making different products.

DNA profiling makes it possible to identify an individual from a small sample of their DNA. This has played an increasing role in forensics with applications in the criminal system, victim identification, immigration, and paternity cases. DNA sequencing is also used in studies of biochemical evolution and human migration.

DNA can be transferred between species

The universal nature of the genetic code makes it possible for DNA from one organism to be expressed by directing protein synthesis when it is transferred into the DNA of a different species. This is the basis of **genetic engineering** which gives rise to **genetically modified organisms** (**GMOs**).

Techniques of inter-species DNA transfer are now advanced, and a wide range of genetically modified foods, known as **GM foods**, are marketed in many countries. The genetic modification may add a gene to yield a new product, inactivate a gene to remove undesired behaviour, or modify a gene for higher yields. Examples include corn which contains a bacterial gene that produces a natural pesticide, rice which produces higher concentrations of vitamin A, and tomatoes that remain fresh for longer. While genetically modified foods such as these promise certain benefits, their development also raises many issues and is the subject of on-going debate. Some aspects of both sides of this debate are summarized here.

TOK

Many countries have DNA databases, with information stored – particularly of individuals who have been convicted of crime. This raises questions of privacy and the right of governments to have access to personal information. Would it be more ethical to record the DNA of every citizen of a country, or would it magnify the concerns about 'big brother'?

Cans of genetically modified tomato puree. The photograph is from the UK where labelling of GM foods is mandatory.

Benefits of GM foods

Some GM foods and GMOs have the following qualities:

- longer shelf-life
- improved flavour, texture, and nutritional value
- increased resistance to diseases and pests, reducing the use of pesticides
- produce a supply of substances such as vitamins and vaccines
- increased crop yields
- tolerance of a wider range of growing conditions, such as drought resistance.

Concerns over GM foods

Many people, however, express the following concerns:

- lack of information about long-term effects
- changes to the natural ecosystem through cross-pollination
- possible links to increased allergies
- risk of altering natural composition of food
- concerns of breeding species that are resistant to control
- in some cases lack of information through food labelling.

Different countries have adopted very different approaches to the marketing of GM foods. In the European Union (EU), strict rules apply to the labelling of GM foods, while in the USA and Canada there is no mandatory labelling of GM content. It has been said that the EU is adopting the 'precautionary principle' in this regard, and this raises several concerns for international trade. At issue is the right for consumers to make informed choices about their food supply.

A genetically modified organism is one whose DNA has been altered, often by the insertion of DNA from a different species.

Exercises

23 **(a)** Outline how nucleotides are linked together to form polynucleotides, explaining the nature of the bonds involved.
(b) Describe the forces that stabilize the DNA molecule.

24 One strand of DNA contains the following base sequence

AATCGCATATAATTCGCTAGC

(a) What is the base sequence in the other strand in the double helix?
(b) What is the sequence of bases in the RNA synthesized using the first strand as a template?
(c) How many amino acids are coded for by this section of RNA?

25 State three perceived benefits and three potential concerns of the use of genetically modified foods.

B.9 Pigments

Understandings:

- Biological pigments are coloured compounds produced by metabolism.
- The colour of pigments is due to highly conjugated systems with delocalized electrons, which have intense absorption bands in the visible region.
- Porphyrin compounds, such as hemoglobin, myoglobin, chlorophyll, and many cytochromes, are chelates of metals with large nitrogen-containing macrocyclic ligands.
- Hemoglobin and myoglobin contain heme groups with the porphyrin group bound to an iron(II) ion.
- Cytochromes contain heme groups in which the iron ion interconverts between iron(II) and iron(III) during redox reactions.
- Anthocyanins are aromatic, water-soluble pigments widely distributed in plants. Their specific colour depends on metal ions and pH.
- Carotenoids are lipid-soluble pigments, and are involved in harvesting light in photosynthesis. They are susceptible to oxidation, catalysed by light.

Guidance

The structures of chlorophyll, heme B, and specific examples of anthocyanins and carotenoids are given in the data booklet in section 35; details of other pigment names and structures are not required.

Applications and skills:

- Explanation of the sigmoidal shape of hemoglobin's oxygen dissociation curve in terms of the cooperative binding of hemoglobin to oxygen.
- Discussion of the factors that influence oxygen saturation of hemoglobin, including temperature, pH, and carbon dioxide.
- Description of the greater affinity of oxygen for fetal hemoglobin.
- Explanation of the action of carbon monoxide as competitive inhibition with oxygen binding.
- Outline of the factors that affect the stabilities of anthocyanins, carotenoids, and chlorophyll in relation to their structures.
- Explanation of the ability of anthocyanins to act as indicators based on their sensitivity to pH.
- Description of the function of photosynthetic pigments in trapping light energy during photosynthesis.
- Investigation of pigments through paper and thin-layer chromatography.

Guidance

- *Explanation of cooperative binding in hemoglobin should be limited to conformational changes occurring in one polypeptide when it becomes oxygenated.*
- *Knowledge of specific colour changes with changing conditions is not required.*

The bright colours of the spectacular flower of *Zingiber spectabile* are the result of different pigment molecules. Photographed in Costa Rica.

Biological pigments are coloured compounds which are produced by metabolism. They include the bright colours in the wings of insects and the feathers of birds, the wide variety of colours of flowers and seaweeds, and the chemicals that give colour to human skin, hair, eyes, and blood. What does this diverse group of molecules have in common?

All pigment molecules have intense absorption bands in the visible region of the spectrum. The colour that we see is the light that is *not* absorbed, but instead is reflected. For example, we saw on page 707 that chlorophyll appears green because it absorbs red and blue light, but reflects the green. The colour seen is the complementary colour to that absorbed, which occupies an opposite position in the colour wheel, as seen in Figure 13.57.

Differences in the colours of hair, eyes, and skin are due to differences in the concentration of the pigment melanin.

Figure 13.57 Complementary colours are opposite each other in the colour wheel.

Pigment molecules absorb visible light because of the nature of their chemical bonds. In most cases they are **highly conjugated** structures, meaning that electrons in p orbitals are delocalized through alternating single and double bonds and through benzene ring structures. As these electrons are not held tightly in one position, they are able to become excited as they absorb certain wavelengths of light energy. The part of the molecule responsible for absorbing the radiation is called the **chromophore**. Important groups of pigments include the porphyrins, carotenoids, and anthocyanins.

NATURE OF SCIENCE

Observations of colour can be made qualitatively by the ability of an individual to detect changes in the environment. But these are subjective measures, which are imprecise and difficult to reproduce. Scientific studies of pigments are more reliably based on quantitative measurements of absorbance data. Improvements in technology involving digital probes have greatly enhanced the speed and accuracy of obtaining this absorbance data.

Porphyrins

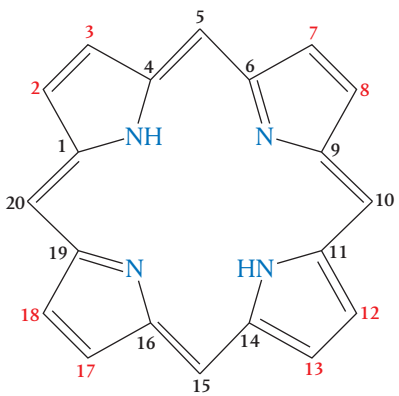


Figure 13.58 The structure of a porphyrin ring.

Porphyrin ring compounds, such as hemoglobin and chlorophyll, contain the planar ring structure shown in Figure 13.58. It is made up of four heterocyclic rings, containing carbon and nitrogen, linked by bridging carbon atoms. The ring acts as a ligand, forming a chelate with a metal involving coordinate bonds. Different porphyrin compounds contain different metals and also differ in the nature of the substituent groups which are attached to the carbon atoms labelled 2, 3, 7, 8, 12, 13, 17, and 18 in Figure 13.58.

Chlorophyll: the primary photosynthetic pigment

Several different forms of chlorophyll exist, but the pigment always contains magnesium. This is why magnesium deficiency in the soil leads to loss of the green colour in leaves.

Yellow patches on leaves caused by chlorophyll deficiency due to lack of magnesium.

Chlorophyll is the primary pigment in photosynthesis, acting to absorb light energy. We saw on page 707 that chlorophyll absorbs light strongly in the blue part of the spectrum and to a lesser extent in the red. Other photosynthetic pigments, known as **accessory pigments**, harvest light in different parts of the spectrum and pass their energy to chlorophyll. As a result, chlorophyll undergoes a redox change, passing electrons to a series of **electron transport carriers**. Ultimately, chlorophyll is reduced back to its original state by gaining electrons from water. The process stores 'reducing power' which is able to reduce carbon dioxide to carbohydrate in reactions that do not depend on light energy.

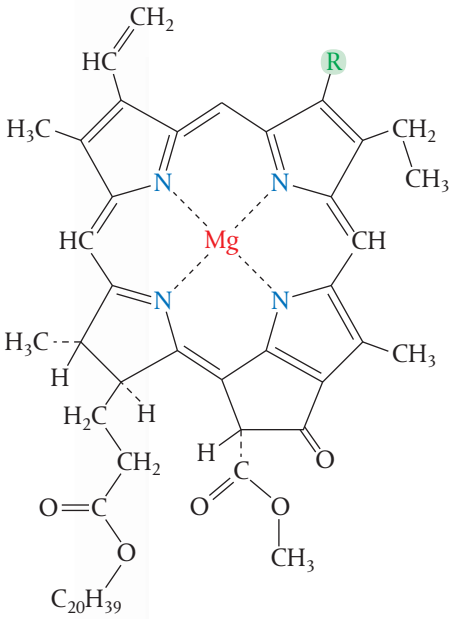


Figure 13.59 The structure of chlorophyll:
R = $-CH_3$ in chlorophyll a
R = $-CHO$ in chlorophyll b

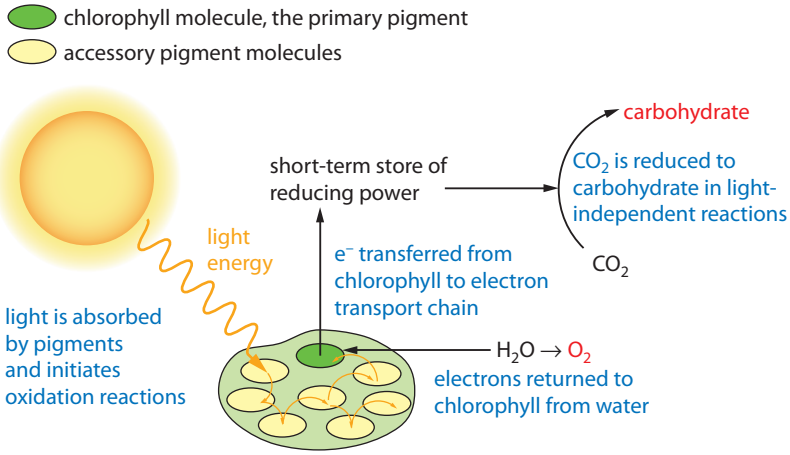


Figure 13.60 The absorption of light by photosynthetic pigments initiates redox reactions that ultimately lead to the storage of chemical energy in carbohydrates.

The thermal stability of chlorophyll depends on the pH. In acidic solution Mg is lost from the porphyrin ring and replaced by two H^+ ions. This causes a colour change from green to olive-brown as the chromophore is altered. Cooking food often breaks cell membranes releasing acids which decrease the pH and bring about this change. Chlorophyll is more stable in alkaline conditions, which is why sodium hydrogencarbonate is sometimes added to water during cooking. The bright green colour of chlorophyll is often used as an indication of the freshness of food.

The structures of chlorophyll and the heme group are given in section 35 of the IB data booklet.

Hemoglobin and myoglobin: the oxygen-carrying team

The heme group, which is common to hemoglobin and myoglobin, contains iron – usually in the +2 oxidation state. Heme is a prosthetic group within protein molecules. Hemoglobin contains four heme groups, each bound within a polypeptide chain. In other words it is a protein with a quaternary structure. Myoglobin, on the other hand, contains one heme group and a single polypeptide chain.

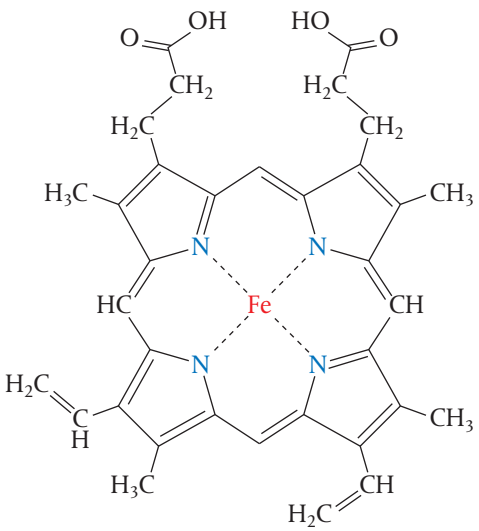
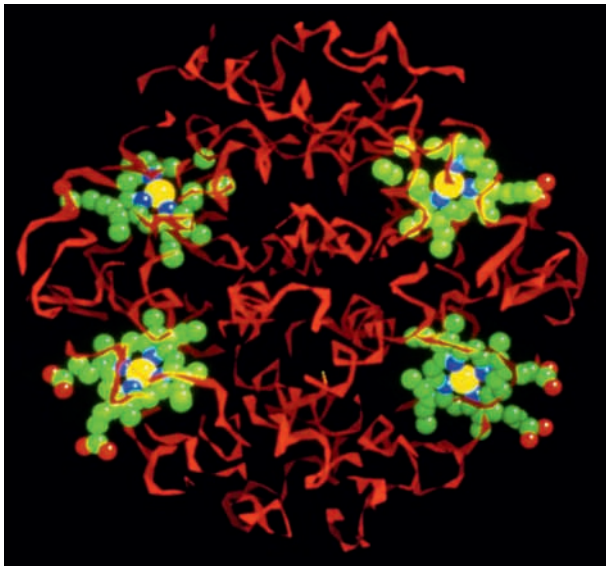


Figure 13.61 The structure of heme.



Computer graphics model of the protein hemoglobin. The yellow balls represent Fe^{2+} ions in the centre of the four porphyrin rings, shown in green. The protein environment of the iron, shown as the red strands, enables oxygen to be bound, carried, and released, without the iron being oxidized.

Iron deficiency is a prevalent micronutrient deficiency globally. Because of iron's role in hemoglobin, its deficiency leads to a serious condition known as anemia – with symptoms of fatigue, poor endurance, and lowered immunity. Iron is found in red meats, green leafy vegetables, nuts, and seeds. While its dietary deficiency is a cause of concern in many parts of the world, the question of how best to alleviate this is complicated by the fact that iron supplementation may increase susceptibility to malaria, which is also widespread.

When the iron atom in hemoglobin or myoglobin is oxidized to Fe(III), this form of the pigment appears brown rather than red and is unable to bind with oxygen. It is the form of myoglobin seen in cooked meat or meat that has 'gone off' through oxidation.

Hemoglobin is designed to carry oxygen in the blood and myoglobin to store it, mostly in muscles. Both molecules do this by binding reversibly with molecular oxygen, O_2, which forms a weak bond with the iron atom. This binding does not change the oxidation of Fe, which remains in the +2 state, so the products are said to be **oxygenated** rather than oxidized, and are known as **oxyhemoglobin** and **oxymyoglobin** respectively. Due to the number of their heme groups, hemoglobin can bind four molecules of O_2, while myoglobin can bind one. The equation below summarizes these reactions, using Hb to represent hemoglobin and Mb to represent myoglobin.

$$Hb + 4O_2 \underset{\text{in respiring cells}}{\overset{\text{in lungs}}{\rightleftharpoons}} Hb{-}(O_2)_4 \quad \text{oxyhemoglobin}$$

$$Mb + O_2 \rightleftharpoons Mb{-}O_2 \quad \text{oxymyoglobin}$$

The reversible nature of these reactions is central to their role. Clearly, hemoglobin needs to associate with oxygen in the lungs, and dissociate to release the oxygen at respiring cells.

The binding of hemoglobin to oxygen is **cooperative** in nature. This means that the ability to bind oxygen is increased by the initial binding of oxygen to a heme group in the molecule. In simple terms this means that the uptake of oxygen by the second, third and fourth heme groups within a molecule gets easier as oxygen binds. This can be seen in the shape of the graph below, which is described as **sigmoidal** or S-shaped.

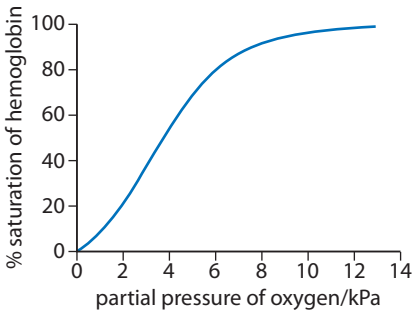


Figure 13.62 The binding curve of hemoglobin and oxygen is sigmoidal, showing cooperative binding. Note that partial pressure is effectively a measure of oxygen concentration. These graphs are often referred to as oxygen dissociation curves.

This cooperative effect is due to subtle changes that occur in the quaternary structure as oxygen binds. A conformational shift caused by the binding of oxygen at one heme group makes the other heme groups more receptive to oxygen. It is an allosteric effect. From the graph we can deduce the following about how this affects hemoglobin's ability to bind O_2.

- At low concentrations of O_2, hemoglobin has a low affinity for O_2.
- At high concentrations of O_2, hemoglobin has a high affinity for O_2.

In other words, the equilibrium

$$Hb + 4O_2 \rightleftharpoons Hb{-}(O_2)_4$$

shifts to the right in the lungs, causing oxygen uptake where oxygen concentration is high, and shifts to the left in respiring cells, releasing oxygen where oxygen concentration is low. This fits exactly with the metabolic requirements.

Other factors that influence the binding of oxygen to hemoglobin are temperature, pH, and carbon dioxide. The effects of these are shown in Figure 13.63.

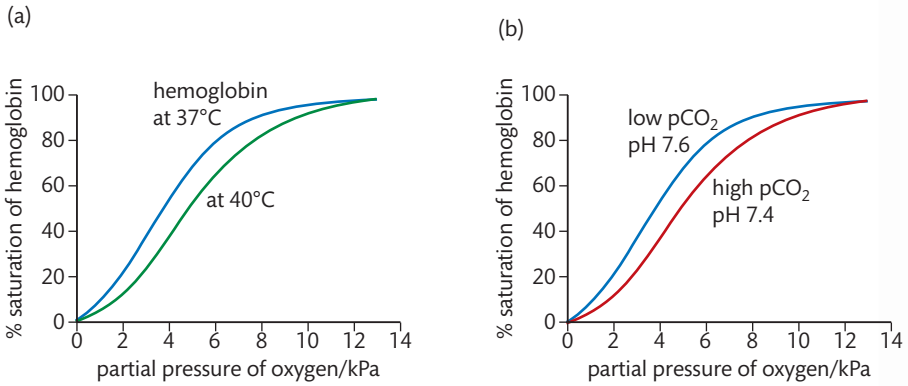


Figure 13.63 Oxygen dissociation curves: (a) the influence of temperature; (b) the influence of pH/carbon dioxide.

Figure 13.63(a) shows that increasing temperature reduces the affinity of hemoglobin for O_2, as the dissociation curve has shifted to the right. This means that oxyhemoglobin more readily releases its oxygen in conditions of higher temperature, for example in cells during high metabolic activity such as exercise.

Figure 13.63(b) shows that decreasing pH reduces the affinity of hemoglobin for O_2, because the curve has shifted to the right as the acidity increases. Note that increases in the concentration of CO_2 have the same effect, as carbon dioxide dissolves to form carbonic acid and increases the acidity of the blood. During respiration, carbon dioxide is produced, decreasing the pH and so causing oxyhemoglobin to dissociate more and release oxygen where it is needed.

The equation below summarizes the factors that influence the equilibrium position in the oxygenation of hemoglobin.

$$Hb + 4O_2 \underset{\text{in respiring cells, low } pO_2,\ \text{higher temperature, low pH / high } pCO_2}{\overset{\text{in lungs, high } pO_2,\ \text{lower temperature, high pH / low } pCO_2}{\rightleftharpoons}} \underset{\text{oxyhemoglobin}}{Hb{-}(O_2)_4}$$

The four polypeptide chains in adult hemoglobin are two α-chains and two β-chains. Before birth hemoglobin in the fetus has a different structure with two α-chains and

Partial pressure is an indication of the concentration of a gas and is shown by the letter p before the name of the gas, e.g. pO_2, pCO_2.

The response of oxygen dissociation curves to changes in pH or CO_2 is known as the Bohr effect. It was first described in 1904 by the Danish physiologist Christian Bohr, the father of Neils Bohr (see Chapter 2).

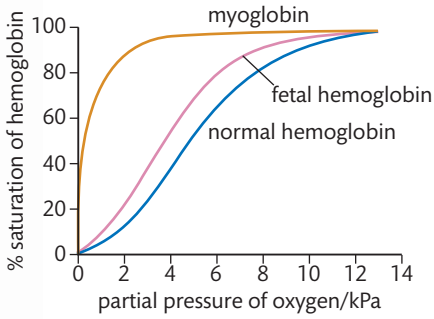


Figure 13.64 Both fetal hemoglobin and myoglobin have a higher affinity for oxygen than adult blood.

two γ-chains. This form has a higher affinity for oxygen, as seen in Figure 13.64. The fact that its oxygen dissociation curve lies to the left of adult blood indicates that it is able to extract oxygen from the maternal blood. Following birth the fetal hemoglobin levels decline and after six months adult hemoglobin becomes the predominant form.

Animals that have evolved to live at high altitude where the oxygen concentration is low have developed hemoglobin with a higher affinity for oxygen.

Myoglobin also has an oxygen dissociation curve to the left of that of hemoglobin, which means it has greater affinity for oxygen and can pick it up from hemoglobin for storage. Note that its dissociation curve is not sigmoidal in shape as there can be no cooperative binding within its one heme structure.

Carbon monoxide is a toxic gas, primarily because of its ability to bind to hemoglobin. Its affinity for hemoglobin is about 200 times that of oxygen, and so it effectively makes the hemoglobin unavailable to carry oxygen to respiring cells. Binding of carbon monoxide to hemoglobin forms carboxyhemoglobin, which does not readily dissociate.

Factors that increase hemoglobin's affinity for O_2 displace the oxygen dissociation curve to the left; factors that decrease hemoglobin's affinity displace the oxygen dissociation curve to the right.

$$\underset{}{Hb + CO} \rightarrow \underset{\text{carboxyhemoglobin}}{Hb{-}CO}$$

Carbon monoxide poisoning can occur from the burning of fossil fuels with insufficient ventilation and from smoking tobacco.

Cytochromes : electron-transport carriers

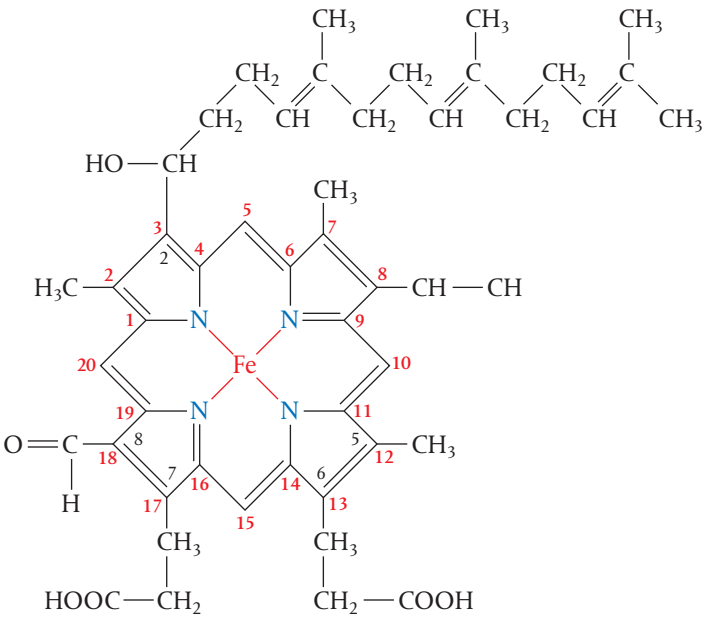
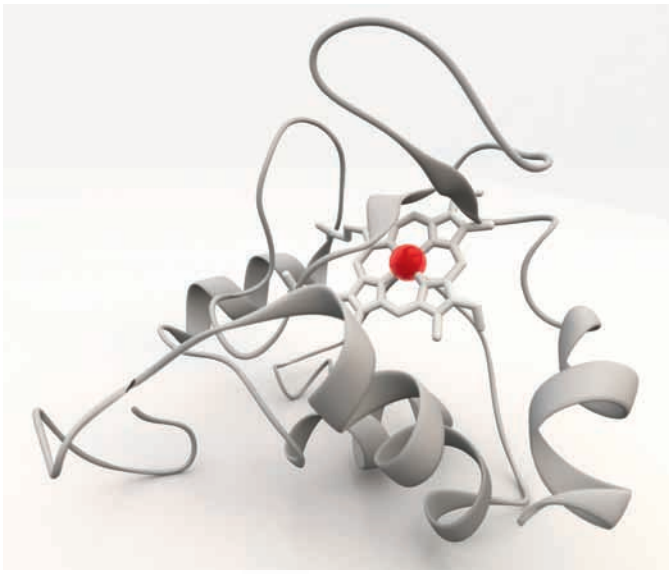


Figure 13.65 Heme structure of cytochromes.

Cytochromes are a varied group of protein molecules that also contain the heme prosthetic group. They are found embedded in membranes and are responsible for electron transport during the redox reactions of aerobic respiration and photosynthesis. During the reactions they become successively reduced and then re-

In hemoglobin and myoglobin, the iron of the heme group remains in the Fe(II) state during O_2 transport. In cytochromes, the iron changes reversibly between Fe(II) and Fe(III) during electron transport.



Computer artwork of the molecular structure of the protein cytochrome c. The secondary structure of alpha helices is shown, along with the heme group surrounding the iron ion, shown in red. Cytochromes act as electron carriers in aerobic respiration, with the iron being successively reduced and re-oxidized in the process. They are found in the membranes of all cells..

oxidized as they in turn accept and then pass on electrons. They are organized in sequence, corresponding to their electrode potentials, so that the electrons effectively flow down an electrochemical gradient.

Interestingly, despite the similarity of their structure with hemoglobin, their carrier mechanism is quite different. In cytochromes the iron of the heme group interconverts its oxidation state between +2 and +3 as the cytochrome undergoes redox change. The final cytochrome involved in aerobic respiration passes its electrons to the terminal acceptor oxygen with the formation of water. This is also the site of inhibition of the poison cyanide: by blocking the chain it prevents aerobic respiration from occurring, which is why it is such a potent poison.

Carotenoids

Carotenoids are a group of pigments containing long hydrocarbon chains with many double bonds. They range in colour from yellow to red, and are fat soluble due to the long non-polar hydrocarbon chain. The structure of β-carotene, a carotenoid commonly found in fruit and vegetables, is shown in Figure 13.66.

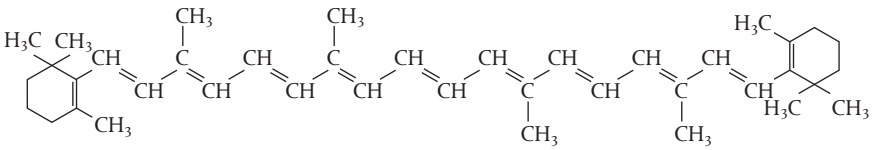


Figure 13.66 The structure of β-carotene.

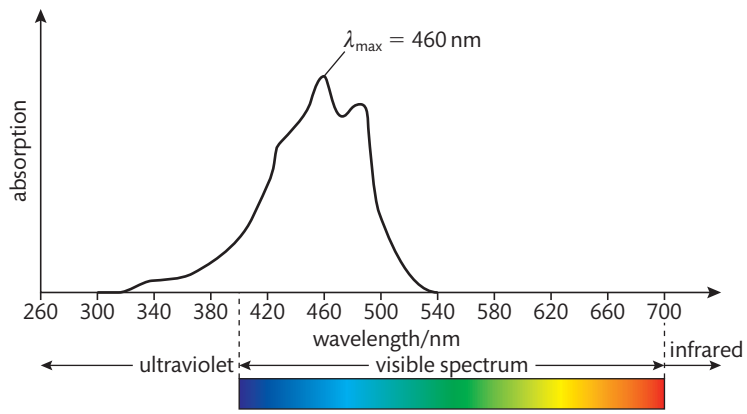


Figure 13.67 The UV-visible spectrum of carotene. Blue/violet light is absorbed, due to conjugation in the long hydrocarbon chain, and so carotene appears orange.

α- and β-carotene are vitamin A precursors, and so play an important role in promoting healthy vision. Carotenoids are found in plant leaves, where they help in the harvesting of light for photosynthesis. They are accessory pigments, as described on page 741, which help to pass light energy to chlorophyll.

Plants use anthocyanins and carotenoids to attract pollinators and agents for seed dispersal. Some animals use the colours as warnings of poisons, and pigments also help to protect from UV damage.

Assortment of fruits that contain zeaxanthin, which is a carotene. Zeaxanthin is important in the physiology of the eye and is thought to play a role in reducing the risk of cataracts.

The multiple conjugated carbon–carbon double bonds which give the carotenoids their colour also make them susceptible to oxidation, including that catalysed by light. This is why they are able to act as antioxidants. Oxidation can lead to a loss of vitamin A activity, as well as loss of colour and 'off' odours due to the release of volatile compounds. The chemical changes that occur during oxidation are not fully understood, but involve conversion of *trans*-carotenoids into *cis*-isomers. Oxidation of carotenoids can be reduced by preventing exposure to air and light, and by decreasing storage time.

Processed food commonly contains artificial colour additives. Countries differ in the compounds approved for this purpose and in the way in which the information is communicated to the consumer. In Europe the compounds are designated as 'E numbers', while in North America they are usually listed by chemical name.

Anthocyanins

Anthocyanins are a widely distributed group of pigments, responsible for many of the pink, red, and blue colours of plants. They absorb strongly in the blue and green parts of the spectrum. They are aromatic compounds with a three-ring $C_6C_3C_6$ structure and conjugated carbon–carbon double bonds.

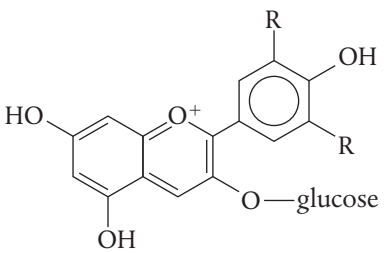


Figure 13.68 Flavylium cation, an anthocyanin.

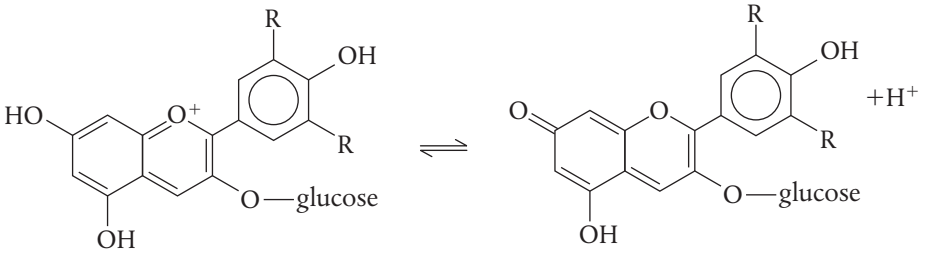
The word 'anthocyanin' is derived from two Greek words: *anthos* meaning flower and *kyanos* meaning blue.

Cranberries and cranberry juice which are rich in anthocyanins. It is thought these products help to reduce the chance of recurrent urinary tract infection.

The polar hydroxyl groups allow the molecules to form hydrogen bonds, which increase their solubility in water. Consequently they are found dissolved in the aqueous cell sap rather than in the lipid-rich membranes.

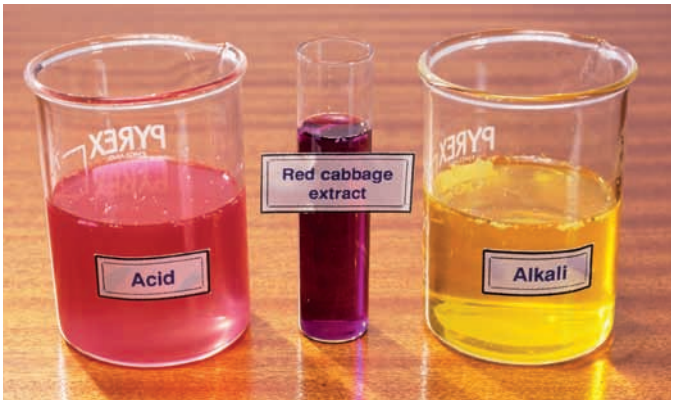
Anthocyanins are formed by a reaction between sugars and proteins that requires light. This is why fruit often changes colour as it ripens and the sugar concentration increases. The colour of anthocyanins also changes as the pH of cell sap changes; they are generally pink in acidic solution, purple in neutral solution and turn greenish-yellow in alkaline solution.

The sensitivity of anthocyanins to pH means that they can be used as pH indicators. The colour changes arise from transfer of H^+ from OH groups, which alters the conjugation and so the absorbance at the chromophore. For example, the red flavylium cation changes to the blue molecule quinonoid as pH increases.



(AH$^+$) flavylium red ⇌ (A) quinonoid blue

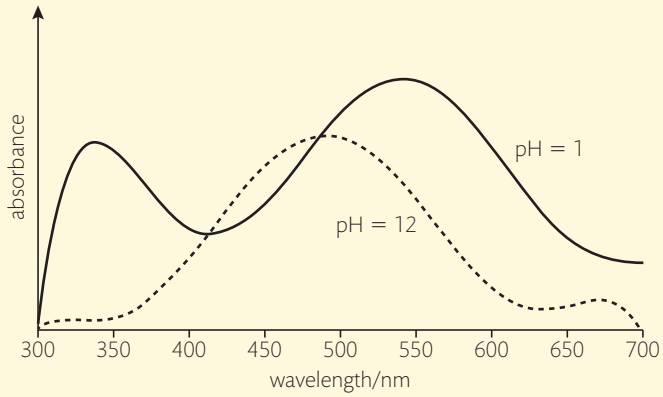
Adding acid will shift the equilibrium back to the left so the red colour reappears.



Red cabbage indicator in acid and alkaline solutions. The colour change is due to changes in ionization of anthocyanins.

Worked example

The anthocyanins can be used as acid–base indicators. Identify the wavelength λ_{max} which corresponds to maximum absorbance at the different pH values shown and suggest the colour of the pigment in acid and in basic conditions.



Solution

pH	λ_{max}	Colour absorbed	Colour of pigment
1	550	green	red
12	475	blue	orange/yellow

The changing colours of leaves with the seasons is due to changes in the relative amounts of chlorophylls, carotenoids, and anthocyanins present. Chlorophyll is often lost first as the temperature decreases, and the bright orange, yellow, and red colours seen in many leaves are due to the higher proportion of carotenes and anthocyanins.

The anthocyanins also form deeply coloured coordination complexes with Fe^{3+} and Al^{3+} ions that are present in metal cans. This sometimes causes a discolouration in canned fruit.

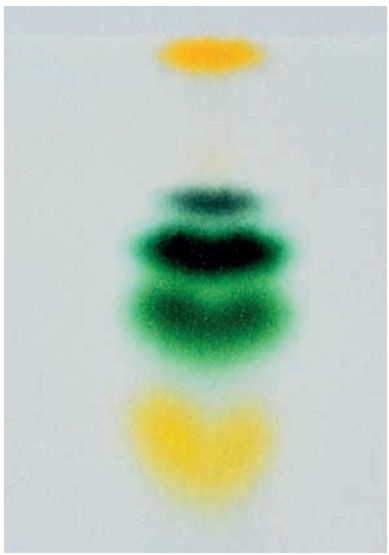
The bright colours of autumn leaves are mainly due to carotenoids and anthocyanins. Photographed in Washington State, USA.

Analysis of pigments

Pigment extracts from plants typically contain a mixture of different pigment molecules. They are suitable for analysis by two different types of chromatography.

- Paper chromatography, which was described on page 701.
- Thin-layer chromatography, described below.

Thin-layer chromatography is an example of adsorption chromatography, which follows the same basic principles as paper chromatography. The stationary phase in this case is a thin layer of adsorbent particles of alumina or silica which may be only 0.2 mm thick. It is supported on glass or thin plastic plate. The mobile phase is the solvent, chosen according to the chemical nature of the pigments. Small spots of the pigment extract are placed on the origin, and capillary action causes the different compounds in the spot to separate, leading to identification from their R_f values.



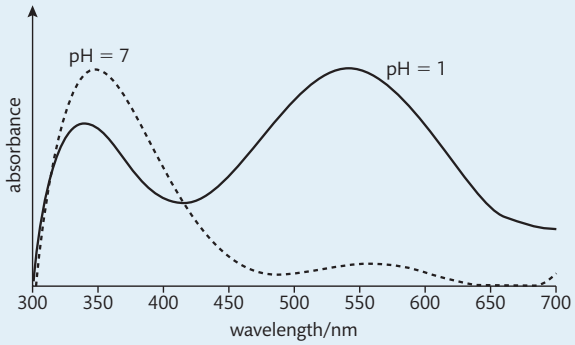
Thin-layer chromatogram (TLC) of an extract of pigments from chloroplasts. TLC plastic sheets are coated with a silica gel and a drop of the extract was placed at the bottom of the origin. The sheet was then placed in a beaker of solvent. The chromatogram shows the different solubilites of pigments in the extract in the solvent. The pigments can be identified as carotene, pheophytin, chlorophyll a, chlorophyll b, and carotenoids. The line across the top of the image is the solvent front.

Thin-layer chromatography is generally a quicker process than paper chromatography, and works efficiently on smaller samples. It is also a more sensitive technique with results which are more easily reproduced.

Exercises

26 With respect to hemoglobin and cytochromes, explain the difference between being *oxidized* and being *oxygenated*.

27 The absorbance spectra of anthocyanins are very sensitive to changes in pH. Identify the wavelength λ_{max} which corresponds to maximum absorbance and suggest the colour of the pigment at the different pHs shown.



28 Explain the following observations in terms of the molecular structure of the pigments involved.

(a) When carrots are boiled, little colouration of the water occurs; when they are fried, the oil changes colour to orange.

(b) When red cabbage is boiled, the water turns purple-red. When vinegar is added the colour changes to bright red.

29 Three compounds were separated using thin-layer chromatography on a silica gel stationary phase.

Compound	Distance travelled / cm
A	2.5
B	7.5
C	10.0
solvent	15.0

Calculate the R_f values and comment on the relative polarity of the components.

30 Explain the difference in the shapes of the oxygen dissociation curves of hemoglobin and myoglobin in terms of their protein structures.

B.10 Stereochemistry in biomolecules

Understandings:

- With one exception, amino acids are chiral, and only the L-configuration is found in proteins.
- Naturally occurring unsaturated fat is mostly in the ***cis*** form, but food processing can convert it into the ***trans*** form.
- D- and L-stereoisomers of sugars refer to the configuration of the chiral carbon atom furthest from the aldehyde or ketone group, and D forms occur most frequently in nature.
- Ring forms of sugars have isomers, known as α and β, depending on whether the position of the hydroxyl group at carbon 1 (glucose) or carbon 2 (fructose) lies below the plane of the ring (α) or above the plane of the ring (β).
- Vision chemistry involves the light activated inter-conversion of ***cis*** and ***trans*** isomers of retinal.

Applications and skills:

- Description of the hydrogenation and partial hydrogenation of unsaturated fats, including the production of *trans* fats, and a discussion of the advantages and disadvantages of these processes.
- Explanation of the structure and properties of cellulose, and comparison with starch.
- Discussion of the importance of cellulose as a structural material and in the diet.
- Outline of the role of vitamin A in vision, including the roles of opsin, rhodopsin, and *cis*- and *trans*-retinal.

Guidance

- *Names of the enzymes involved in the visual cycle are not required.*
- *Relative melting points of saturated and cis/trans unsaturated fats should be covered.*

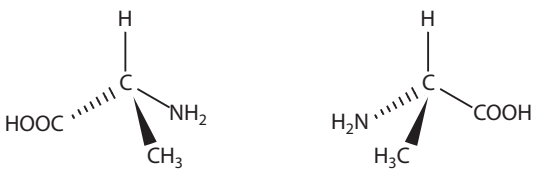
NATURE OF SCIENCE

The existence of stereospecific environments is universal in biological systems. This is in contrast to reactions that take place outside living cells, which more commonly involve racemic mixtures rather than one enantiomer. This difference has sometimes been used as a broad guiding distinction between living and non-living matter.

As we learned in Chapter 10, stereoisomers represent different spatial arrangements of the atoms in a molecule, and their study involves a three-dimensional approach. Many biopolymers can exist as stereoisomers, and metabolic reactions are usually stereospecific, meaning that only one form of the isomer has the required activity. In this section we will consider examples of stereochemistry in each of the main groups of biomolecules we have studied.

Stereochemistry in proteins

In amino acids, the amino group, the carboxylic acid group, a hydrogen atom, and a variable group R are all attached to the same carbon atom, known as the α-carbon or carbon-2 in the numbering of the chain. So with four different groups attached, this carbon atom is chiral. This means that amino acids are optically active and can exist as two different stereoisomers, known as enantiomers. The stereochemistry of alanine is shown in Figure 13.69.



D-amino acid L-amino acid

Figure 13.69 L and D forms of the amino acid alanine where R = CH_3.

The only exception is glycine, in which R = H and so there is no chiral carbon atom.

The different stereoisomers of the amino acids are most commonly known as the L and D forms, applying a convention that compares their absolute configuration to that of glyceraldehyde. The L and D forms of amino acids have identical physical properties and chemical reactivities *apart from* the direction in which they rotate plane-polarized light and their reactions with reagents that are chiral. This last point is crucial in biochemistry. As enzymes, themselves made of proteins, are chiral molecules, they distinguish completely between the L and D forms of amino acids. Biological systems have evolved to use only the L forms of amino acids.

All naturally occurring amino acids in proteins are the L form.

D-amino acids are scarce in nature but are widely used in the pharmaceutical industry, occurring in many drugs, including antibiotics such as penicillin and some anti-cancer agents. Preparation of D-amino acids industrially involves the use of enzymes to resolve racemic mixtures of amino acids. Recent research on preparing D-amino acids by direct synthesis is a promising development that could provide much cheaper synthetic routes for many drugs.

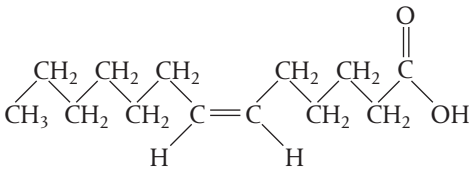
CHALLENGE YOURSELF

8 How many stereoisomers exist for threonine and isoleucine? Refer to section 33 of the IB data booklet.

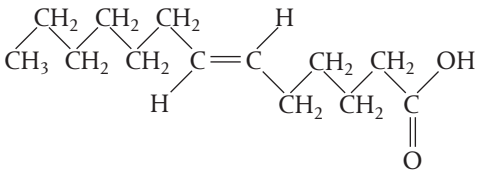
Stereochemistry in lipids

Unsaturated fatty acids in fats and oils contain carbon–carbon double bonds. These exist in two forms, known as *cis–trans* isomers, which arise due to the restriction on rotation around the double bond. They were introduced in Chapter 10, page 516.

- The *cis* form occurs when the same group, for example hydrogen, has the same orientation relative to the double bond.

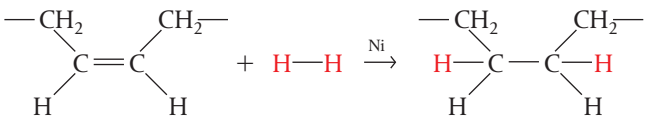


- The *trans* form occurs when the same group has opposite orientation across the double bond.



With the exception of some dairy products, most naturally occurring unsaturated fats are in the *cis* form. Molecules of the *cis* isomer cannot easily arrange themselves side by side to solidify, so they tend to have lower melting points than the corresponding *trans* isomer.

Hydrogenation of fats takes place in the food industry when hydrogen is added across the carbon–carbon double bonds using a finely divided metal catalyst such as nickel.



The product is a fat that, being more saturated, has a higher melting point, and therefore is a more convenient form for packing and storage as a solid or semi-solid. Fats made in this way also break down less easily under conditions of high temperature frying and usually have a longer shelf-life than liquid oils. Most margarines and shortening come into this category.

There is, however, a problem with the process. In **partial hydrogenation**, only some of the carbon–carbon double bonds in a fat are broken, and those that remain often get chemically modified from the *cis* position to

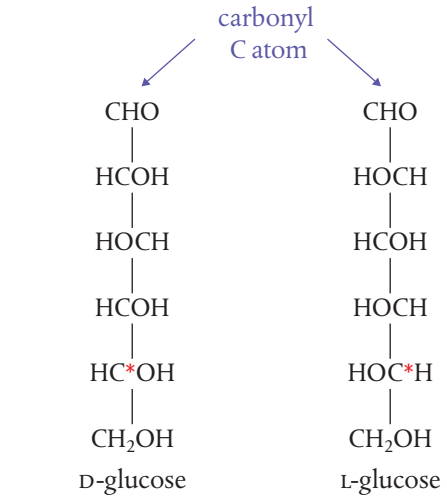
Foods containing *trans* fats produced by hydrogenation. The process is used to solidify fats and extend their shelf-life, but has been linked to increased risk of heart disease.

The growing awareness of the link between *trans* fats and cardiovascular and other diseases has caused different responses around the world. Canada was the first country to introduce mandatory labelling of *trans* fats on food products, while Denmark, Iceland, and Switzerland have banned their use. The World Health Organization has called for the elimination of *trans* fatty acids from the global food supply. The cost of this is, however, considerable as it will involve major changes in food formulations.

the *trans* position. The resulting fatty acids are therefore known as ***trans* fats** and are particularly prevalent in processed foods. Evidence shows that consuming *trans* fats raises the level of LDL cholesterol, which is a risk factor for heart disease. *Trans* fats also reduce the blood levels of HDL cholesterol, which protects against heart disease.

Stereochemistry in carbohydrates

All simple sugars are chiral molecules as they contain at least one chiral carbon atom. The stereoisomers are described as D and L, again in reference to their configuration relative to glyceraldehyde. For sugars having two or more chiral carbon atoms, the prefixes D and L refer to the configuration of the chiral carbon atom furthest away from the carbonyl carbon.

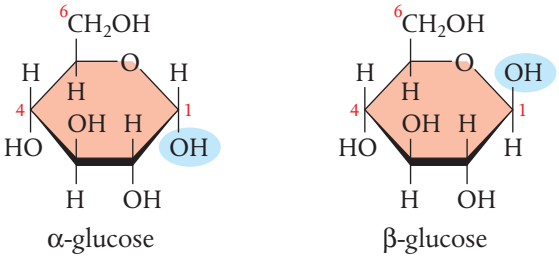


*this C atom determines notation D or L

D sugars are the most abundant form in nature.

The conversion of sugars in the straight-chain form to the ring form, described on page 722, creates an additional type of isomer, known as α and β forms. These are distinguished by the relative position of the groups attached to the carbon atoms that close the ring by forming an ether link with oxygen.

In glucose, α and β forms are determined by the positions of the –OH group at C_1.

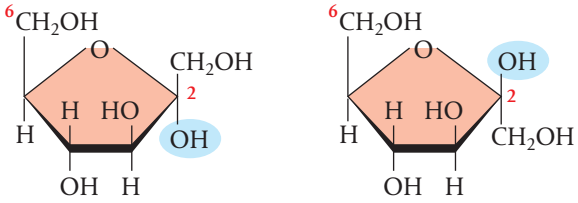


The alpha form has the —OH at C_1 on the opposite side of the ring to C_6. In the Haworth structure this is represented as a downwards projection. The beta form has the —OH on the same side of the ring as C_6, giving an upwards projection.

CHALLENGE YOURSELF

9 Which form of glucose, α or β, would you expect to be the more stable and why?

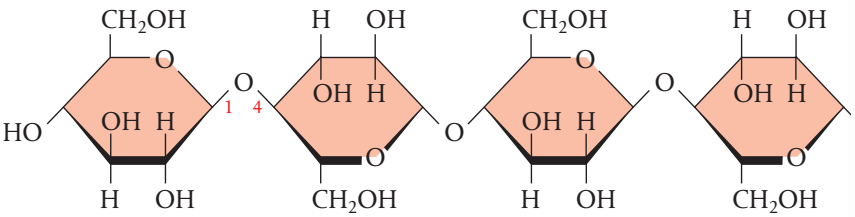
In fructose, the designations are similar, but with the focus on C_2 instead of C_1.



α-fructose β-fructose

The structural differences between α and β glucose have a large effect on the properties of their polymers.

- Starch and glycogen are polymers of α-glucose. Starch forms a relatively compact spiral structure and is stored as starch grains in plant cells. The structure of starch shown on page 724.
- Cellulose is a polymer of β-glucose. It is a linear polymer with 1–4 links known as β-glycosidic links. These position the sugars at a different angle from the α-glycosidic links found in starch so the cellulose chain forms an uncoiled linear structure with alternate glucose monomers 'upside down' with respect to each other. This enables the hydroxyl groups to form hydrogen bonds with the hydroxyls of other cellulose molecules lying parallel.



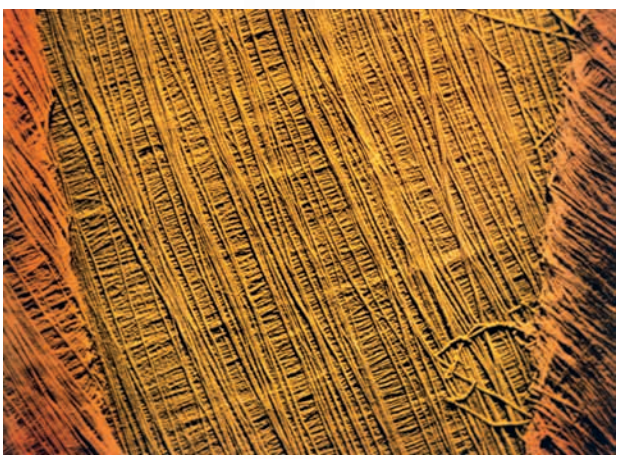
cellulose: 1-4 linkage of β-glucose monomers

Consequently, cellulose forms cables, known as microfibrils, of parallel chains that give it a rigid structure. Cellulose is found in all plant cell walls and is one of the main sources of support in plant cells. This is why wood, which is rich in cellulose, is such a useful building material.

Starch and glycogen can be relatively easily hydrolysed into glucose by the action of digestive enzymes, but the human body does not produce an enzyme to hydrolyse the β-glycosidic links in cellulose. (The required enzyme, known as **cellulase**, is secreted by some bacteria, such as those living in the gut of ruminants.) Cellulose therefore passes through the gut largely chemically intact, contributing to the bulk of the faeces. Medical studies have indicated that this **dietary fibre** is of benefit to the health of the large intestine. The cellulose fibrils abrade the wall of the digestive tract and stimulate the lining to produce mucus which helps in the passage of undigested food through the gut. Fibre in the diet helps to reduce conditions such as constipation, haemorrhoids, and, possibly, colorectal cancer. In general, foods derived from plants with little or no processing are likely to be a good source of fibre.

The World Health Organization cites a low fruit and vegetable intake as a key risk factor in chronic diseases such as diabetes, obesity, and some cancers. Our growing understanding of the significance of fibre in the diet has led to an increase in the marketing of 'whole foods' such as grains and plant foods such as vegetables and salads. But fruit and vegetable intake varies considerably among countries, largely reflecting the prevailing economic, cultural, and agricultural environments. In developed countries fresh fruit and vegetable intake has decreased with increasing dependence on fast foods that are highly processed. It is estimated that globally 2.7 million deaths per year can be traced to low fruit and vegetable intake.

Starch and glycogen are polymers of α-glucose, cellulose is a polymer of β-glucose.



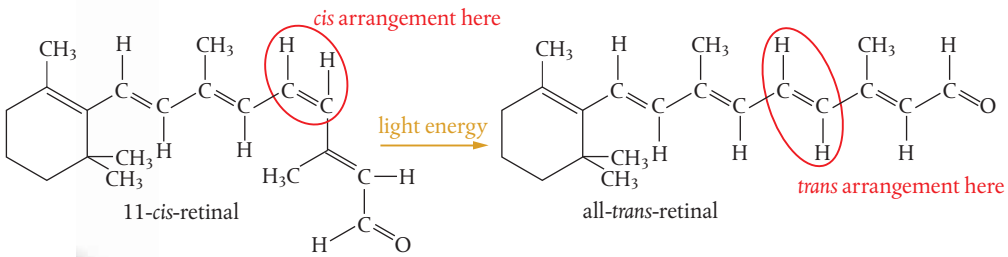
Coloured scanning electron micrograph of cellulose microfibrils in a plant cell wall. Microfibrils measure between 5 nm and 15 nm in diameter.

Stereochemistry in vitamins

Vitamin A, also known as retinal, is involved in the so-called **visual cycle**, the photochemical changes associated with our ability to detect light.

The retina of the eye contains two types of light-sensitive cells, known as **rods** and **cones**. The rods are stimulated by light of lower intensity and do not provide colour vision. The major photoreceptor pigment in rods is a large conjugated protein molecule called **rhodopsin**. This consists of a protein, **opsin**, tightly bound to **11-*cis*-retinal**, which is derived from vitamin A. When rhodopsin is exposed to light, a transformation of 11-*cis*-retinal occurs, changing it to **all-*trans* retinal**.

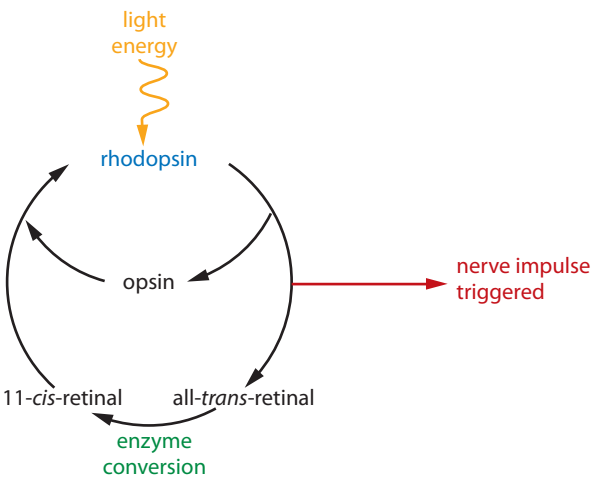
Figure 13.70 The isomerization of the 11-*cis* form of retinal to the all-*trans* form is induced by light energy. Note the change in stereochemistry at C_{11} alters the shape of the molecule, and so its ability to bind to opsin.



The structures of 11-*cis* and all-*trans* retinal are given in section 35 of the IB data booklet.

This causes the all-*trans* isomer to dissociate from the opsin, which triggers a nerve impulse. Rhodopsin is regenerated from opsin and 11-*cis*-retinal after the all-*trans* form isomerizes back to the 11-*cis* form in a series of steps catalysed by enzymes. A summary of the visual cycle is given in Figure 13.71.

Figure 13.71 The visual cycle in rod cells.



Rods are responsible for our vision in low light intensity, and at the sides of the eye. This is why we see outlines of objects but not colours in the dark and at the peripheries of our vision. Most people find this hard to believe as our brain 'fills in' the information from other sources, but if you hold an object of unknown colour at the corner of your eye while looking straight ahead, you will not be able to distinguish its colour.

Exercises

31 Explain why partial hydrogenation of fats is associated with a problem that does not arise from the complete hydrogenation of fats.

32 Cellulose and starch are both polymers of glucose but they have very different properties. Explain this by making reference to their structural units.

33 What is meant by photochemical isomerization in the visual cycle?

B.6 Biochemistry and the environment

Understandings:

- Xenobiotics refer to chemicals that are found in an organism but which are not normally present there.
- Biodegradable/compostable plastics can be consumed or broken down by bacteria or other living organisms.
- Host–guest chemistry involves the creation of synthetic host molecules that mimic some of the actions performed by enzymes in cells, by selectively binding to specific guest species, such as toxic materials in the environment.
- Enzymes have been developed to help in the breakdown of oil spills and other industrial wastes.
- Enzymes in biological detergents can improve energy efficiency by enabling effective cleaning at lower temperatures.
- Biomagnification is the increase in concentration of a substance in a food chain.
- Green Chemistry, also called sustainable chemistry, is an approach to chemical research and engineering that seeks to minimize the production and release to the environment of hazardous substances.

Applications and skills:

- Discussion of the increasing problem of xenobiotics such as antibiotics in sewage treatment plants.
- Description of the role of starch in biodegradable plastics.
- Application of host–guest chemistry to the removal of a specific pollutant in the environment.
- Description of an example of biomagnification, including the chemical source of the substance. Examples could include heavy metals or pesticides.
- Discussion of the challenges and criteria in assessing the 'greenness' of a substance used in biochemical research, including the atom economy.

Guidance

- *Specific names of 'green chemicals' such as solvents are not expected.*
- *The emphasis in explanations of host–guest chemistry should be on non-covalent bonding within the supramolecule.*

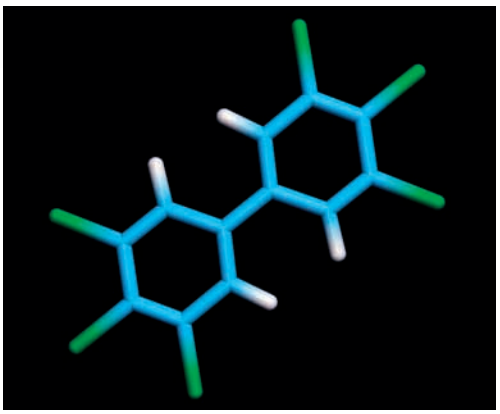
Xenobiotics: strangers to life

Xenobiotics are chemical compounds that are found in a living organism, but which are foreign to that organism. The term is also used to describe chemicals found in higher-than-normal concentrations, or compounds that are not produced naturally but only by synthetic processes – in other words chemicals that are foreign to the biosphere.

Examples of xenobiotics are:

- drugs, including antibiotics such as penicillin
- food additives
- pollutants, such as PCBs and dioxins
- insecticides, such as DDT
- heavy metals, such as mercury and lead ions
- hormones, such as estrogens
- plastics, such as PVC.

Scientist monitoring a water reservoir by sampling fish to analyse the mercury concentrations in their bodies. A nearby gold mine uses mercury to amalgamate the gold and this mercury is a xenobiotic, polluting the surrounding areas. Photographed in French Guiana.



Molecular model of a polychlorinated biphenyl (PCB) molecule. The carbon atoms in the phenyl rings are shown in blue, chlorine is shown in green, and hydrogen in white. PCBs are organic compounds that are highly stable, good insulators, and have low electrical conductivity. They were widely used in capacitors and transformers. The discovery in the 1960s that PCBs are toxic and carcinogenic and can act as **xenoestrogens** led to tight regulation of their use, but due to their stability they persist in the environment and in living tissue.

Within an organism, xenobiotics can have a wide variety of effects. Non-polar molecules pass relatively easily across the hydrophobic cell membranes. They enter cells where they may be modified by enzymes and then detoxified. This is how many drugs are broken down in the body. In agriculture, pesticides may be metabolized by similar processes, sometimes leading to resistance to the effect of the chemical.

However, if the xenobiotic cannot be modified in the organism it may build up in the cells. The increasing concentration of the substance in an organism is known as **bioaccumulation**. For example, mercury compounds in the form of methylmercury, which is non-polar, cross into the brain, where they build up causing mercury poisoning. Similarly, lead, in the form of tetraethyl lead from petroleum, can accumulate in fat tissues.

An increasing area of concern is the existence of **pharmaceutically active compounds** such as antibiotics, painkillers, and chemotherapy drugs in the environment. These may be discharged from industries or hospitals, or passed through the human body and released unmodified or partially metabolized in urine. Sewage treatment plants may break the xenobiotic down through bacterial action, but too often this process is incomplete. As a result, the compounds are released into the effluent, where they can be taken up by fish living downstream.

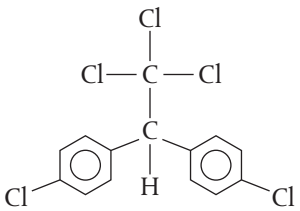
The waste water from sewage treatment may also contain sex hormones such as the female estrogens that have been released in human urine, particularly due to use of the synthetic contraceptive pill. There is some concern that male fish may absorb sufficient quantities of estrogens to be 'feminized' and unable to breed. Related concerns about disruptions to reproductive cycles arise from the widespread presence of so-called xenoestrogens, chemicals that imitate the effects of estrogen. Examples include polychlorinated biphenyls (PCBs) and bisphenol A.

Bisphenol A is an organic compound widely used in making polymers such as the polycarbonate plastics that are used in reusable water bottles, food containers, and water pipes. Bisphenol A has become controversial because it may mimic hormones, especially estrogens, and so give rise to a range of health problems. Studies have shown that risk of the chemical leaching from the plastic is increased when it is heated, and so particular concern has been expressed about its use in babies' bottles that are routinely heated during sterilization. While debate continues about safe levels, many industries have withdrawn these products and several governments have legislation pending to limit their use.

Biomagnification

Although biological processes can produce many harmful substances, such as snake venom and irritants, natural toxins do not build up in the environment as they are broken down by enzymes. By contrast, some of the synthetic chemicals produced by humans are not broken down naturally as there are no enzymes to achieve this. Consequently, they build up in air, water, soil and living cells, and in some cases their concentration can increase in food webs to potentially harmful levels.

Biomagnification refers to the increase in concentration of a xenobiotic substance in a food web. It occurs when a xenobiotic cannot be metabolized, and so is taken up directly when one organism feeds on another, causing the greatest effect for animals that feed at the top of a food chain. A well-studied example of this is the insecticide DDT, dichlorodiphenyltrichloroethane.



DDT is a complex aromatic molecule, which was used to great effect starting during World War II to control the mosquitoes that are responsible for the spread of diseases such as malaria and typhus. The World Health Organization suggests that five million human lives were saved in the early years of its use. DDT is readily soluble in fat, and does not undergo metabolic breakdown. It therefore bioaccumulates in tissues and passes unchanged through food chains.

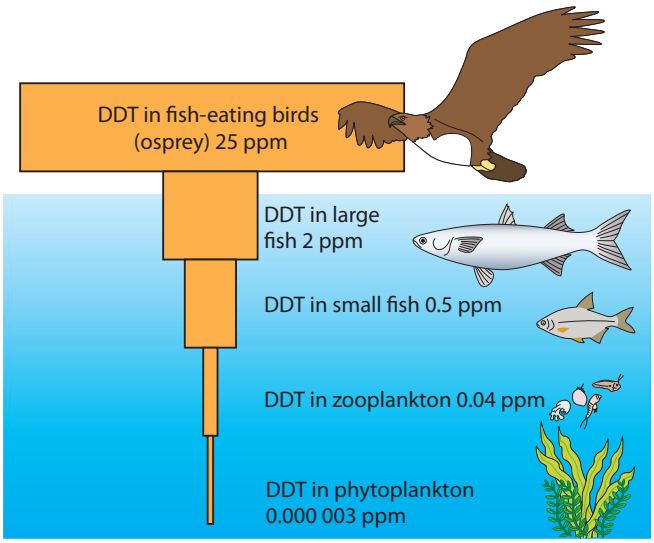


Figure 13.72 Data on DDT concentrations taken from an estuary in Long Island Sound, USA. Ospreys at the top of the chain have 10 million times the concentration of DDT present in the phytoplankton.

Water bottle label showing that it does not contain bisphenol A, BPA. The chemical has been withdrawn from many applications due to concern over its xenoestrogenic properties.

The bald eagle feeds high in the food chain and is therefore susceptible to the effects of biomagnification.

Concerns over the environmental impact of DDT and other pesticides, particularly on bird life, were brought to the public attention through the publication of the book *Silent Spring* by American biologist Rachel Carson in 1962. This book is largely credited as launching the environmental movement that is now global in scope.

Who should determine the balance between the individual's right to freedom from disease and society's right to freedom from environmental degradation?

Bioaccumulation refers to the build up of a substance within an organism; biomagnification refers to the build up of a substance in a food web.

The Stockholm Convention on Persistent Organic Pollutants (POPs) identified twelve chemicals, known as 'the dirty dozen', whose use should be banned. Although legislation to this effect now exists in many countries, many of the chemicals are stable and persist in the environment. There is concern that climate change causing melting of the polar ice is re-mobilizing some of these banned chemicals into the Arctic atmosphere.

Figure 13.72 illustrates the biomagnification of DDT at each feeding level, known as a **trophic level**. As, for example, the small fish feed on a vast number of zooplankton, the concentrations of DDT are multiplied at each level. So the large fish and birds that feed high in the food chain accumulate elevated concentrations of DDT, as much as 10 million times the original concentration.

In the 1960s it was noticed that birds of prey, such as ospreys and peregrine falcons, suffered a serious decline in their numbers. The cause was traced to a thinning of their eggshells that made the eggs break under their parent's weight, and this was found to be due to the toxic effects of the high levels of DDT in their tissues. This and other negative environmental impacts eventually led to a ban on the use of indiscriminate spraying of DDT in many countries by the 1970s. Nonetheless, some continued use of DDT persists in vector control, particularly in countries where malaria is a serious health risk, and this remains controversial.

Other examples of bioaccumulation include organic molecules such as dioxins and PCBs, as well as heavy metals such as mercury and uranium. Data collected on these effects have been used to influence government policies and health advisories.

NATURE OF SCIENCE

It is the responsibility of scientists to consider the ways in which their research and findings impact the environment, and to find ways of amelioration. This involves risk assessment and long-term data collection. The issues presented raise ethical issues which cross national boundaries and demand international collaboration by scientists from different disciplines.

Amelioration: responses to xenobiotics

Clearly, the widespread existence of xenobiotics in organisms and in the environment is a major cause of concern for human health and for biodiversity. *Amelioration* refers to approaches that seek to lessen the problems and improve the outlook. We will consider some examples here.

Host–guest chemistry

We saw in section B.3 that enzymes and substrates form complexes as a result of molecular recognition and specific binding. The complexes are held together by forces other than covalent bonds which depend on the three-dimensional shape of the molecules. This and similar examples of molecular recognition in biological systems have inspired the development of so-called host–guest chemistry. This involves the synthesis of a **host** molecule which is able to bind non-covalently to a guest molecule, and form a **supermolecule**. In essence the host, the larger molecule, is analogous to the enzyme, and the guest to the substrate.

$$\underset{\text{host}}{H} + \underset{\text{guest}}{G} \rightleftharpoons \underset{\text{host–guest complex}}{H\text{–}G}$$

The forces within the supermolecule, like those in enzyme–substrate complexes, include ionic bonds, hydrogen bonds, van der Waals' forces, and hydrophobic interactions.

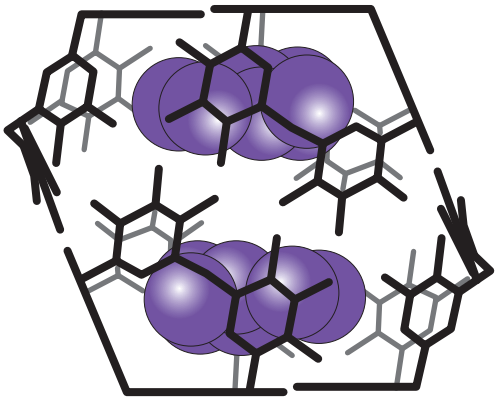


Figure 13.73 Host–guest complex showing two guest molecules (pyrene butyric acid) in blue trapped within the structure of the host molecule which is a six-chain nanocapsule held together by hydrogen bonds.

Host–guest chemistry can be applied to the removal of some xenobiotics in the environment. The xenobiotic is the guest, and its chemical features determine the synthesis of the host, which is designed to bind to it. Many host molecules have a cage-like or tube-like structure which traps the guest molecule. The technique has been used in the removal of radioactive ions such as cesium-137 from nuclear waste.

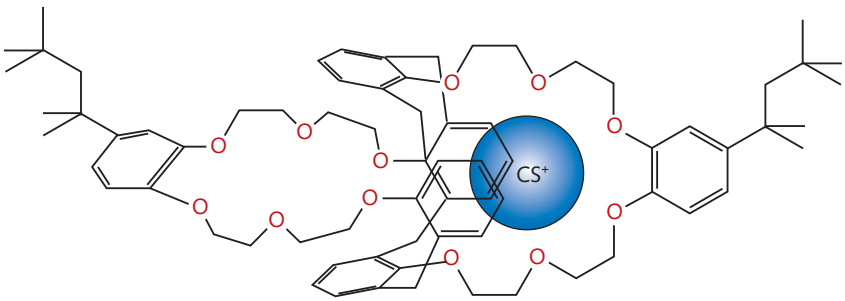


Figure 13.74 Host molecule (known as BOBCalix6) shown with a positively charged cesium ion held inside one of its cavities. This host molecule was specifically developed to bind to the cesium-137 ion, by application of the thermodynamics of ion binding.

Another example is the removal of aromatic amines, which are derived from the cosmetics industry and are known to be carcinogenic. Drugs and drug metabolites are also targets of research for suitable host molecules for their removal.

The associations between the host and guest molecules in a supermolecule are all non-covalent.

Biodegradable substances

Substances that cannot be broken down by natural processes, which mostly involve microbial action, are said to be **non-biodegradable**. Compounds in this category often contain carbon–halogen bonds or stable aromatic structures, which enzymes are not able to break. This is why many plastics, such as PVC and polystyrene, and compounds like DDT persist in the environment indefinitely.

On the other hand, a compound is **biodegradable** if it can undergo bacterial degradation into end products that are found in nature, and therefore are not harmful to the environment. Much research attention has been given to the development of biodegradable plastics, and two main types exist.

1 Plant-based hydro-biodegradable plastic

This plastic has a high starch content and is often obtained from corn. Genetic modification of grasses may help to produce similar plastics. The breakdown is initiated by hydrolysis and produces carbon dioxide and water. Swelling of starch grains can help to break up the plastic. At high temperatures it decomposes relatively quickly, but when buried in a landfill it may take much longer to decompose and may produce methane.

Biodegradable food packaging. Apples sealed in a wrapper made from starch-based plastic. The starch used to make this packaging was sourced from a non-genetically modified crop, as seen in the 'Non GM' label.

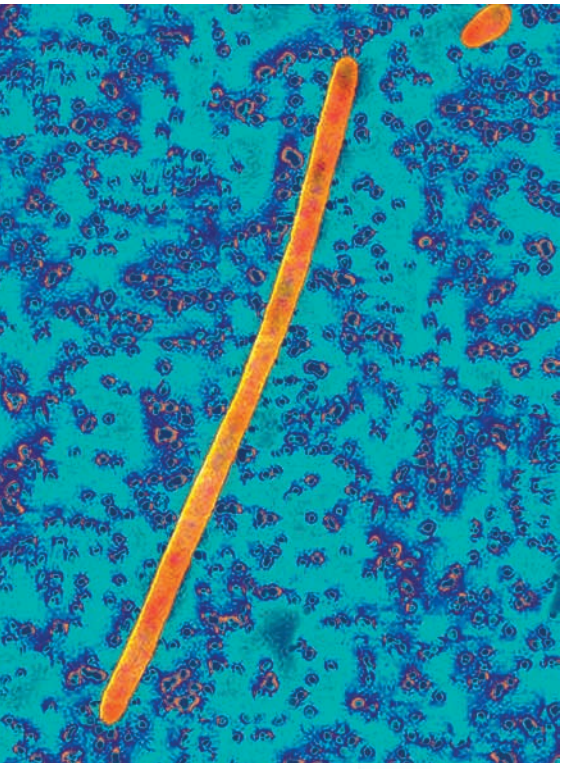
Biodegradable substances are those that can be broken down in natural processes.

2 Petroleum-based oxo-biodegradable plastic

This is derived from a by-product of the oil industry. Additives, often cobalt, are used to act as catalysts for the breakdown process, which can be programmed for different times depending on the use of the plastic. The plastic degrades into microfragments which are dispersed and eventually broken down by bacteria.

Plastics are sometimes described as **compostable**, which means they can be broken down in a compost pile, along with natural food products and garden waste.

Waste plastic is a major international problem, and estimates suggest only 3% of plastic is recycled globally. Use of oxo-biodegradable plastics is becoming mandated in many rapidly-developing countries in the Middle East and Africa to help mitigate the problems of plastic disposal. This puts pressure on other countries to use this plastic in their export goods. Some critics though express concerns over the use of artificial additives in these plastics.



Coloured transmission electron micrograph (TEM) of an oil-degrading bacterium, shown in yellow, collected from a deep-sea oil plume in the Gulf of Mexico after the Deepwater Horizon oil spill in April 2010. This is a previously unknown species of bacteria that helped in the breakdown of oil from the spill around one kilometre below the ocean's surface.

Bioremediation – the use of enzymes

Although as a fossil fuel, crude oil is a natural product, it can be present in sufficient concentrations to be considered xenobiotic. This is the case, for example, when oil spills occur on land or at sea, depositing millions of litres into the adjacent environment. Organisms such as sea birds and marine mammals are then at risk from physical damage and from the toxic effects of the crude oil spill.

Ways to ameliorate the impact of oil spills include the use of microorganisms which are able to break the oil down by using it as a food source and oxidize it in respiration. This is known as **bioremediation**.

As crude oil contains a large number of different compounds, many different chemical reactions are involved in its breakdown. Microorganisms have evolved different enzymes that are specific for the degradation of different hydrocarbons in the oil so breakdown of the oil takes the combined action of a community of bacteria and fungi. Most of these are found naturally in the environment.

Over time, microbes are a dependable means of breaking down oil in the environment, although some of the larger and more highly branched molecules seem resistant to breakdown by enzymes. Environmental conditions such as temperature, supply of nutrients,

and availability of oxygen all influence the efficiency of the process, and in many cases the process may be too slow to prevent ecological damage. Research is ongoing to find ways to enhance the enzymic processes involved.

The use of enzymes in bioremediation also includes processes for the clean-up of industrial waste such as effluents from paper mills, textile industries, and the leather industry.

Green Chemistry

Concerns over health and environmental impact have put the chemical industry under a lot of pressure for greater accountability. New regulations and rising oil prices have contributed to the need for significant changes to be made. This has become formalized to some extent with the establishment of **Green Chemistry**, also known as **sustainable chemistry**, as a field in its own right. The term 'Green Chemistry' was first coined in 1991, and is now a recognizable and rapidly growing area of study.

In essence, Green Chemistry is chemistry for the environment. Twelve principles have been developed which cover such concepts as mimimizing production of waste, the use of safe solvents, the design of energy-efficient processes, and increasing the atom economy of processes. In other words, it seeks to reduce the footprint of chemical manufacturing processes while improving product and environmental safety. In the context of biochemical research, Green Chemistry has contributed to some innovative processes. A few examples are discussed below.

Food and drink

Carbon dioxide under pressure is known as supercritical carbon dioxide, and can penetrate into substances and act as a solvent. It is cheap and non-toxic. It is used in the extraction of caffeine in the preparation of decaffeinated coffee, replacing previously used toxic solvents. It is also used to remove fungicide contaminants from wine-bottle corks, and to pull pungent oil out of sesame seeds.

Bioplastics

Plastics derived from corn starch which has been converted into a resin by bacteria can replace traditional oil-based plastics. Genetically engineered plants such as tobacco may be able to harvest useable plastics.

Cosmetics

Production of esters for face creams has traditionally used sulfuric acid at high temperatures, but can be done using enzymes at room temperature.

Clothing industry

Enzymes can replace polluting detergents and improve energy efficiency by enabling effective cleaning at lower temperatures.

Wool dyed with natural products. Green Chemistry seeks to limit the use of toxic dyes and solvents which are widely used in the textile industry.

Renewably sourced textile fibres such as bamboo and eucalyptus may replace synthetic materials from the petrochemical industry or fabrics such as cotton which rely on heavy use of fertilizers.

Many of the applications of Green Chemistry are still developing, and are the focus of active biochemical research. With better education, it is hoped that individuals will exert consumer choice in helping to promote some of these innovative practices. Green Chemistry provides an opportunity for the chemical industry to be both forward-looking and environmentally responsible.

Exercises

34 State three types of association found in host–guest supermolecules.

35 Explain what is meant by biomagnification and why it is such a problem.

36 State three examples of the use of enzymes in helping to ameliorate environmental problems.

Practice questions

1 The structures of the amino acids cysteine and serine are shown in section 33 of the IB data booklet. They can react with each other to form a dipeptide.

(a) State the type of reaction occurring when amino acids react together and identify the other product of the reaction. (2)

(b) Draw the structures of the two possible dipeptides formed in the reaction between one molecule each of cysteine and serine. (2)

(c) Six tripeptides can be formed by reacting together one molecule of each of the amino acids arginine, histidine, and leucine. Predict the primary structures of these six tripeptides using the symbols shown in Section 33 of the IB data booklet to represent the amino acids. (3)

(d) When many amino acid molecules react together a protein is formed. These proteins have primary, secondary, and tertiary structures.

(i) State the type of intermolecular force responsible for maintaining the secondary structure. (1)

(ii) State **two** other ways in which the tertiary structure of the protein is maintained. (2)

(*Total 10 marks*)

2 **(a)** Maltose is a disaccharide of α-glucose.

(i) Use information from section 34 of the IB data booklet to draw the structure of maltose. (3)

(ii) Use molecular formulas to show the reaction for the breakdown of maltose into its monomers. (2)

(iii) State the name of this type of metabolic process. (1)

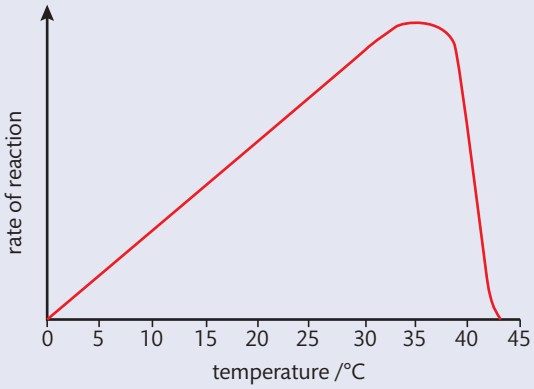
(b) Cellulose is a polysaccharide of β-glucose.

(i) Draw a monomer of cellulose. (1)

(ii) Describe and explain the differences in the properties of cellulose and starch. (3)

(*Total 10 marks*)

3 The graph below shows the effect of temperature on the rate of a reaction catalysed by an enzyme in the human body.



(a) With reference to the structure and function of enzymes, explain the shape of the graph. (5)

(b) Describe how the graph might be different if it showed the effect of temperature on the rate of a reaction catalysed by an enzyme in a plant adapted to a very cold climate. (2)

(*Total 7 marks*)

4 **(a)** The structures of some fatty acids are given in section 34 of the IB data booklet.

(i) Determine the iodine number of linoleic acid. (3)

(ii) A different fatty acid known as X is isolated, and it is found to have a similar molecular mass to that of linoleic acid, but a lower iodine number. What can you conclude about its structure? (1)

(iii) State and explain which fatty acid, linoleic acid or X, would have the higher melting point. (2)

(b) Rancidity in fats can occur as a result of two separate processes. Name these processes, and describe how they differ in the site of reactivity and the conditions that favour the reaction. (6)

(c) Compare and contrast the use of lipids and carbohydrates as a source of energy. (3)

(*Total 15 marks*)

5 Hemoglobin is a protein with a quaternary structure.

(a) Justify the statement that hemoglobin is a globular protein. (3)

(b) Sketch a graph that shows the cooperative nature of the binding of hemoglobin to oxygen. (4)

(c) Explain how the cooperative nature of the hemoglobin–oxygen binding is a consequence of the quaternary structure. (2)

(d) Annotate your graph in part (b) to show how hemoglobin's oxygen dissociation curve is changed by the presence of a higher concentration of carbon dioxide. Explain how this change affects the oxygen saturation of the blood when it is close to cells that are actively respiring. (3)

(e) Uncooked meat on display for sale is often wrapped in a plastic film which is permeable to air. Deduce why air is allowed to reach the meat. (2)

(f) State and explain two ways in which the oxygen dissociation curve for myoglobin differs from that of hemoglobin. (4)

(*Total 18 marks*)

6 Foods derived from genetically modified organisms were introduced in the early 1990s. State **one** benefit and **one** concern of consuming genetically modified foods. (2)

7 Deoxyribonucleic acid (DNA) is a double helical molecule made of two polynucleotide chains.

(a) State the components of a nucleotide, and the type of reaction by which they join together. (2)

(b) Explain in terms of the structure of the nucleotide, why DNA fragments move towards the positive electrode during electrophoresis. (2)

(c) A sequence of bases in one strand of DNA is given below:

GCCTACTTAGCTA

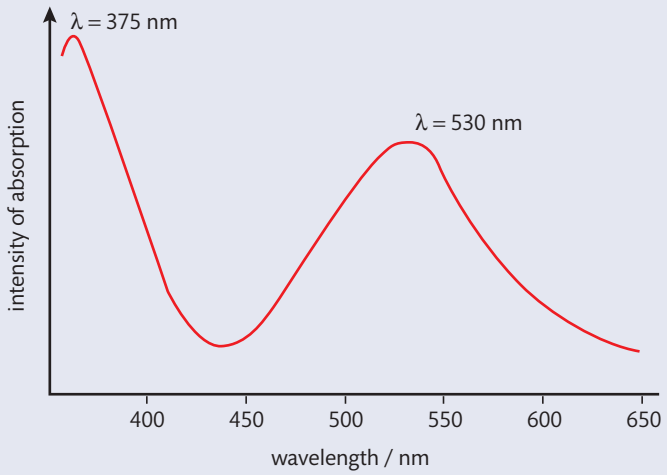
State the corresponding base sequence in:

(i) the complementary strand of DNA;

(ii) the RNA derived from this complementary strand of DNA. (2)

8 **(a)** Explain why pigments such as anthocyanins are coloured. (2)

(b) The wavelength of visible light lies between 400 and 750 nm. The absorption spectrum of a particular anthocyanin is shown below.



(i) Explain what effect, if any, the absorption at 375 nm will have on the colour of the anthocyanin. (1)

(ii) Explain what effect, if any, the absorption at 530 nm will have on the colour of the anthocyanin. (1)

(c) List **two** factors which could alter the precise colour of a particular anthocyanin. (2)

(*Total 6 marks*)

9 The structures of some anthocyanins and carotenoids are given in section 35 of the IB data booklet. Deduce and explain whether anthocyanins and carotenoids are water soluble or fat soluble. (4)

10 (a) Predict and explain the solubilities in water of vitamins A and C from their structures given in section 35 of the IB data booklet. (2)

(b) If excess amounts of vitamins A and C are consumed, predict with a reason which one is more likely to be stored in the body. (1)

(c) Suggest three reasons why vitamin deficiency diseases are prevalent in many parts of the world. (3)

(*Total 6 marks*)

11 (a) Outline the meaning of the term xenobiotic. (2)

(b) Describe the role of starch in biodegradable plastics. (2)

(c) Explain why a plastic such as PVC, poly(chloroethene), is non-biodegradable. (2)

(*Total 6 marks*)

To access weblinks on the topics covered in this chapter, please go to www.pearsonhotlinks.com and enter the ISBN or title of this book.