A Level Science Applications Support Booklet: Chemistry

Contents

Acknowledgements

Thanks are due to Richard Harwood, Peter Cann and Brian Ratcliff for writing this booklet, to D[illegible] Bevan for editing it and to the Royal Society of Chemistry for generous provision of source materi[illegible]

Figure 1.1 NASA Images http://www.nasa.gov/multimedia/imagegallery/reproduction_guidelines.html

Figures 1.2, 1.3, 1.7, 1.10, 1.11, 1.12, 1.13, 1.16, 1.18, 1.21, 1.23, 1.26, , 1.32, 1.33, 1.34, 1.35, 1.36, 1.38, 1.39, 1.40, 1.41, 1.42, 1.43 and 1.44 © Richard Harwood (2002) *Biochemistry* in the *Cambridge Advanced Sciences* series, Cambridge University Press, ISBN: 0521797519

Figure 1.30 reproduced with permission from Joseph W. Lauher, SUNY Stony Brook

Figures 1.31 and 2.20 © Science Photo Library, image G110/636 www.sciencephoto.com

Figures 2.1, 2.9, 2.31, 2.33, 2.34 and 2.36 © Ben Faust (1997) *Modern Chemical Techniques*, Royal Society of Chemistry, ISBN 1870343190

Figures 2.2 and 2.3 National Institute of Advanced Industrial Science and Technology *Spectral Database for Organic Compounds*, *SDBS*. A free website organized by the National Institute of Advanced Industrial Science and Technology (AIST), Japan, www.aist.go.jp/RIODB/SDBS/cgi-bin/cre_index.cgi

Figure 2.4 *Spectroscopy for Schools and Colleges CD-ROM*, Royal Society of Chemistry

Figures 2.19, 2.40 and 2.41 © Ralph Levinson (2001) *More Modern Chemical Techniques*, Royal Society of Chemistry, ISBN 0854049290

Figure 2.24 Myoglobin, http://en.wikipedia.org/wiki/Image:Myoglobindiffraction.png

Figure 2.27 © Philip Matthews (2002) *Gases, Liquids and Solids* in the *Cambridge Advanced Sciences* series, Cambridge University Press, ISBN: 0521797500

Figures 2.29, 2.30, 2.35 and 2.44 © Anne McCarthy (2001) *Methods of Analysis and Detection* in the *Cambridge Advanced Sciences* series, Cambridge University Press ISBN 0521787246

Figure 2.43 © Thomas D Gelehrter, Francis Collins, David Ginsburg (Editor) (1998) *Principles of Medical Genetics*, Lippincott Williams and Wilkins, ISBN: 0683034456

Figure 2.45 © Stephen M Carr, http://www.mun.ca/biology/scarr/DNA_fingerprinting.html

Figure 3.2 reproduced with permission from Dr Corinne Baumgartner, ETH Zürich

Figure 3.5 reproduced with permission from the University of Bristol, http://www.chm.bris.ac.uk/motm/spider/page3h.htm

Figure 3.6 reproduced with permission from Brian Ratcliff

Figure 3.7 reproduced with permission from Jessops

Figure 3.8 Wikipedia, Fullerene, http://en.wikipedia.org/wiki/Image:C60a.png

Figure 3.9 Carbon Nanotechnologies Incorporated, Image Gallery, Bucky vials http://www.cnanotech.com/pages/resources_and_news/gallery/3-2_buckytube_gallery.html

Figure 3.11 Carbon Nanotechnologies Incorporated, Image Gallery, adapted from Buckytube with open end http://www.cnanotech.com/pages/resources_and_news/gallery/3-2_buckytube_gallery.html

Figure 3.12 Carbon Nanotechnologies Incorporated, Image Gallery, adapted from Buckytube with 1 open end http://www.cnanotech.com/pages/resources_and_news/gallery/3-2_buckytube_gallery.html

Figure 3.13 Carbon Nanomaterials Group, Chemical Reactions in Carbon Nanotubes http://www.nottingham.ac.uk/nanocarbon/index2.html

Figure 3.14 reproduced with permission, Dr Simon Benjamin, University of Oxford

Figure 3.16 reproduced with permission, Professor Ben Feringa, University of Gröningen, Light-driven unidirectional molecular motor, http://feringa.fmns.rug.nl/

Figure 3.17 Toshiba Research Europe Limited

Figure 3.19 mining-technology.com, Bingham Canyon, http://www.mining-technology.com/projects/bingham/bingham1.html

Table 3.1 and Nanotechnology worksheet © Vanessa Kind (2004), *Contemporary Chemistry for Schools and Colleges*, Royal Society of Chemistry, ISBN: 0854043829

Nanotechnology worksheet © Mike Ross, IBM Almaden Research Center, California USA

APPLICATIONS OF CHEMISTRY

INTRODUCTION

Too often the study of chemistry at A Level can seem like a disorganised collection of facts. The problem is that to understand many of the ways in which chemistry is used to enhance our lives, we need to understand the basic principles of chemistry, and this can take time. As students approach the end of their A Level studies they have a better understanding of the way chemistry works.

This section of the syllabus examines some of the important areas in which chemistry is being used in biological systems and medicine, in analysing and designing new materials, and in monitoring and helping protect the environment.

The first topic deals with the chemistry of proteins and the mechanism of enzyme catalysis. A knowledge of these is critical to understanding the way in which living organisms function. Next the chemistry of DNA is examined, both in terms of its structure, and the way in which genetic information is encoded and passed on. The provision of energy within cells using ATP is explained, and finally the importance of metals both in trace amounts, and in terms of their toxicity is examined.

The second topic looks at the various analytical techniques that chemists use to help them determine the structure of compounds and to follow the course of chemical reactions. These include electrophoresis and DNA 'fingerprinting', NMR and mass spectroscopy, X-ray crystallography and chromatographic techniques. The emphasis in this topic is on understanding the techniques and interpreting data rather than the recall of particular examples or the detailed theory behind the different techniques.

The final topic looks at ways in which chemical techniques are used to design new materials for specific purposes. These include the targeting and delivery of modern drugs to fight disease, the development of new polymers with properties similar to traditional structural materials, the use of nanotechnology to assemble chemical structures, and finally using chemistry to extend the life of known resources and to protect the environment. Once again the emphasis is on understanding the principles involved rather than acquiring a detailed knowledge of particular examples.

This book is designed so that it can be used by teachers, alongside the Applications of Chemistry syllabus and the Scheme of Work that can be found on the CIE Teacher Support Website (contact international@cie.org.uk to find out how to gain access to this learning resource). This book should help teachers to design effective learning programmes to teach this material, which makes up 16% of the total assessment at A Level and should thus make up just over 30% of the total teaching time available during the A2 part of the course.

This booklet is also designed so that it can be used by students, to promote their own learning, and for this purpose contains self-assessment questions for students to use in helping to determine how effective their learning has been. At the end of each section there is also a glossary to help students remember any new terms introduced. Finally, at the end of the booklet there are some specimen questions to give an idea of what they might expect to see in the examination.

Website addresses are given for many topics in this book, particularly those in the final chapter. All of these have been tested and are working at the time of writing, but such sites are notoriously quick to change their url addresses, so by the time that you get to try and use them, they may well be found not to work. The author strongly recommends that a good search engine be used to seek up-to-date and reliable websites for information on each of the topics on the syllabus.

Finally, this section of the syllabus aims to help students develop the skill of applying their chemical knowledge to novel situations. For this reason, the contexts and case studies given in this booklet are not for rote learning, unless specified in the syllabus. Rather, the contexts given are to illustrate the concepts in the syllabus and it is hoped that this will stimulate students to investigate further examples that are of interest to them. Students may be asked in the examination to give examples of contexts that they have studied, but examination questions on this part of the syllabus will primarily be testing core knowledge applied to novel situations, rather than rote learning of examples.

INTRODUCTION TO THE 2010 REVISION

Because of imminent debate on the future pattern of this part of the syllabus, this revision [illegible] exhaustive as might otherwise have been necessary. However, it has managed to address m[illegible] the concerns expressed by users, and has, we hope, corrected all the errors and omissions brou[illegible] to the attention of CIE (and also some errors that were not picked up by users!).

With nine years of past papers now available on the teachers' support website (http://teachers.cie.org.uk), it should now be increasingly clear to teachers and students as to how this part of the syllabus is interpreted by Examiners, and hence what aspects of this support booklet need to be learned and understood thoroughly.

This booklet is intended mainly as a support booklet for teachers, and as such contains a significant amount of relevant, but 'background' information. However, it is recognised that it is also used by students, especially perhaps during their periods of revision.

With this in mind, each section now concludes not only with the summary, as contained in the original booklet, but also an interpretation of ***what the student needs to know*** with regard to each of the learning objectives of the syllabus. These summaries are in boxes with the same background colour as this one. We hope students will find these useful.

1 – THE CHEMISTRY OF LIFE

1.1 – Introduction

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

Figure 1.1 – the Earth as seen from space

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

Water – Life's matrix

What makes water so vital for life? Animals and plants living in 'normal' environments cannot survive without a regular supply of water. Without water, life could never have evolved. Some of the most significant biological roles of water are summarised in Figure 1.2.

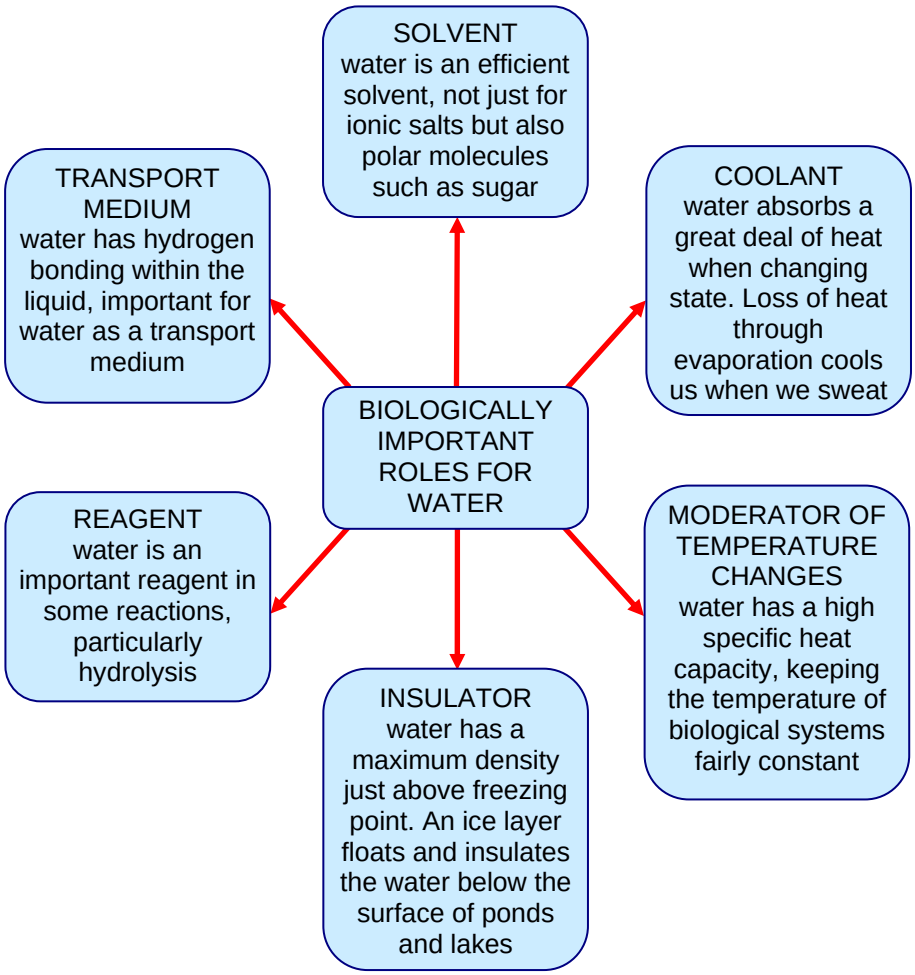


Figure 1.2 – biologically important roles of water

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

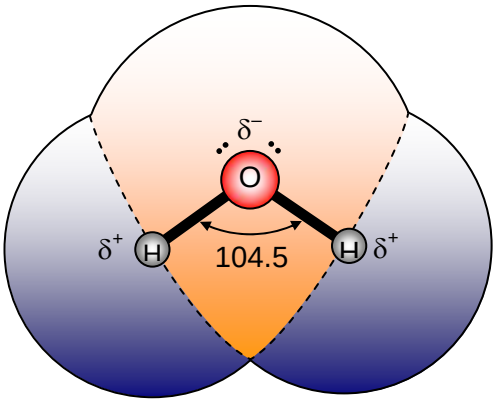


Figure 1.3 – the polarised nature of the water molecule

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

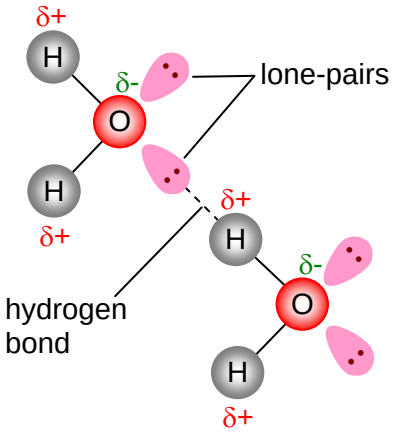


Figure 1.4 – hydrogen bonding between water molecules

The importance of hydrogen bonding to life

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

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

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

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

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

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

Condensation polymerisation

Given the importance of a water environment to the emergence of life on Earth it is not surprising that condensation polymerisation (see section 10.8 of the core syllabus) is 'the method of choice' in the natural world when it comes to making the macromolecules important to life. It is within an aqueous environment that small biological molecules can associate and polymerise through condensation

reactions (Figure 1.5). The reverse process of hydrolysis can also break down macrom[illegible]t are no longer required by the organism.

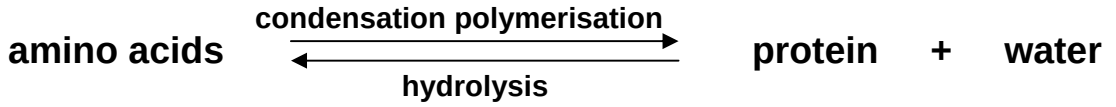


Figure 1.5 – the processes of condensation and hydrolysis

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

- nucleic acids,
- proteins,
- carbohydrates.

Another important class of compounds involved in life's processes are the lipids. These are esters, so are also formed by condensation reactions and are split by hydrolysis reactions.

Shaping up

Hydrogen bonding mentioned earlier is one of the forces that play a significant role in the three-dimensional folding of proteins. Whether it is an enzyme, an antibody or a structural protein, the correct functioning of a protein depends upon its shape. Life depends on molecules recognising and interacting with each other in specific ways. Such interaction is most efficient if the molecules involved have complementary structures – they fit together as would a lock and key. Thus an enzyme recognises its substrate, a DNA strand associates with its complementary strand, and a hormone recognises its receptor. The interaction between biological molecules begins with recognition of complementary molecules, followed by short-lived binding through forces such as hydrogen bonding, and ends in those activities that support and propagate life.

1.2 – Protein chemistry

By the end of this section, students should be able to:

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

1.2a Proteins – the workhorses of life

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

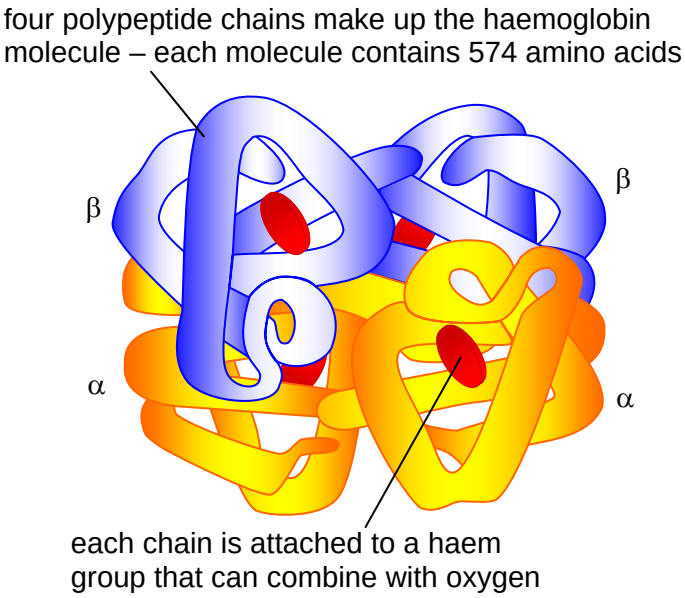


Figure 1.6 – diagram of haemoglobin

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

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

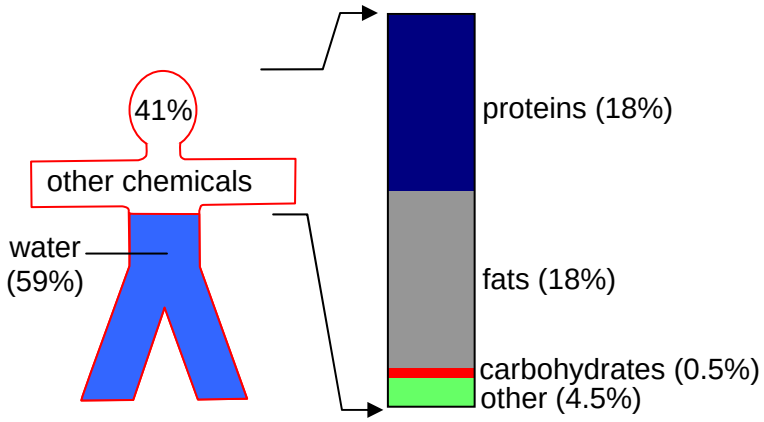


Figure 1.7 – a summary of the chemical components of the average person

Table 1.1 – some proteins and their functions

protein(s)	function	location
myosin actin	muscle contraction	muscle tissue
chymotrypsin	digestive enzymes	small intestine
pepsin		stomach

insulin	hormone	blood
immunoglobulins	antibodies	blood
collagen	structural proteins	skin, tendon
keratin		hair
haemoglobin	transport	blood
ferritin	storage	bone marrow, liver, spleen

Proteins are unbranched polymer chains made by linking together large numbers (from hundreds to several thousands) of amino acid monomer units by amide bonds. Such chains are often referred to as polypeptide chains. An amide bond formed between two amino acids is called a *peptide bond*.

$NH_2–CHR–CO_2H$

+

$NH_2–CHR'–CO_2H$ → $–NH–CHR–CONH–CHR'–CONH–CHR''–CO–$

+

$NH_2–CHR''–CO_2H$

three separate amino acids $\xrightarrow{-3H_2O}$ part of a polypeptide

(The peptide bond is described in detail in section 1.2c.)

1.2b Amino acids – the building blocks of proteins

Common features of amino acids

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

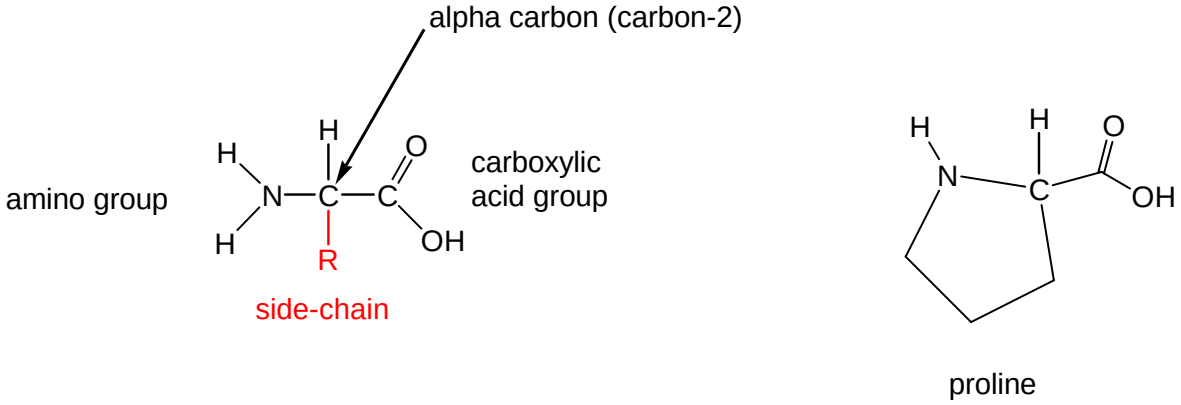
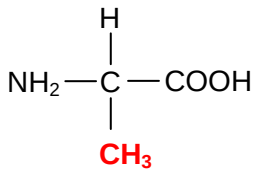
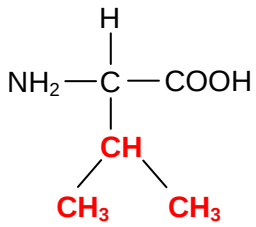


Figure 1.8 – general structure of an amino acid, highlighting the key features, and the structure of proline

When naming organic compounds systematically, the carbon atom of an acid group is alw... d as the first in the structure. This means that in these molecules the amino group is always a... the second carbon atom (C-2), the one immediately adjacent to the carboxyl group. This carbo... is also sometimes known as the α-carbon atom (spoken as “alpha carbon atom”). These impo... molecules are therefore all 2-amino acids (or α-amino acids).

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

The 20 different amino acids can usefully be categorized into separate sub-groups according to the nature of the R-group. There are three broad categories depending on whether the side-chain group is non-polar, polar, or can be ionised (charged) under appropriate conditions (see Figure 1.9). This figure also includes the three-letter abbreviations that are often used in place of the full names.

sub-group (based on type of R-group)	example	structure
non-polar	alanine (ala)	
	valine (val)	

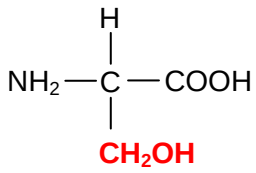
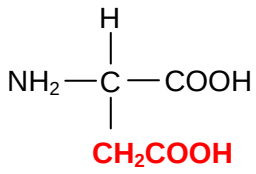
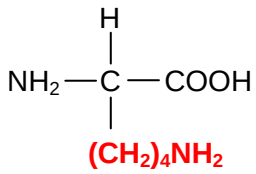
sub-group (based on type of R-group)	example	structure
polar	serine (ser)	
electrically-charged (acidic or basic side-chains)	aspartic acid (asp)	
	lysine (lys)	

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

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

The ionisation of amino acids

As we have seen, each amino acid molecule contains an acidic group (–COOH) and a basic group, ($–NH_2$). As a consequence, these molecules show both the properties of an acid and those of a base. Such compounds are said to be amphoteric. The acid group has about the same acidity as ethanoic acid, and so in water it can lose a proton (H^+ ion) to form the carboxylate ion.

$$NH_2CH(R)COOH\ (aq) \longrightarrow NH_2CH(R)COO^-\ (aq) + H^+\ (aq)$$

Likewise, the amino group has about the same basicity as ammonia, and so, like ami[illegible] gain a proton from water.

$$NH_2CH(R)COOH\ (aq) + H_2O(l) \longrightarrow {}^{+}NH_3CH(R)COOH\ (aq) + OH^{-}(aq)$$

In this latter case the proton attaches to the lone pair of electrons on the N atom.

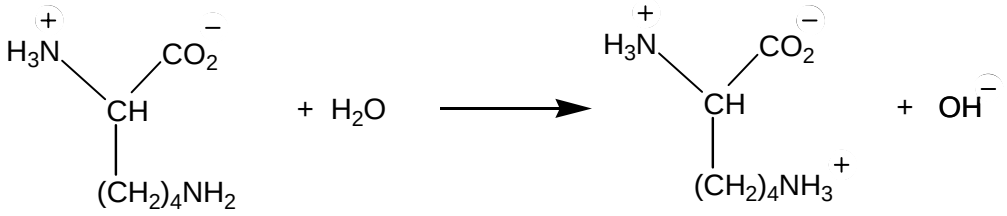
At physiological pH, which is just above neutrality, the acid and amino groups react with each other, as shown below.

$$NH_2CH_2CO_2H \xrightarrow{-H^+} NH_2CH_2CO_2^- \xrightarrow{+H^+} {}^{+}\mathbf{NH_3CH_2COO^-}$$

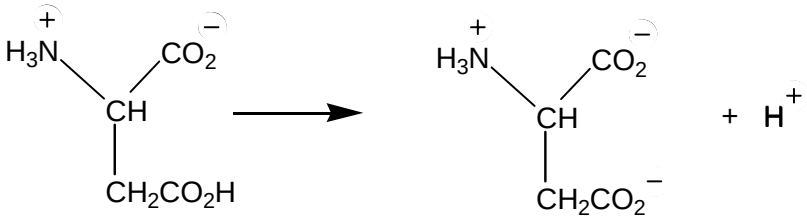
Therefore molecules of amino often exist as dipolar ions, or *zwitterions*. X-ray crystallography of crystalline amino acids has shown that amino acids also exist in the zwitterionic form in the solid. Because of this they show the physical properties characteristic of ionic compounds: they are white solids with fairly high melting points that are soluble in water.

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

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



lysine lysine cation (overall charge = +1)



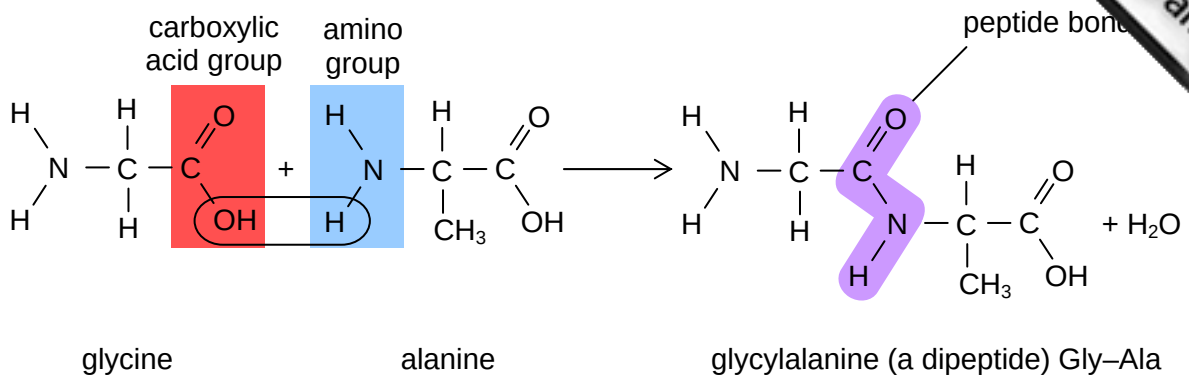
aspartic acid aspartate anion (overall charge = –1)

1.2c Structure – the key to protein function

Condensation polymerisation of amino acids

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

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



glycine alanine glycylalanine (a dipeptide) Gly–Ala

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

Additional amino acids can react with the dipeptide to form first a tripeptide and then eventually a polypeptide. In this way a protein can be put together.

The peptide bond consists of the group –CONH– in which the four atoms lie in one plane, with all bond angles being about 120°.

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

Levels of protein structure

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

- Primary structure. This is the sequence of amino acids in a polypeptide chain, and is the direct result of the coding sequence in the gene.
- Secondary structure. This is the regular structural arrangements of the polypeptide chain that result from hydrogen bonding between peptide bond regions of the chain.
- Tertiary structure. This is the overall folding of a polypeptide chain that is the result of interactions between the amino acid side-chains.

The primary structure of proteins

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

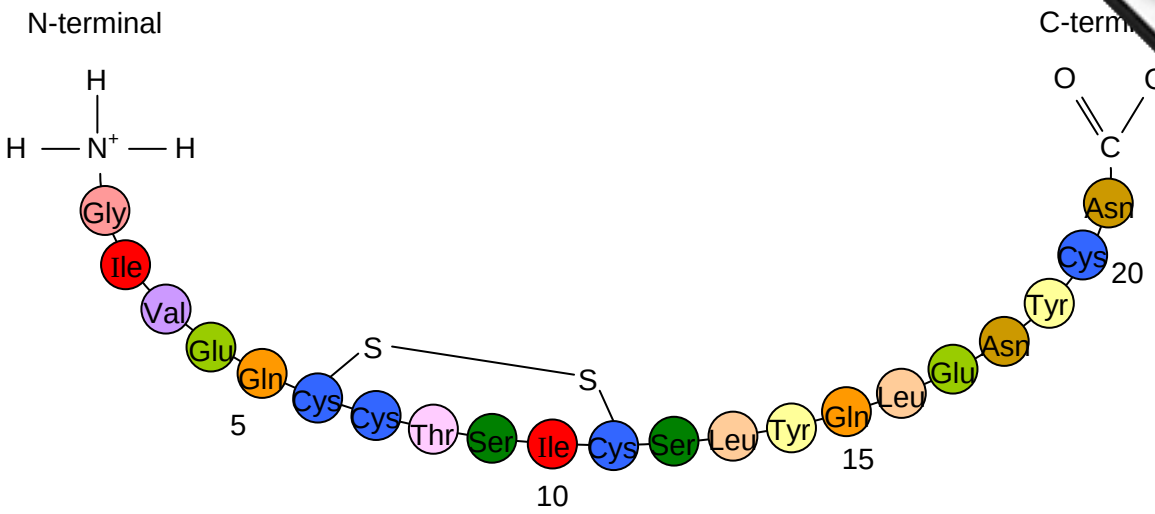


Figure 1.11 – the primary sequence of the insulin A chain, a short polypeptide of 21 amino acids

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

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

The secondary structure of proteins

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

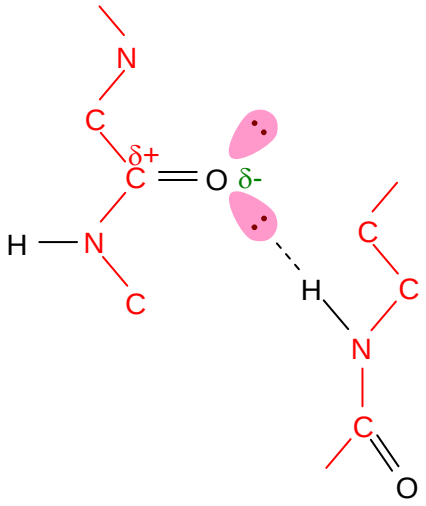
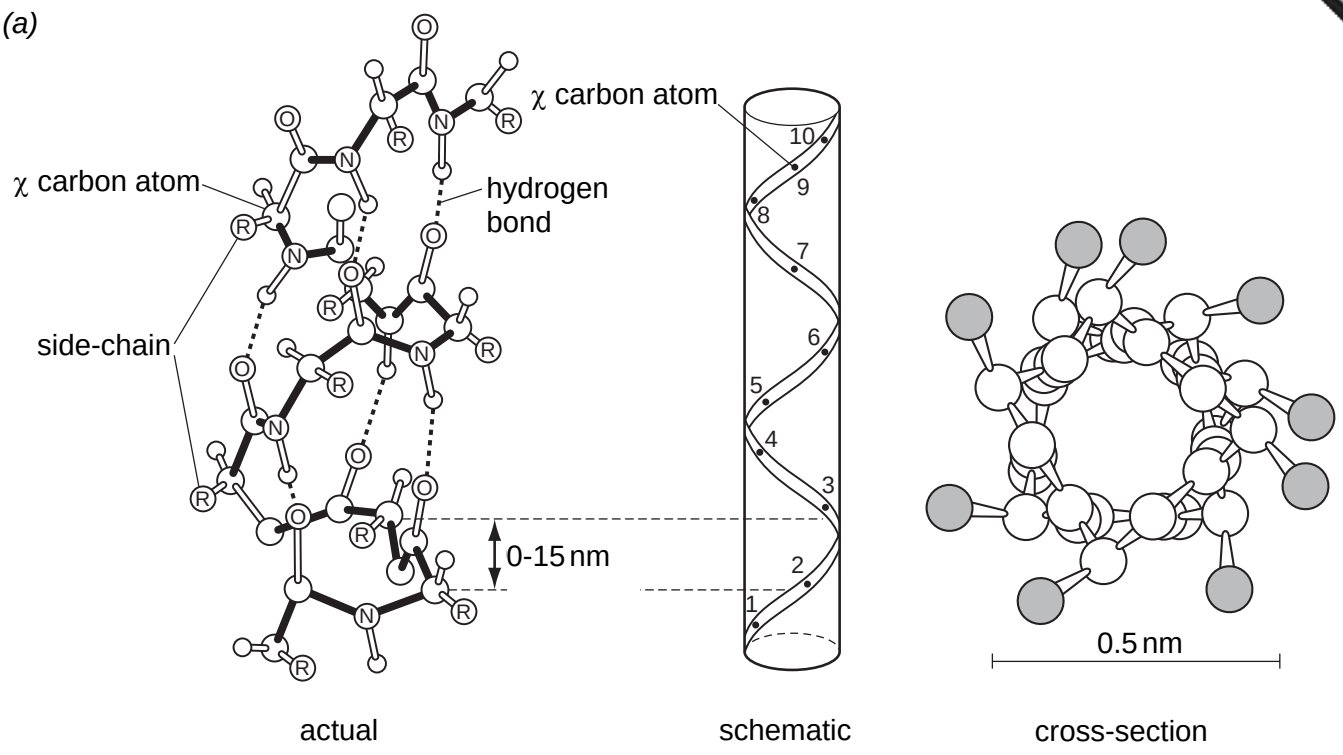


Figure 1.12 – illustration of hydrogen bond between two polypeptide link regions

This type of folding, stabilised by the intramolecular hydrogen bonding, gives rise to certain structural features which are found in many different types of protein. Collectively this level of structural

organisation is known as secondary structure. Two of the most stable arrangements at [illegible]f protein folding are the α-helix ("alpha helix") and the β-pleated ("beta pleated") sheet (Figure [illegible]). In both these structures the chain folds on itself in a very stable arrangement because of the hydrogen bonds formed between adjacent peptide bond regions.



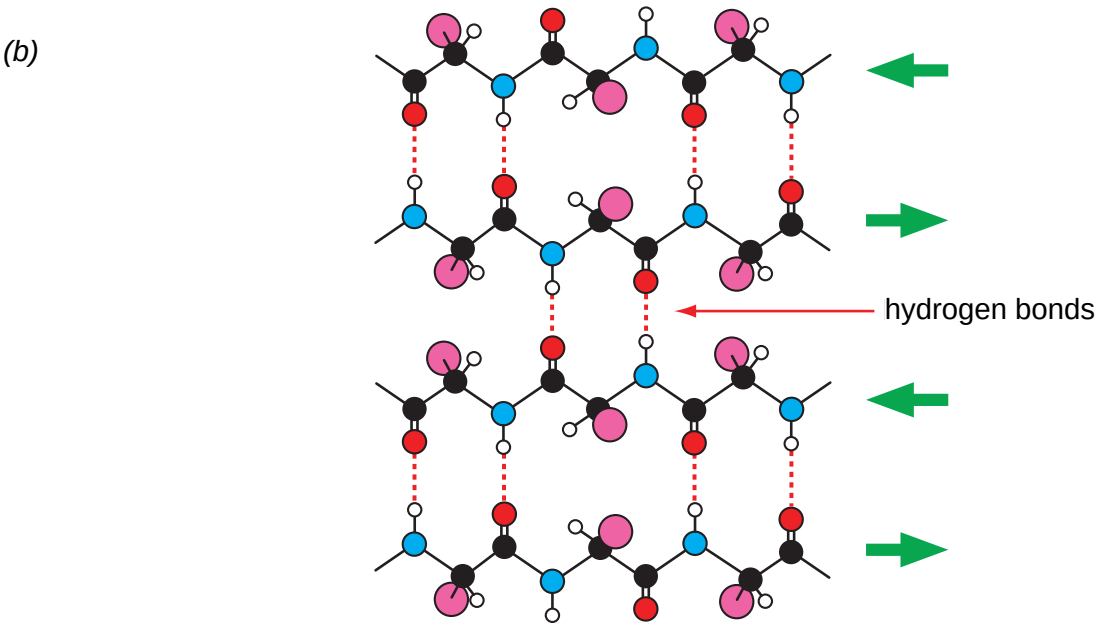


Figure 1.13 – the structure of (a) the α-helix [as found in keratin]; (b) the β-pleated sheet [as found in silk fibroin]

Regions of regular secondary structure occur in many proteins. Figure 1.14 shows a computer-generated graphic of the structure of pepsin – a protease present in our stomach which helps digest

our food. The structure has distinct α-helical regions, represented by the 'cylindrical' [illegible] β-pleated sheet regions, represented by the 'arrows'.

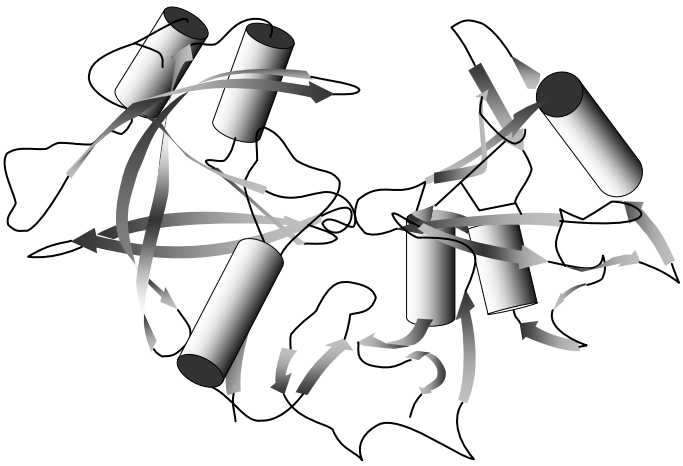


Figure 1.14 – a computer graphic model of the structure of pepsin

Protein tertiary structure

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

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

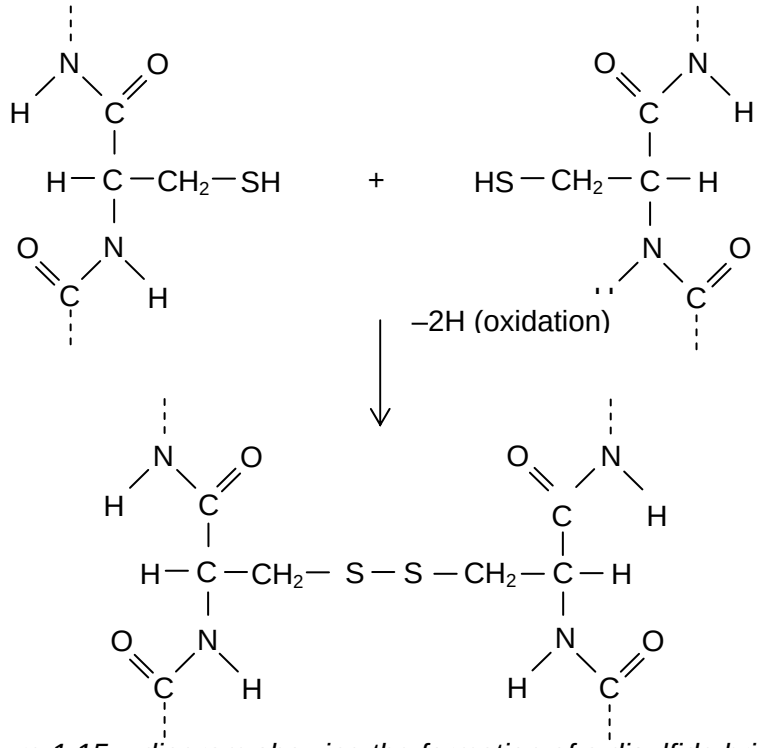


Figure 1.15 – diagram showing the formation of a disulfide bridge

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

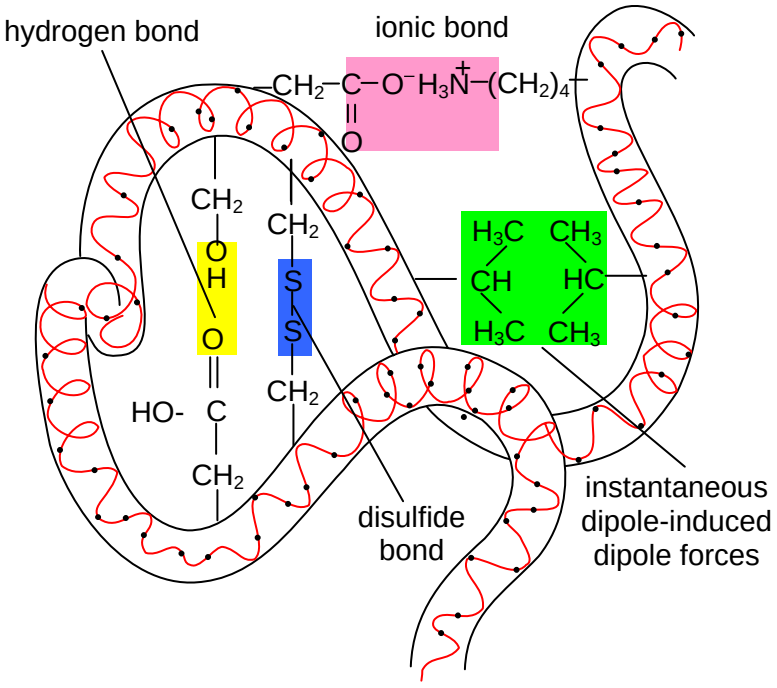
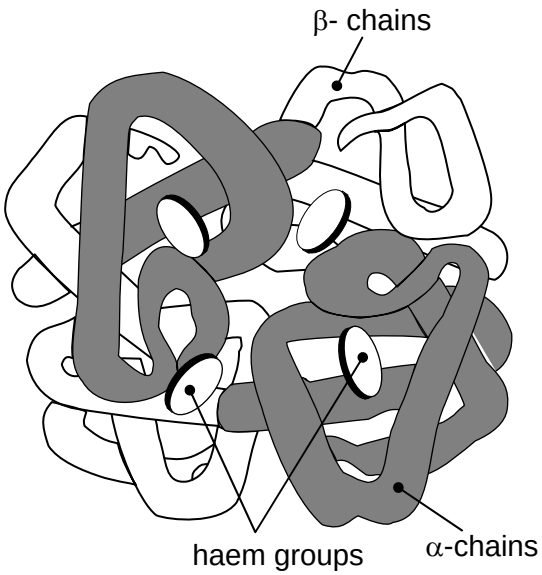


Figure 1.16 – diagram illustrating the nature of the interactions responsible for protein tertiary structure

Haemoglobin – the molecular 'breathing machine'

Human haemoglobin consists of two pairs of identical protein chains (sub-units) known as α-chains and β-chains. [The use of the terms α- and β- as applied by biochemists in this context might be confusing to the reader! The terms α- and β- are used here just to differentiate between the two different chains and this use is not at all related to how we used them before, to describe the α-helix and the β-pleated sheet secondary structures of proteins on pages 13 and 14.] These chains assemble together to form the functional, oxygen-carrying protein in our red blood cells (Figure 1.17). Although similar in size to each other (141 and 146 amino acid residues respectively), the human globin α- and β-chains are different polypeptide chains coded for by different genes. When it is folded, each of the globin chains has a high degree (about 70%) of α-helical secondary structure. On the other hand there are very few, if any, β-pleated sheet regions to the structure. Each of the protein sub-units in one molecule is also bound to a non-protein haem group containing an iron(II) ion (Fe^{2+}).

(a)



(b)

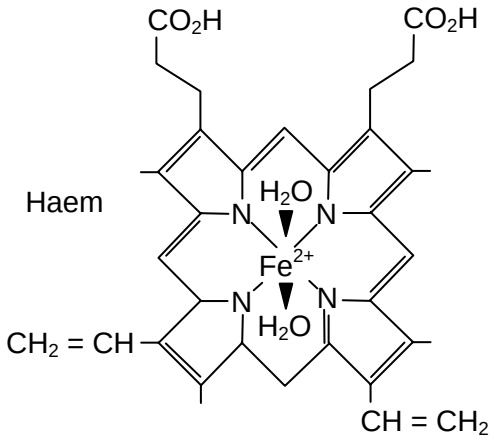


Figure 1.17 – (a) the structure of haemoglobin showing the α and β chains; (b) the structure of the haem group.

In haemoglobin, the iron(II) ions in the haem groups bind oxygen molecules and enable the haemoglobin molecules to perform their function of transporting oxygen around the body. The haem groups are bound to the globin chains by a dative covalent bond to one side of the Fe(II) ion, and also by non-covalent interactions with the $-CO_2H$ and $CH=CH_2$ groups around the haem ring. These non-covalent interactions (ionic and van der Waals') are similar to those involved in the tertiary structures

of proteins. The interconnected nature of the whole structure is emphasised by the [illegible] overall structure of the whole haemoglobin molecule changes as oxygen binds to the haem [illegible]. Remarkably the protein structure actually co-operates in the attachment of O_2, subtly chang[illegible] shape as the oxygen molecules attach themselves. The molecule undergoes a 'breathing mo[illegible]'. The positive co-operative effect achieved by these changes in structure means that there is a favoured sequence for the attachment and detachment of oxygen from the molecule. Either all four sub-units have oxygen attached to their haem groups, or none of them do. This leads to the most efficient transfer of oxygen from the lungs to the other tissues of the body. The haemoglobin molecule is not simply an inert carrier, but a functioning piece of machinery of molecular dimensions. Students will come across the idea of molecular machinery again in the section on nanotechnology.

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

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

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

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

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

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

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

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

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

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

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

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

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

Summary

- Proteins are condensation polymers. The amino acid monomers are linked by peptide bonds to form the polypeptide chain.
- α-amino acids are the biological monomers from which protein molecules are built.
- α-amino acids contain both a carboxylic acid group and amino group. These groups are both attached to the same carbon atom. They have the general formula $NH_2CH(R)COOH$.
- α-amino acids exist in a zwitterionic form in the solid state and when in aqueous solution around pH 6 to 7. In this state both the amino group and the acid group are ionised
- α-amino acids have different R groups, which are categorised according to their polar or non-polar nature. There are 20 different amino acids used by cells to build proteins.
- The primary structure of a protein chain is the sequence of amino acids in the chain.
- Polypeptide chains have direction. They have an amino- (or N-)terminal end and a carboxy- (or C-)terminal end.
- The secondary structure of a protein involves the folding of the polypeptide backbone and is stabilised by hydrogen bonding between peptide bond regions of the chains.
- The two most stable types of secondary structure are the α-helix and the β-pleated sheet.

- The tertiary structure of a protein consists of the folding of the polypeptide chain [illegible] interactions between the R groups (side-chains) of the amino acids in the cha[illegible] interactions include van der Waals' forces, ionic bonds, hydrogen bonding and c[illegible] disulfide bridges.
- The functioning of proteins in their biological role is very much linked to their three dimensional shape.
- Proteins can be hydrolysed back to their constituent amino acids by treatment with hot concentrated hydrochloric acid.

1.2a-c What the student needs to know	
11.1a	The general formula of an amino acid, NH_2CHRCO_2H. The structures of zwitterions and how they are formed from uncharged amino acids. The overall charge at pH 7 of amino acids with acidic or basic side-chains.
11.1b	How amino acids join together to form the primary structure of proteins. The displayed formula of the peptide bond.
11.1c	The hydrogen bonding that occurs in the secondary structure. The two stable forms of secondary structure (α-helix and β-sheet). The four types of bonding that can stabilise the tertiary structure.

1.2d Enzymes – nature's catalysts

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

Catalytic efficiency

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

Most inorganic catalysts catalyse a range of reactions involving similar substances. They are able to function over quite a wide range of temperatures and pressures. Although they are able to produce significant effects in speeding up reactions, the price of this relatively broad specificity is that the increases in reaction rate achieved are not as spectacular as those produced by enzymes. Enzymes are very effective catalysts, despite many being restricted to operating under very mild conditions. Table 2.1 compares the efficiency (turnover number) of several enzymes.

Table 2.1 – comparison of the catalytic efficiency of certain enzymes

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

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

$$CO_2 + H_2O \rightarrow H_2CO_3$$

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

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

- Higher reaction rates – the rates of enzyme catalysed reactions are typically increased by factors of 10^6 to 10^{12} times compared to the uncatalysed reaction and are several orders of magnitude greater than those of the corresponding chemically catalysed reaction.
- Milder conditions – enzyme catalysed reactions occur under relatively mild conditions: temperatures below 100 °C, atmospheric pressure, and at pH's around neutrality.
- Greater reaction specificity – enzymes are much more choosy with regard to their substrate and products: enzyme-catalysed reactions are 'clean', that is they do not produce side products.
- Capacity for regulation – the catalytic activities of many enzymes can be varied by the concentrations of substances other than the substrate: the mechanism of these regulatory processes can be complex, but some are described later in this section (see non-competitive inhibition on page 24).

Shapely molecules

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

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

The 'lock and key' model

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

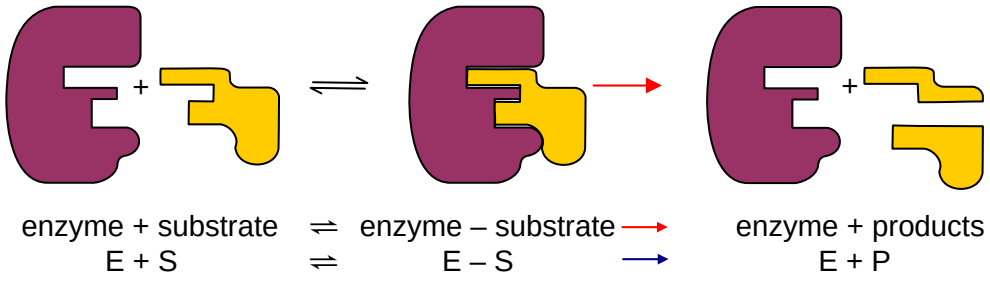


Figure 1.18 – the 'lock and key' mechanism

Enzymes, being homogeneous catalysts (see core syllabus section 8k), function by [illegible] alternative reaction pathway that requires a lower activation energy (E_a). Thus more [illegible] interactions possess sufficient energy to produce products. The energy profile shown in Figu[illegible] shows how the formation of the enzyme-substrate complex reduces the energy requirement fo[illegible] reaction to proceed.

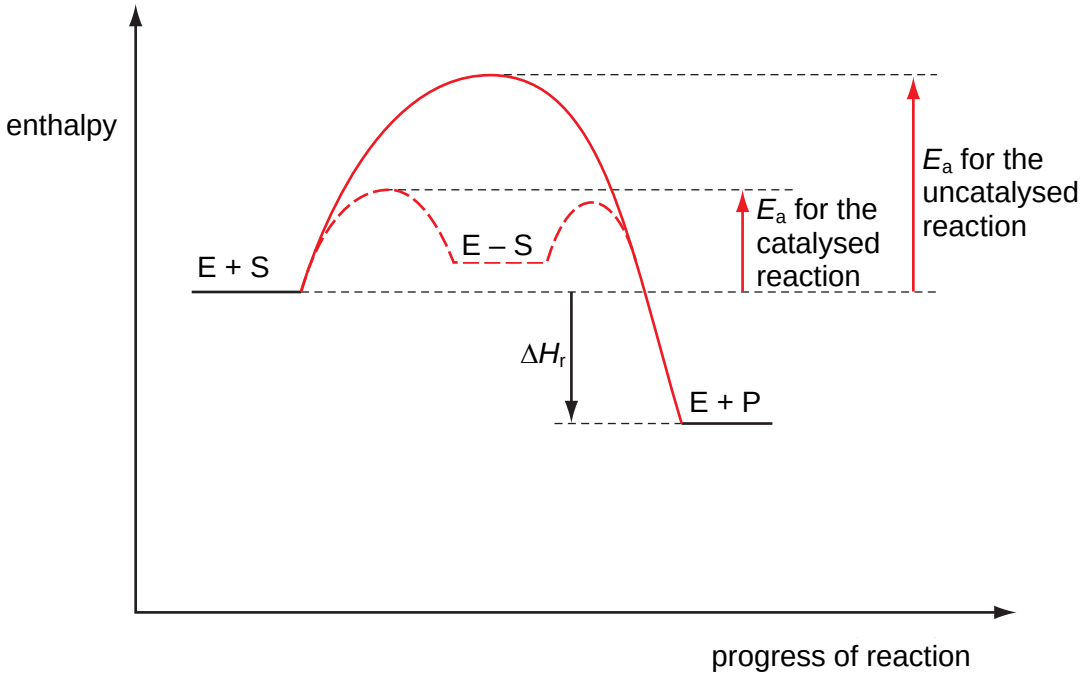


Figure 1.19 – energy profile for enzyme-catalysed and non-catalysed reaction

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

$$\underset{\text{(lock)}}{\text{ENZYME}} + \underset{\text{(key)}}{\text{SUBSTRATE}} \rightleftharpoons \underset{\text{(key in lock)}}{\text{ENZYME-SUBSTRATE}} \rightarrow \text{ENZYME} + \text{PRODUCTS}$$

E – S complexes are often almost as stable as the separated enzyme + substrate, because of the extra bonding that binds the substrate to the active site.

The first stage of the reaction is reversible, since the activation energy for the dissociation of the E – S complex back to E + S is similar in size to that for the breakdown of E – S into E + P. In some cases the second stage is also reversible, making the whole enzyme catalysed process capable of proceeding in either direction depending on the cell's metabolic requirements. Once the products have been formed, they leave the active site of the enzyme. The enzyme is then free to combine with a new substrate molecule. Enzymes, like inorganic catalysts, are not used up in the reaction they catalyse so they can be used again and again.

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

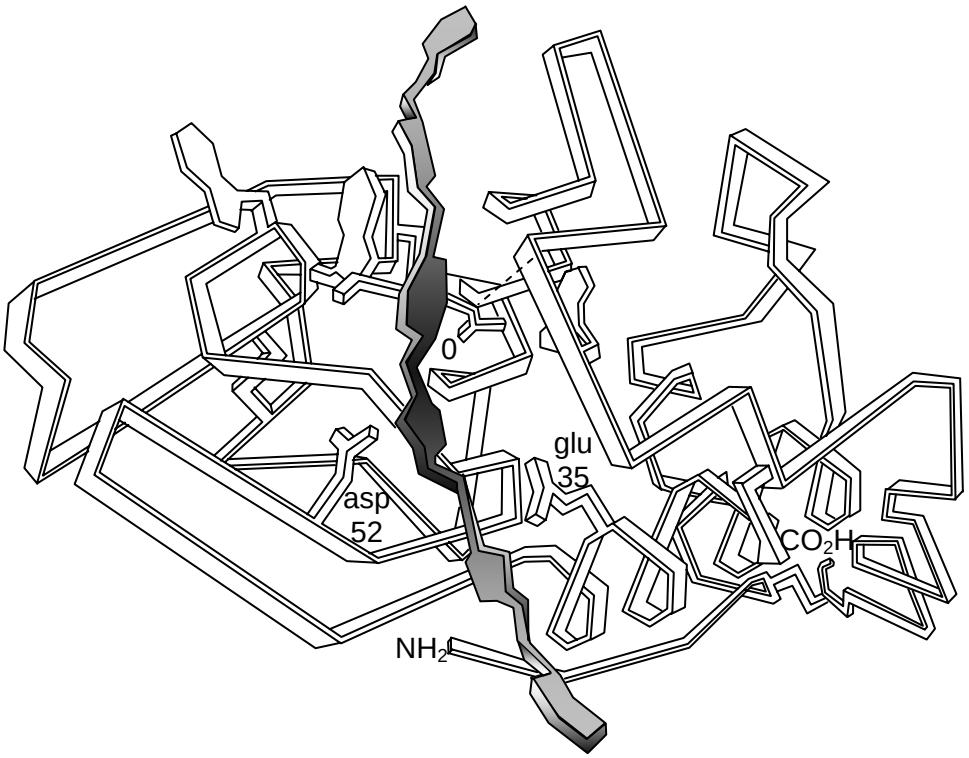


Figure 1.20 – the structure of the enzyme lysozyme, showing the carbohydrate substrate lying in the active site.

Competitive inhibition of enzyme activity

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

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

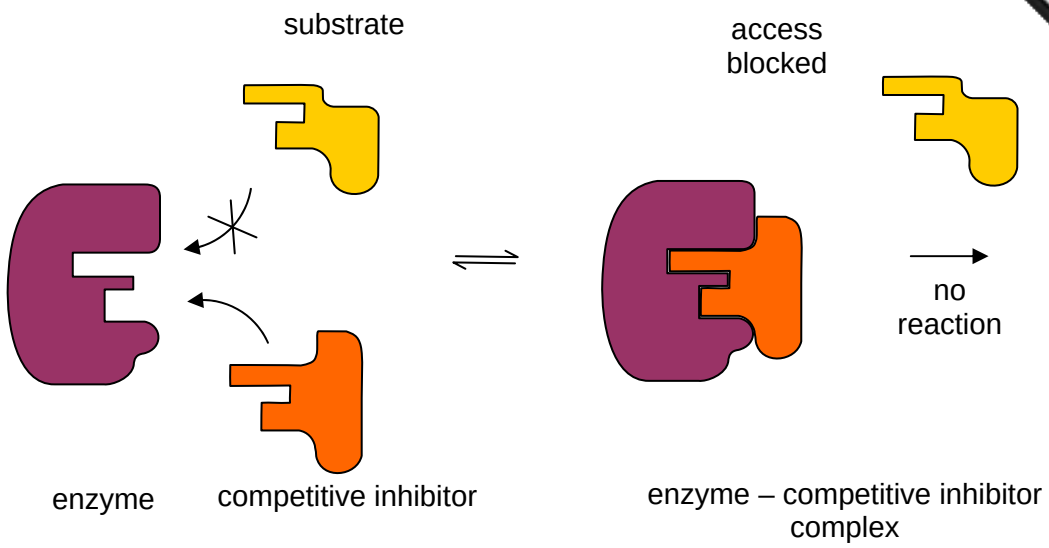


Figure 1.21 – model of action of a competitive inhibitor

The classic example of competitive inhibition is the inhibition of succinate dehydrogenase by various ionic species that structurally resemble the substrate, succinate (remember, carboxylic acids will be mostly in their anionic form at pH 7) (Figure 1.22).

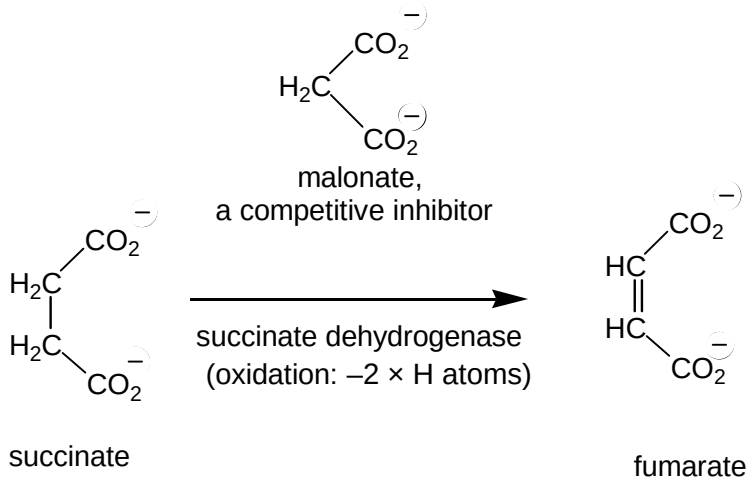


Figure 1.22 – succinate, the reaction, and one of its competitive inhibitors

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

Non-competitive inhibition of enzymes

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

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

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

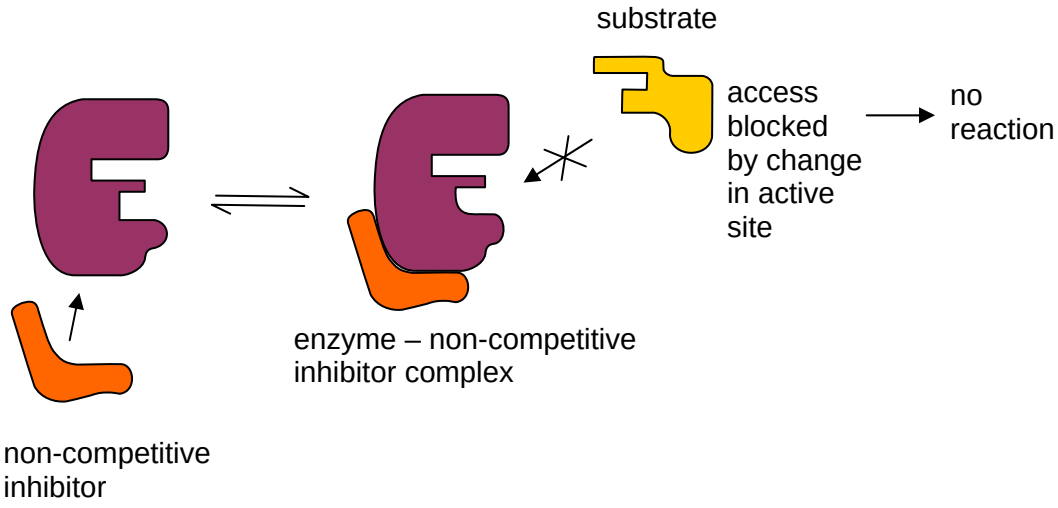


Figure 1.23 – scheme for non-competitive inhibition

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

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

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

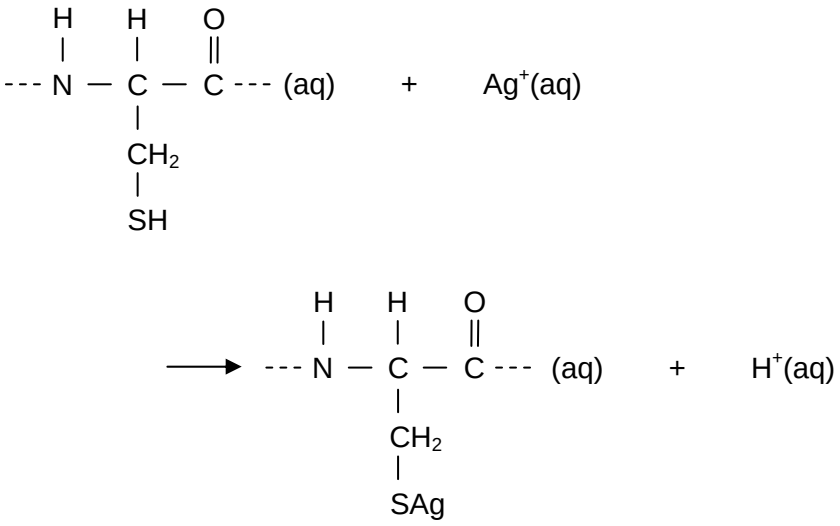


Figure 1.24 – the reaction between heavy metal ions and an –SH group

The resulting modification does not allow disulfide bridges to form, and so changes the [illegible] enzyme sufficiently enough to prevent the catalysed reaction taking place. The enzyme is t[illegible] denatured.

Factors affecting enzyme activity

The catalytic activity of any enzyme is highly dependent on those relatively weak, but highly significant, interactions that give rise to the three-dimensional tertiary structure of the protein. The same non-covalent interactions are also involved in the interaction between the substrate and the enzyme. Recognition and molecular 'fit' are the key ideas behind enzyme function. Even subtle changes in pH or temperature can modify the interactions involved in molecular shape and recognition, resulting in an enzyme working at less than maximum efficiency. Interactions with other chemical substances that cause irreversible changes in structure can also result in loss of the correct structure and the destruction of enzyme activity known as denaturation.

The effect of temperature

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

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

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

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

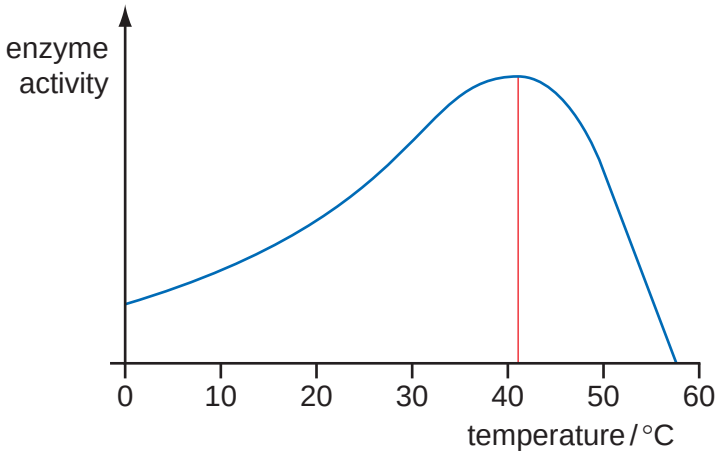


Figure 1.25 – profile of enzyme activity with temperature

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

However, there are thermophilic organisms that show adaptation to the high tempera[illegible]t springs and deep-sea thermal vents. Enzymes from such 'extremophiles' retain activity a[illegible] higher. They provide fascinating models for studying the modifications of protein structure t[illegible] needed in order to maintain enzyme function at such temperatures.

The effect of pH changes

Enzyme activity is also dependent upon pH.

- Extremes of pH (high acidity or alkalinity) will denature proteins by disrupting the precise three-dimensional arrangement of the protein chains.
- Even small changes around neutral pH can affect the ionisation of amino acid side-chains (see section 1.2b) in the active site and/or the substrate itself (Figure 1.26). If enzyme activity depends on particular residues in the active site being charged or not, then a shift of just one pH unit (remember that this represents a ten-fold change in H^+ ion concentration) can change the enzyme activity significantly.

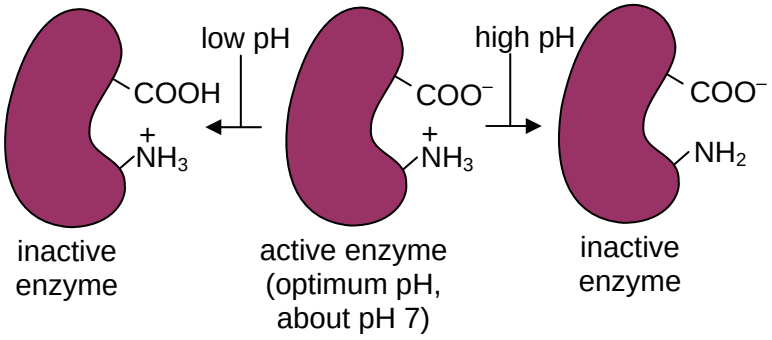


Figure 1.26 – pH changes can affect the ionization of the active site

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

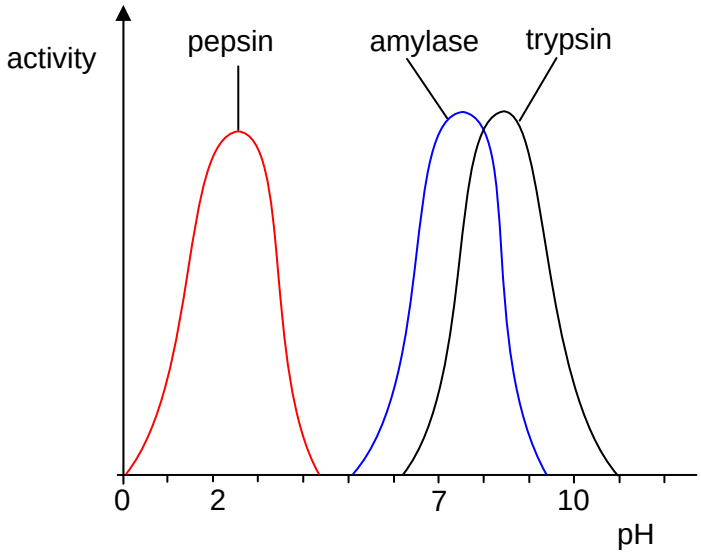


Figure 1.27 – curves showing pH optima for several enzymes

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

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

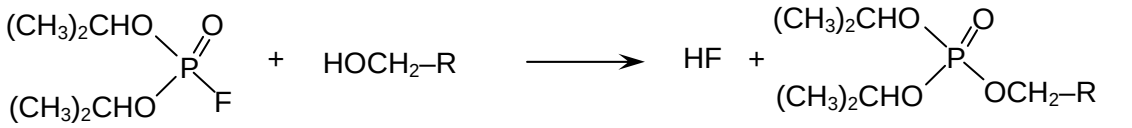
Chemical denaturation

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

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

Certain chemical inhibitors totally inactivate enzymes; their effects are irreversible. Some of these inhibitors have been useful in establishing the key amino acid residues responsible for the catalytic function of the active site of a particular enzyme.

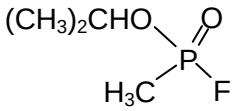
DFP (systematic name: di-(prop-2-yl)fluorophosphate) is a reagent that binds to serine residues in enzymes.



DFP

a serine residue at the active site

esterified serine – no longer able to bond to substrate



sarin

It has been used to show the importance of serine in the active sites of chymotrypsin and acetylcholinesterase. Its effect on this latter enzyme explains its function as a nerve gas. The horrific attack on the Tokyo underground in March 1995 (Figure 1.28) used the nerve gas, sarin, which, as the above diagram shows, is very similar in structure to DFP.


MASSACRE BY POISON GAS

Many die, 1,200 collapse in Tokyo Tube attack

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

Prosthetic groups and coenzymes – 'little helpers'

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

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

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

Other cofactors, known as coenzymes, are not integral parts of the enzyme structure. They are complex organic molecules, often derived from a vitamin. These coenzymes, such as NAD^+ (nicotinamide-adenine dinucleotide), work together with the enzyme to bring about the required reaction. They bind temporarily to the active site of the enzyme and effectively function as a co-substrate, providing for the transfer of groups or electrons not readily available from the side-chains of the enzyme protein (Figure 1.29). They are released from the enzyme at the end of the catalysed reaction.

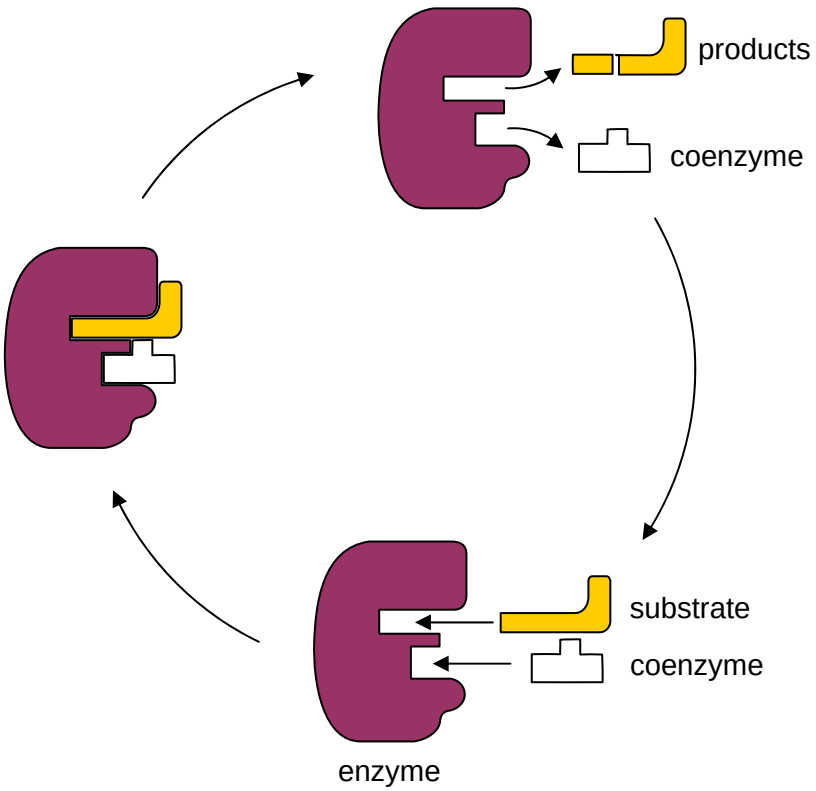


Figure 1.29 – binding of substrate and coenzyme to active site – the catalytic cycle

NAD^+, $NADP^+$, and FAD are important coenzymes because of their ability to accept H^+ ions and electrons and therefore take part in redox reactions. They are sometimes referred to as 'hydrogen carriers' because of their ability to transfer hydrogen atoms between reactions. Coenzyme A acts as a carrier of $CH_3CO–$ (acetyl, or ethanoyl) groups and is important in the metabolism of fatty acids (the long chain carboxylic acids in food). The vitamin precursors of many of these coenzymes are essential components of the diet of an organism. For instance, both NAD^+ and FAD are derived from the water-soluble B group vitamins. In humans, an inadequate supply of these precursors gives rise to deficiency diseases.

Ion channels in biological membranes

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

One recent focus of crucial importance has been the proteins that form the water and ion channels in cell membranes. These channels have been found to consist of protein sub-units which sit across the plasma membrane – they are trans-membrane proteins. They have enzyme-like function and, by their interaction with ions such as Na^+ and K^+ and relatively small molecules such as ATP (adenosine triphosphate), these proteins are able to selectively control the transport of ions into, and out of, the cell. In a manner similar to enzyme inhibitors, certain plant alkaloids (oubain, for instance) can interfere with ion transport by interacting with the protein units of the ion channels (see section 1.5c).

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

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

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

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

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

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

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

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

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

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

Summary

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

1.2d	***What the student needs to know***
11.1d	The characteristic properties of enzymes as homogeneous catalysts (i.e. reactions go via an alternative pathway with lower activation energy; the enzyme forms an intermediate complex which then re-forms the catalyst in the last step in the reaction). The 'lock and key' mechanism for forming enzyme-substrate complexes. Why enzyme activity is dependent on pH and temperature. The key features of competitive and non-competitive inhibition of enzymes, namely: competitive - The inhibitor is similar in shape to the substrate and competes with the substrate for the active site. The inhibition is reversible – it can be reversed by adding more substrate. Adding more substrate increases the rate again. non-competitive - The inhibitor attaches to the enzyme at a different 'allosteric' site from the active site. Once attached to the enzyme, the inhibitor changes the shape of the active site. This effectively reduces the concentration of the (active) enzyme. The rate cannot now be increased by adding more substrate.
11.1e	Cofactors and vitamins are often needed to allow enzyme-catalysed reactions to occur. Small molecules (e.g. sarin) can inhibit enzymes or disrupt ion channels.

1.3 – Genetic information

By the end of this section, students should be able to:

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

1.3a DNA – the source of heredity

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

Watson & Crick, Nature, 1953

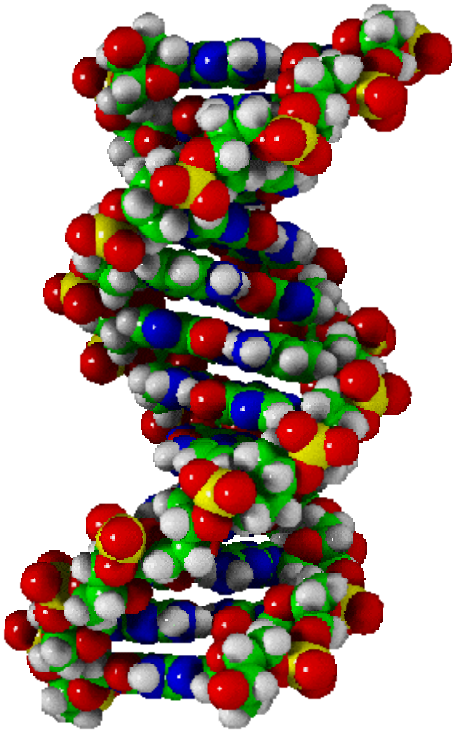


Figure 1.30 – computer-generated picture of the double helix

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

The recent dramatic spiralling of information about the origins of life, evolutionary development, and the transfer of genetic information from one generation to another starts on a one-page letter by

James Watson and Francis Crick. This letter was published on the 25th April 1953 in [illegible] journal *Nature*. The advent of molecular biology and genetic engineering stems from this [illegible] arguably the most important scientific development of the twentieth century.

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

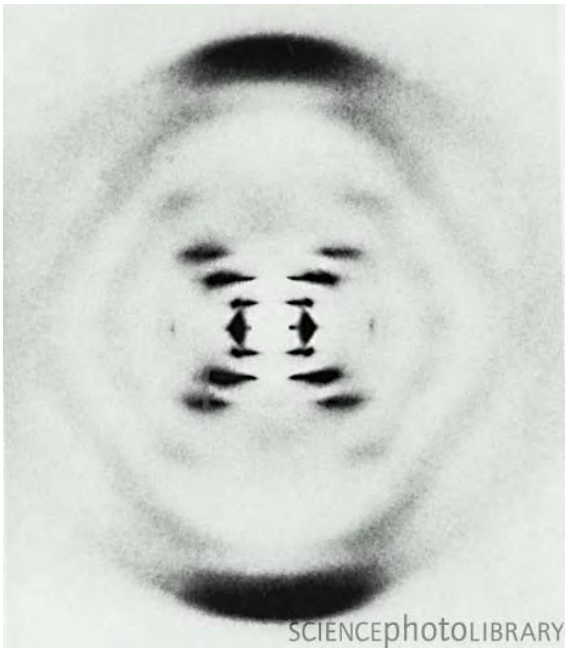


Figure 1.31 – X-ray diffraction pattern of crystalline DNA

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

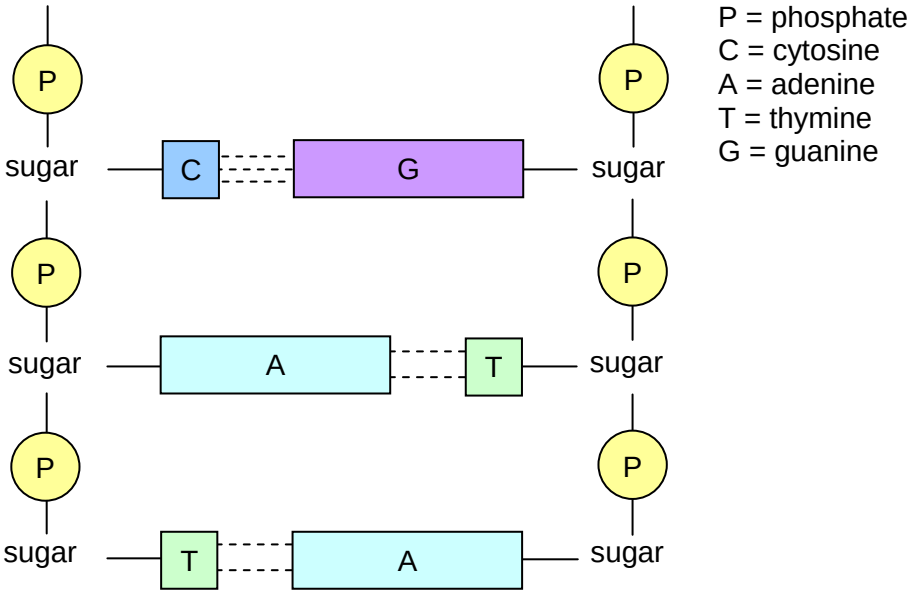


Figure 1.32 – the sugar-phosphate-base structure of DNA

1.3b The structure of DNA

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

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

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

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

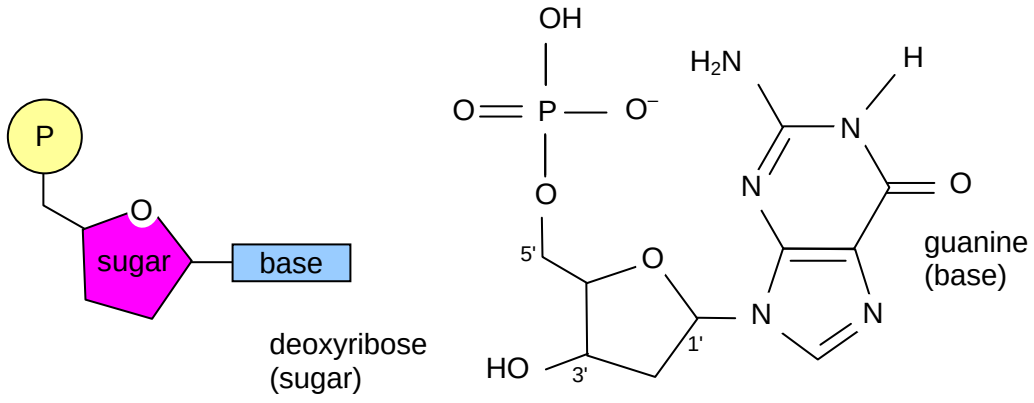
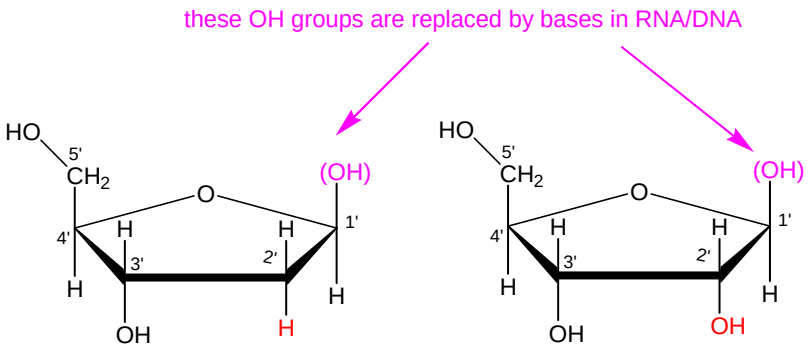


Figure 1.33 – the three components that make up a nucleotide

The sugar molecule in the nucleotides that make up DNA is deoxyribose (a pentose sugar with a five-membered ring). At this point it would be useful to compare the structure of 2-deoxyribose with that of ribose, which is the sugar molecule in the backbone of RNA (see section 1.3c).



2-deoxyribose (in DNA) ribose (in RNA)

The phosphate group is attached by an ester link to the deoxyribose. The final components of the nucleotides in DNA are the four different bases:

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

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

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

The DNA double helix

A DNA molecule consists of two strands. The backbone of each strand is made up of molecules of the sugar deoxyribose, connected by phosphate esters groups that join the 3'-OH of one deoxyribose molecule to the 5'-OH group of the adjacent deoxyribose (see Figure 1.33 for the way the OH groups are numbered). The two strands sit next to each other running in *opposite directions*, i.e. in one strand the 5'-OH phosphate is at the top, whereas in the other strand the 5'-phosphate is at the bottom. The strands are linked together by hydrogen bonding between the bases.

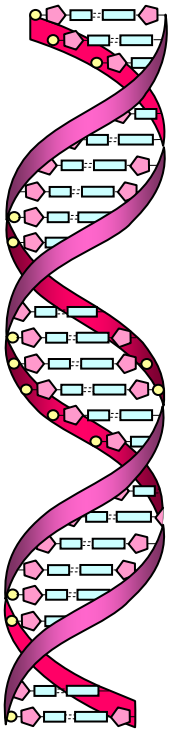


Figure 1.34 – the DNA double helix

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

The bases always pair up as follows:

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

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

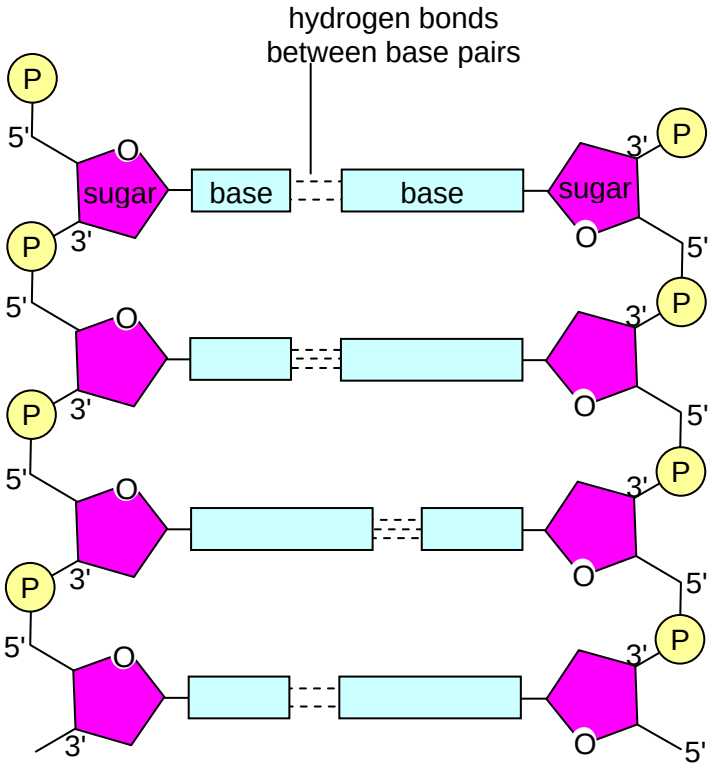


Figure 1.35 – the outline structure of DNA

1.3c The structure of RNA

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

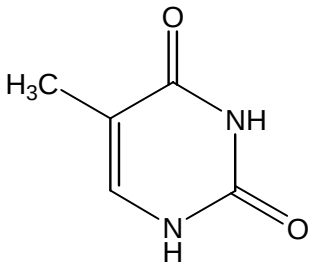
- messenger RNA (mRNA);
- ribosomal RNA (rRNA);
- transfer RNA (tRNA).

The different forms of RNA are also polynucleotides like DNA but there are significant [illegible]s (Table 1.3).

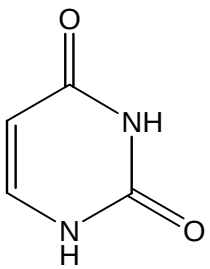
Table 1.3 – comparison of the structures of DNA and RNA

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

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



thymine (in DNA)



uracil (in RNA)

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

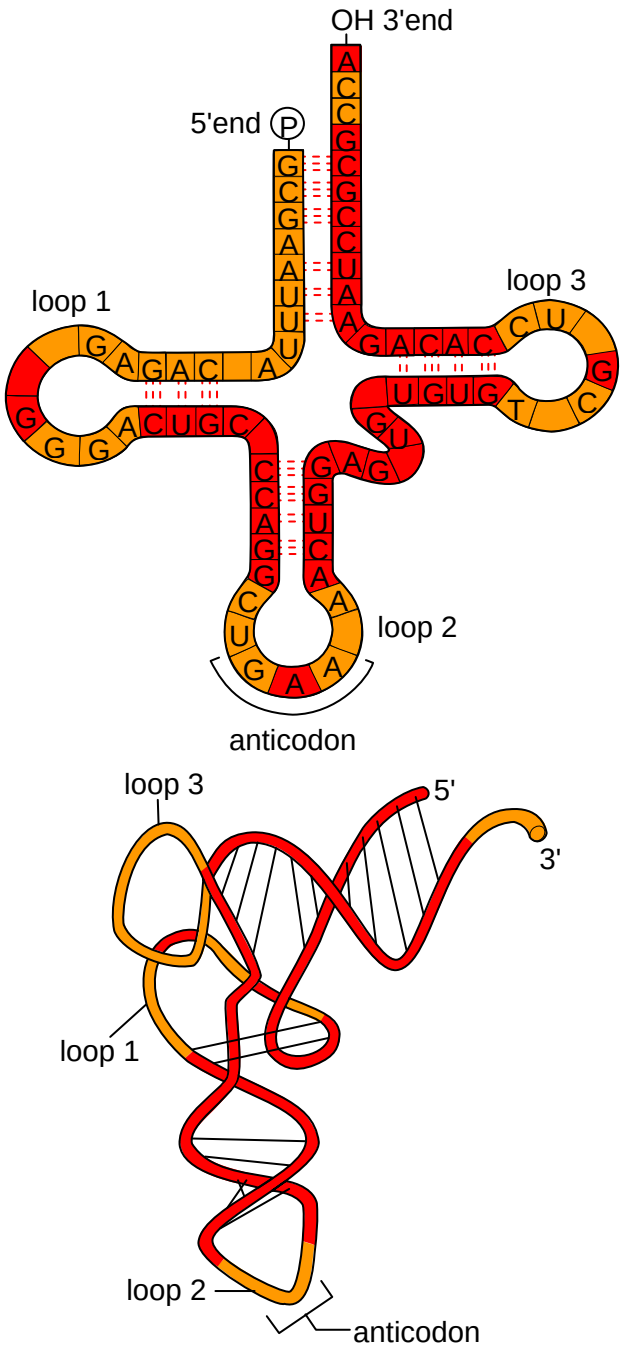


Figure 1.36 – the cloverleaf structure of tRNA molecules

1.3d Gene expression

The role of DNA - chains of information

The astounding and compelling 'neatness' of the discovered structure of DNA was that it contained a built-in mechanism for information transfer. DNA strands are capable of self-recognition and self-replication. Duplication of the genetic information can therefore readily take place every time a cell divides. This means that information encoded in the structure can be passed from one generation to another. Parent DNA molecules can produce identical daughter copies in a process known as replication.

The major interest of DNA, however, is not as the material itself, but as the blueprint for the synthesis of proteins – the enzymes, antibodies, and structural proteins which determine the nature and function

of an organism. Broadly speaking the amino acid sequence of each polypeptide chain is [illegible] specific stretch of DNA, or gene. The 'message' that is coded in a gene generates co[illegible] particular polypeptide chain through a two-stage process.

- **Transcription:** The DNA template is first copied, or transcribed, into an intermediary nuc[illegible] acid molecule, messenger ribonucleic acid (mRNA).
- **Translation:** mRNA molecules, copied from the gene, direct the assembly of the polypeptide chain. The translation process involves ribosomes attaching to, and moving along, the mRNA as the chain is synthesised. (Ribosomes are small organelles contained within a cell, about 20 nm across, which are made up from about 50 different proteins and three large RNA molecules of their own)

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

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

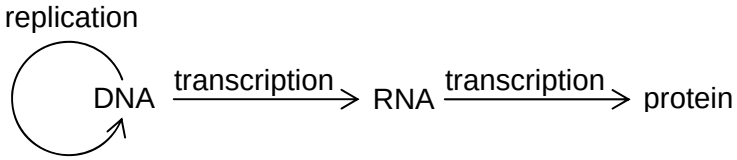


Figure 1.37 – summary diagram of replication, transcription and translation

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

Replication – the biological assembly of new DNA

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

(Note that Figure 1.38 is a simplification: in reality, after the helix unzips, one chain is replicated one nucleotide at a time, but the replication of the other chain has to wait until a number of bases in the strand become free. This is because each strand is replicated in the 5'-to-3'direction. Showing this would make the figure unnecessarily complicated, and students don't need to know it at this stage in their studies.)

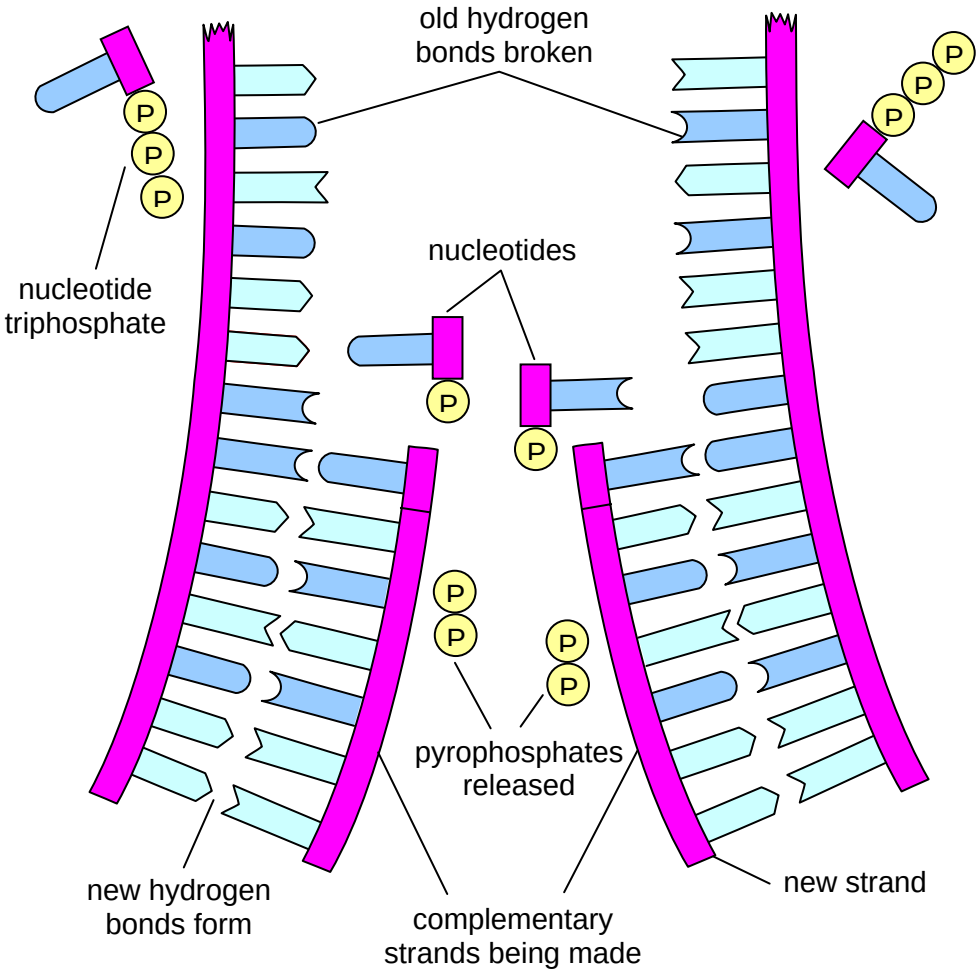


Figure 1.38 – strand of DNA acting as template for replication

Semi-conservative replication

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

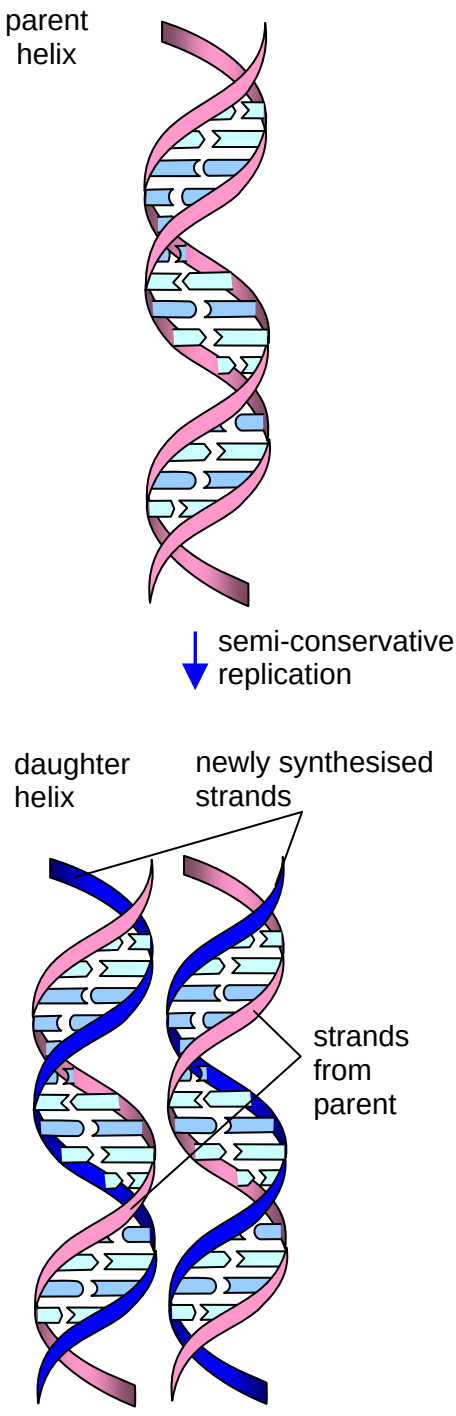


Figure 1.39 – semi-conservative replication

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

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

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

Expressing the message – the role of RNA

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

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

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

Delivering the message – transcription

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

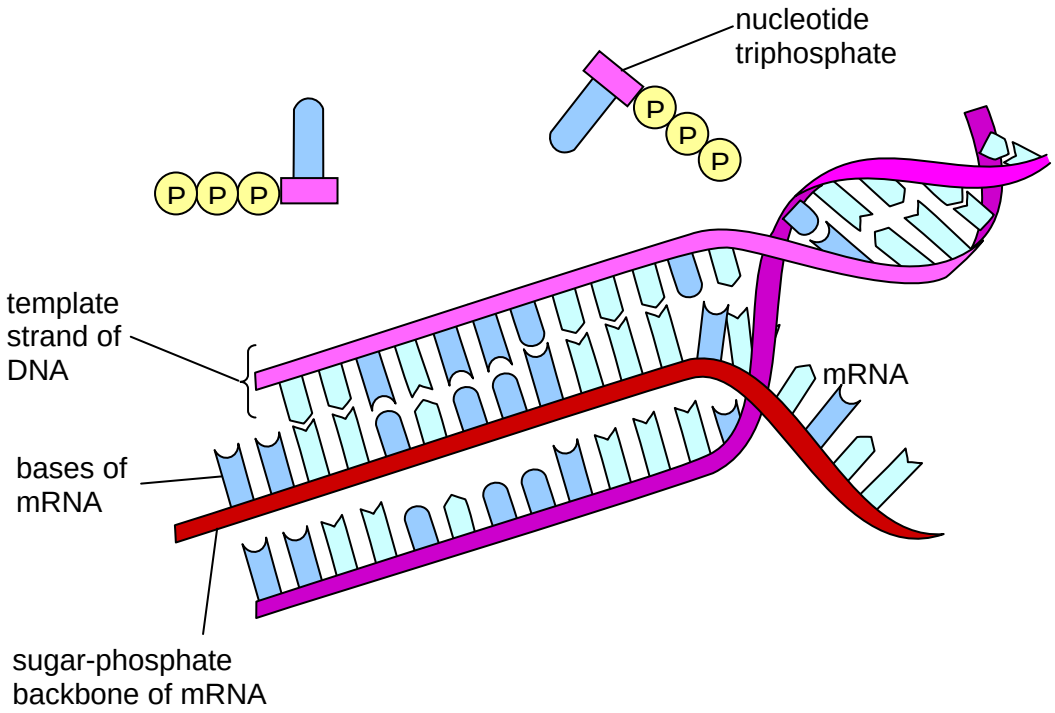


Figure 1.40 – diagram of the transcription process

Protein synthesis – translating the message

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

The genetic code

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

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

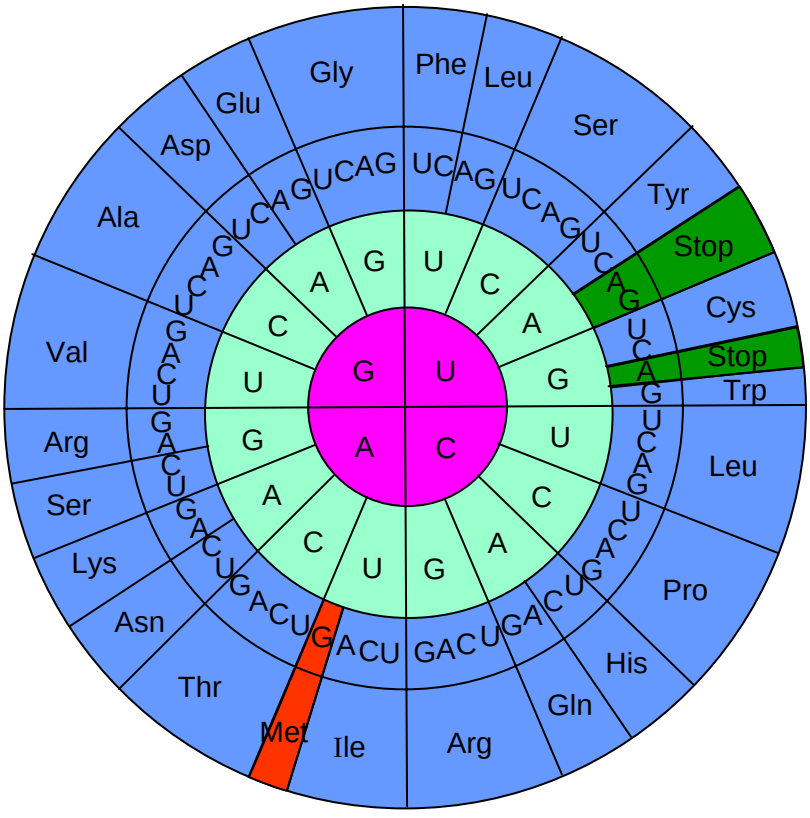


Figure 1.41 – the genetic code

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

All polypeptide chains have a defined length and sequence and so there must be a codon [illegible] amino acid in the chain (the amino-terminal end). This 'START' signal is 5'-AUG-3', which [illegible] methionine. The 'start' signal ensures that the series of triplet codons is read in the correct gr[illegible] three. Consequently the first amino acid in any newly-synthesised protein chain is always methio[illegible] though in many cases it is removed after translation is completed. There are also three codons th[illegible] do not code for any amino acid. These codons act as 'STOP' signals to end the assembly of a polypeptide chain.

Translating the message

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

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

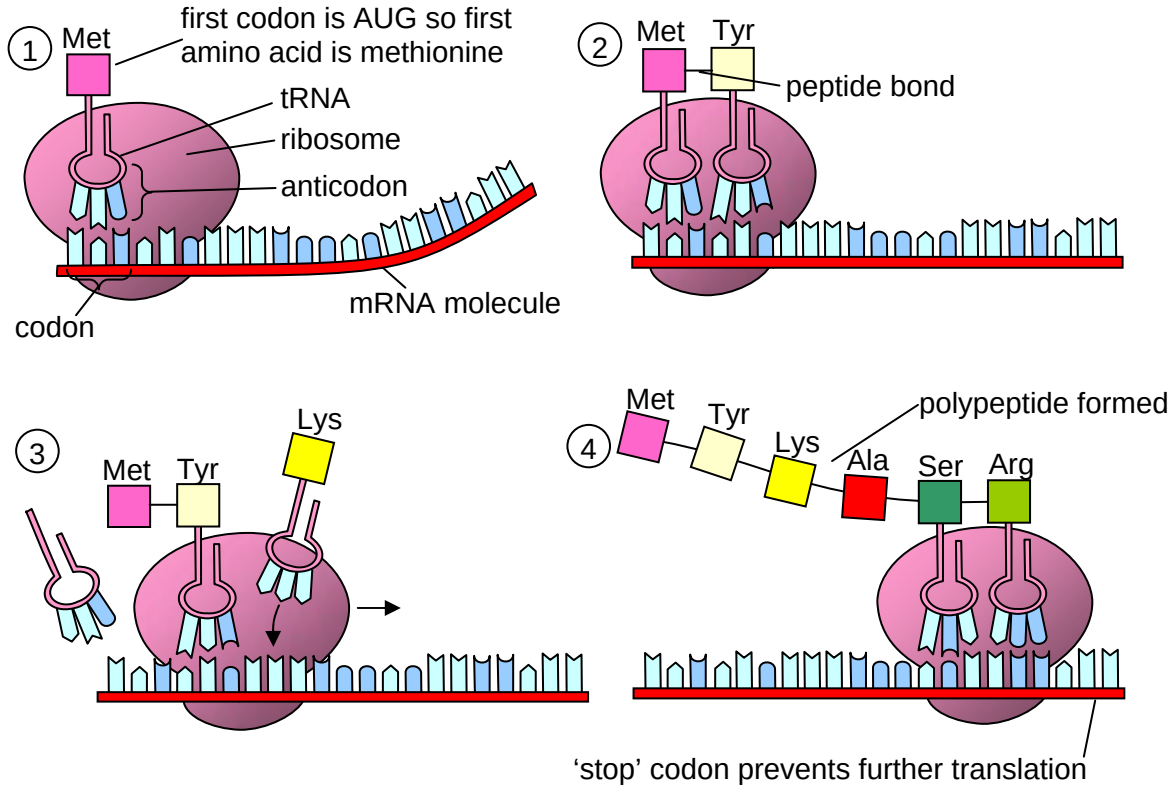


Figure 1.42 – the process of translation

1.3e Mutations

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

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

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

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

Sickle cell anaemia

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

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

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

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

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

Cystic fibrosis

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

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

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

The chloride ions diffuse out of the cell through a channel created by a protein. The protein channel is only open in the presence of ATP. The name of this membrane protein is CFTR protein (short for cystic fibrosis transmembrane regulatory protein). In a person with cystic fibrosis the CFTR protein

may be missing or, if present, it does not work properly. It does not allow chloride ion[illegible]e being pumped into the cell to leave (Figure 1.43). The chloride ion concentration in the cell [illegible] The high solute concentration in the cell causes water to move into the cell instead of out [illegible] osmosis. As a result the mucus covering the cells lining the airways becomes thick and sticky.

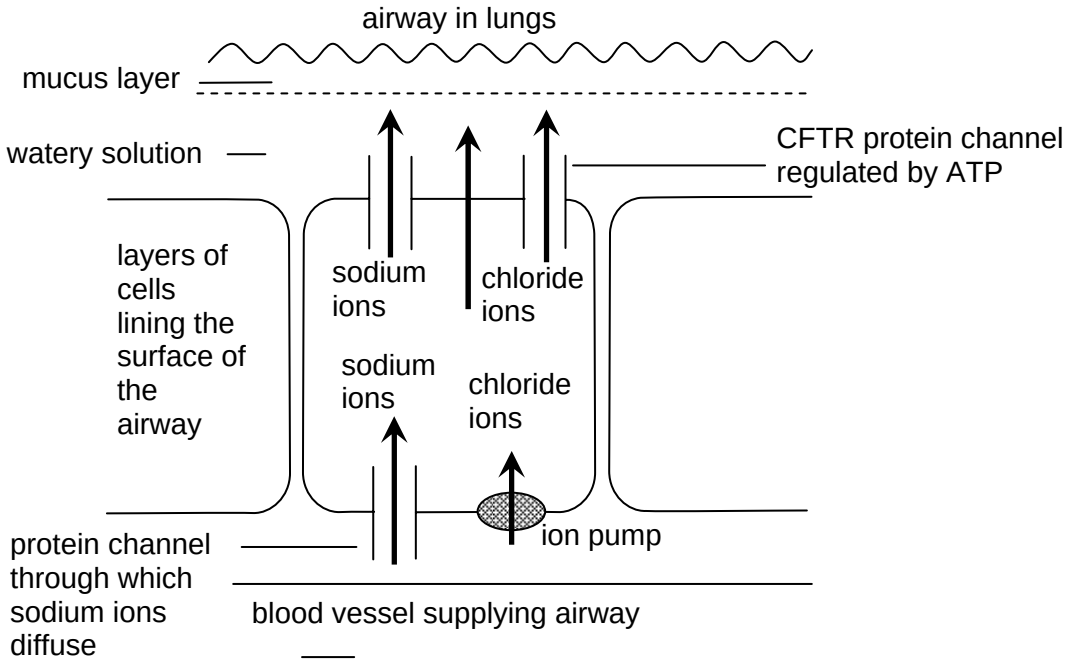


Figure 1.43 – the movement of ions across cell membranes in the lungs

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

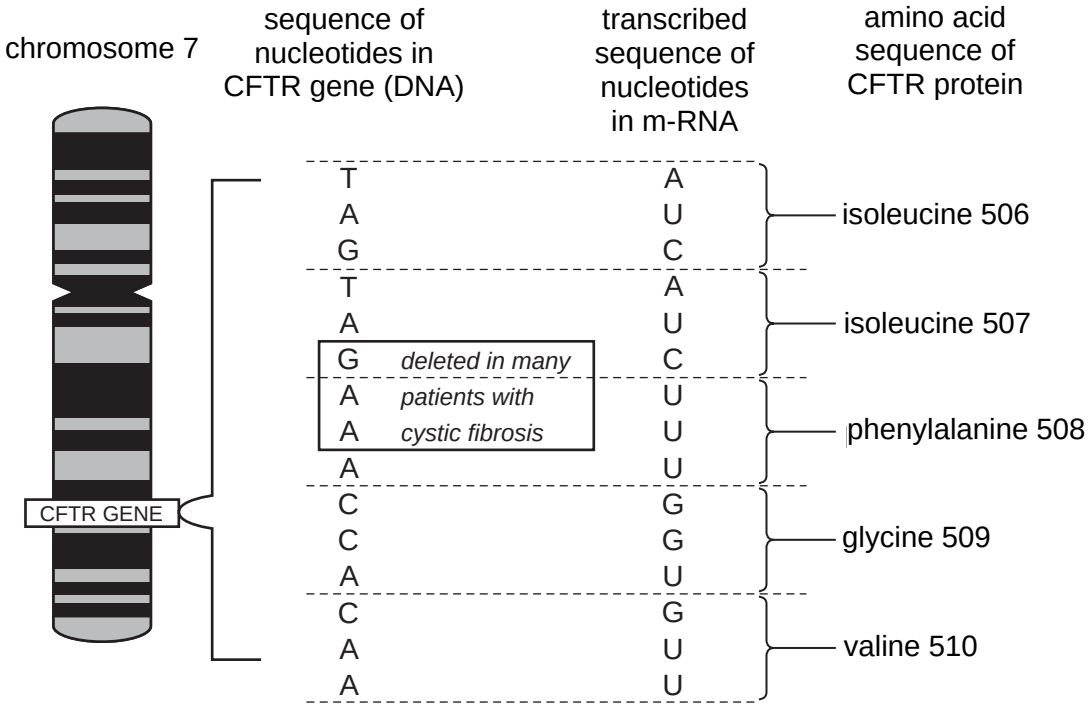


Figure 1.44 – the site of the commonest mutation that causes cystic fibrosis

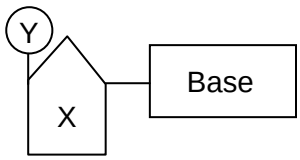
Notice that the three nucleotides that are lost - G, A and A - come from two *separ*[illegible]: TA***G*** +***AA***A – ***GAA*** = TAA. It so happens that the "new" codon, TAA, (= AUU in mRNA) co[illegible] *same* amino acid, isoleucine, as the original TAG (= AUC in mRNA). So the only difference[illegible] resulting protein is the absence of phenylalanine at position 508. The fact that the loss of just[illegible] non-polar amino acid in the middle of a sequence of similarly non-polar amino acids (isoleucin[illegible] glycine and valine are all non-polar) has such a devastating effect on the function of a protein, shows how important the protein's *exact* shape is.

Improving nature – re-designing enzymes

Despite the complexity of biological enzyme systems it is interesting to note that evolution has not necessarily developed particular enzymes to perfection. There was no need for that. Having reached a certain level of efficiency in the natural world, there was no reason to develop the structure of a particular enzyme any further. A research team in Cambridge led by Sir Alan Fersht has been able to synthesise an 'improved' enzyme by modifying its gene.

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

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



(a) *State the name of:*

(i) *the whole chemical unit shown;*

(ii) *the component labelled X;*

(iii) *the component labelled Y.*

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

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

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

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

(i) *explain how the two strands differ;*

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

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

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

Adenine 23 mole %

Guanine 27 mole %

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

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

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

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

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

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

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

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

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

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

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

Summary

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

- tRNA brings specific amino acids to the ribosomes and binds to specific codons [illegible] molecule. Protein synthesis then takes place in three stages: initiation, elong[illegible] termination.

1.3 What the student needs to know

11.1f The structure of DNA:
- The two strands made up from 3',5'-deoxyribose+phosphate backbones.
- The four types of bases (students don't need to know their structures, just their names, symbols (A,T,C,G), and which ones pair with which).
- How the bases keep the strands together through hydrogen bonding.

11.1g How the A-T and C-G pairing allows exact replication of the DNA molecule.

11.1h The structure of single-stranded mRNA.
A list of the differences between DNA and mRNA.
How mRNA uses one strand of DNA as a template during its biosynthesis (transcription).
The overall structure of a tRNA (attaching a specific amino acid, depending of which three-base codon it contains).
The hydrogen bonding that occurs between the codon on tRNA and a triplet of bases on mRNA.
The *general* role of ribosomes in the process of translation, namely:
- Bringing together the relevant tRNAs and the corresponding triplets on mRNA.
- Joining the amino acids at the other end of the tRNAs to form peptide bonds.

The overall logic of the genetic code (students do not have to learn it!)

11.1i Mutations in DNA are the result of the addition or deletion of bases, via misreadings that can occur during replication.
A mutation in DNA can (and often does) cause an incorrect protein to be made during the transcription and translation process.

11.1j+k Sickle cell anaemia is caused by the change of just one amino acid in the β-chain of haemoglobin, from Glu (glutamate) to Val (valine).
There are many mutations that affect the functioning of the CFTR protein in cystic fibrosis, but the most common one is the loss of three bases causing the absence of a Phe (phenylalanine) residue.

1.4 – ATP, Life's energy currency

By the end of this section, students should be able to:

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

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

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

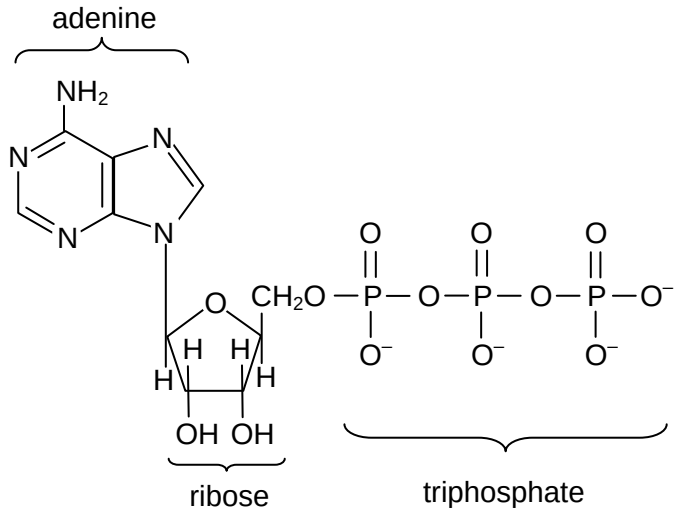
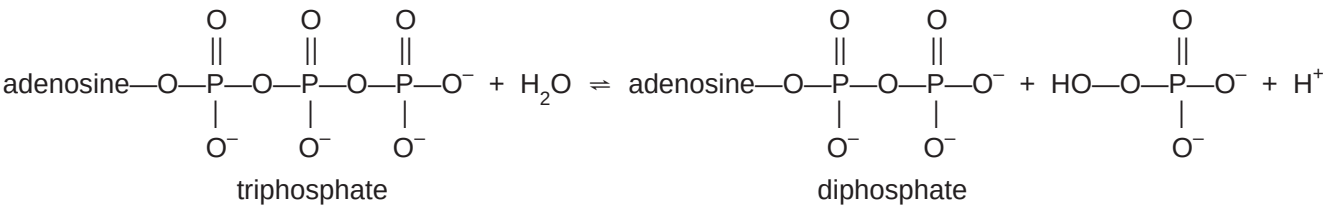


Figure 1.45 – the structure of ATP

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

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

$$ATP + H_2O \rightleftharpoons ADP + P_i \qquad \Delta H = -22 \text{ kJ mol}^{-1}$$



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

Metabolic pathways use large amounts of ATP all the time, so its synthesis in the mit[illegible]f cells is an important process. Ultimately each ATP-dependent process produces ADP, and [illegible] in metabolism is the regeneration of ATP from this. This reaction is of course energ[illegible] endothermic and this is where the energy available from the oxidation of food is needed.

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

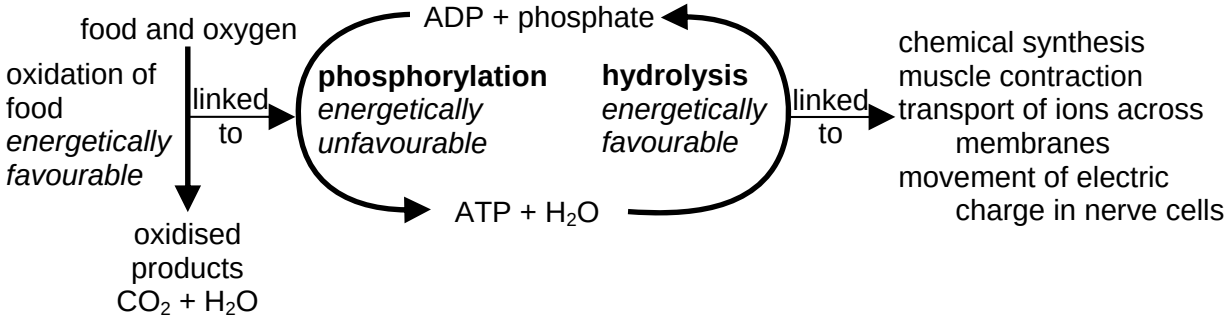


Figure 1.46 – the role of ATP in metabolism

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

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

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

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

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

(d) *Draw the structure of ADP.*

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

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

Summary

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

1.4	***What the student needs to know***
11.1l	The hydrolysis of ATP to ADP + inorganic phosphate is exothermic ($\Delta H = -22$ kJ mol^{-1}) and provides the cell with an energy source. Some of the ways in which this energy is used. For example: biosynthesis of new cells and new compounds within the cell; muscle power and motion; active transport of ions and molecules through cell walls; conduction of impulses through nerves.

1.5 – Metals in biological systems

By the end of this section, students should be able to:

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

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

1.5a Iron and the haem proteins

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

Hb	+	$4O_2$	$\rightleftharpoons$	HbO_8
haemoglobin		oxygen		oxyhaemoglobin

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

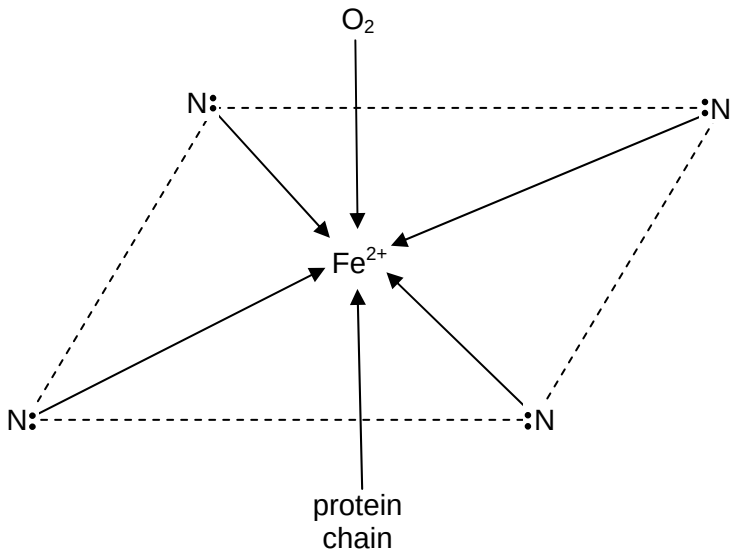


Figure 1.47 – the complex ion in haemoglobin

As in other complexes the oxygen ligand can be replaced by another ligand that binds m[illegible]. Thus, in carbon monoxide poisoning the CO molecules replace the oxygen in oxyhaemogl[illegible] binding is 200 times stronger and irreversible, resulting in the haemoglobin molecule lo[illegible] function.

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

1.5b Zinc as an enzyme cofactor

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

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

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

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

1.5c Sodium and potassium ion transfer across cell membranes

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

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

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

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

ions are stripped of the associated water molecules as they enter the channel, linking [illegible] oxygen atoms in certain R-groups of the protein.

The enthalpy required to lose the hydration shell around the ions is compensated for by that giv[illegible] when the new association is formed with the protein. The K^+ ions pass through the channel and t[illegible] re-associate with water on the other side. The hydration shell is re-formed around the ion and energy is released. The selectivity of the channel depends on the distances between the oxygen atoms in the protein side-chains and the K^+ ions. The smaller Na^+ ions will not 'fit' the channels as the distances are too great for the complex to form.

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

1.5d Toxic trace metals in the environment

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

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

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

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

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

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

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

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

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

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

(b) *Give two important features of the nature of the attachment of oxyg[illegible]h haemoglobin molecule.*

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Summary

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

- The effects of these heavy metals are evident at even relatively low co[illegible] However, the problems associated with them are made worse by the fact that the[illegible] can be concentrated within the food chain.

1.5	***What the student needs to know***
11.1m	The roles of some of the metals essential for life: How iron helps haemoglobin molecules to transport O_2 molecules in the blood. How the sodium-potassium pump works (in outline only). The importance of metal ions such as zinc acting as cofactors for some enzymes.
11.1n	Some metal ions, e.g. Ag^+ and Hg^{2+}, are toxic. The outline mechanism of this toxicity (e.g. complexing with $–CO_2^–$ and –SH groups on the side-chains of amino acids, thus disrupting the tertiary structure of enzymes).

Conclusion

The tremendous increase in our understanding of the biochemical basis of life has been developed over the past 60 years or so. The beauty of the DNA story and its universality across the different levels of the evolutionary spectrum of life is of great intellectual interest and excitement. The overall mechanisms involved in the generation of proteins, and the manner in which they function to support the different processes involved in sustaining life, are intriguing. Although the molecules and the processes are complex, they can be seen to rely on certain basic chemical principles and forces of interaction. The factors that determine and maintain the three-dimensional structures of these complex polymers are absolutely crucial to the correct functioning of life's biochemical reactions.

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

1.6 – Revision self-assessment questions

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

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

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

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

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

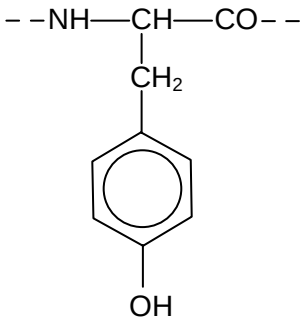
(a) *Which insoluble fibrous protein is hair made of?*

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

(c) *What types of reaction are* ***(i)*** *the thioglycollate solution, and* ***(ii)*** *the hydrog[illegible] solution carrying out?*

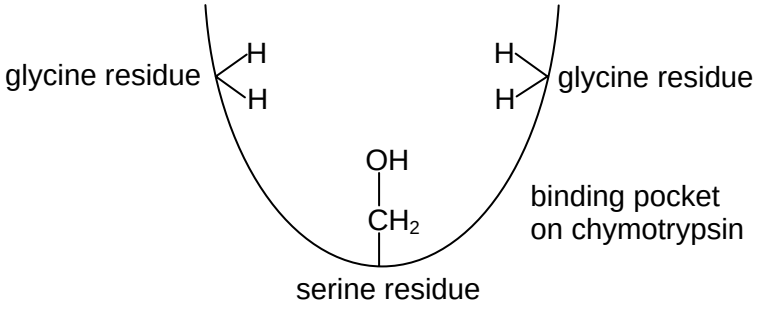
(d) *What causes the shaping of the hair to eventually be lost?*

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



The tyrosine residue binds the polypeptide to the enzyme chymotrypsin.

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



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

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

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

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

. . . . .AGAAGAGAAGCU . . . .

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

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

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

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

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

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

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

Normal 5'- AGUCCAUCACUUAAU - 3'

Mutant 5'- GUCCAUCACUUAAUG - 3'

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

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

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

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

1.7 – Key definitions

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

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

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

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

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

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

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

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

cytoplasm: the fluid within cells surrounding the nucleus and organelles

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

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

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

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

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

extracellular: outside cells

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

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

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

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

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

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

intracellular: inside cells

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

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

membrane: the boundary surrounding all cells and also surrounding organelles within eukaryotic cells, average thickness 7 nm

metabolism: the chemical reactions taking place in a living organism

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

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

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

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

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

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

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

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

peptide bond: the link present between amino acids in a polypeptide (protein) chain; the link is formed by a condensation reaction between the amino group ($–NH_2$) of one amino acid and the carboxylic acid group (–COOH) of another amino acid

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

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

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

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

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

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

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

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

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

tertiary structure (of proteins): the third level of protein structure involving the overall fol[illegible] polypeptide chain; the chain is stabilised by ionic interactions, van der Waals' forces, hyd[illegible] bonding and covalent disulfide bond formation

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

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

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

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

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

1.8 – Resources

Texts & booklets:

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

Resources available for practical and project work on biochemical topics

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

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

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

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

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

2 – APPLICATIONS OF ANALYTICAL CHEMISTRY

2.1 – Introduction

What do the following branches of science have in common?

- archaeology
- environmental science
- food science
- forensic science
- medical research
- perfume science
- astrobiology

The answer is that they all involve the identification of chemical compounds that occur in very small quantities in mixtures containing possibly several hundred different substances. This chapter looks at how we can separate these highly complex mixtures into their individual components, and how we can then find out the chemical structures of these compounds.

The techniques of separation, purification and identification described in this chapter have been the subject of continuous development over the last 80 years. During that time they have become increasingly more powerful, accurate, sensitive, and miniaturised.

2.1a Astrobiology

The extreme portability of some of these techniques was demonstrated by the fact that both the Viking Mars landers, and the abortive Beagle 2 Mars lander, included on-board gas chromatographs and mass spectrometers that weighed just a few kg, but their sensitivity was such that they could detect nanogram (1×10^{-9} g) quantities of substances.

Why did these craft contain these instruments?

Life on Mars?

The purpose of the Mars probes was to search for signs of life (past or present) on the Red Planet. Once the presence of water (albeit frozen) had been confirmed in the surface rocks of Mars, scientists wondered whether there might be some form of life there. But how could they be sure? The easiest way was to try to discover whether the small molecules which are the hallmarks of life could be found in the Mars environment.

The 1976 Viking lander was able to shovel up soil samples from the Martian surface and use an automated gas chromatograph-mass spectrometer (GCMS) to investigate them. The GCMS baked a small soil sample in an oven to drive off any volatile gases present, separated these volatiles using a gas chromatograph, and analysed their composition with a mass spectrometer. The experimental set-up worked perfectly, but disappointingly found none of the small amino and carboxylic acids that might have been expected if living organisms had been present in the soil.

The Beagle 2 Mars lander unfortunately never sent back any signals after it left the orbiting craft in December 2004. It is assumed that it crashed into the planet's surface and was damaged or destroyed. It had on board a sophisticated mass spectrometer that could analyse the ^{12}C:^{13}C ratio in samples of CO_2. By heating up soil samples with pure oxygen, any "organic" compounds derived from life forms would be oxidised to CO_2. By comparing the ^{12}C:^{13}C ratio in this sample of CO_2 to that of a CO_2 sample from the Martian atmosphere (which contains 95% CO_2), it would have been possible to determine whether the carbon in the soil sample was of biological origin. This is because it has been found that in every biosystem on Earth, organisms concentrate ^{12}C at the expense of ^{13}C, and it was assumed Martian organisms would do likewise.

Scientists still disagree and have been unable to confirm if there is life on Mars. Some results give a strong indication that there could be simple forms of life there, whereas other experiments have given negative results. One thing is certain, however: the use of chemical instrumentation and analysis, like

the methods described in this chapter, will play a crucial part in providing a definite [illegible] question.

2.2 – Determining structures

By the end of this section, students should be able to:

- *outline the principles of the mass spectrometer*
- *explain the meaning of the terms base peak, molecular ion and fragmentation pattern as applied to mass spectra*
- *suggest the identity of a molecule based on its fragmentation products*
- *deduce the number of carbon atoms in a compound using the M+1 peak*
- *deduce the presence of chlorine or bromine in a molecule by use of the M+2 peak*
- *outline the principles of ^{1}H nuclear magnetic resonance*
- *explain how the chemical environment of a proton affects the magnetic field it experiences, and hence the frequency of radiation absorbed at resonance*
- *explain the use of the δ scale, and the use of TMS as an internal standard*
- *describe the effects of adjacent protons on the magnetic field experienced by a given proton*
- *predict, from the integration of the peak areas in an NMR spectrum, the number of protons in each group present in a given molecule*
- *predict, from the splitting pattern of a particular peak in an NMR spectrum, the number of protons adjacent to a given proton*
- *suggest, from an NMR spectrum, possible structures for a given molecule*
- *describe how the addition of D_2O can be used to identify protons on oxygen or nitrogen atoms*
- *appreciate the application of X-ray crystallography in the elucidation of the structures of large molecules of biochemical importance*

2.2a Mass spectrometry

The principle of mass spectrometry is very simple, although modern mass spectrometers are very sophisticated, precision-made instruments, capable of determining molecular masses to an accuracy of 1 part in 100 000.

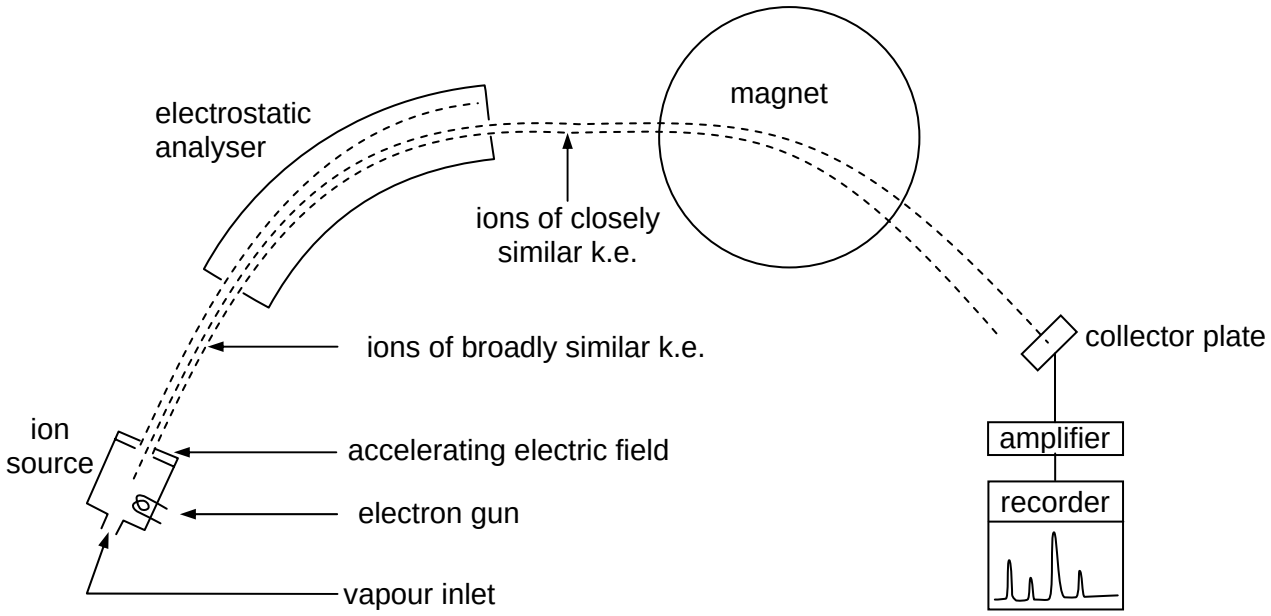


Figure 2.1 – diagram of a mass spectrometer

Six processes occur in a mass spectrometer (see Figure 2.1).

1. If not already a gas, the compound is vaporised in an oven. Only a small vapour pr[illegible] required, since the interior of the instrument is kept under a high vacuum.
2. Electrons are fired at the gaseous molecules. These knock off other electrons from some of the molecules:

$$M + e^- \longrightarrow M^+ + 2e^-$$

3. The gaseous ions are accelerated by passing through an electric field (at a voltage of 5-10 kV). At this stage they can be travelling at up to 2×10^5 m sec^{-1} (about 1/1000 the speed of light).
4. They then pass through an electrostatic analyser, which selects ions with a narrow range of kinetic energies by using an electric field.
5. The fast-moving ions now pass through the poles of an electromagnet, where they are deflected (see below).
6. The deflected ions pass through a narrow slit and are collected on a metallic plate connected to an amplifier. For a given strength of magnetic field, only ions of a certain mass pass through the slit and hit the collector plate. As the (positive) ions hit the plate, they cause a current to flow through the amplifier. The more ions there are, the larger the current.

The ions may travel a metre or so through the spectrometer. In order for them to do this without hitting into too many air molecules (which would deflect them from their course), the inside of the spectrometer is evacuated to a very low pressure.

The equation governing the deflection of ions in the magnetic field is as follows:

$$r = \sqrt{\frac{2mV}{eB^2}}$$

where r = radius of circular path in the magnetic field

m = mass of ion
V = accelerating voltage
e = electrical charge on the ion
B = strength of magnetic field

The accelerating voltage V is usually kept constant. We thus see that the radius of curvature is proportional to $\sqrt{m/e}$, but inversely proportional to B. To obtain a mass spectrum, therefore, the current through the electromagnet is changed at a steady rate. This causes the magnetic field, B, to change its strength, and hence allows ions of different mass/charge values to pass successively through the slit. A mass spectrum is produced, which plots (ion current) against (electromagnetic current), which is equivalent to (relative abundance) against (mass/charge *(m/e)* ratio). In practice, most ions that are formed in a mass spectrometer have a charge of +1, and so the *x*-axis is a measure of the masses of the ions. The *y*-axis normally shows the abundances of the peaks as a percentage of that of the most abundant peak (known as the base peak). The base peak usually corresponds to a particularly stable fragment of the molecules under investigation.

There are three main ways in which mass spectrometry is applied to the determination of the structures of organic compounds.

1. By measuring the relative heights of the molecular ion (M) peak and the (M+1) peak we can determine the number of carbon atoms in a molecule, and by using the (M+2) and (M+4) peaks (if any) we can identify halogen-containing compounds.
2. By measuring the accurate mass of a molecular ion we can determine its molecular formula.
3. By identifying the fragments produced when an ion breaks up inside a mass spectrometer we can often piece together the structure of the parent molecule.

We shall look at each of these techniques in turn.

^{12}C:^{13}C ratio

Naturally-occurring carbon is composed of 98.9% ^{12}C and 1.1% ^{13}C (along with extremely small, and variable, amount of ^{14}C). Although the ^{12}C:^{13}C ratio is very small for compounds like methane which contain just one carbon atom, the ratio increases in proportion to the number of carbon atoms, as the

chances of a molecule containing at least one ^{13}C atom increase. The formula relating t[illegible] ratio the number of carbon atoms is:

$$n = \frac{100}{1.1}\left(\frac{A_{M+1}}{A_M}\right)$$ where n = number of carbon atoms

A_{M+1} = the abundance of the M+1 peak

and A_M = the abundance of the molecular ion, M, peak.

Example: Compound A has a molecular ion at an *m/e* value of 120, and relative abundance 23%, and a peak at *m/e* 121 with a relative abundance of 2%. How many carbon atoms are in a molecule of A?

Answer: $n = (100/1.1) \times (2/23) =$ 7.91

The nearest whole number is 8. Therefore compound A has 8 carbon atoms per molecule.

SAQ 1. *Compound B contains carbon, hydrogen and oxygen only. Its mass spectrum contains a molecular ion peak at m/e = 102 (relative abundance 35%) and an M+1 peak at m/e 103 (1.5%).*
Calculate the number of carbon atoms in the molecule, and hence deduce the number of oxygen atoms it contains, and its molecular formula.

(M), (M+2) and (M+4) peaks

Both chlorine and bromine naturally occur as mixtures of two isotopes, with the relative abundances shown in the table.

element	isotope	relative abundance	approximate ratio
chlorine	^{35}Cl	75.8%	3:1
	^{37}Cl	22.4%	
bromine	^{79}Br	50.5%	1:1
	^{81}Br	49.5%	

The mass spectrum of a compound containing one of these elements should therefore show two molecular ions, one with an *m/e* value two mass units higher than the other. If the molecule contains **two** chlorine atoms, (or two bromine atoms, or one of each) we should expect to see **three** molecular ions, at *m/e* values of M, M+2 and M+4. The ratio of the M/(M+2) peak should reflect the natural abundances given in the table (i.e. 3:1 for chlorine; 1:1 for bromine). The situation is more complicated for molecules containing two halogen atoms. For example dibromomethane, CH_2Br_2, shows three molecular ion peaks at *m/e* 172, 174 and 176 in the ratio 1:2:1, corresponding to $CH_2{}^{79}Br_2$, $\{CH_2{}^{79}Br^{81}Br + CH_2{}^{81}Br\ {}^{79}Br\}$ and $CH_2{}^{81}Br_2$. The three molecular ion peaks in dichloromethane, CH_2Cl_2, however, appear in the ratio 9:6:1, due to the greater natural abundance of the ^{35}Cl isotope.

SAQ 2. *Calculate the M : M+2 : M+4 ratio for CH_2BrCl*

Molecular formulae from accurate masses

Using *very high-resolution* mass spectrometry, we can measure *m/e* ratios to an accuracy of 5 significant figures (1 part in 100 000). By this means, it is not only possible to measure the M_r value of a compound, but also to determine its molecular formula. We can do this because the accurate relative atomic masses of individual atoms are not exact whole numbers.

For example:
The following three compounds all have an approximate M_r of 70:

name	structure	molecular formula
pentene	$CH_3CH_2CH_2CH{=}CH_2$	C_5H_{10}
aminopropanonitrile	$CH_3CH(NH_2)CN$	$C_3H_6N_2$
but-1-ene-3-one	$CH_2{=}CHCOCH_3$	C_4H_6O

By using the following accurate atomic masses we can calculate their accurate M_r values.

element	accurate relative atomic mass
H	1.0078
C	12.000
N	14.003
O	15.995

Answer: The accurate masses are as follows:

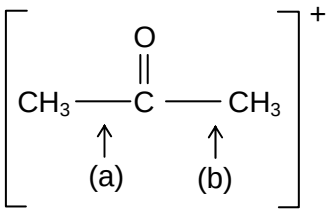
C_5H_{10} = 5 × 12.000 + 10 × 1.0078 = 70.078
$C_3H_6N_2$ = 3 × 12.000 + 6 × 1.0078 + 2 × 14.003 = 70.054
C_4H_6O = 4 × 12.000 + 6 × 1.0078 + 15.995 = 70.045

The last two are quite close together. They differ by 9 parts in 70 000, or about 0.13%. This is well within the capabilities of a high-resolution mass spectrometer.

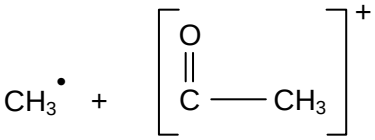
SAQ 3. *Explain whether a molecule having an accurate mass of 60.0574 is 1,2-diaminoethane, $C_2H_8N_2$, or propan-1-ol, C_3H_8O.*

The use of fragmentation patterns

If the ionising electron beam in a mass spectrometer has enough energy (anything from 25 – 70 electron-volts), the molecular ions formed by the loss of an electron can undergo bond fission, and molecular fragments are formed. Some of these will carry the positive charge, and therefore appear as further peaks in the mass spectrometer. Take propanone as an example:



cleavage at (a) gives:



m/e = 43

cleavage at (b) gives:

m/e = 15

We therefore expect the mass spectrum of propanone to contain peaks at *m/e* 15, and 43, as well as the molecular ion peak at 58 (see Figure 2.2).

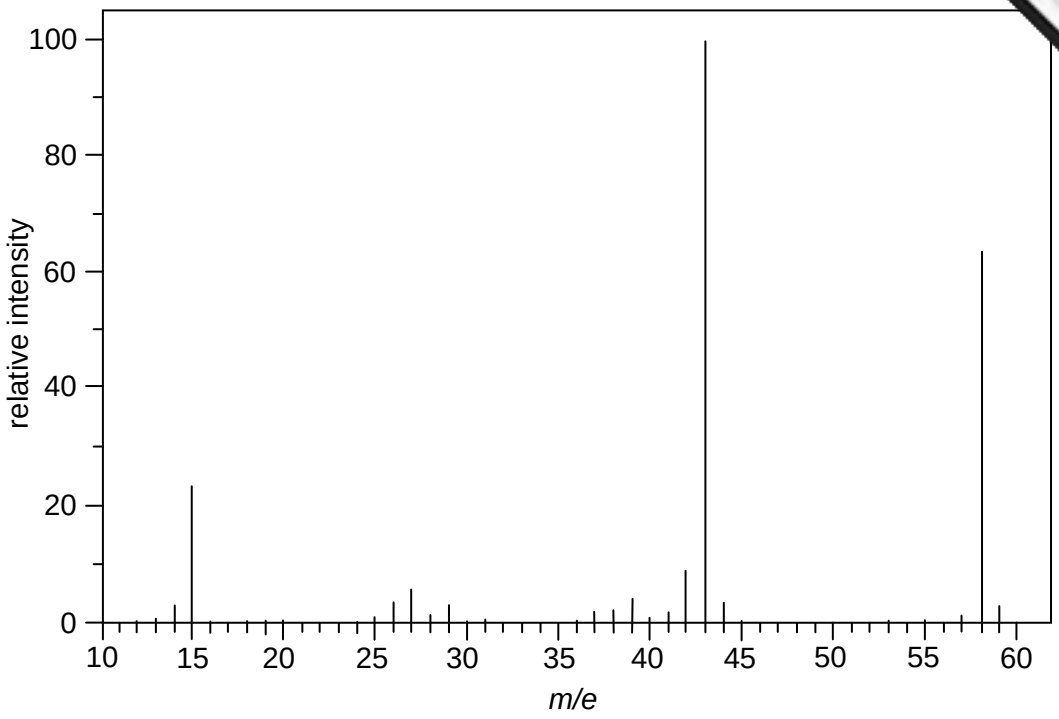


Figure 2.2 – mass spectrum of propanone

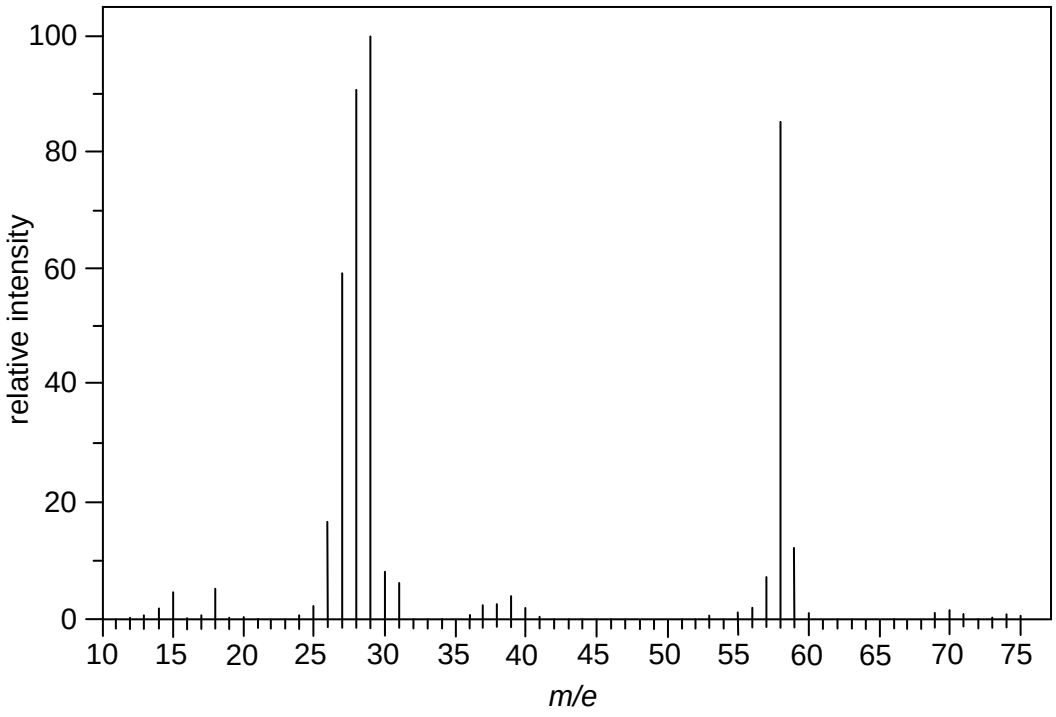


Figure 2.3 – mass spectrum of propanal

The fragmentation pattern can readily distinguish between isomers. Compare Figure 2.2 with Figure 2.3 which shows the mass spectrum of propanal. Here there is no peak at *m/e* = 15, nor one at *m/e* = 43. Instead, there are peaks at *m/e* = 57 and several from *m/e* = 26 to *m/e* = 29. This is readily explained by the following fragmentations.

$[CH_3CH_2CHO]^+ \longrightarrow [CH_3CH_2]^+ + {}^{\bullet}CHO$ *or* $\longrightarrow CH_3CH_2^{\bullet} + [CHO]^+$

m/e = 29 *m/e* = 29

$[CH_3CH_2CHO]^+ \longrightarrow [CH_3CH_2CHO]^+ + H^{\bullet}$ *or* $\longrightarrow CH_3CH_2CHO^{\bullet} + H^+$

m/e = 57 *m/e* = 1

SAQ 4. *Use the values of accurate relative atomic masses in the table above to see whe[illegible]d be possible to decide whether the peak at m/e = 29 is due to $CH_3CH_2^+$ or CHO^+.*

An example of how these three techniques can help us to determine the structure of an unkn[illegible] compound is given by the following spectrum of compound B.

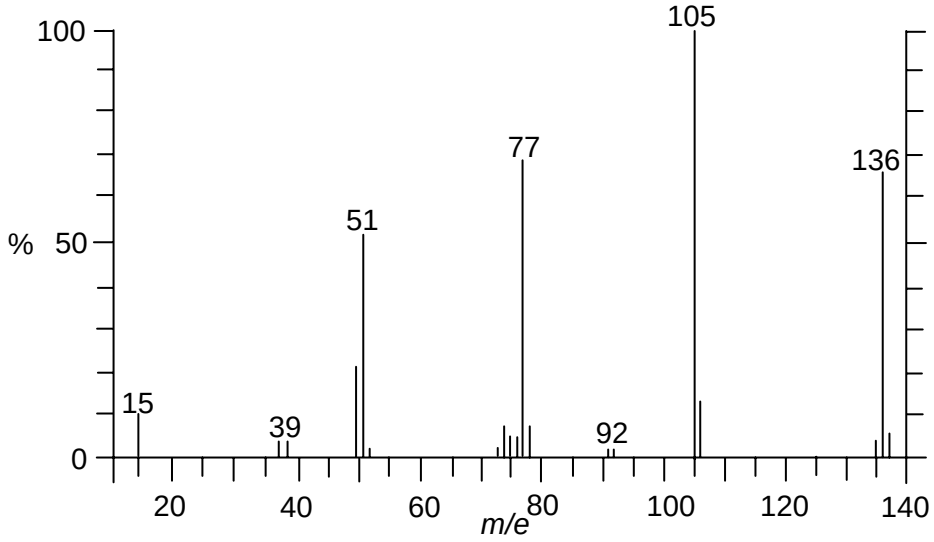


Figure 2.4 – mass spectrum of compound B

The M:M+1 ratio of 11.5:1 suggests that a molecule of compound B contains 8 carbon atoms, and an accurate determination of its relative atomic mass suggests its molecular formula is $C_8H_8O_2$. The high carbon:hydrogen ratio suggests a compound containing an aryl ring. There are, however, many isomers with this formula: phenyl ethanoate, methyl benzoate and methyl benzoic acid are just three examples.

The base peak (= most stable fragment) in the spectrum is at *m/e* = 105. The other significant peaks are at *m/e* 15, 51 and 77. We can analyse these peaks by either taking their M_r values and guessing at their molecular formulae, or by calculating the molecular formula of the fragment(s) that have been lost from the molecular ion at *m/e* = 136)

Thus: the peak at *m/e* = 105 represents the loss of 136-105 = 31

$$\underset{136}{C_8H_8O_2} - \underset{31}{CH_3O} \longrightarrow \underset{105}{C_7H_5O}$$

and the peak at *m/e* = 77 represents the loss of 136-77 = 59

$$\underset{136}{C_8H_8O_2} - \underset{59}{C_2H_3O_2} \longrightarrow \underset{77}{C_6H_5}$$

The peak at *m/e* 105 (28 units higher than C_6H_5 at *m/e* 77) can be identified with $C_6H_5CO^+$ (acyl cations are particularly stable, so are often the base peaks in mass spectra). This suggests that compound B is methyl benzoate:

$$\begin{array}{l} C_6H_5\text{——}CO\text{——}OCH_3 \\ \lfloor\ \text{--77--}\rfloor\ \lfloor\ \text{------59------}\rfloor \\ \lfloor\ \text{-----105-----}\rfloor\ \lfloor\ \text{---31--}\rfloor \end{array}$$

Just as important as the assigning of structures to the fragments of the molecule is the awareness of the *absence* of fragments that might have been expected in the mass spectrum of alternative structures. Thus, bearing in mind the stability of acyl cations mentioned above, we might have expected phenyl ethanoate to have produced a fragment at *m/e* 43,

$$\underset{m/e = 136}{CH_3\text{–}CO\text{–}O\text{–}C_6H_5^{+}} - OC_6H_5 \longrightarrow \underset{m/e = 43}{CH_3\text{–}CO^{+}}$$

and methylbenzoic acid to have produced a fragment at *m/e* 119.

$$\underset{m/e = 136}{CH_3–C_6H_4–COOH^{\cdot+}} \; – \; OH \longrightarrow \underset{m/e = 119}{CH_3–C_6H_4CO^{+}}$$

The absence of both these peaks confirms that B cannot be either of these compounds.

2.2b Nuclear Magnetic Resonance (NMR) spectroscopy

Magnetic properties of the nucleus

The nuclei of hydrogen atoms spin about an axis. Because the nuclei are positively charged, this spin is associated with a circulation of electric charge. Circulating charges give rise to magnetic fields, so the spinning ^1H nucleus has a magnetic moment, rather like the magnet of a compass needle. When put in an *external* magnetic field the nuclei tend to turn to a preferred orientation in which the nuclear magnet is aligned **with** the external field (like a compass needle does in the Earth's magnetic field). Another, less favoured, orientation is when the nuclear magnet is aligned **against** (i.e. opposed to) the external field. Nuclei obey the laws of quantum mechanics, and according to quantum laws these are the only two orientations allowed for nuclei such as ^1H, which have a nuclear spin of half a unit. This situation is illustrated in Figure 2.5.

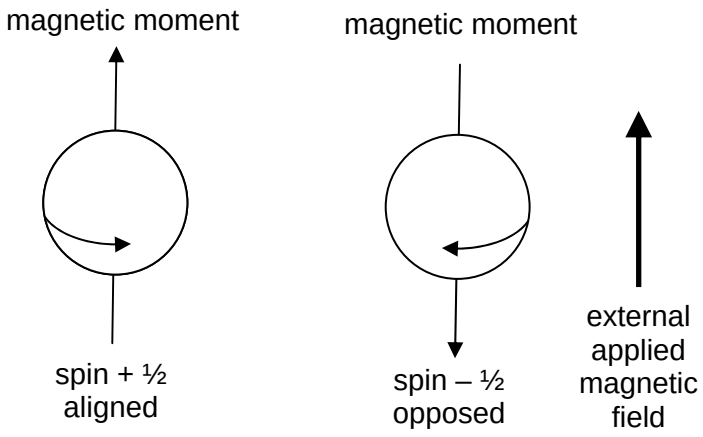


Figure 2.5 - nuclear spin

In the absence of an applied field, the spin states of a given nucleus are of equal energy. As can be seen in Figure 2.6, however, the spin states in an external magnetic field are no longer of equivalent energy and the two spin states occupy two different energy levels.

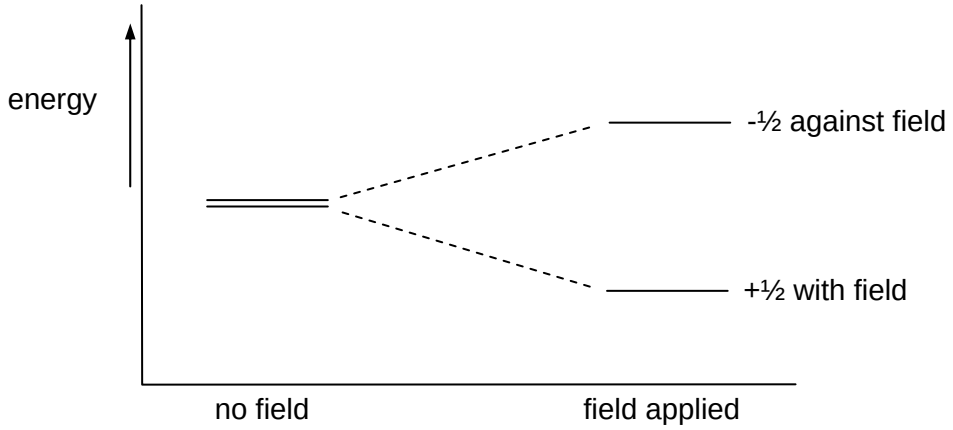


Figure 2.6 – spin states in an external magnetic field

The phenomenon of nuclear magnetic resonance occurs when nuclei aligned with the [illegible]d absorb energy (ΔE in Figure 2.6) and change their spin orientation with respect to the [illegible] energy difference between the two spin states is related to the frequency of electromagnetic ra[illegible] by the Planck equation:

$$\text{energy absorbed} = \Delta E = hf$$

where f = frequency of radiation, and h = Planck's constant.

The energy gap separating the two levels increases as the strength of the external field increases. The strength of the magnetic field is measured in teslas, T (one tesla is about 10 000 times as strong as the Earth's magnetic field). For a typical field of 9.4 Tesla, T, $\Delta E = 3 \times 10^{-25}$ J per nucleus, and the frequency of radiation absorbed is about 400 MHz, which is in the UHF radio-frequency region of the spectrum. Because the energy gap, ΔE, is so small (only 0.2 J mol^{-1}, compared to the average thermal energy of 1300 J mol^{-1} at room temperature) the two spin states are approximately equally populated, the difference in population being only 1 part in 30 000. That is, for every 30 000 nuclei in the lower energy spin state, there are 29 999 nuclei in the upper spin state. The consequence of this is that the net absorption of radiation is very small, and so a very sensitive detection system is required. In addition, as we shall see later, the frequencies at which different protons within molecules absorb energy differ only very slightly from each other, so the detection system has to be able to measure frequencies to an accuracy of 1 part in 100 million, and the electromagnet has to provide a magnetic field of comparable stability.

The NMR spectrometer

The basic elements of a typical NMR spectrometer are shown in Figure 2.7. About 20 mg of the sample is dissolved in about 0.5 cm^3 of a solvent that contains no protons, e.g. tetrachloromethane, CCl_4, or deuterochloroform, $CDCl_3$, or heavy water, D_2O, and a drop of tetramethylsilane (TMS), $(CH_3)_4Si$, is added to act as an internal reference. (TMS is used as a reference because it is volatile, inert and produces a strong singlet peak at a higher field that most ^1H absorptions in organic molecules, so its peak does not interfere with the other peaks.) The solution is placed in a precision glass tube of 0.5 cm diameter, to a depth of 2-3 cm. The sample tube is then suspended between the poles of the permanent magnet. The effect of slight variations in the magnetic field is minimised by spinning the sample at 20-30 revolutions per minute. Also in the gap between the pole pieces of the magnet is a coil attached to a radio-frequency (UHF) oscillator which supplies the energy for the nuclear spins to "flip". At right angles to this energising or oscillator coil, there is a second coil coupled to an amplifier and recorder. It is this second coil which acts as the detector. When no energy is being absorbed by the sample, the detector coil picks up none of the energy from the oscillator coil. When the sample absorbs energy, however, the re-orientation of nuclear spins induces a signal in the detector coil; this signal is recorded as a resonance signal. Most modern NMR spectrometers are Fourier Transform machines, which are designed to operate with a fixed external field and a range of frequencies.

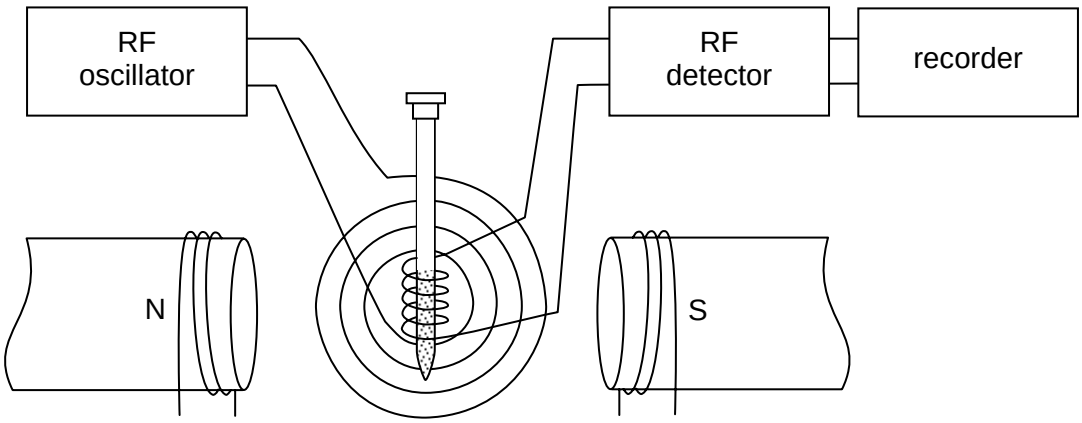


Figure 2.7 – the NMR spectrometer

The ^{1}H NMR spectra of organic compounds

When the frequency of radiation supplied corresponds to the energy difference ΔE, the population of the higher energy state increases as radiation is absorbed. The equilibrium population ratio is re-established by redistributing the energy amongst other energy states in the molecule (e.g. the translational movement of molecules). This takes a fraction of a second.

Not all the ^{1}H atoms within a molecule absorb energy at the same frequency. This is because the frequency at which a proton absorbs radiation depends on the strength of the *local* magnetic field around it. Even in a *constant external* field, protons in different chemical environments within a molecule absorb at different frequencies, because the local magnetic field they experience depends on the electrical and magnetic environment around them.

The electrons within molecules are usually 'paired' (i.e. they occur as pairs of electrons spinning in opposite directions). When a molecule is placed in an external field, the electron pairs rotate in their orbits in such a way that they produce a magnetic field which opposes the external field. This phenomenon is called *diamagnetism*. The effect is to 'shield' nearby protons from the external field. This in turn reduces the frequency at which they absorb energy when they flip back from their lower to their higher energy state (see Figure 2.8).

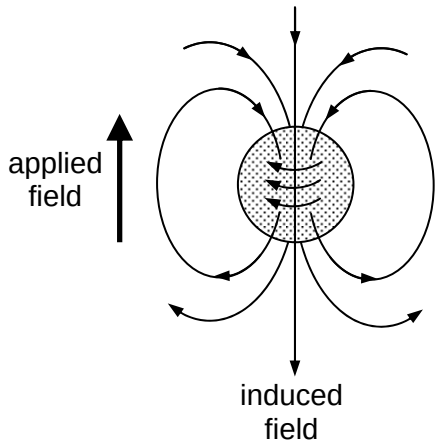


Figure 2.8 – shielding

When, however, a proton is near an electronegative atom within a molecule, the bonding electrons are drawn away from the proton to the electronegative atom. The proton is less shielded from the external magnetic field, and hence it absorbs radiation at a higher frequency. The effect is very pronounced if the proton is attached to a benzene ring. In this situation the mobile delocalised π electrons in the ring can create a strong diamagnetic effect, opposing the external field. This has the effect of *strengthening* the magnetic field within the vicinity of the protons (see Figure 2.9).

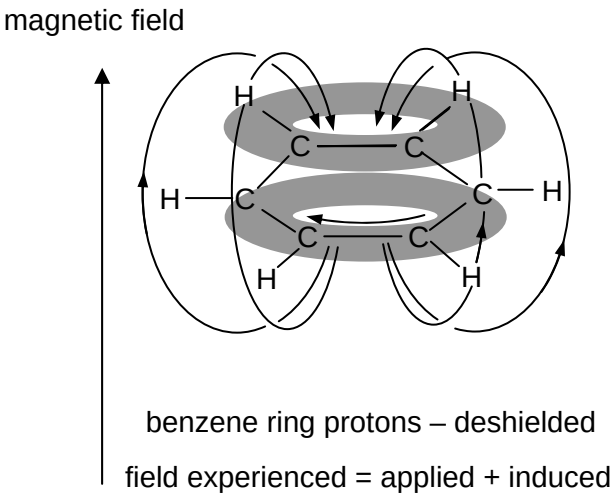


Figure 2.9 – deshielding

The ^{1}H NMR spectrum of ethanol

An NMR spectrum consists of a graph of absorbance against frequency. The frequency scale is measured in units of **chemical shift** (symbol δ). The chemical shift of a proton is the difference between its absorption frequency and that of TMS, measured in parts per million (ppm). Most protons in organic molecules resonate within 10 ppm of TMS, and by convention, the zero point of the scale (TMS) is on the right hand side (Figure 2.10).

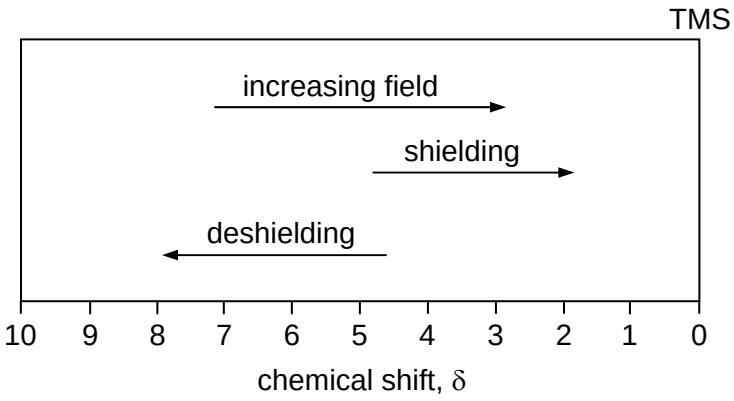


Figure 2.10 – chemical shift

The low resolution NMR spectrum of ethanol contains three different absorption peaks occurring at slightly different values of the applied field (Figure 2.11). The areas under these peaks are found to be in the ratio 1:2:3. These different peaks are explained by assuming that the single –OH proton, the two $–CH_2–$ protons and the three $–CH_3$ protons all resonate at slightly different values of applied field: the nearer a proton is to the electronegative oxygen atom, the less shielded it is from the external magnetic field. This is because the oxygen atom attracts the bonding electrons, and their diamagnetic effect on the resonating proton is reduced.

To a certain extent the changes in chemical shifts caused by adjacent electronegative atoms or aryl rings are additive, and transferable from one type of proton to another. To illustrate this, take a look at the table of proton chemical shift values in the *Data Booklet*. Alkyl CH_2 protons (δ 1.3) resonate at a lower field than alkyl CH_3 protons (δ 0.9) by about 0.4 δ units. The table shows that a proton in a CH_3 group on an aryl ring resonates at δ 2.3. We would therefore expect a CH_2 proton on an aryl ring to resonate about 0.4 δ units away from this, at δ 2.7. This is found to be the case. It is important to bear this in mind when interpreting spectra, especially if the type of proton thought to be present is not

listed precisely in the table in the *Data Booklet*. What the expected δ value should be [illegible] worked out, based on similar estimates to that used in this example.

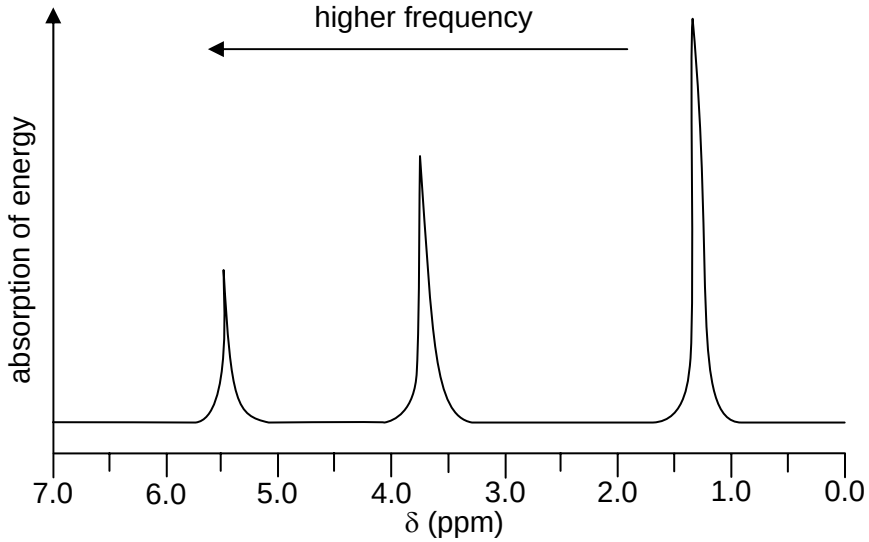


Figure 2.11 – low resolution NMR spectrum of ethanol

A high-resolution spectrum of ethanol (see Figure 2.12) shows the same three peaks, but this time the peak at δ 1.2 is split into three peaks, and that at δ 3.7 is split into four. This is due to a phenomenon known as **spin-spin coupling**.

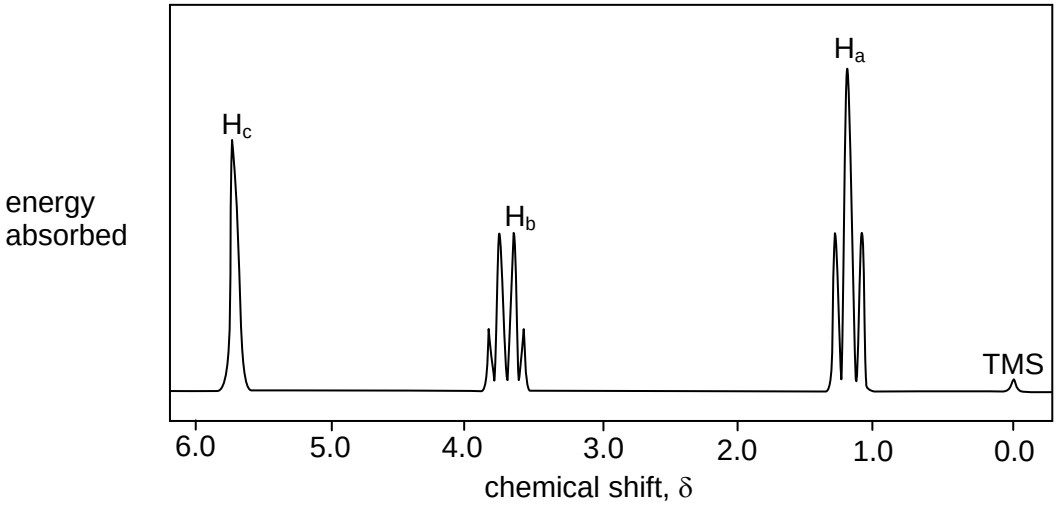


Figure 2.12 – high resolution NMR spectrum of ethanol

The local magnetic field experienced by a resonating proton is not only influenced by the presence of the electrons in the bonds around it, but is also influenced by the orientation (with respect to the external field) of the magnetic moments of nearby protons. If the magnetic moment of a nearby proton is aligned **with** the external field, the local magnetic field experienced by the resonating proton will be increased. The opposite will occur if the magnetic moment of the nearby proton is aligned **against** the external field. The resonating proton will therefore absorb energy at two different frequencies. Because the populations of the two energy levels (+½ and –½) are almost equal, the absorbances at the two frequencies of the resonating proton will be equal, and a doublet of lines will be seen. The situation is a little more complicated if there are **two** nearby protons. The magnetic moment of each proton can be either with or against the field, and the possible combinations are shown in Figure 2.13.

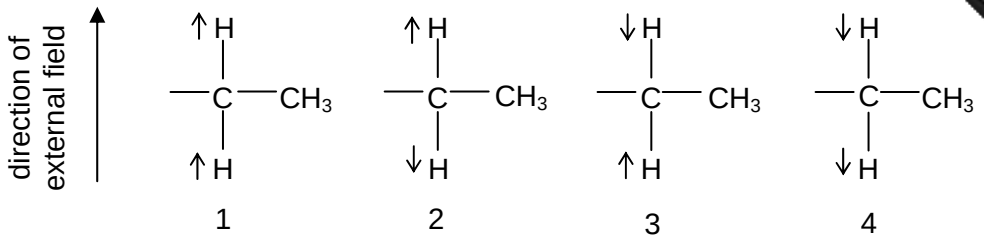


Figure 2.13 – combinations of magnetic moment of two nearby protons

In situations 2 and 3 the magnetic moments of the two $–CH_2–$ protons cancel each other out, so the field experienced by the $–CH_3$ protons will not change. In situations 1 and 4, however, the magnetic moments of the $–CH_2–$ protons reinforce each other, and so the field experienced by the $–CH_3$ protons will be higher or lower. We would therefore expect the $–CH_3$ protons to absorb at three different frequencies. Because the probabilities of the four states 1 to 4 are all equal, there is twice the chance of the $–CH_3$ protons experiencing **no** change in field (states 2 and 3) as there is for the protons to experience either an enhanced or reduced field (states 1 or 4). We therefore expect the absorbance at the central (unchanged field) frequency to be twice as great as the other two, forming a 1:2:1 triplet of lines. This is exactly what is observed for the absorption centred on δ 1.2.

We can work out the splitting pattern of the peak at δ 3.7 in a similar way: the magnetic moments of the three protons in the $–CH_3$ group can be oriented in the following combinations:

	↑↑↑	↑↑↓ ↑↓↑ ↓↑↑	↑↓↓ ↓↑↓ ↓↓↑	↓↓↓
net moment:	+3	+1	–1	–3
number of combinations:	1	3	3	1

We can now see why the peak at δ 3.7 is split into a 1:3:3:1 quartet.

The general rules governing splitting patterns are as follows:

1. Protons in identical chemical environments (e.g. the three H atoms in the $–CH_3$ group) do not split their own absorption peak.
2. The absorption peak of resonating protons is only split by nearby protons on the **adjacent** carbon atom – protons further away (usually) have little effect
3. The absorption of protons adjacent to *n* protons is split into (*n* + 1) peaks.
4. The intensities of the peaks in a multiplet are as in the following table.

number of protons adjacent to the resonating proton	number of lines in multiplet	relative intensities of lines
1	2	1:1
2	3	1:2:1
3	4	1:3:3:1
4	5	1:4:6:4:1

SAQ 5. ***(a)*** *In the molecule $CH^A{}_3–CH^B{}_2–CH^CCl–CH^DCl_2$, which protons would be split by the H^B protons?*

(b) *Which would be split by the H^C proton?*

(c) *Predict the splitting pattern you would observe for (i) the H^C proton, and (ii) the H^B protons.*

The peak in Figure 2.12 due to the –OH proton is a singlet, whereas it may have been expected to be split by the adjacent $–CH_2–$ protons into a triplet. Likewise, it might have been expected that the $–CH_2–$ protons would appear as a quintet, split by the three $–CH_3$ protons and also by the –OH proton. This splitting does not occur because protons on –OH groups undergo rapid exchange with each other, and with protons on other –OH groups such as those in water. This means that adjacent

protons only experience an averaged-out field of all the exchanging protons. (The –OH [illegible] **ultra-dry** ethanol does in fact appear as a triplet).

$$CH_3CH_2O\boldsymbol{H} + H_2O \rightleftharpoons CH_3CH_2OH + HO\boldsymbol{H}$$

A useful application of this ready exchange of –OH protons is the disappearance of their absorption peak when an NMR sample is shaken with D_2O (D is deuterium, ^2H). The deuterium nucleus has a nuclear spin of ± 1 rather than $\pm ½$, so does not absorb in the same frequency range as ^1H.

$$CH_3CH_2OH + D_2O \rightleftharpoons CH_3CH_2OD + HOD$$

One further point can be seen from Figure 2.11. The area under each peak is proportional to the number of protons responsible for that absorption. Thus from the left hand side, the areas of the peaks are in the ratio 1:2:3. This is also true for the high-resolution spectrum in Figure 2.12, but in this case the sum of the areas underneath *all* the split peaks in a group is proportional to the number of protons. This information is usually presented on the spectrum as the indication *n*H printed near the peak, where *n* is the number of protons (see Figure 2.14).

Some examples of ^1H NMR spectra

To illustrate the application of these principles let us look at the NMR spectrum of two unknown compounds X and Y, and see if we can use the spectra to identify X and Y. We shall refer to the *Data Booklet* for δ values. Students who use these spectra for revision should try to identify them by themselves, and then look at the analysis of each spectrum given below.

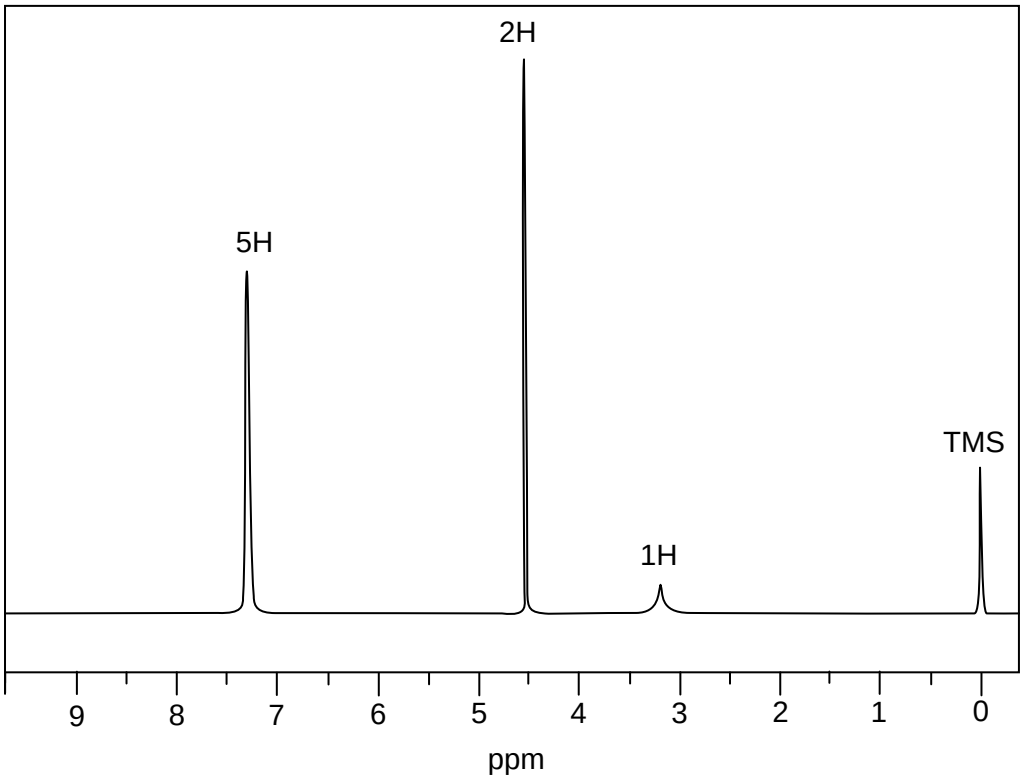


Figure 2.14 – ^1H NMR spectrum of X

Further information on X: $M_r = 108$
the 1H peak disappears on shaking with D_2O

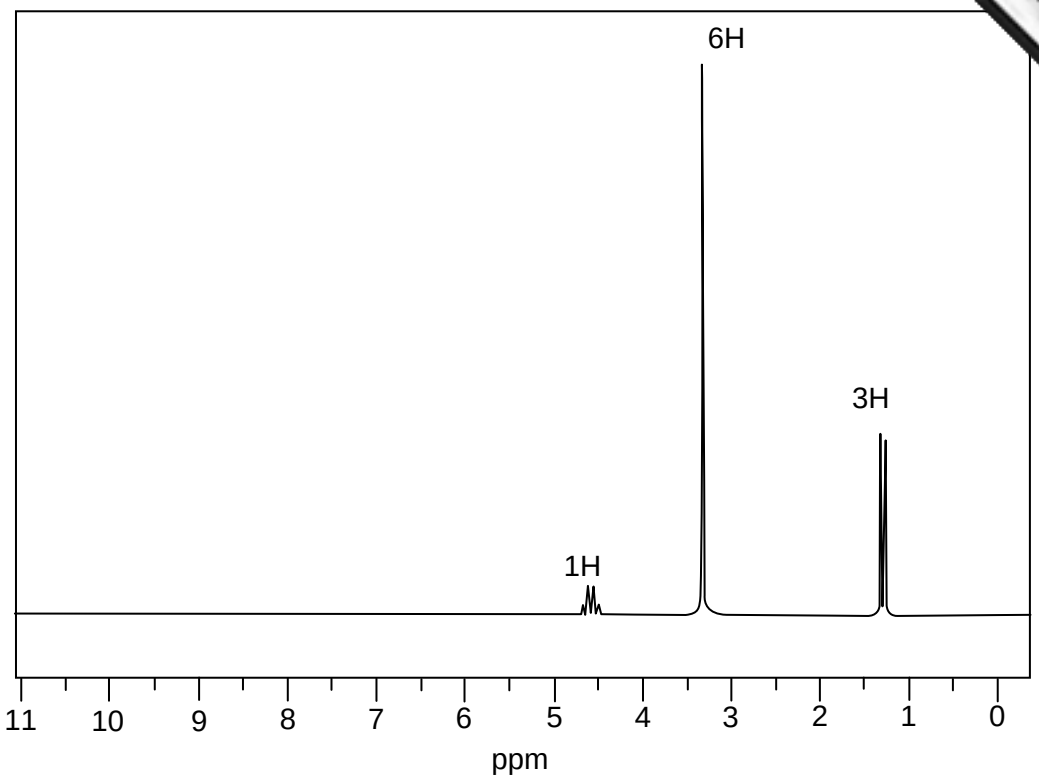


Figure 2.15 – ^{1}H NMR spectrum of Y

Further information on Y: M_r = 90
no change in the spectrum on shaking with D_2O

Analysis of the data

Compound X

- None of the three peaks in the spectrum (don't forget to ignore the TMS peak, no matter how high it is) is split into a multiplet, so compound X does not contain any non-equivalent protons on adjacent carbon atoms.
- The 1H peak at δ 3.2 (identical to *Data Booklet* value), which disappears in shaking with D_2O, is likely to be due to an –OH (alcohol) group (as can be seen from the *Data Booklet*, the δ value for an OH proton is very variable – anywhere from δ 0.5 – 6.0).
- The 5H peak at δ 7.3 is in the aromatic C–H region of the spectrum (*Data Booklet* value δ 6.0 – 9.0). Its integration of 5 protons suggests an aryl ring with just one group attached to it.
- The 2H peak at δ 4.5 is consistent with a CH_2 group joined to an oxygen atom, brought downfield by its proximity to an aromatic ring. (*Data Booklet* for a CH_3 group next to oxygen is δ 3.3 – 4.0: CH_2 instead of CH_3 will shift this by about 0.4 δ, and the proximity to the aryl ring will shift it downfield by a further 1.0 δ)

Compound Y

- The 3H doublet at δ 1.3 is in the region of the spectrum we expect to see C–CH_3 protons. (The slightly higher δ value compared to the *Data Booklet* value of 0.9 could be due to the proximity of an aromatic ring or an electronegative atom such as oxygen). The fact that this peak is **split into a doublet** suggests it is adjacent to a carbon atom with one hydrogen on it.
- The fact that there are **six** protons responsible for the singlet at δ 3.3 suggests two equivalent CH_3 groups. The δ value is near that of an –O–CH_3 group (*Data Booklet* 3.3-4.0).
- The 1H peak at δ 4.5 is **split into a quartet** due to its proximity to the C–CH_3 protons mentioned above (if one peak in a spectrum is split into a multiplet, another split peak is

bound to be seen also). Although its δ value is in the wide range expected for an [illegible] this cannot be the case because D_2O has no effect on the spectrum. It is, consistent with a C–H proton adjacent to two electronegative oxygen atoms (*Data B* the change from R–CH_3 (δ 0.9) to R_3C–H (δ ~1.6) increases the δ value by about –O–CH_3 is between δ 3.3 – 4.0, so –O–CH– should be between 4.0 – 4.7; this proton is the more-deshielded end of this range due to being joined to **two** electronegative atoms.

Answers:	Compound X is phenylmethanol, C_6H_5–CH_2–OH Compound Y is 1,1-dimethoxyethane, $(CH_3O)_2CH$–CH_3

Magnetic Resonance Imaging (MRI)

The application of NMR in medicine is becoming increasingly commonplace. The technique most commonly used is magnetic resonance imaging. A patient is placed on a steady table and the part of the body to be scanned (often the head) is immobilised. (Occasionally the patient is anaesthetised.) The table is slowly passed through the magnetic field, where a fine beam of UHF radiation is passed through and analysed. This shows the ^1H absorption pattern of a thin section (about 1 cm thick) of the body. The table continues to move gently into the magnetic field, and after each 1 cm another absorption pattern is recorded. A typical brain scan containing 20-30 slices takes only 10 minutes to complete.

Most of the hydrogen atoms within cells are in water molecules or lipid molecules, and it is the environment of the water molecules that gives an indication of the medical state of the cell. MRI is useful in pinpointing brain tumours and sites of injury, and diagnosing hydrocephalus, multiple sclerosis, Alzheimer's and other brain diseases.

2.2c X-ray crystallography

X-rays are very short-wavelength electromagnetic rays. Their wavelength of about 0.1 nm is comparable to the interatomic distances in solids. If a beam of monochromatic X-rays (i.e. rays of a single wavelength) is passed through a crystal, some X-rays will be diffracted by the electrons that surround the atoms in the crystal. (The electron clouds interact with the electric field of the X-rays.) The single incoming beam produces many diffracted beams. The angle between the incoming beam and each diffracted beam depends on the geometry of the crystal lattice, and hence on the arrangement of the atoms and their distances apart.

All atoms except hydrogen contain enough electrons to diffract the X-rays, and the 'heavier' the atom (i.e. the greater the number of electrons around it), the more intense is the reinforcement of the beam.

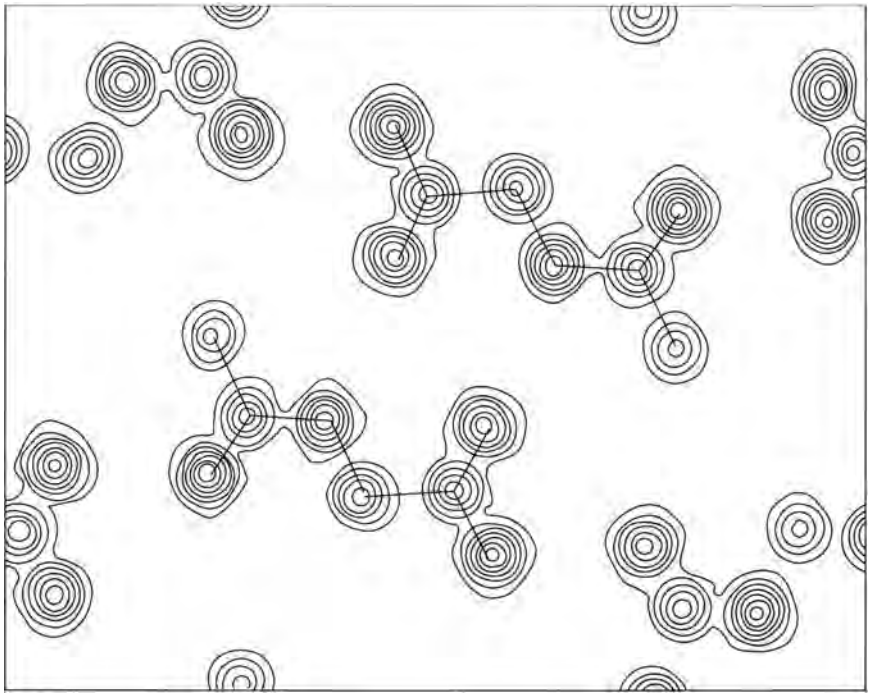
Thus, by measuring the angles between the incoming beam and the diffracted beams, and the relative intensities of those beams, we can piece together a picture of the electron density at all points in the unit cell. This process requires an enormous number of calculations to be done, requiring many hours of high speed computer time. It also relies on some good luck and informed judgement on the part of the X-ray crystallographer. The technique is often to guess at a structure, and get the computer to predict the diffraction pattern that would be formed by that structure. This prediction is then compared to the observed pattern. Repeated modifications of the suggested structure eventually produce a diffraction pattern that is identical to the one observed. With molecules containing many hundreds of atoms (e.g. a protein), the process can be quite time consuming. Often, though, the limiting factor is the production of a sufficiently pure crystal in the first place.

The structures of simpler molecules such as ethanoylaminoethanoic acid (see Figure 2.19) can be worked out much more quickly. The stages involved were as follows:

1. The product was synthesised and purified by recrystallisation from water.
2. Batches of crystals were grown from aqueous solution.
3. An X-ray diffraction pattern was collected. The unit cell size was estimated. The spots were indexed and the lengths of the three sides of the unit cell were measured.
4. The density of the substance was measured which enabled the number of molecules per unit cell to be calculated.

5. The reflections from one set of planes were found to be exceptionally intense, which [illegible]d that all the atoms in the molecule were concentrated in those planes.
6. The perpendicular distance between the planes was calculated to be 3.24×10^{-10}m. Th[illegible] distance suggested that the molecules within the layers had a planar arrangement.
7. Trial structures were suggested based on the known chemistry of ethanoylaminoethanoic acid.
8. The electron density map was interpreted and related to the arrangement of atoms within the molecules.

The result is shown in Figure 2.19.



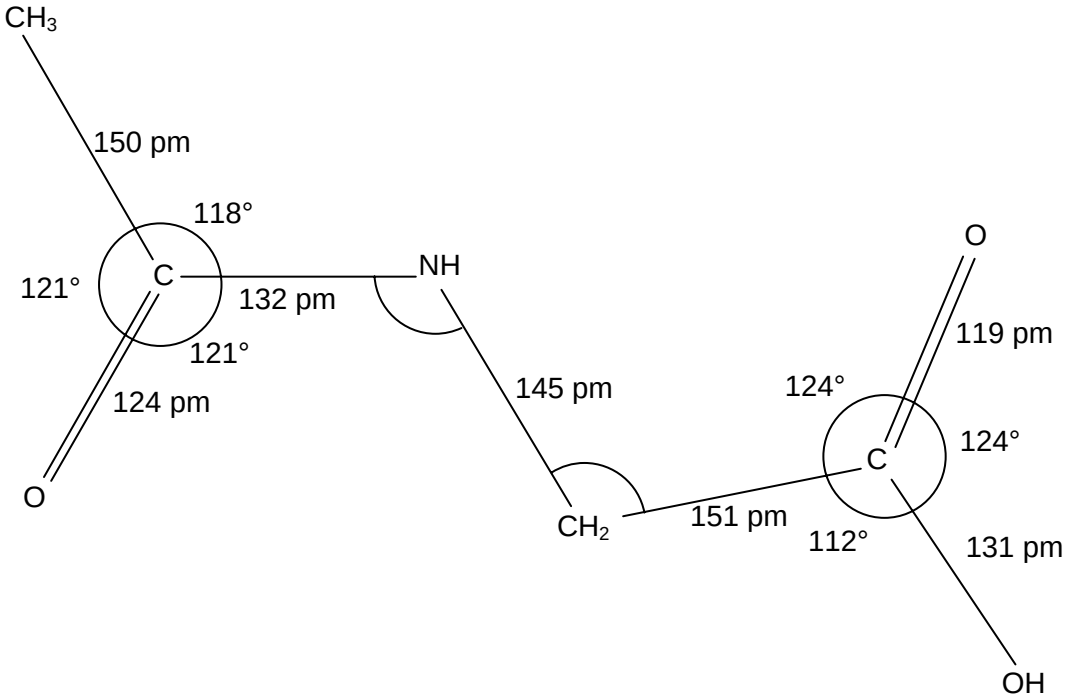


Figure 2.19 – electron density map of N-ethanoylaminoethanoic acid

SAQ 6. *Work out which atoms are at the centres of the electron density contours in Figu[illegible] Using a protractor, estimate the C–N–C and the N–C–O bond angles.*

The role of X-ray crystallography in the discovery of the structure of DNA by Watson and Cr[illegible] 1953 is well known. The DNA molecule has a good deal of symmetry, with the planar base-pa[illegible] taking up positions parallel to each other. The relatively simple X-ray diffraction photograph of a hydrated DNA fibre (Figure 2.20) shows the central spots arranged in a cross pattern – typical of a helical structure. Key dimensions obtained by X-ray crystallography are shown in Figure 2.21.

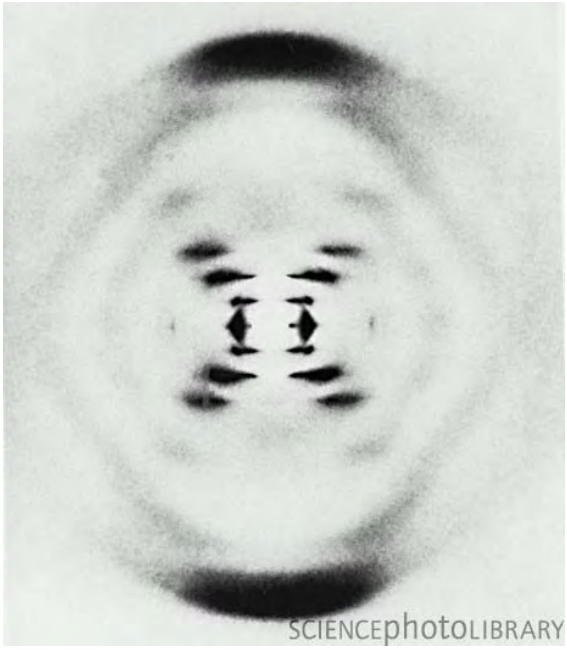


Figure 2.20 – X-ray diffraction photograph of a hydrated DNA fibre

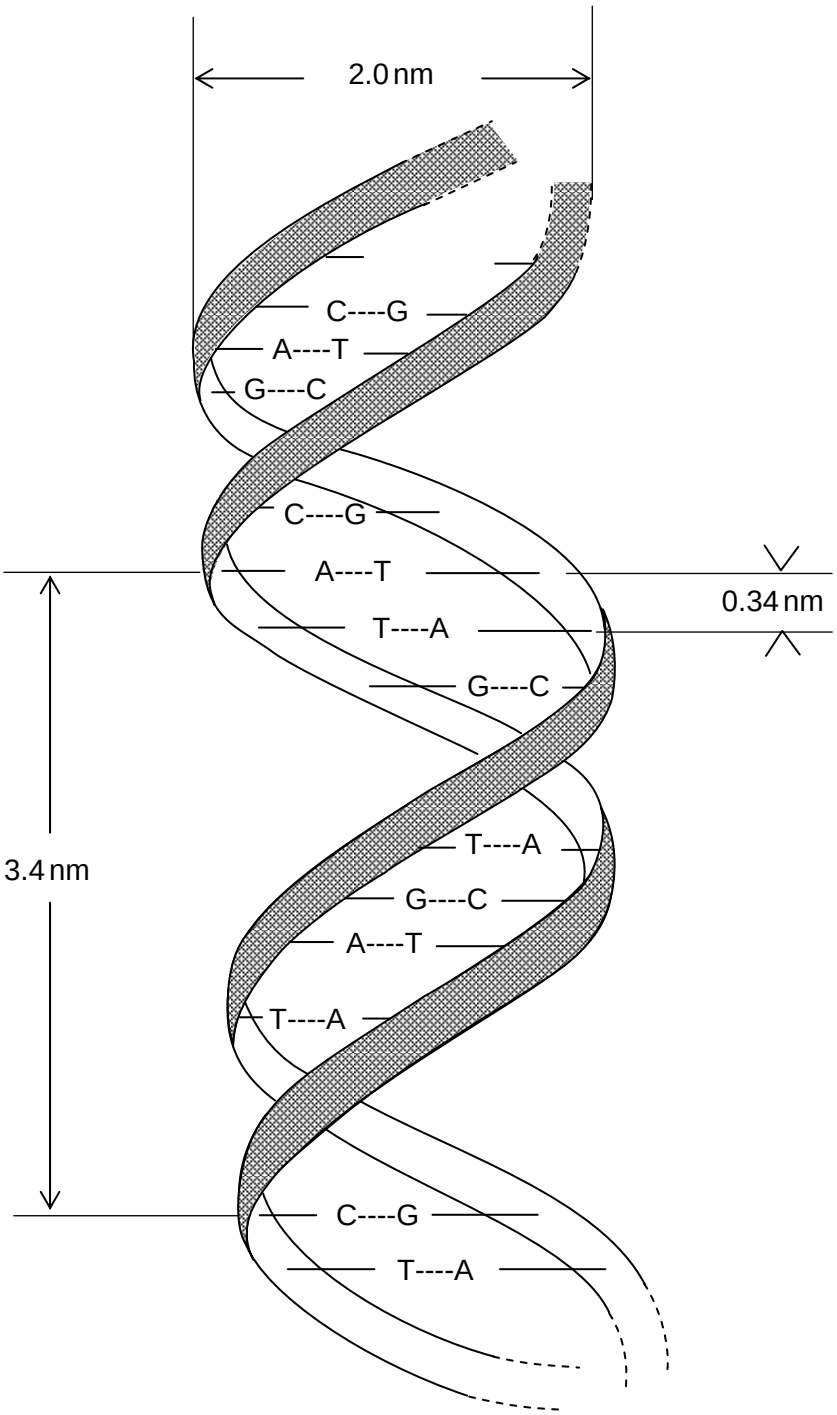


Figure 2.21 – key dimensions obtained for DNA by X-ray crystallography

Case study: The structure and function of myoglobin

Myoglobin is an oxygen transporting agent and oxygen reservoir found in the muscular tissues of vertebrates. It achieves its function by the use of a haem group. This consists of a Fe^{2+} ion surrounded by a protoporphyrin ring. The iron is bound to the polypeptide chain by an Fe–N covalent bond to a histidine residue. The oxygen molecule is transported by forming a dative bond to the Fe^{2+} ion (Figure 2.22).

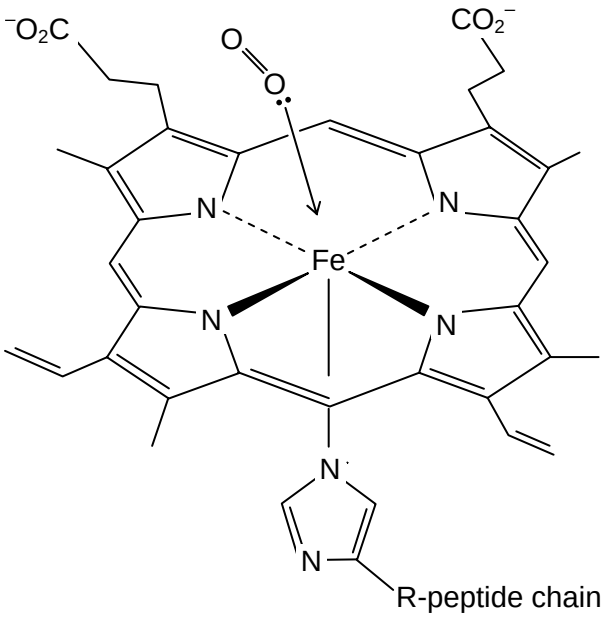


Figure 2.22 – iron protoporphyrin with O_2 and histidine

One problem that has fascinated chemists for some time is why, in the presence of oxygen, the Fe^{2+} ion does **not** become oxidised to Fe^{3+}. If iron(II) protoporphyrin is dissolved in water, and oxygen is bubbled through the solution, an almost immediate oxidation to Fe^{3+} occurs. A key intermediate in this oxidation is a complex of an oxygen molecule sandwiched between two haem groups (Figure 2.23).

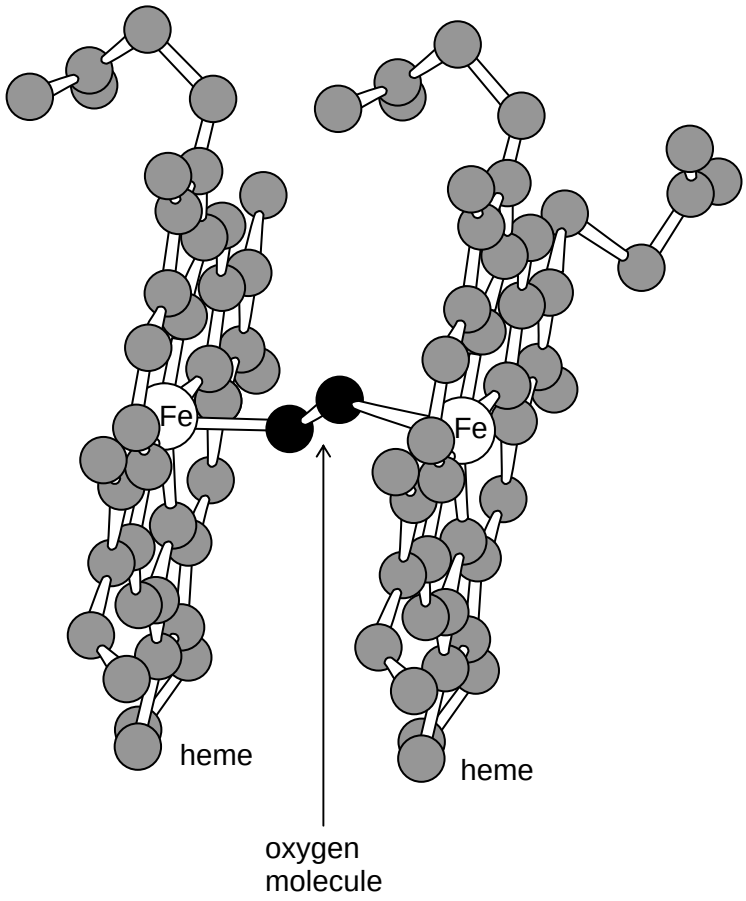


Figure 2.23 – two iron protoporphyrins with O_2 sandwiched in between

Like many iron complexes, haem has a large affinity for carbon monoxide. This displac[illegible] by CO is why carbon monoxide is so poisonous. But another curious property of the haem[illegible] both myoglobin and its blood counterpart, haemoglobin, is that their affinities for carbon mo[illegible] although large, are very much smaller than that of an isolated haem group in solution.

Both these observations were readily explainable once the three-dimensional structure of myoglobi[illegible] had been determined.

This was achieved by John Kendrew and his colleagues between 1957 and 1962. It was the first protein to have its structure determined by X-ray crystallography. Kendrew chose myoglobin for two important reasons:

- it is a comparatively small protein ($M_r \approx 18\,000$) which has the advantage of containing a heavier (iron) atom. Since this is surrounded by a large number of electrons (26) it should diffract X-rays well,
- it readily forms good quality crystals, and is available in quantity.

After several years of intense work and many hundreds of hours of calculation, involving the measurement and analysis of the intensities of 25 000 spots in the X-ray photograph (see Figure 2.24), Kendrew succeeded in working out its structure (Figure 2.25). As can be seen, the myoglobin molecule is extremely compact, with a large percentage of its amino acids joined in the α-helical configuration. This produced several stable, stiff “rods”, which are joined by small lengths of more flexible parts of the amino acid chain.

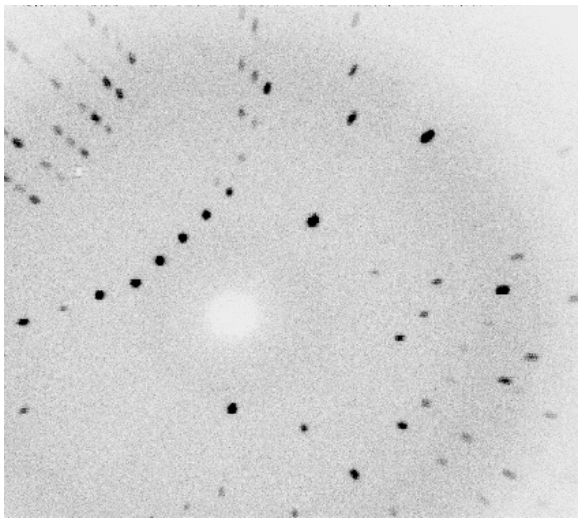


Figure 2.24 – X-ray diffraction photograph of myoglobin

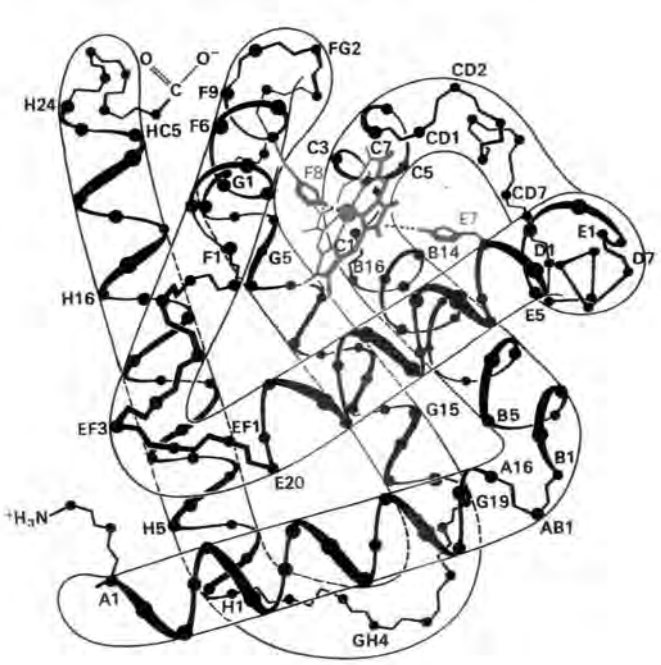


Figure 2.25 – structure of myoglobin molecule

In particular, the oxygen-binding site is very hindered by the amino acids surrounding it, and the haem group is on the inside of the molecule. Furthermore, there is a second histidine group positioned just over (but slight to one side of) the 6th coordination position of the Fe^{2+} iron, where the O_2 molecule sits.

This structure now explains the two observations described above. The fact that the haem group is on the inside of the molecule makes it impossible for two such groups to come together with an oxygen molecule bridge between them, as in Figure 2.23. In addition, the second histidine group, above the Fe^{2+} ion, forces the complexing molecule to bond with the Fe^{2+} ion at an angle. This favours O_2, but hinders the bonding of CO (see Figure 2.26).

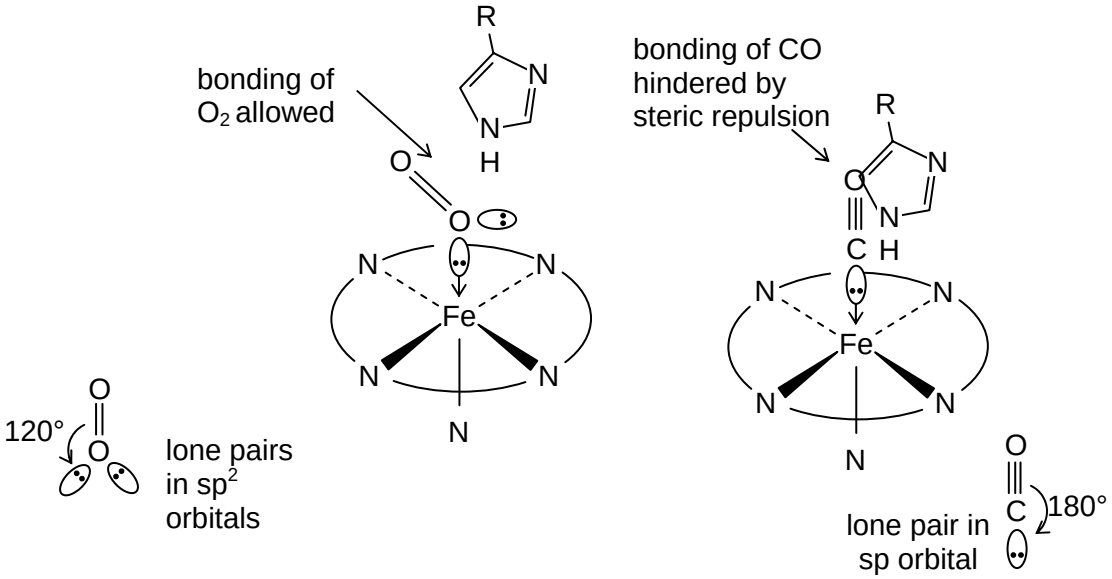


Figure 2.26 – arrangement of molecules within the myoglobin complex

Thus the specific nature of the function of myoglobin can be readily understood once its structure is known.

Summary

- Mass spectrometry, nuclear magnetic spectroscopy and X-ray crystallography a[illegible] methods of determining the structures of molecules.
- Mass spectrometry can tell us the number of carbon atoms in a molecule (from the M:M[illegible] ratio); the molecular formula (from an accurate determination of the M_r of the molecular ion) and something about its structure (from the fragmentation pattern).
- Nuclear magnetic spectroscopy tells us the number of hydrogen atoms in each chemical environment in a molecule (from the integration trace and the chemical shift (δ) values), and the number of their nearest neighbours (from the splitting patterns). If a particular type of proton has *n* nearest neighbours, its peak is split into (*n*+1) lines.
- X-ray crystallography tells us how the atoms are arranged in a crystal of the compound. It can determine the positions of all atoms except hydrogen.

2.2 What the student needs to know

11.2h A mass spectrometer allows us to measure the relative masses of molecules and molecular fragments, and their relative abundances. (Students do **not** need to know the details of how a mass spectrometer works.)
How to use the abundances of the M and M+1 peaks to calculate the number of carbon atoms in a molecule.
How to use the abundances of the M, M+2 and M+4 peaks to identify which halogen atoms are in a molecule, and how many there are.
The *m/e* values of molecular fragments can be used to piece together the structure of a complete molecule.

11.2d The principles of ^{1}H NMR spectra – i.e. proton 'magnets' spinning in opposite directions have different energies when placed in a strong external magnetic field; they absorb radio frequency electromagnetic radiation when they flip from one state ('direction') to the other; the frequency depends on the energy gap between the two states, which in turn depends on their chemical environment. (Students do **not** need to know how an NMR spectrometer works).
How to use the table of chemical shift values in the *Data Booklet* to identify the environment of a proton in a molecule.
How to use the 'integration' numbers showing the number of protons responsible for each peak.
The reason that lines in an NMR spectrum can be split into doublets, triplets, quartets and multiplets.
How to use these splitting patterns to work out the structure of a molecule
What the disappearance of NMR peaks on the addition of D_2O means.

11.2e NMR can be used to determine the structures of large protein molecules, and also the cellular environment of water molecules in live patients (MRI).
X-rays are diffracted by the electrons that surround atoms.
Light atoms such as hydrogen, which have few electrons around them, do not diffract X-rays.
X-ray crystallography can be used to determine the structures of complicated molecules. (Students do **not** need to know anything about the principles or the methods of X-ray crystallography.)

2.3 – Separating and identifying substances

By the end of this section, students should be able to:

- *understand what is meant by the term partition coefficient*
- *calculate a value for the partition coefficient, and use it in calculations of concentrations of solute in various solvents*
- *appreciate the use of solvent extraction in the extraction of plant products and the analysis of drug and pesticide residues*
- *outline the principles of paper chromatography (PC), thin layer chromatography (TLC), gas-liquid chromatography (GLC) and high performance liquid chromatography (HPLC)*
- *understand the meaning of the terms retention ratio (R_f) and retention time.*
- *outline the principles of electrophoresis*
- *appreciate the uses of the various forms of chromatography and of electrophoresis in the separation and analysis of proteins, drugs, pesticides, explosives, stimulants and other food additives.*

2.3a The partition of a solute between two immiscible solvents

We are all familiar with the most common liquid on Earth, water. Water is an excellent solvent for many substances. Some, like salt, sugar, and ammonia, dissolve well in water; whereas others, like oxygen or iodine, dissolve only partially; still others, like silica, benzene or helium, are virtually insoluble in water.

If we take another solvent, such as hexane, we find a different pattern of solubilities. Silica and helium are still insoluble, but so also are salt and sugar. Iodine and benzene dissolve well, but ammonia and oxygen are only slightly soluble.

It is worth digressing for a moment to look at the reasons why some substances dissolve and others do not. In general we can say that a substance dissolves because the total energy of the solute+solvent system is less when the solute is in solution than when it is not. This tends to be the case if the interactions between the solute molecules and those of the solvent are **similar** to those between the solvent molecules themselves.

In general, there are four different ways that molecules attract each other. These are:

- ion-ion attractions
- ion-dipole attractions
- van der Waals' (induced dipole) attractions
- hydrogen bonding

Although there are van der Waals' attractions between water molecules, these are small because of the small number of electrons in the H_2O molecule. The main force of attraction is due to hydrogen bonding. This explains why sugar ($C_{12}H_{22}O_{11}$) dissolves well – the 11 oxygen atoms in the sucrose molecule are all hydrogen bonded to water molecules, as are the eight hydrogen atoms on the OH groups. The H_2O molecule also possesses a dipole, which is attracted to cations and anions by ion-dipole interactions. This explains why salt (NaCl) dissolves in water. Ammonia dissolves well due to extensive hydrogen bonding between NH_3 and H_2O molecules.

The main interactions between adjacent hexane molecules are van der Waals' (induced dipole) forces. Molecules such as benzene and iodine, which also have induced dipole forces between them, dissolve well in hexane. But benzene and iodine are insoluble in water because the energy gain by forming van der Waals' attractions between their molecules and H_2O molecules does not exceed the energy penalty of breaking the hydrogen bonds between two adjacent H_2O molecules in the solvent. Likewise, sugar and salt are insoluble in hexane because there is no possibility of hexane either hydrogen bonding with sugar molecules, or undergoing ion-dipole interactions with the Na^+ and Cl^- ions in salt.

Solvents can, in general, be listed according to their **polarity**. Polar solvents are more [illegible]e able to dissolve ions, hydrogen bonded molecules and molecules with dipoles; non-polar so[illegible] more likely to dissolve those solutes whose molecules are only attracted to each other by [illegible] Waals' (induced dipole) attractions.

SAQ 7. ***(a)*** *Methanol, CH_3OH, is very soluble in water, whereas heptanol, $C_7H_{15}OH$, is almos[illegible] insoluble in water. Why is this?*

(b) *Predict, with a reason, which alcohol would be more soluble in hexane.*

If a solute such as iodine is soluble in two immiscible solvents (e.g. water and hexane), its solubility (usually measured in $g\,cm^{-3}$ or $mol\,dm^{-3}$) is very unlikely to be the same in both solvents. One solvent will be better at dissolving it than the other.

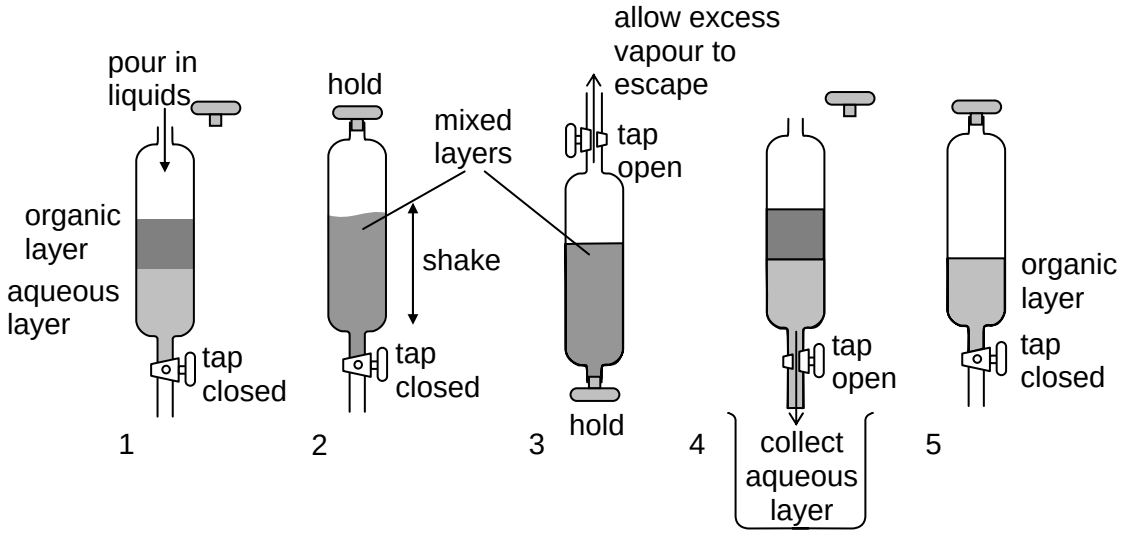


Figure 2.27 – solvent extraction

When some iodine crystals are shaken with a mixture of hexane and water until no further change takes place, and the two layers allowed to separate (see Figure 2.27), we find that the ratio of the concentrations of iodine in each layer is a constant, no matter how much iodine we start with. This constant is the equilibrium constant for the change:

$$I_2(aq) \rightleftharpoons I_2(hexane)$$

$$K_c = [I_2(hexane)]/(I_2(aq)]$$

K_c is called the **partition coefficient** of iodine between hexane and water. The usual symbol for a partition coefficient is K_{pc}. Like all equilibrium constants, the value of a partition coefficient changes with temperature. Notice that, in general partition coefficients have no units.

Partition coefficients are useful in working out how much solvent we need in order to extract a minimum amount of solute from one solvent into another. An example will make this clear.

Example 1:

Question:

When 100 cm^3 of an aqueous solution containing 2.0 g of the organic dye X was shaken with 20 [illegible] hexane, it was found that 1.6 g of the dye had been extracted into the hexane.

(i) Calculate the partition coefficient of X between hexane and water.

(ii) Calculate the minimum volume of hexane needed to reduce the amount of X in the aqueous layer to less than 0.1 g.

Answer:

(i) If 1.6 g of X had been extracted by the hexane, 0.4 g must have been left in the aqueous layer. The two concentrations are therefore:

In the hexane layer: 1.6/20 = **0.08 g cm^{-3}.** In the water layer: 0.4/100 = **0.004 g cm^{-3}**

K_{pc} = [X(hexane)]/[X(aq)] = 0.08/0.004 = **20**

(ii) If the mass left in the aqueous layer is to be less than 0.1 g, then the mass extracted must be at least 1.9 g. The corresponding concentrations are as follows:

in the hexane layer: $1.9/v$ g cm^{-3} in the water layer: 0.1/100 = **0.001 g cm^{-3}**

(where v = volume of hexane needed)

K_{pc} = [X(hexane)]/[X(aq)]

$\therefore$ 20 = $(1.9/v)/0.001$

v = $1.9/(20 \times 0.001)$ = **95 cm^3**

Successive extractions

Solvents are often expensive and flammable, and can also be polluting to the environment. It is sensible, therefore, to use the minimum amount of solvent that is needed to achieve the intended goal. Thus in the example above, the use of 20 cm^3 of hexane allowed 1.6 g of the dye X to be extracted from its aqueous solution, i.e. (1.6/2.0) × 100 = **80%.**

We can extract more than this if we use two separate 10 cm^3 portions of hexane, rather than one 20 cm^3 portion. The following calculation will make this clear.

Example 2:

1st extraction: Let us assume that x grams of X have been extracted by the first 10 cm^3. The equilibrium concentrations will therefore be:

in the hexane layer: $x/10$ g cm^{-3} in the water layer: $(2.0 - x)/100$ g cm^{-3}

$$K_{sp} = \frac{x/10}{(2.0-x)/100} = 20$$

$\therefore$ $x/10 = (40 - 20x)/100$

$\therefore$ $100x = 10(40 - 20x)$

$\therefore$ $100x = 400 - 200x$

$\therefore$ $\boldsymbol{x = 1.33}$ **g**

Thus the amount extracted = 1.33 g, so the amount remaining in the aqueous layer = 0.67 g

2nd extraction: We now separate the 10 cm^3 of hexane solution of X from the aqueous layer, and add another 10 cm^3 of hexane and shake again. Let us assume that the second 10 cm^3 will extract y grams of X, the equilibrium concentrations will be:

in the hexane layer: $y/10$ g cm^{-3} in the water layer: $(0.67 - x)/100$ g cm^{-3}

$$K_{sp} = \frac{y/10}{(0.67-y)/100} = 20$$

By a calculation similar to the one shown above, we find that

$$\boldsymbol{y = 0.44}\text{ g}$$

The total amount of X extracted by two successive 10 cm^3 portions of hexane is therefore 1.33 + 0.44 = **1.77 g**. This represents **89%** of the original 2.0 g amount of X in the 100 cm^3 of water.

Similar calculations show that if we had split the 20 cm^3 of hexane into four 5 cm^3 [illegible] amounts of X extracted at each stage would have been as follows:

1st extraction by 5 cm^3: 1.00 g
2nd extraction by 5 cm^3: 0.50 g
3rd extraction by 5 cm^3: 0.25 g
4th extraction by 5 cm^3: 0.125 g
total extracted = 1.875 g (94%)

All these results are collected together in the following table:

extraction method	percentage extracted
1 × 20 cm^3 of hexane	80%
2 × 10 cm^3 of hexane	89%
4 × 5 cm^3 of hexane	94%

As can be seen, the amount extracted steadily increases as we split the amount of extracting solvent into smaller and smaller portions. However, it is impossible to extract **all** of a solute, no matter how many portions of solvent we use, since it is never possible to move any equilibrium **completely** to one side or the other. But, if the solvent is a volatile one, and if the solute is involatile and stable to heat, it is possible to 'automate' the process by using a continuous extraction apparatus as shown in Figure 2.28.

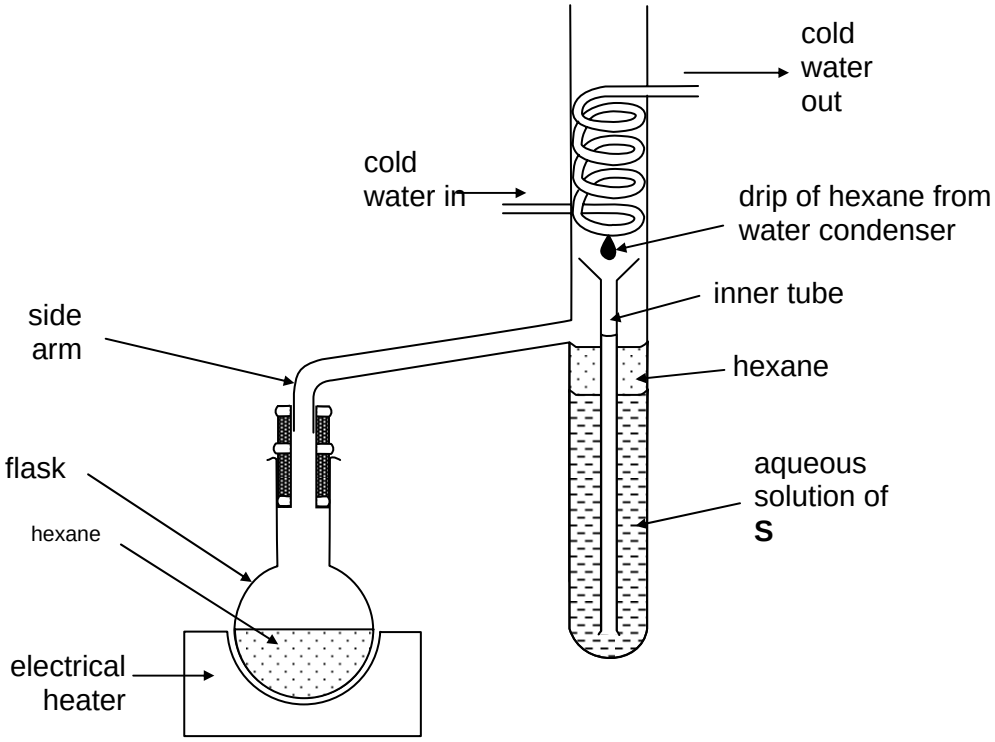


Figure 2.28 – continuous hexane extraction

SAQ 8. *Explain in a few sentences how the continuous extraction apparatus shown in Figure 2.28 works.*

Solvent extraction is used to extract perfumes and pharmaceutical precursors from plants, and in the analysis of insecticide residues and other pollutants in drinking water supplies, blood and milk. Various applications of solvent extraction are given throughout this chapter.

SAQ 9. *A solution of iodine in trichloroethane (TCE) was shaken with water. The iodine content of the two layers was determined by titration with aqueous sodium thiosulfate. A 25 cm^3 portion of the aqueous layer required 9.5 cm^3 of 0.02 mol dm^{-3} thiosulfate to reach the end point, whereas 5 cm^3 portion of the TCE layer required 17.5 cm^3 of 0.2 mol dm^{-3} thiosulfate. Calculate the K_{pc} for iodine between TCE and water.*

2.3b Chromatography

The name “chromatography” comes from two Greek words meaning “colour picture”, and sho[illegible] origin of the technique in the separation and analysis of coloured dyes and plant pigme[illegible] Nowadays, however, we do not have to rely on colour to detect the compounds we separate using th[illegible] various techniques of chromatography, since there are many other detection methods available. Four techniques will be described here:

- paper chromatography (PC),
- thin layer chromatography (TLC),
- gas/liquid chromatography (GLC),
- high performance liquid chromatography (HPLC).

The basic principles and techniques

All chromatographic methods use the same principle of a mobile phase (a liquid or a gas) moving past a stationary phase. The stationary phase may be a solid onto which the solutes are adsorbed (TLC, and some GLC and HPLC) or a liquid which is held in a thin film of the surface of an inert solid (PC, GLC and HPLC). The solute molecules partition themselves between the stationary phase and the moving phase. Unlike solvent extraction, however, the partition is not a true equilibrium, since the mobile phase is constantly moving past the area on the stationary phase where a particular solute is adsorbed/dissolved. Thus a fresh partition with pure mobile phase is continually being set up. The mobile phase should never become saturated with the solute.

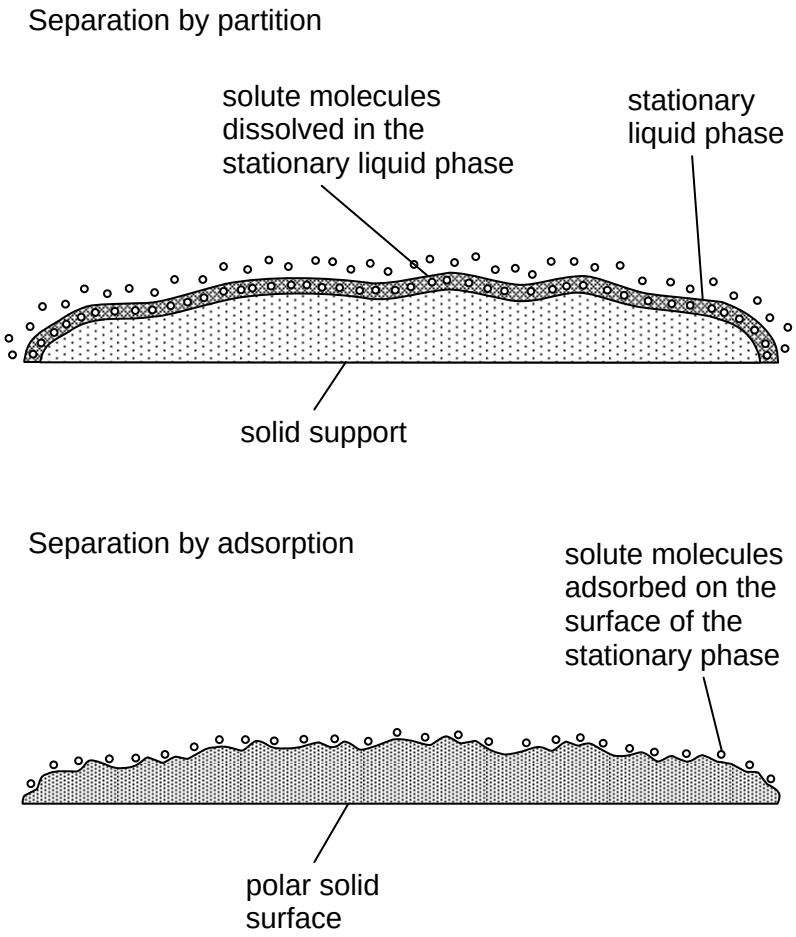


Figure 2.29 – the two types of chromatographic separation – liquid-liquid partition and adsorption

Paper chromatography

The mixture to be separated and analysed is dissolved in a solvent, such as water or ethanol, a small spot of the solution is placed about 1 cm from the edge of a rectangular sheet of chromatography paper (which is like filter paper, but is made to a more exact specification). Other spots of "reference" compounds may also be applied at the same distance from the edge. The solvent is allowed to evaporate. This allows the solutes to become adsorbed into the fibres of the paper (often by means of hydrogen bonding to the OH groups of the cellulose fibres, or more likely to the H_2O molecules that are still associated with cellulose, even when the paper feels dry).

The edge of the sheet is then immersed in the chromatography solvent (the moving phase), being careful at all times to keep the spots above the surface of the solvent. Capillary action draws the liquid up the sheet, and as it passes the point where the spot has been adsorbed, the compounds in the mixture will partition themselves between the cellulose surface and the moving solvent. The more strongly a compound is adsorbed, the less likely it is to be drawn off the cellulose surface. On the other hand, those solutes that are weakly adsorbed, or which have a high solubility in the moving phase, will be drawn up the paper as the solvent front ascends. The solvent is usually less polar than the cellulose surface and its associated water layer, so polar, strongly hydrogen bonded compounds only travel to a small degree, whereas the less polar and less hydrogen bonded compounds travel a larger distance.

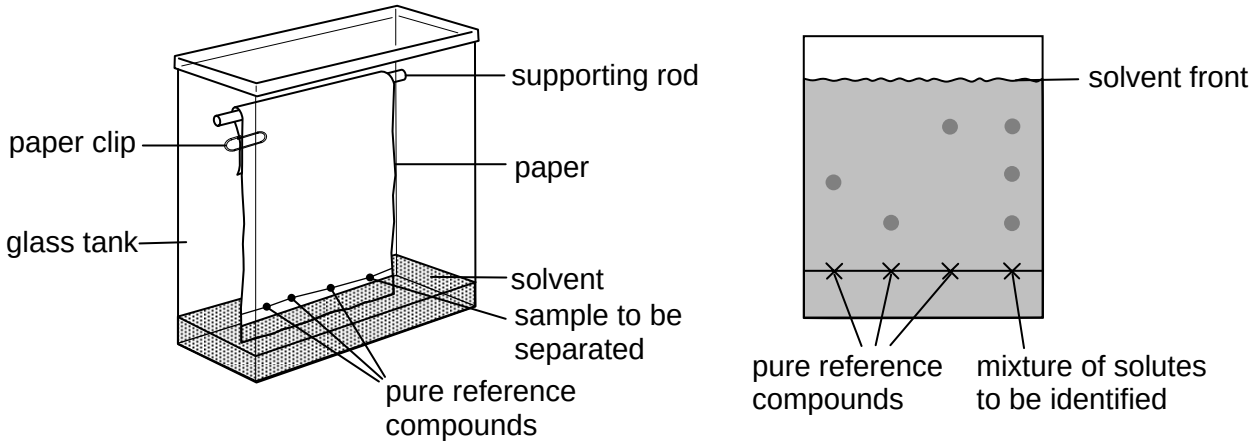


Figure 2.30 – paper chromatography and the resulting chromatogram

Once the solvent has reached the top end of the sheet of paper, the paper is removed and the solvent is allowed to evaporate. All solutes will now be re-adsorbed onto the area of the paper where they had reached, and their presence can be detected by their colour, or by their UV absorbance, if appropriate. For colourless compounds, an easy way of visualising the spots is to place the paper in a beaker with a few iodide crystals. The iodine evaporates slightly, and the vapour is preferentially absorbed by the less polar solute spots, making them appear brown, against a white paper background. Alternatively, the paper can be sprayed with a dilute solution of a reagent that forms colours with the compounds contained in the spots. Ninhydrin is used to visualise amino acids and small peptides; Molisch's reagent (a mixture of 1-naphthol and hydrochloric acid) can be used for sugars in general; Tollens' reagent can be used for reducing sugars such as glucose and maltose.

In Figure 2.31 the distance x that is travelled by a component in the solute mixture and distance y is that travelled by the solvent front. The ratio R_f, called the retention ratio (or retardation factor), is defined by

$$R_f = \frac{x}{y} = \frac{\text{distance moved by solute}}{\text{distance moved by solvent}}$$

Each solute or component has a characteristic R_f value for a given solvent. This means that the value of R_f can be used to identify a possible component. However, for this to be so, the chromatogram

needs to be carried out carefully, under controlled conditions. For example, the solvent [illegible] be buffered, and draughts must be excluded by placing the chromatographic paper inside a [illegible] closed with a lid.

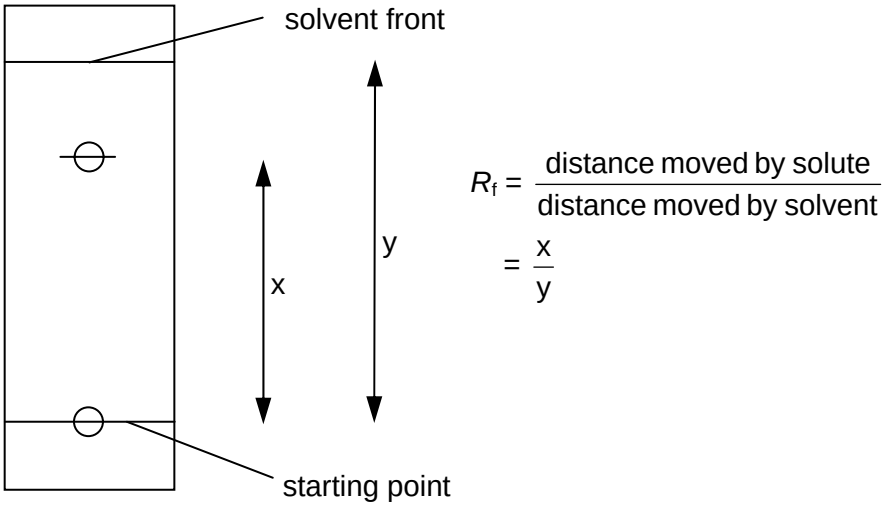


Figure 2.31 – calculating the retention ratio, R_f

Sometimes, different compounds may have very similar R_f values in a particular solvent, and so no separation can be achieved. If a different solvent is tried, the partition coefficients of the two compounds between the new solvent and the stationary phase may be sufficiently different to allow a separation.

A complex mixture can be more thoroughly analysed by two-way paper chromatography. A spot of the mixture is placed towards one corner of a square sheet of chromatography paper (only one spot can be applied – no references can be used). One solvent is used first, and the solvent front allowed to reach the far edge of the paper, after which the paper is allowed to dry completely, rotated by 90° and placed into another solvent for a second development. The spots are then more widely dispersed and can be recognised from their positions on the final chromatogram. This two-way technique is commonly used to identify the amino acids obtained from the hydrolysis of a protein (see Figure 2.32).

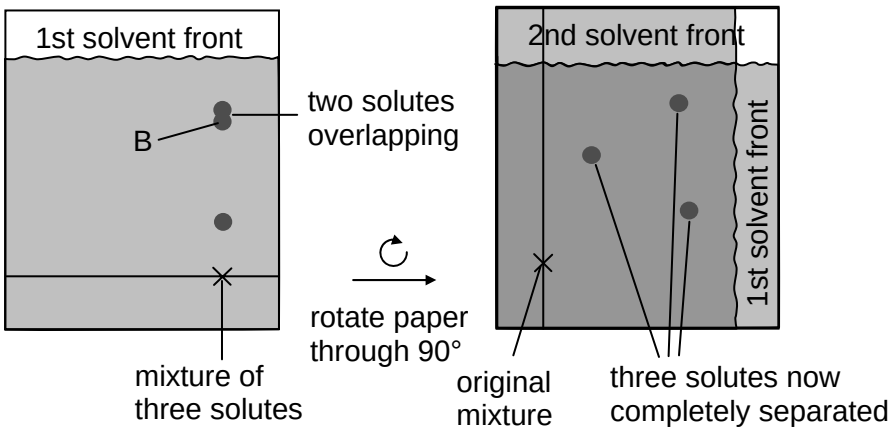


Figure 2.32 – two-way paper chromatography

SAQ 10. *Calculate the R_f value in the **second** solvent of the spot labelled B.*

Thin layer chromatography

Thin layer chromatography (TLC) is similar to paper chromatography, but the stationary phase is a thin layer of a solid such as alumina or silica supported on an inert base such as glass, aluminium foil or insoluble plastic. The plate can be prepared in the laboratory by using a slurry of the powder and then carefully oven-dried; practice is necessary to make standard and uniform coverings of the plates.

They can also be obtained from laboratory suppliers, which is the normal practice. [illegible]s relatively small, e.g. about 15 cm × 5 cm.

The mixture to be analysed is 'spotted' at the bottom of the TLC plate and allowed to dry. The p[illegible] placed in a closed vessel containing solvent (the mobile phase) so that the liquid level is below spot.

TLC has advantages over paper chromatography in that its results are more reproducible, and separations are very efficient because of the much smaller particle size of the stationary phase.

The solvent ascends the plate by capillary action, the liquid filling the spaces between the solid particles. This technique is usually done in a closed vessel to ensure that the atmosphere is saturated with solvent vapour and that evaporation from the plate is minimised before the run is complete. The partition here is between the solute adsorbed onto the SiO_2 or Al_2O_3 particles (via polar/hydrogen bonded interactions, but also some acid/base attractions, since SiO_2 is slightly acidic, and Al_2O_3 slightly basic) and the solute dissolved in the moving phase. The plate is removed when the solvent front approaches the top of the plate and the position of the solvent front is recorded before it is dried (this allows the R_f value to be calculated).

TLC has applications in industry in determining the progress of a reaction by studying the components present; and in separating reaction intermediates. In the latter case a line of the reaction mixture is 'painted' across the TLC plate instead of a single spot, and the line of product after separation is cut out of the plate and dissolved in an appropriate solvent.

The ways in which spots of colourless compounds on a TLC plate can be visualised are similar to those used for PC, but an additional technique is often used for compounds containing aromatic rings, or other systems that absorb UV radiation at 254 nm: the silica or alumina is impregnated with a fluorescent insoluble compound that absorbs UV light and emits it as visible light. When placed under a UV lamp, the plate emits a bright white light except where a UV absorbing compound is situated. Here, a dark spot is observed.

Gas/liquid chromatography (GLC)

This technique uses a gas as the mobile phase, and a non-volatile liquid coated onto small inert particles as the stationary phase. The particles are packed into a narrow glass column a few mm in diameter and between 1 m and 3 m long. The column is mounted inside an oven whose temperature can be controlled, or even programmed to increase as the separation is underway.

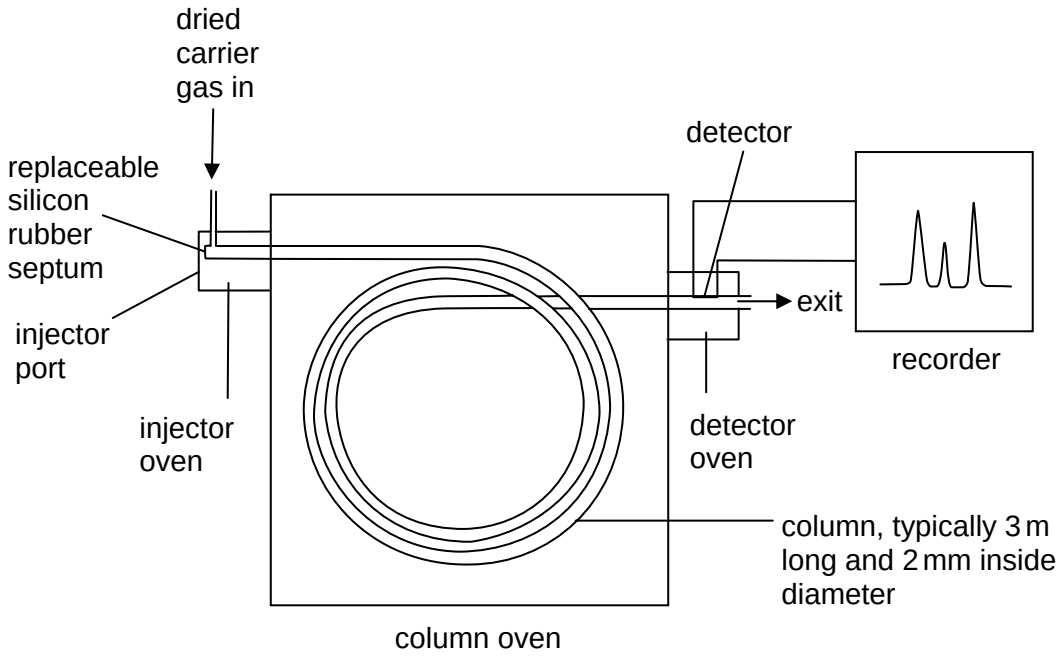


Figure 2.33 – GLC apparatus

For separation or identification the sample must be either a gas or have an appreci[illegible]r pressure at the temperature of the column - it does not have to be room temperature. The [illegible] injected through a self-sealing disc (a rubber septum) into a small heated chamber whe[illegible] vaporised if necessary. Although the sample must all go into the column as a gas, once it is ther[illegible] temperature can be below the boiling point of the components as long as they have appreciab[illegible] vapour pressures inside the column. This ensures that all the fractions pass through the column over a reasonable time span. The injector oven is usually 50-100 °C hotter than the start of the column.

The sample is then taken through the column by an inert gas (known as the carrier gas) such as helium or nitrogen, which must be dry to avoid interference from water molecules. It can be dried by passing it through anhydrous copper(II) sulfate or self-indicating silica (silica impregnated with cobalt(II) chloride). Unwanted organic solvent vapours can be removed by passing the gas through activated charcoal. The column is coiled so that it will fit into the thermostatically controlled oven.

The temperature of the oven is kept constant for a straightforward separation, but if there is a large number of components, or if they have similar affinities for the stationary phase relative to the mobile phase, then it is common for the temperature of the column to be increased gradually over a required range. This is done by using computer control, and gives a better separation if the boiling points of the components are close. It also gives a faster separation if some components are relatively involatile.

The fractions progress to the end of the column, and then to a detector. Two types of detector are commonly used: thermal conductivity detectors and flame ionisation detectors. Only the flame ionisation detector (FID) will be described here (see Figure 2.34). It is the one that is most commonly used, and is particularly useful for detecting organic compounds. The gas from the column is mixed with hydrogen and air, and is then burned. Some $CH^{\bullet}$ radicals, which are formed on combustion, are then oxidised to CHO^{+} ions. These ions produce a current of about 1 μA through a cathode to a sensitive amplifier and then to a computer or chart recorder.

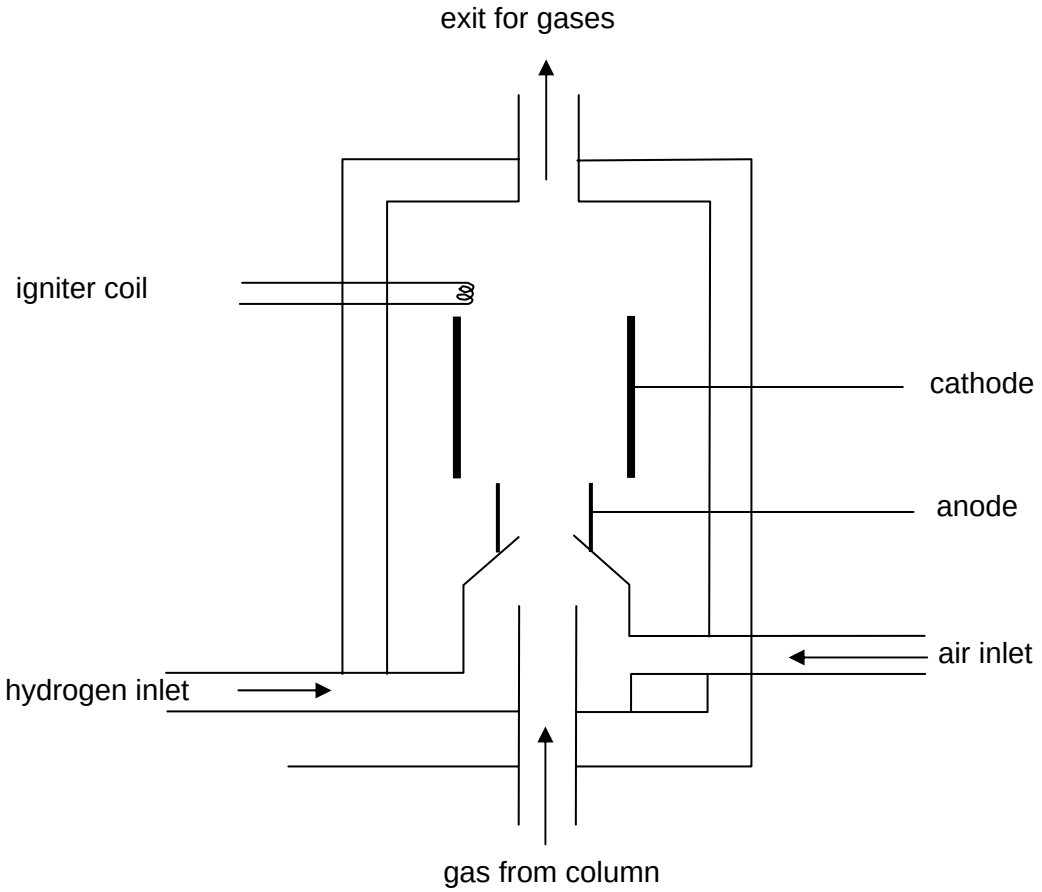


Figure 2.34 – flame ionisation detector

Once a mixture has been separated by GLC its components need to be identified. [illegible] substances this can be done from a knowledge of the time it takes for the components to [illegible] detector once they have been injected into the column. These are known as **retention times** [illegible] vary depending on each of the following:

- the flow rate of the carrier gas,
- the temperature of the column,
- the length and diameter of the column,
- the nature of and interactions between the component and the stationary and mobile phases,
- the volatility of the solute.

Each substance to be identified by GLC is run through the column so that its retention time can be determined. For compounds of completely unknown structure or composition the components must be collected individually and then analysed by using another method – e.g. mass spectrometry.

Figure 2.35 shows a GLC chromatogram of a mixture of organic chemicals.

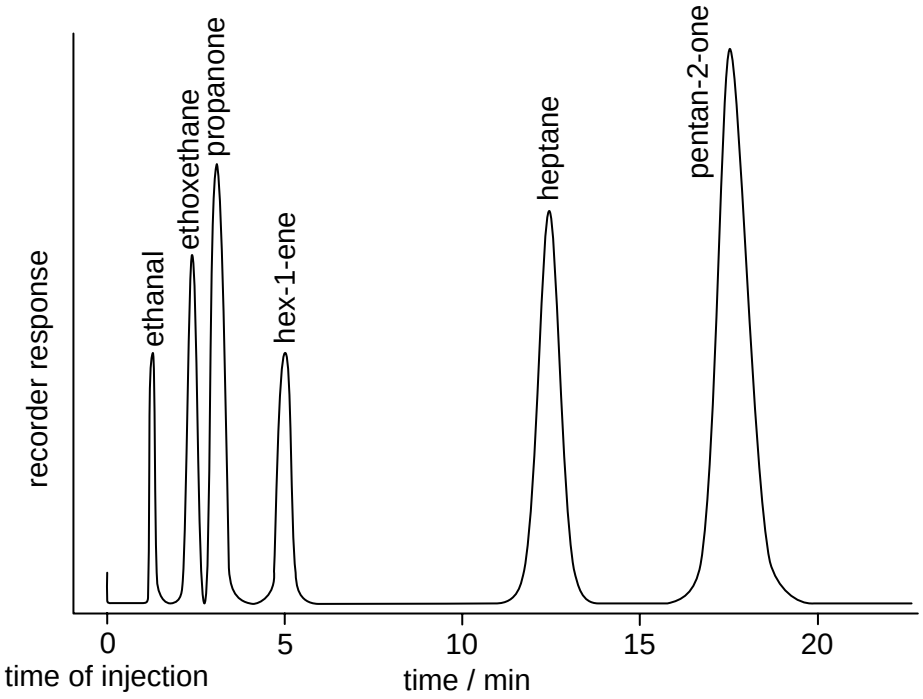


Figure 2.35 – a GLC chromatogram of a mixture of organic chemicals

SAQ 11. ***(a)*** *Referring to Figure 2.35, which compound is present in the largest amount?*

(b) *Which compound has a retention time of 2.5 minutes?*

High performance liquid chromatography (HPLC)

The principle behind HPLC, and the equipment used, is very similar to GLC. The main differences are that the moving phase is a high-purity solvent rather than a gas, the columns are shorter (10-30 cm) and the components are usually detected by measuring the absorbance of UV radiation through a microcell at the end of the column (see Figure 2.36). The efficiency of a separation increases if the particles in the stationary phase are made smaller. This is because the solute can equilibrate more rapidly between the two phases. The stationary phase in an HPLC column normally consists of uniform porous silica particles of diameter 1×10^{-6} m, with surface pores of 1×10^{-9} m diameter to increase the surface area. Sometimes the particles are coated with molecular fragments of a solvent having particular polar groups, joined to the silica by covalent bonds. This ensures they do not dissolve in the mobile phase. Thus it is debatable whether this is an example of the solid adsorption-liquid partition or the liquid-liquid partition of the solute.

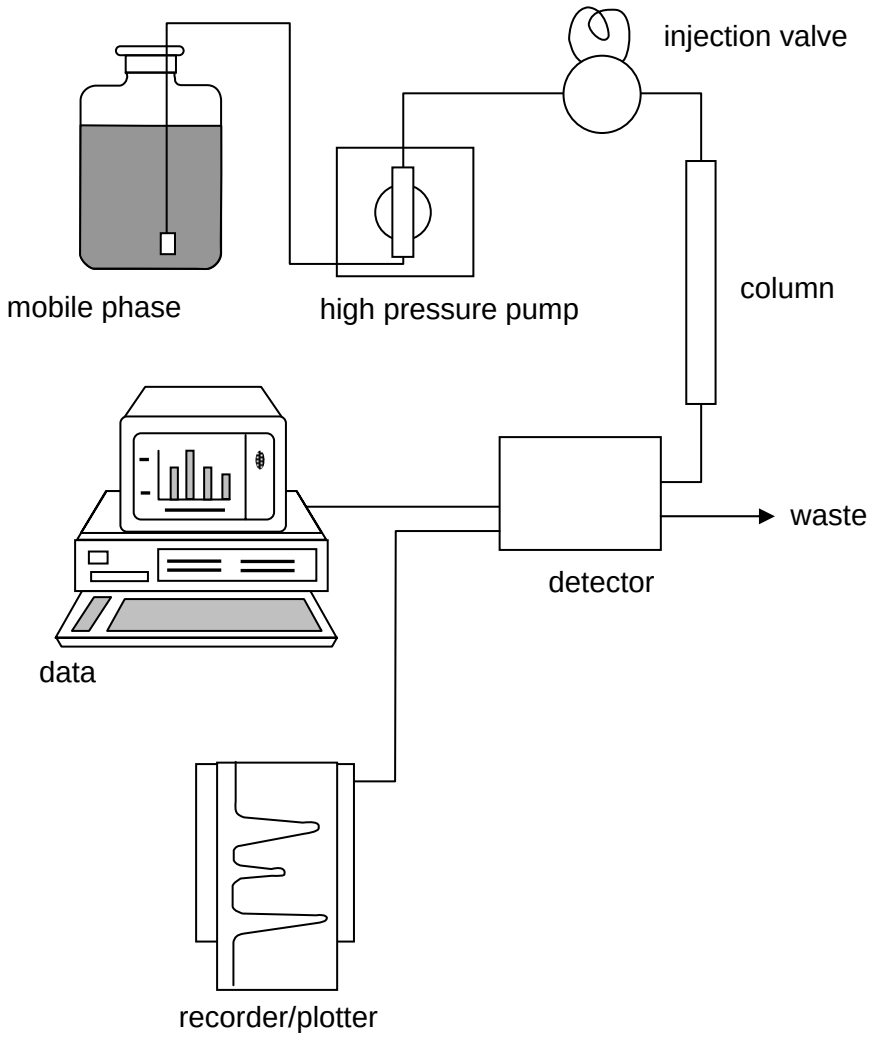


Figure 2.36 – HPLC apparatus

To ensure reproducibility of retention times, various factors have to be controlled exactly. A constant flow rate is maintained by special twin cylinder reciprocating pumps, generating steady precise pressures of up to 100 atmospheres. Even with such high pressures, the flow rate through the column is small (about 2 $cm^3\,min^{-1}$), due to the tightly packed small particles of the stationary phase. Nevertheless, because the column volume is small, the timing of the injection of the sample must be precise if an accurate retention time is to be measured. A rough calculation will show this.

Imagine a 20 cm column of diameter 0.4 cm. This has a volume of about 2.5 cm^3 (=20 × π × 0.2 × 0.2). Now assume that the tightly-packed particles take up about ¾ of this volume. The volume available for the mobile phase to flow through is therefore about 0.63 cm^3 (=2.5/4). At a flow rate of 2 $cm^3\,min^{-1}$, it will take only 19 sec =(60 × 0.63/2) to travel through the whole of this column.

Applications of Chromatography in Analysis

The relative use of HPLC and GLC varies from industry to industry and very much depends on the compounds to be separated. Many compounds decompose at the temperatures required for efficient GLC separation, while HPLC separation can be achieved readily. However, GLC is particularly useful in detecting residual solvents in formulations and is also invaluable in looking for degradation products. Amines and acids are not separated well by GLC because they tend to be too polar, but can be efficiently separated using HPLC.

Other methods are often used in conjunction with HPLC for determining the presence of components in the output stream, and/or identifying the compounds in those components. These include:

- mass spectrometry

- infra-red spectroscopy
- visible spectroscopy
- ultraviolet spectroscopy

Some examples are:

- The residual caffeine content in “decaffeinated” coffee and tea can be analysed by extraction with hot water and analysis of the solution by HPLC, comparing the retention times and the peak heights with those of standard solutions of caffeine.
- Dope testing of racehorses. Each year about 2/3 of all race winning horses are tested for the presence of performance-enhancing drugs (or even sedatives, that can cause the horse to knowingly *under*-perform). The procedure is as follows.
  1. A sample of the horse’s urine is first separated into acidic components, basic components and neutral components by passing through various silica chromatography columns.
  2. After elution from the columns, the various components are analysed as follows:
     strong acidic components by TLC on silica or alumina
     neutral components by HPLC
     basic components by GLC followed by MS

  To each component various standards are added for comparison of retention times, etc.
- Detection of explosive residues on skin, clothing and in vehicles is often achieved by dissolving the residue in an organic solvent and analysing the solution by GLC or HPLC.
- Detection of pesticide residues in fruit and vegetable is routinely carried out in much the same way as the detection of explosive residues: the food is macerated in a buffer solution and extracted with an organic solvent (most pesticides are more soluble in organic solvents than in water). The solution is then analysed by HPLC, using standard samples of pesticides to compare retention times.

2.3c Electrophoresis

Although the appearance of a gel electrophoresis plate after development is similar to that of a thin layer chromatography plate, the principles behind the two methods of separation are very different. In electrophoresis there is no mobile phase, and only ions, not neutral molecules, move through the buffer solution along the plate.

The principle can be simply demonstrated as follows:

- soak a strip of filter paper with pH 7 buffer and clip it to a microscope slide with two crocodile clips
- place a drop of a dilute solution containing the three amino acids glycine, glutamic acid and lysine (see Figure 2.37) in the middle of the paper
- apply a voltage of 50-100 V DC across the crocodile clips, and leave for an hour or so
- disconnect the clips, and develop the spots with ninhydrin spray

The original spot will have been separated into three spots, as in Figure 2.38.

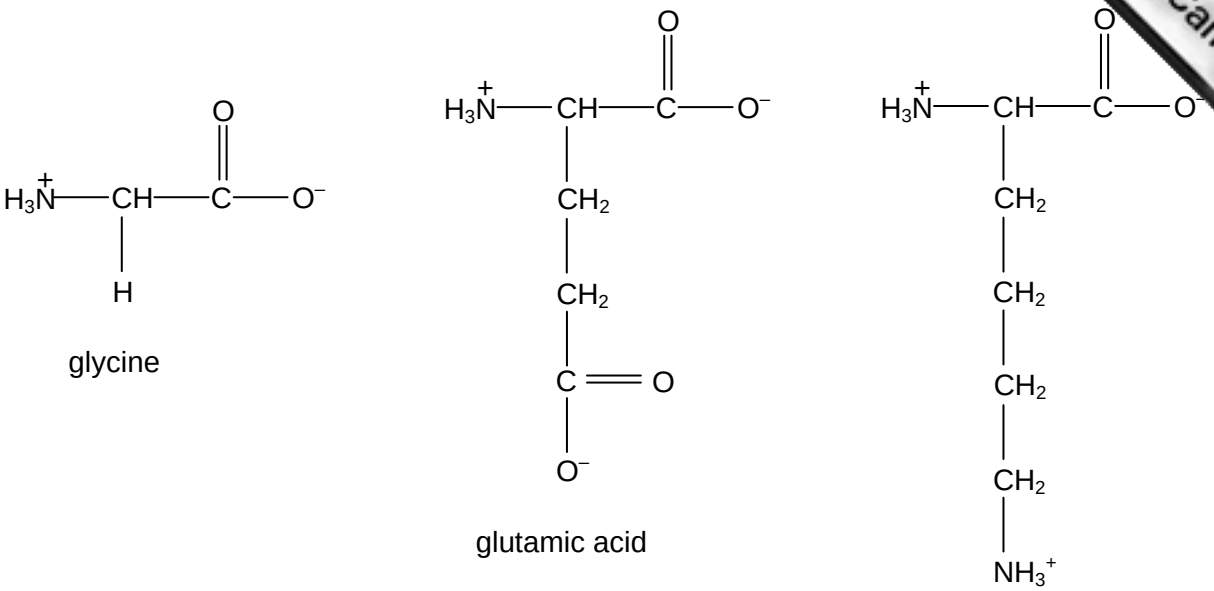


Figure 2.37 – structures of glycine, glutamic acid and lysine at pH 7

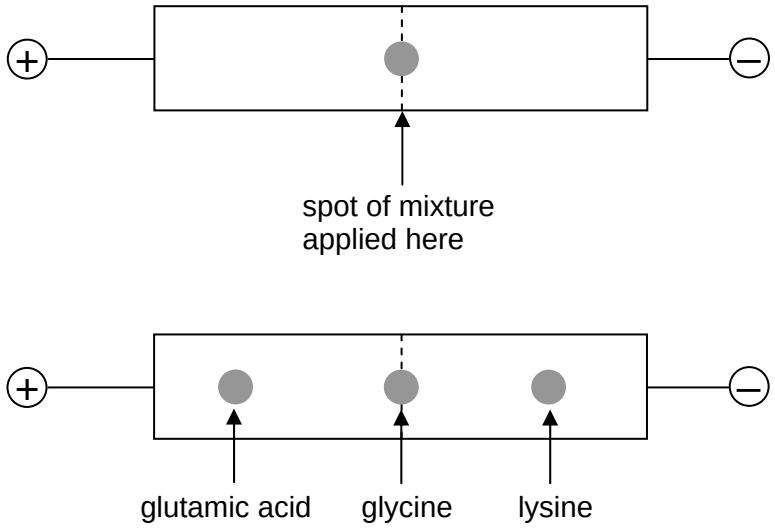


Figure 2.38 – results of electrophoresis of glycine, glutamic acid and lysine at pH 7

SAQ 12. *Suggest and explain how the result of electrophoresis might differ if the experiment was carried out at pH 12 instead of pH 7.*

In industrial, analytical or research environments, paper electrophoresis as described above is too imprecise a method to use. Either gel or capillary electrophoresis is employed. Gel electrophoresis is more common, and is the one described here.

The term **gel electrophoresis** covers a range of techniques that is used to separate, analyse and purify mixtures of biological molecules such as proteins and nucleic acids. These techniques can be adapted to:

- measure the relative masses of macromolecules,
- prepare nucleic acids and polypeptides for sequencing the component monomers, e.g. purine and pyrimidine bases and amino acids,

- separate proteins, so antibodies can be raised.

The sample mixture is placed in a gel and is separated into its constituents by applying an electric field to the gel, which is soaked in a liquid buffer. The gel is a sponge-like structure based on a three-dimensional polymeric network. It has the texture of a jelly.

Gels are used because their properties can be precisely controlled during their preparation, and they are more chemically stable as a support medium.

The components in the sample mixture have an electrical charge because proteins, like amino acids, carry either an overall positive or an overall negative charge, depending on the pH of their environment. Nucleic acids such as deoxyribonucleic acid (DNA) are usually negatively charged (at the phosphate groups in the chain) at the pH used for their separation using electrophoresis. The molecules move in response to an electrical field applied across the mixture. The rate of progress of the molecules depends on their size, charge and shape. On separation, the components are concentrated into bands or zones.

Each band or zone can be quantified using a variety of methods. Using sensitive techniques, such as bioassays or silver staining, amounts as small as 10^{-18} g can be detected. Gel electrophoresis has a variety of applications such as checking the adulteration of foods, chromosome sequencing, DNA fingerprinting and characterising the chemicals responsible for allergic symptoms. It is used to screen infants' milk for α-lactoglobulin, a protein which is lethal for small babies.

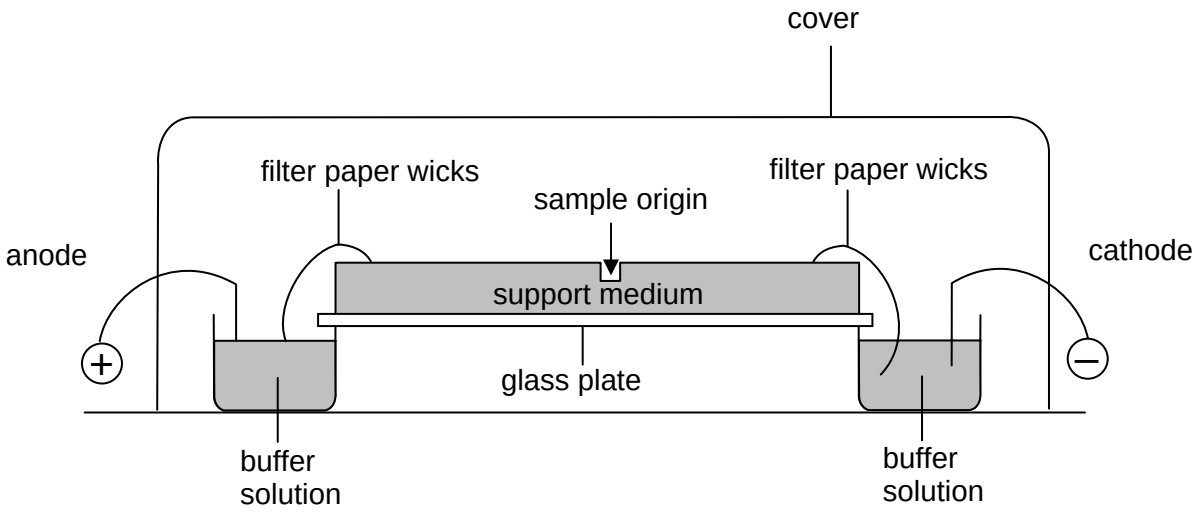


Figure 2.39 – gel electrophoresis apparatus

The main components used in a gel electrophoresis analysis (see Figure 2.40) are:

- an electrophoresis chamber;
- a gel support medium soaked in conducting buffer;
- a means of generating an electric field – e.g. a power pack;
- probes for detecting and/or measuring the separated molecules,
- a means of extracting the individual products (if necessary).

The mixtures are placed in small wells in the gel medium and when the electric field is applied the components of the mixture separate out. Their rate of migration across the gel becomes constant when the force of attraction between the electrode and the oppositely charged component is equal to the frictional force of the gel medium resisting the motion for that particular species. The relative rates at which components move and therefore the extent to which they separate are influenced by the strength of the field, the nature of the gel, and the surrounding buffer. Each species responds to these factors according to what is called its electrophoretic mobility (μ). This is a constant for each species.

Factors affecting mobility

1. Voltage: the velocity of a molecule is directly proportional to the voltage gradient across the
2. Size: smaller molecules migrate quicker than larger molecules carrying the same charge.
3. Shape: a molecule with lots of side-chains experiences more frictional resistance than a linea molecule of the same mass and charge, and will therefore move more slowly.
4. Buffer pH: proteins exist as zwitterions like amino acids, and can be either positively or negatively charged because they contain both acidic and basic groups. The extent, and direction, of ionisation depends on the pH of the buffer.
5. Temperature: a rise in temperature can speed up electrophoresis, but can also denature the proteins. The temperature is normally controlled by cooling the plate with a flow of water underneath.

If a number of components in a sample have similar electrophoretic mobilities, a complete separation may not be achieved. In this case, either the buffer medium or the gel support can be altered. In addition, just as in paper chromatography, two-way electrophoresis in different buffers can often separate components.

Some medical applications of gel electrophoresis

Defects in newborn babies. Neural tube malformations can be detected from proteins leaking from the central nervous system of a foetus into the amniotic fluid (the fluid enveloping the foetus). Analysing these proteins can indicate spinal problems in new-born babies.

Sweating polypeptides. Over 400 polypeptide spots show up in a two-dimensional electropherogram of human sweat. Many of these have been previously unidentified.

Alcohol abuse. The extent of alcohol abuse can be investigated by analysing blood. This is because excess alcohol is associated with changes in acidic proteins and glycoproteins in blood plasma.

'Fish eye' disease. This is an inherited condition in which lipid is laid down in the eyes making them appear opaque like fish's eyes. Its medical name is dyslipoproteinanaemia. Protein samples from people with this condition have been analysed, suggesting that there is a deficiency in an enzyme system associated with lipid metabolism.

Heart attacks. Blood samples are taken at regular intervals (once a day for three days after initial chest pain). Studies are being done to identify more sensitive *'*marker*'* proteins which will indicate early blockage in blood vessels.

Assessing fitness. When a person is very much out of condition there can be a serious increase in the amount of protein in urine, which can be detected by two-dimensional electrophoresis. As well as being a pointer to lack of physical fitness, increased protein content in the urine can also be a possible indicator of the onset of diabetes.

Other types of electrophoresis

Different techniques of electrophoresis include *capillary gel electrophoresis* and *capillary zone electrophoresis*. These have found useful applications in

- the analysis of haemoglobins to identify congenital diseases such as sickle-cell anaemia, thalassemia, and conditions associated with diabetes
- the detection of chemical weapons products – for example, the alkylphosphonates that arise from the breakdown of sarin, one of the nerve gases
- detecting gunshot residues on skin and clothing
- detecting drugs by analysing hair samples
- separating milk proteins (see Figure 2.40)

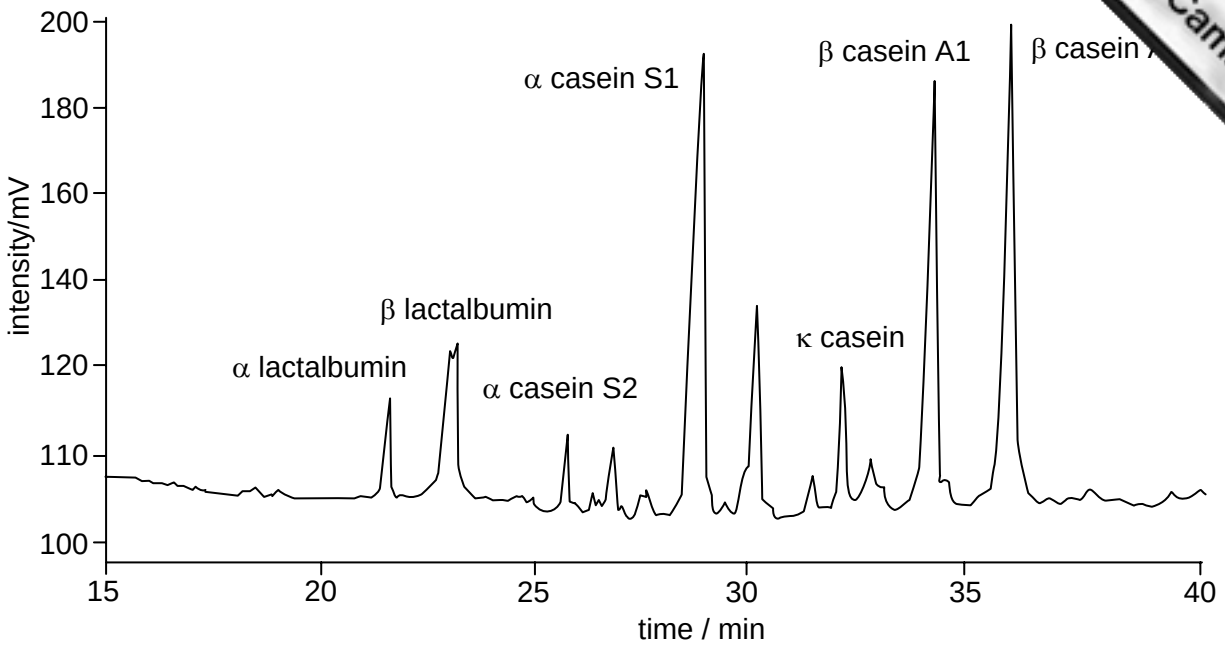


Figure 2.40 – separation of milk proteins

2.3d Genetic fingerprinting

The principle and method

Genetic fingerprinting is perhaps the most high profile and widely used application of the electrophoresis of nucleic acids.

DNA, or deoxyribonucleic acid, is the biological macromolecule that is contained within the chromosomes of cells. It is a two-stranded polymer and each strand is composed of a sugar (deoxyribose)-phosphate backbone, with a cyclic organic base attached to each sugar residue. The strands are held together by hydrogen bonds between the organic bases and twine around each other in a helical fashion (see Figure 2.41). All DNA molecules contain the same molecular backbones, and the same four bases. They differ only in their overall length (even the shortest has many thousand units in its chain) and in the order in which the bases are arranged along the backbone.

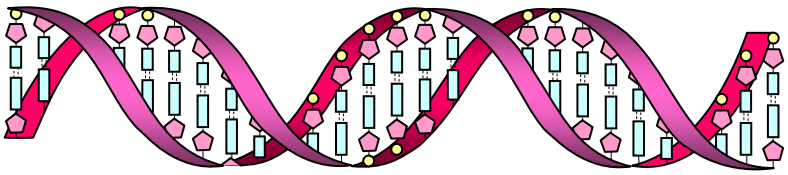


Figure 2.41 – structure of DNA

Genes are sections of a DNA molecule that become templates for the synthesis of strands of RNA (ribonucleic acid) which in turn form the code that translates into the arrangement of amino acids along the chains of the enzymes and other protein molecules in the organism. Of the DNA contained within the chromosomes, only about 10% is used in the genes. This small portion of “useful” DNA will be almost identical in most members of a particular species (and much will be identical for different species too). This is because any large change in this genetic material would be mirrored by a change in the amino acid sequence of the final protein, which would affect their efficiency as enzymes and could prove fatal to the organism.

The other 90% of chromosomal DNA (the “useless” DNA) lies in between the genes, and is highly variable in its base sequences. These portions often contain sequences of bases (about 10-50 base-

pairs in length) that are repeated several times. All members of the same species [illegible] repeats, but individuals vary in the number of times each sequence is repeated. The areas a[illegible] minisatellites, or VNTRs (variable number of tandem repeats). There are also quite short seq[illegible] (2-5 base-pairs) that are repeated many times. These are termed short tandem repeats (STRs) microsatellites (see Figure 2.42). The key to genetic fingerprinting is that an individual's pattern VNTRs or STRs is entirely unique – no one else (except an identical twin) will have the same pattern (see Figure 2.43). Furthermore, each individual inherits half their pattern from their mother, and half from their father. Members of a family throughout many generations will have similar minisatellite patterns to each other.

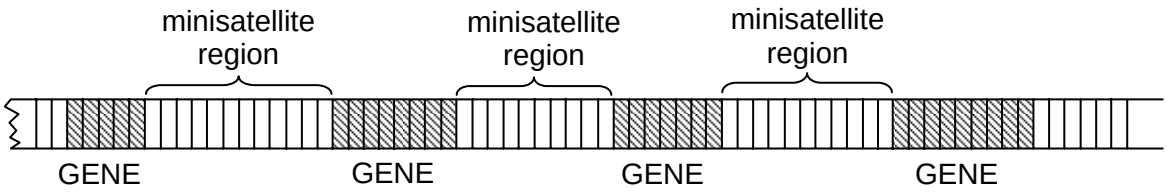


Figure 2.42 – strand of DNA showing genes and minisatellites

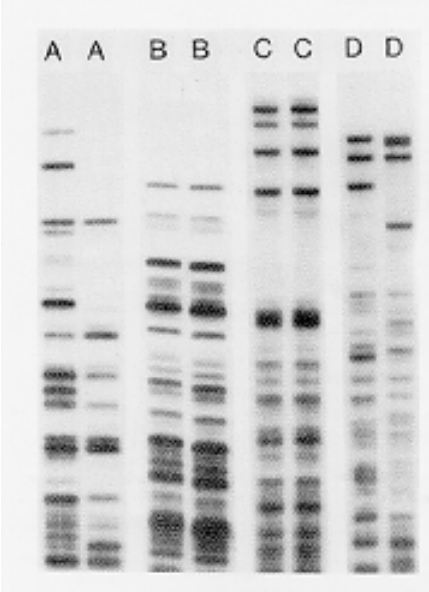


Figure 2.43 – DNA from twins

The technique of genetic fingerprinting starts with the extraction of the DNA from a sample of chromosomal material, such as blood (only the white cells contain DNA), hair, inner cheek cells, semen or skin. About 25 mg of DNA is required. The DNA is then broken into fragments by a restriction enzyme. There is a range of these enzymes available, each one breaking the DNA at several different, but known and specific, places. This produces fragments that can now be subjected to electrophoresis.

Modern techniques using the polymerase chain reaction (PCR) can be used when the amount of available DNA is smaller, in which case an analysis can start with as little as 0.2 nanograms (2×10^{-10} g). In this technique, after the DNA has been broken up into pieces in the usual way, an individual segment is extracted. Heating a solution of the segment separates the two DNA strands from each other. The sample is then cooled and a short "primer" length of DNA which matches the end of the segment is added, together with the enzyme DNA polymerase and the four monomer nucleotides A, G, C, T. On warming the solution the four nucleotide bases bond to the two single strands in the precise order as determined by the original base sequence, to form two new doubly-stranded DNA molecules. The process is repeated many times (it can be automated readily) until a sufficient amount of DNA has been made - 1 billion (1×10^9) copies would be produced after only 30 repeats.

The samples of DNA to be analysed are placed in wells near the cathode of an agaro[illegible] Since the phosphate groups along the chains of DNA are negatively charged, all DNA sa[illegible] move towards the anode, but the smaller fragments will move faster than the larger ones.

Discovering the positions of the separated fragments on the developed electropherogram require[illegible] special technique, since the amounts of DNA are so small. Spraying and staining with a dye would give only a faint image. The DNA fragments are transferred from the gel onto a nylon membrane and are broken into single strands by soaking in dilute alkali. The membrane is then soaked in a solution containing short lengths of DNA identical to the minisatellite sequence, but labelled with radioactive ^{32}P. The ^{32}P DNA thus sticks to the membrane where the DNA fragments are situated, and if an X-ray sensitive film is placed over the plate, it will record the pattern originally on the gel plate (see Figure 2.44).

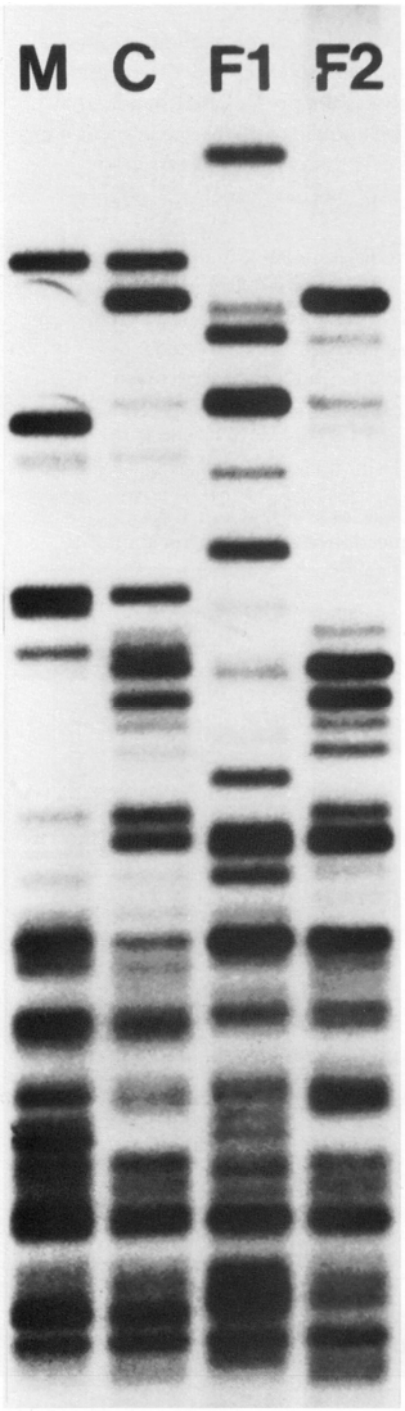


Figure 2.44 – genetic fingerprints of a child, C, and its mother, M, and two possible fathers, F1 and F2

The uses of genetic fingerprinting

Genetic fingerprinting has been used in

- paternity testing
- establishing other familial relationships – between both the living and the dead
- establishing the relationship between archaeological artefacts
- forensic testing
- medicine

One example of each of these will suffice to illustrate the application.

Figure 2.44 demonstrates how genetic fingerprinting could be used to establish paternity. A child inherits one set of chromosomes from its mother, and the other set from its father. In the figure it can be seen that the child's fragments correspond with those of its mother, together with those of possible father F2, but show little correspondence to those of possible father F1.

The same technique can be used to establish familial links of children to parents – for example in immigration disputes. An interesting historical example is the investigation by STR analysis of the DNA from the bones of five bodies dug up in Russia in 1991. It was suggested that this was the family of the last Tsar of Russia, assassinated by a Bolshevik firing squad in 1918. The DNA of the three female children had clear correspondences to that of the two adults (who were not genetically related to each other), clearly showing the presence of a family group. Further work linked the mitochondrial DNA to that of the present Duke of Edinburgh, who was known to be related to the late Tsarina through a maternal line (mitochondrial DNA is only inherited from the mother, not the father). Thus the identities of the exhumed bodies were established.

The archaeological investigation of the Dead Sea Scrolls, written on goat skins two thousand years ago, has been made easier by the genetic fingerprinting of the DNA from small scrapings on the old skin cells. The various scrolls were written on the skins of different animals, but had over the years been broken into many fragments. Being able to piece together the original skins by matching the DNA has helped archaeologists make sense of the texts.

A high profile use of genetic fingerprinting was in the trial of the basketball player OJ Simpson in the USA in 1995. A sample of DNA from his cheek cells was identical to samples of blood found at the scenes of two murders he was accused of committing.

Another useful application of genetic fingerprinting in forensic science is its ability to *exclude* suspects, as well as give strong evidence to *include* one particular suspect. Figure 2.45 shows the DNA fingerprints of seven suspects, and of a bloodstain at the scene of a crime.

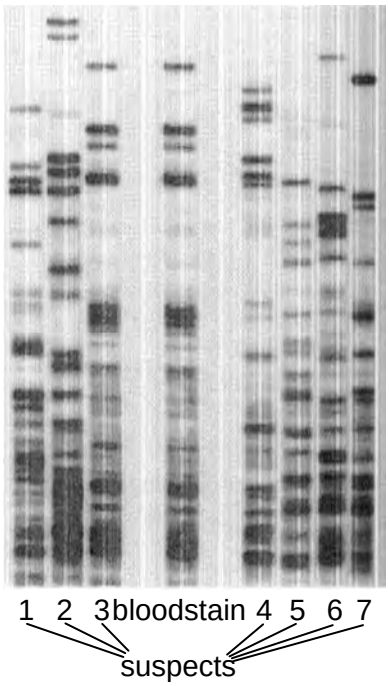


Figure 2.45 – DNA fingerprints of seven suspects and blood taken from the scene of a crime

SAQ 13.(a) *Which suspect's DNA best matches the DNA of the bloodstain?*

(b) *Are there DNA samples from any of the other suspects that show a correspondence?*

Medical applications have included the investigations of whether cases of tuberculosis or cancers are re-infections from previous illnesses, or brand new infections. This can help in their treatment. Recently a genetic fingerprinting analysis of various samples of the MRSA "superbug" *Staphylococcus aureus* from around the world has shown that the strain can be traced back to the 1950s, and even early examples were resistant to penicillin, streptomycin and tetracyclin antibiotics. This should help the development of new types of antibiotic.

Summary

- The ratio of concentrations of a solute in equilibrium with two immiscible solvents is gi[illegible] the partition coefficient: K_{pc} = [solute(solvent A)]/[solute (solvent B)].
- Partition of different components between a mobile phase and a stationary phase is the basis of chromatography, of which the four main type are paper chromatography (PC), thin layer chromatography (TLC), gas-liquid chromatography and high performance liquid chromatography (HPLC).
- These four techniques are used to separate and identify compounds in a number of applications in medicine, forensics, environmental monitoring.
- Electrophoresis is a useful technique for separating amino acids, proteins and nucleic acids
- A major application of electrophoresis is for genetic fingerprinting.

2.3	***What the student needs to know***
11.2f	How to write the K_c expression for the partition coefficient. How to use the K_c expression to calculate K_c from concentration data, [X] in one solvent given [X] in another, the amount of X that is extracted, and hence the amount remaining, after 1 or more successive extractions.
11.2g	The principles of paper and thin-layer chromatography, HPLC and GLC in terms of absorption and/or liquid-liquid or liquid-gas partition. (Students do **not** need to know any details about the design of HPLC or GLC machines, but you should know the principles of how they operate, and, in outline, how the components are detected as they come off the column). The principles of 2-way paper chromatography, and how to interpret the resulting chromatogram. How to calculate R_f values and retention times. How to interpret the results of the four different forms of chromatography expressed pictorially, graphically or in the form of a table. The factors that affect the mobility of a compound in the various chromatography techniques (e.g. the volatility of molecule X, or the relative polarity of molecule X with respect to the solvent and the support). In outline only – one or two applications of each form of chromatography.
11.2a	The principle of electrophoresis, and the diagram of the simple set-up. The factors that affect the mobility of ions in electrophoresis (charge, pH, voltage)
11.2b	How electrophoresis is used in genetic fingerprinting (outline only), to include: the concepts of VNTRs and STRs, the use of restriction enzymes, the use of PCR to increase the amount of the sample. How to compare and interpret genetic fingerprints.
11.2c	A couple of examples of the medical application of electrophoresis.

2.4 – Examples of applications of analytical chemistry

By the end of this section, students should be able to:

- *appreciate the use of the techniques of gas liquid chromatography, mass spectro[illegible] solvent extraction in the analysis of environmental, archaeological and geological samples,*
- *understand the conclusions that can be drawn from such analyses,*
- *predict techniques suitable for the separation and analysis of samples taken from environmental sources.*

2.4a Three methods of dating using mass spectrometry

Carbon dating

The use of the radioactive isotope ^{14}C in the dating of once-living artefacts is well known, but because of its fairly short half-life (5730 years) it becomes less accurate for ages greater than about 10 000 years. Using a "tandem accelerator" to purge samples of ^{14}N and $^{12}CH_2$ (both of which have an *m/e* value of 14), and then analysing the ^{14}C:^{12}C ratio in a conventional mass spectrometer, it is possible to extend the range to 100 000 years.

Potassium-argon dating

Naturally occurring potassium contains (now) 0.01167 per cent of the radioactive isotope ^{40}K with a half-life of 1.3×10^9 years. When it undergoes radioactive decay, one of the products is ^{40}Ar. If an igneous rock sample contains potassium, and the argon formed by the decay of ^{40}K is unable to diffuse out of the crystal lattice, it is possible to determine the age of the rock sample – that is, the time when the rock crystallised from the magma.

The argon content of a sample is determined by using a mass spectrometer to analyse the gas released when the sample is heated and melted in a vacuum. (It is assumed that none of the argon produced within the crystal diffuses out of the crystal lattice of the rock, and that there has been no loss or gain of potassium since it was formed.) A known amount of ^{38}Ar is added, and the ^{40}Ar content is measured as a proportion of the ^{38}Ar, in a method known as *isotope dilution*. The total remaining potassium can be measured by using several methods, including flame photometry, atomic absorption spectrometry and isotope dilution. By using the half-life of ^{40}K and the ^{40}K:^{40}Ar ratio, the age of the rock can be determined.

Oxygen isotopic ratios in the dating of ice cores

Naturally-occurring oxygen consists of three isotopes, ^{16}O, ^{17}O and ^{18}O, in the percentage abundances of 99.76%, 0.04% and 0.20% respectively. Most of the ice in the ice sheets at the poles has been formed over many thousands of years by water molecules evaporating from the oceans and travelling in the vapour state towards the polar regions, where they precipitate as snow or hail. $H_2{}^{16}O$ evaporates, and travels through the atmosphere, more quickly than $H_2{}^{18}O$. Thus the snow precipitating in the polar regions is less rich in ^{18}O-water than is the water in the oceans. This is a temperature-dependent phenomenon, however. Snow that falls in the winter months is even more depleted than that which falls in the summer. Thus by measuring the $^{16}O/^{18}O$ ratio continuously throughout its depth, and counting the annual changes (in a similar way to aging a tree by counting its annual growth rings), the age of an ice field can be determined.

For example, the Vostock Ice-Core was collected in east Antarctica by a Russian expedition over the years 1970 – 1983. It measured over 2000 m long, representing just over half the thickness of the ice sheet from which it was collected. The age of the bottom of the ice core was calculated at about 150 000 years – the origins of the ice sheet itself must be considerably older.

2.4b The use of combined GLC-MS in environmental monitoring

Biomarkers in the petrochemical industry

Biomarkers are organic compounds whose carbon skeletons provide an unambiguous link with a known natural product, and these are particularly useful in the petrochemical industry.

When the marine organisms that formed crude oil died, their bodies were subjected to extreme heat and pressure, resulting in the formation of the oil. In the chemically reducing environment present under the sea many of the natural products – e.g. steroids and terpenoids - became saturated (i.e. highly reduced to form cycloalkanes).

These saturated products are now still present in crude oil, and serve as biomarkers. It is thus possible to identify the organisms from which the oil was formed, enabling the terrain at the time of formation to be deduced. Once the biomarkers from a particular oilfield have been identified it is also possible to determine where an oil sample originated. Thus, if an oil spillage occurs at sea and nobody admits responsibility, the biomarkers could be used to identify its source. It may then be a simple matter to find out which tankers were carrying oil from that field at the time of the incident, and so identify the culprit.

Identifying the biomarkers is done by separating the components of the crude oil using GLC and feeding them directly into the chamber of a mass spectrometer. The components can be identified by their retention times on the GLC column and the fragmentation patterns of their mass spectra.

Figure 2.46 shows GLC traces of the steroidal cycloalkanes fractions of two samples of crude oil.

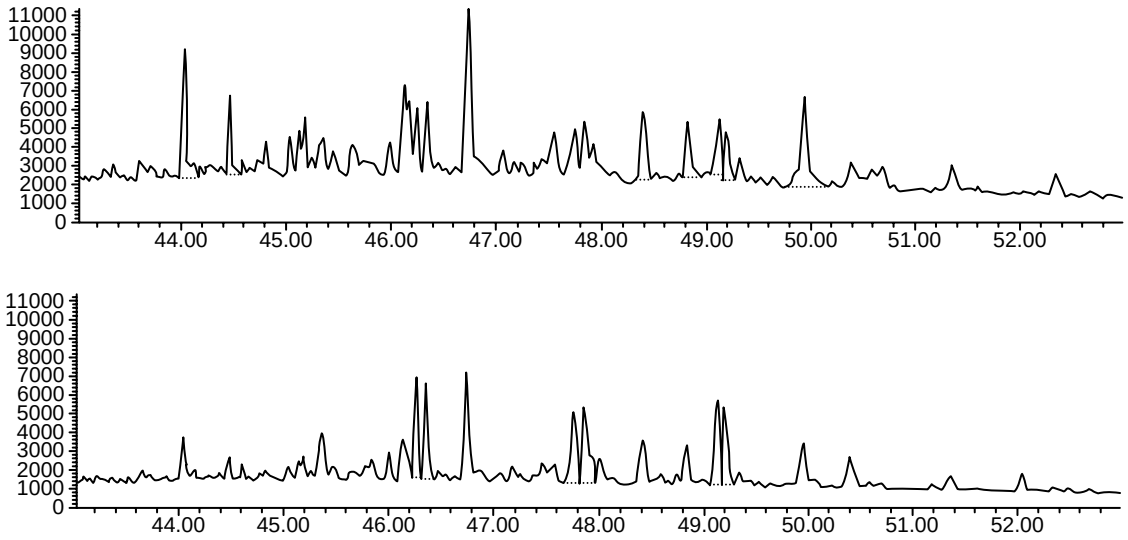
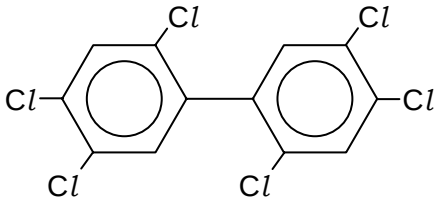


Figure 2.46 – GLC traces of the steroidal cycloalkane fractions of two samples of crude oil

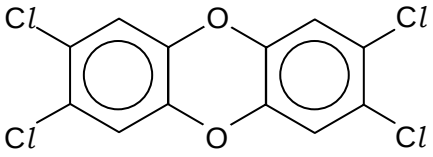
SAQ 14. *Are the two samples identical or not? Choose a number of peaks from the two traces to support your choice.*

Detecting PCBs and dioxins

Although concentration methods are sometimes required to supply the spectrometer with sufficient sample to analyse, GLC-MS is sensitive enough to detect pollutants such as polychlorinated biphenyls (PCBs), which are found in surface water in concentrations on the nanogram per dm^3 (1×10^{-9} g dm^3) scale. Monitoring for dioxin can also be done by using mass spectrometry. Dioxin (more correctly, tetrachlorodibenzodioxin) can be found in the effluent stacks of incinerators where PVC and other chlorinated polymers have been burned. Both PCBs and dioxin are stable, fat-soluble compounds that do not degrade in nature, but build up in food chains. Some seals, for example, have been found to have several per cent of PCBs in their total fat content.



one of many possible PCBs



tetrachlorodibenzodioxin

Figure 2.47 – a PCB and a dioxin

The Mars Viking lander

Mass spectrometers have been adapted for use on board spaceships, where less attention has to be paid to the inclusion of vacuum pumps because the atmosphere is already at low pressure. The Viking spacecraft that landed on Mars was equipped with GLC-MS, and both the Martian atmosphere and the Martian soil were analysed. It found that the partial pressures of water, nitrogen and carbon dioxide in the atmosphere are very low, and that the soil contained a lot of chlorine and sulfur.

2.4c Two stories of detection by mass spectrometry

The solving of a historical mystery

In the spring of 1845 Sir John Franklin and a crew of 128 men set sail to try to complete the North West Passage – a hoped-for route from Europe to Asia around the north of Canada. They never returned, and despite several search parties being sent out, the first evidence of their demise did not come to light until 1850. It was not until 1981, and later in 1986, that an investigation by physical anthropologist Owen Beattie was carried out on the still-frozen remains of crew members. An analysis of a sample of hair showed it contained 600 ppm (parts per million) of lead, which is 120 times the amount one might expect in hair. Nearby, Beattie found a pile of old empty tin food cans, which had been the expedition's provisions during their last winter. He observed that there were lumps of solder on the inside of the seams. Until the late 19th century the solder used to seal food tins consisted of 50% lead and 50% tin.

The isotopic composition of lead varies from one ore to another, due to lead being the final stable element from various different radioactive decay chains. No two samples of lead will contain exactly the same proportions of the four main isotopes ^{204}Pb, ^{206}Pb, ^{207}Pb and ^{208}Pb. By analysing the lead from the hair and that from the solder on the tins Beattie found them to be identical.

The crew had been poisoned by the lead that had dissolved in the weak acids in the food contained in the cans.

An ecological detective story

This account illustrates the use of both solvent extraction and GLC-MS in the separation and identification of drug samples.

In the spring of 1991, various incidents were reported of eagles being found dead or dying near the small town of Kodiak in the USA. There was no clear evidence of any disease, and nothing to suggest violent injury. Poisoning was suspected. Inside the eagles' stomachs, there were still pieces of undigested meat from their last meal. They had clearly eaten well, but died soon afterwards. This suggested that any poison that they may have eaten was not slow-acting, but had worked quickly.

A sample of the meat was macerated, acidified with HC*l*(aq) and filtered. The aqueous extract was then shaken with ether and the layers separated. Any substances that were basic would remain in the HC*l*(aq) as ionic salts (many drugs and poisons are amines, which are bases), whereas neutral or acidic covalent compounds would dissolve in the ether layer. The aqueous layer was then made alkaline with NaOH(aq) to liberate the free un-ionised bases, and extracted with trichloromethane.

The ether and the chloroform layers were separately analysed by GLC-MS using a capi[illegible]n 30 metres long. The retention times and the fragmentation patterns of the various compon[illegible] compared to those of known poisons, pesticides, etc.

No identifiable drugs or poisons were found in the chloroform extract, but the mass spectrum of on[illegible] the components in the ether extract contained fragments at *m/e* values of 71, 155, 197 and 226. On searching its computerised memory, the mass spectrometer identified this component as **pentobarbital**, a popular short-acting barbiturate derivative used primarily as a sedative.

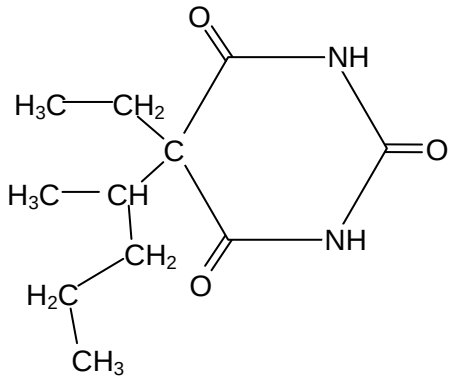


Figure 2.48 – pentobarbital

SAQ 15. *Identify the molecular formulae and possible structures of the 4 peaks in the mass spectrum of pentobarbital*

The next stage was to identify the source of the meat. Extracting the proteins from the meat and testing them with samples of serum from various species of animal produced a positive reaction with cat serum. Further investigations found that the local animal shelter had been inundated with stray cats for which they could find no homes, so they had painlessly "put them to sleep" with large doses of pentobarbital. Unfortunately they had not buried the cats' bodies deep enough for the eagles not to find them. Once the source of the poison had been identified, the animal shelter was much more careful about the disposal of the dead bodies, and the eagle population is now back to its normal level.

Summary

- Partition between solvents, various types of chromatography and electrophoresis find many applications in environmental monitoring.
- Mass spectrometry – either on its own or coupled to GLC or HPLC, is of great use in archaeological and geological dating, and in the analysis of pesticides and other pollutants.

2.4	***What the student needs to know***
11.2i	It is possible to date objects using a mass spectrometer to determine $^{14}C/^{12}C$, $^{40}Ar/^{38}Ar$, and $^{18}O/^{16}O$ ratios. Combined GC-MS (gas chromatography-mass spectrometry) can be used to detect pollutants in the environment.

2.5 – Key definitions

absorption: the reduction in the intensity of radiation when passed through a sample

adsorption: the attraction of a molecule to the surface of a solid

base peak: the peak in a mass spectrum that corresponds to the most abundant fragment

capillary: a glass tube with an internal diameter of about 1 mm, used to make columns for GLC, HPLC and capillary electrophoresis

carbon dating: a method of dating once-living artefacts containing carbon or its compounds by measuring the $^{12}C/^{14}C$ ratio

chemical shift: (symbol δ) the frequency of absorption of an NMR peak compared to the frequency of the TMS peak, expressed as parts per million

chromatography: the separation of the components of a mixture by their different speeds of movement through/over the stationary phase

δ scale: see chemical shift

diamagnetism: the phenomenon of the electrons in a molecule moving round their orbits so as to produce an induced magnetic field that opposes an applied field

dioxins: stable, poisonous compounds with the formula $C_{12}H_4O_2Cl_4$, produced when chlorinated plastics are incinerated or as a by-product in the preparation of some pesticides

DNA: deoxyribonucleic acid - the double helix polymer that contains the hereditary information of the genes

electropherogram: the visual representation of the results of electrophoresis

electrophoresis: the separation of ions/charged molecules by their different rates of movement in an electric field

electrophoretic mobility: a measure of the extent of movement of a particular charged molecule under standard electrophoresis conditions

fragmentation: the breaking up of a molecule into smaller parts by the cleavage of covalent bonds

gel electrophoresis: electrophoresis using a thin layer of gel supported on a glass plate

genetic fingerprinting: the technique of using gel electrophoresis to create a profile of an individual's pattern of STRs or VNTRs

GLC: gas-liquid chromatography

GLC-MS: a techniques where a mass spectrometer is coupled directly to the outlet of a gas chromatograph, to allow the mass spectrum of each peak to be obtained as soon as it is eluted

heavy water (D_2O): water containing the isotope ^2H (given the symbol D) instead of ^1H

HPLC: high performance liquid chromatography

hydrogen bonding: a relatively strong intermolecular force between a δ^+ hydrogen atom of one molecule and a δ^- nitrogen, oxygen or fluorine atom of another molecule

immiscible: two liquids that do not mix/dissolve in each other

induced dipole forces: see van der Waals' forces

integration trace: the trace on an NMR spectrum that indicates the relative number of hydrogen atoms responsible for each peak in the spectrum

isotopes: atoms of the same element with different numbers of neutrons, and hence different masses

isotopic dilution: a method of estimation of the amount of an isotope present in a samp[illegible] a known amount of a different isotope of the same element and comparing their relative ab[illegible] in a mass spectrum

K_{pc}: see partition coefficient

lattice: a regular arrangement of particles (atoms, molecules or ions) in a crystalline solid

M+1 peak: an ion that has a mass one unit greater than the molecular ion in a mass spectrum, due to the presence of the ^{13}C isotope in naturally occurring carbon

M+2 peak: an ion that has a mass two units greater than the molecular ion in a mass spectrum, due to the presence of the ^{37}Cl isotope in chlorine, or the ^{81}Br isotope in bromine

mass/charge ratio: the ratio of the mass of an ion (in atomic mass units) to its charge (in units of the charge on the electron)

MHz: megahertz (1×10^6 Hz). A unit of frequency of electromagnetic radiation

mobile phase: the solvent in the chromatography process, which moves through the column or over the paper or thin layer

molecular ion: the ion formed by the loss of an electron from the original complete molecule during mass spectrometry

monochromatic: radiation that consists of a single wavelength, or a very small range of wavelengths

MS: mass spectrometry

NMR: nuclear magnetic resonance spectroscopy

partition: the division of a solute between two phases

partition coefficient (K_{pc}): the ratio of the concentrations of a solute in two different immiscible solvents when an equilibrium has been established

PC: paper chromatography

PCBs: polychlorobiphenyls – inert compounds used as insulators in transformers, and plasticisers in PVC, paints and adhesives. They persist in the environment and can cause cancer and other illnesses

PCR: polymerase chain reaction – a method of making many thousands of copies of specific fragments of DNA using the enzyme DNA polymerise

pesticide: a chemical applied to crops to kill weeds, insects or fungi

polarity of a solvent: the extent to which a solvent dissolves polar molecules (i.e. those with dipole moments)

potassium-argon dating: a method of dating rocks by measuring the $^{40}Ar/^{40}K$ ratio

relative abundance: the height of a peak in a mass spectrum compared to the height of the base peak, expressed as a percentage

restriction enzyme: an enzyme that splits the DNA molecule in specific places

retardation factor: see retention ratio

retention ratio: a number between 0 and 1 which is a measure of how far a component has travelled compared to the distance travelled by the solvent front during PC or TLC

retention time: the time taken for a component in a mixture to travel through the column in GLC or HPLC

***R_f*:** see retention ratio

shielding of protons: the screening of a proton from the applied external field by the induced diamagnetic field of the electrons surrounding it

solute: the substance that is dissolved in a solution

solvent: the liquid in which the solute dissolves

solvent extraction: the process whereby a solvent is used to dissolve solutes from a mixture (e.g. leaves or seeds), or whereby a second (immiscible) solvent is used to extract a solute out of another solvent

spin-spin coupling: the interaction between the spins of protons on adjacent atoms, causing their respective NMR absorbances to be split into a number of lines

splitting pattern: the pattern or lines produced when spin-spin coupling occurs

stationary phase: the immobile phase in chromatography, e.g. the paper surface in PC, the surface of the thin-layer particles in TLC, or the involatile liquid adsorbed onto the column in GLC or HPLC

STR: short tandem repeats – very short stable segments of DNA used in genetic fingerprinting

tesla: a unit of magnetic field strength, about 10 000 times as strong as the Earth's magnetic field

TLC: thin layer chromatography

TMS: tetramethylsilane - an inert volatile liquid used as an internal reference in NMR

two-way chromatography: a technique used in PC or TLC wherein one spot of a mixture is placed at the corner of a square sheet and is developed in the first solvent as usual. The sheet is then turned through 90° and developed in the second solvent. This allows a better separation of components having similar R_f values

unit cell: the smallest three-dimensional repeating unit in a crystal lattice

van der Waals' forces: weak intermolecular forces arising from the oscillation of electrons around the nuclei in the atoms in a molecule, causing an instantaneous dipole in one molecule which then induces an attractive dipole on an adjacent molecule

VNTR: variable number of tandem repeats – 5-10-base lengths of DNA that repeat themselves along the chain, used in genetic fingerprinting

X-rays: electromagnetic waves of very short wavelength, about 0.1 nm (visible light has a wavelength about 500 nm)

XRC: X-ray crystallography

2.6 – Resource list

Modern Chemical Techniques; An Essential Reference for Students and Teachers, by Ben Faus[illegible] Royal Society of Chemistry, ISBN 0 85404 286 5

More Modern Chemical Techniques by R Levinson, Royal Society of Chemistry, ISBN 0 85404 929 0

Methods of Analysis and Detection (Cambridge Advanced Sciences) by Anne McCarthy, Cambridge University Press, ISBN 0 52178 724 6

Cutting Edge Chemistry by Ted Lister, Royal Society of Chemistry, ISBN 0 85404 914 2

For an online list of resources in analytical chemistry, see LearnNet, Royal Society of Chemistry, http://www.chemsoc.org/CFLEARNNET/dets.cfm?subj=a

3 – MATERIALS AND DESIGN

3.1 – Introduction

Students will be aware of the many wonders of modern electronic gizmos and gadgets, such as the 'must have' Apple iPod, that are spreading worldwide. These only exist as a result of the work done by scientists. This chapter explains more about the role of chemists in developing new materials and designing new molecules. Taking the iPod as an example, among other features, this has an LCD screen and a rechargeable battery, all in a very small plastic case. LCD screens, rechargeable batteries and the plastic casing all originate from the work of chemists.

This section covers new developments in these technologies, such as the bright new Organic Light Emitting Diodes and the work on fuel cells to replace traditional dry batteries. It also looks at nanotechnology – science on a very small scale and the ways in which chemists are working to improve medical treatments, both in designing new drugs and in making sure these drugs reach the cells in the body where they need to act. Finally it will take a look at the ways in which chemists are working to combat environmental problems, developing new ways of cleaning up contamination and researching alternative fuels.

Designing new drugs to cure diseases

The developments of modern medicine have undoubtedly had a great positive impact worldwide – the development in the last century of antibiotics, transplant surgery and many other life-saving advances have transformed the way we live. But there are ever more challenges for the medicinal chemist. As average life expectancies improve, diseases such as cancer become more common. As bacteria become immune to antibiotics, new antibiotics must be developed. And there is the challenge of new diseases such as HIV/AIDS and bird flu that spread worldwide.

This chapter looks at two of the ways in which chemists are trying to combat disease. Firstly much chemical research is directed at developing and synthesising new drugs to target specific diseases. One way of doing this is to start from natural products – molecules that are synthesised by plants and other organisms – that have been shown to have beneficial effects, for example, the drug Taxol®, found in yew tree leaves. Taxol® has been found to be effective against some cancers. As only small amounts of the Taxol® can be extracted from yew, chemists around the world have established the structure of Taxol® and have sought, and successfully found, synthetic pathways to produce the drug in quantity.

Secondly, chemists are investigating means of getting drugs to the specific part of the body where they need to act. Such methods can reduce side-effects and the quantities of drugs needed. Two of these methods will be studied. One involves liposomes, where, in effect, the drug is delivered in a "bag" to the diseased body site.

Can chemists make artificial silk?

Not as yet, but they are getting closer! A good deal of interest has been directed at silk because it is incredibly strong. Silk is a natural polymer containing amino acids. It is a condensation polymer. This chapter will revisit both addition and condensation polymerisation and look at how chemists have made use of both types of polymerisation to create new products such as conducting polymers as well as a condensation polymer known as Kevlar. Kevlar resists fire and is five times as strong as steel. An early use of Kevlar was to replace the steel in motor tyres. The tyres were lighter and lasted longer than the steel reinforced tyres.

Chemists have become particularly interested in spider silk because spiders produce the strongest silks known. There is considerable potential for strengthening bullet-proof jackets or fireproof clothing using related materials.

Nanotechnology

A modern soccer-ball is a sphere made from pentagons and hexagons; it is the same shape as the "buckyball" molecule, the third allotrope of carbon with the molecular formula C_{60}. Yes, a buckyball contains just 60 carbon atoms! Think of the number of atoms in the extended structures of graphite and diamond. The size of the 60 carbon molecule is just less than a nanometre, and nanotechnology involves the applications of molecules of this sort of size.

Nanotechnology has been described as the science of the very small with big potential. Already many new molecules have been created from carbon that have a similar structure to buckyballs and other materials have been used to create other molecules of similar sizes that have interesting properties.

If nanotechnology involves such small molecules, how can we be certain about their structures? Two new types of microscope are involved, the Atomic Force Microscope (AFM) and the Scanning Tunnelling Microscope (STM). With such equipment it has become possible to move individual atoms around. Researchers at IBM were the first to demonstrate this, in 1990. They used an STM to move Xenon atoms on a nickel surface to produce the letters IBM.

Green chemistry and sustainability

The media makes much of stories of environmental contamination, whether this is of soil, water or air, and chemists often find themselves under fire from the public for causing this pollution. While the previous parts of this section discuss how chemists are involved in advances in technology and medicine, this section encourages students to think about the problems that can occur as a result of these advances, how chemists seek to solve these problems and how they research alternative technologies that are less polluting and make the most of our resources.

Chemists are very much to the fore when it comes to cleaning up sites that have been chemically contaminated. Sites may be contaminated in different ways. For example oil slicks at sea may kill many sea birds as well as other animals. Chemists have now improved existing methods for removing oil slicks by over 100% using materials known as sorbents.

On land, soils are often contaminated when waste products are dumped or accidentally spilled. The method of clean-up can take advantage of the contaminant's physical or chemical properties. For example, organic molecules can be flushed out with solvents in which they are soluble, or they can be degraded by adding other chemicals.

Chemists have also found ways to clean water supplies. For example, water contaminated with arsenic has been cleaned and made fit to drink by using a local plant that absorbs the arsenic. The improvement of soils and water supplies is called remediation.

While there will always be a need to combat pollution, it is clearly preferable if pollutants can be replaced with alternatives. One of biggest pollution stories of recent decades has been the release of CFCs into the atmosphere and the resulting loss of the ozone layer in the stratosphere. This layer of ozone is the Earth's protective sunscreen that shields us from high-energy ultraviolet rays that cause skin cancer. Later in this chapter we will discuss why it is important that chemists understand the mechanism by which these reactions occur, in order to design alternatives.

Concern about the environment has resulted in sustainability becoming a very important part of the work of many chemists. The intention is to reduce the impact on the environment by reducing and recycling materials used. To assist this, the American Chemical Society and the US Environmental Protection Agency have agreed a set of 12 principles to provide a framework for guidance to achieving sustainability.

Possibly the biggest problem facing mankind is the supply of oil. Oil is currently crucial to a developed world, both as a supply of energy and as a raw material. So far we have not come up with alternatives that are sufficient to supply our current or future needs, but there is much research into sustainable alternatives as well as cleaner fuels.

And finally…

You will probably find that many parts of this chapter raise more questions in students' mi[illegible] they answer. If your students want to find out more, they can try visiting the websites listed at th[illegible] of this chapter, do their own websearches, or visit their local library. Many of the examples in [illegible] chapter are from Western companies as they actively advertise their technological advances to demonstrate their value to society, but innovation in science is both an international and a local endeavour and students should look for examples in their own region or local university.

As students are exploring, they should think about how what they find relates to the chemistry that they have learned in the rest of their course. Lastly, they should take a second to step back and think about the huge variety of activities that chemists are involved in worldwide. Consider what their environment might be like without the work of chemists, and how different their understanding of the world around them would be without the understanding they are gaining through the study of chemistry.

3.2 – Medicinal chemistry and drug delivery

By the end of this section, students should be able to:

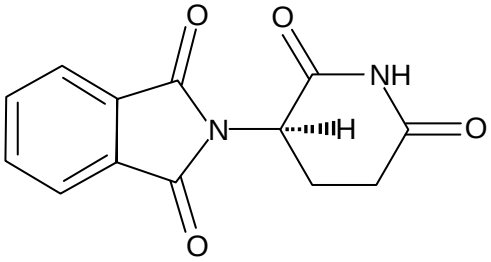
- *discuss the challenges of drug design and explain in simple terms how molecules may be identified and developed to overcome these problems,*
- *discuss the challenges of drug delivery and explain in simple terms how materials may be identified, designed and developed to overcome these problems.*

3.2a Designing drugs

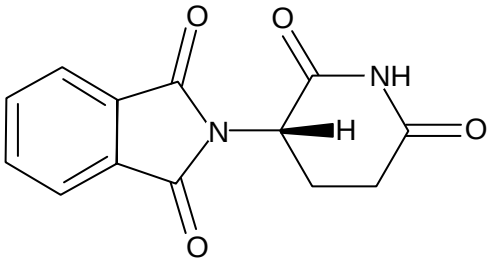
When a molecule is to be used as a drug it is important that it is effective in achieving its desired effect *and* that undesirable side-effects are avoided.

Undesirable effects proved to be a particular problem with the drug thalidomide that was prescribed to pregnant women as a sedative in the early 1960s. Thalidomide was, at this time, the preferred sedative during pregnancy as the alternatives, such as Valium, were addictive.

At that time the thalidomide produced was a mixture of two optical isomers. Unfortunately one of the isomers of thalidomide proved to have disastrous side-effects, causing babies to be born with congenital deformities such as shortened limbs (the drug is said to be teratogenic).



(R) – thalidomide desirable properties: sedative and antinausea drug



(S) – thalidomide teratogenic: causes birth defects

Figure 3.1 – the two enantiomers of thalidomide

Nowadays, there are many reasons why the same mistake is much less likely to be made. One of these is the methods that are available to chemists to design drugs and gain an understanding of their action. Think back to the section of this booklet on the chemistry of life and the action of enzymes. Students will have learned about the lock and key model of enzyme catalysis and about competitive inhibition. Drug molecules act by binding to receptors, and in many cases these receptors are enzymes. Drugs can be competitive inhibitors of enzymes and if a drug has optical isomers, only the isomer that is complementary to the shape of the enzyme active site will fit.

In order to bind to its receptor a drug must not only have the shape to fit, but must also be able to interact with the groups on the receptor molecule by hydrogen bonds, ionic interactions or dipole-dipole interactions. Chemists are able to use computer simulations to model how their drugs will fit into the receptor site. They can also search databases to see if their drug will interact with other enzymes to get an idea of possible side-effects. These computational methods have proved very powerful both in designing new medicines and in understanding how drugs act.

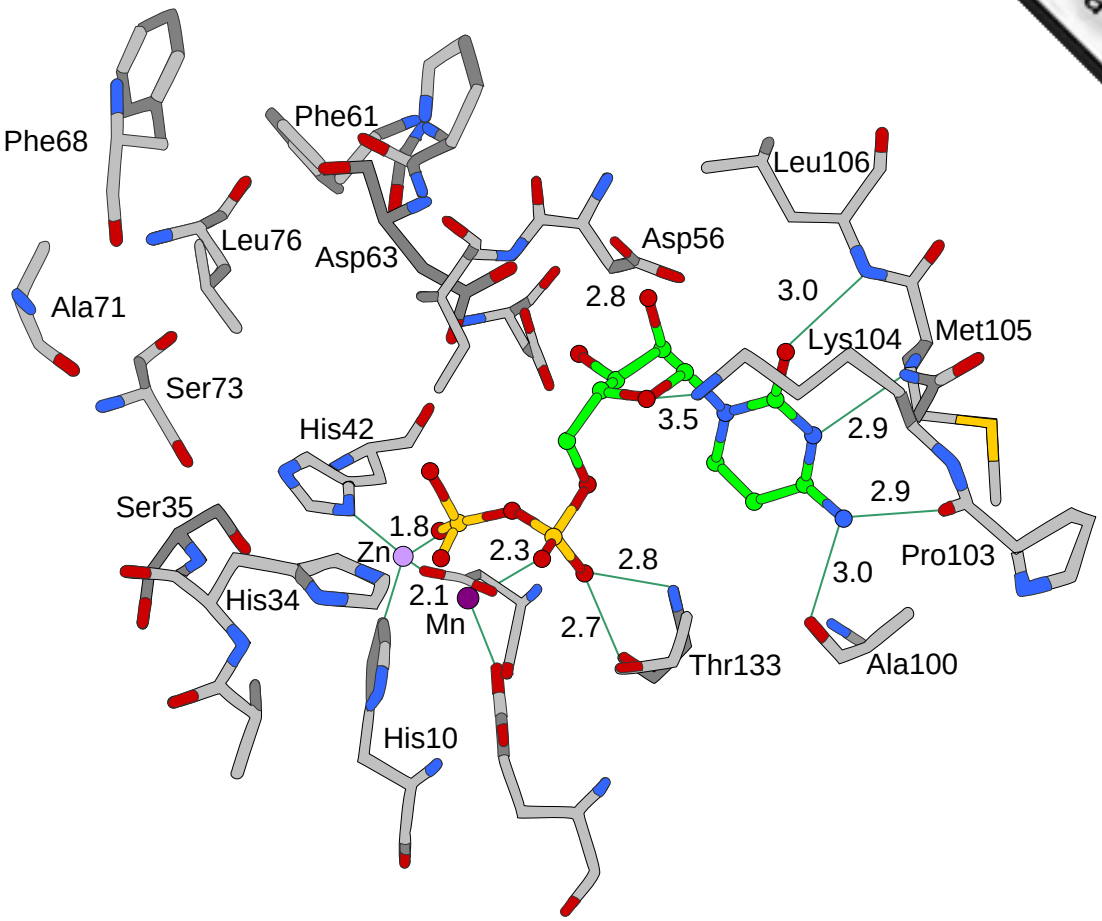
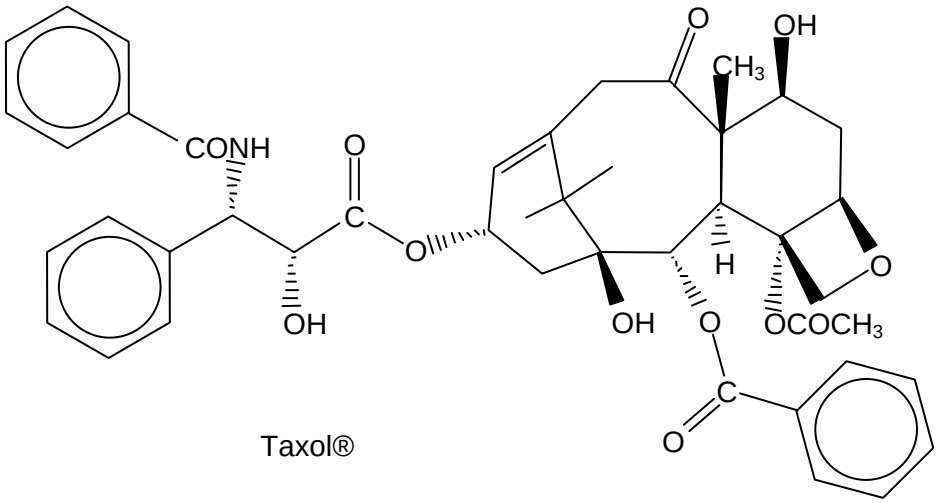


Figure 3.2 – computer-generated model of a drug in the active site of an enzyme, showing hydrogen bonds

Another reason for only using one optical isomer is economics. Where a synthetic route to a pharmaceutical produces two chiral molecules, chemists seek a route that produces only the active molecule that is desired. This process is called **asymmetric synthesis**. A key reason for this development is that when two chiral molecules are produced from a non-chiral starting material, they are normally produced in equal quantities, so half of the reactants are wasted. Asymmetric synthesis therefore saves on resources and costs.

Actually, these changes alone would not have had an effect on the thalidomide story, as the molecule can switch between enantiomers at the pH of the blood. So, even if the desirable enantiomer is administered, it would turn into a mixture in the bloodstream within hours. A successful example of asymmetric synthesis, however, is in the treatment of Parkinson's disease. This disease causes much suffering and is characterised by tremors in the hands and loss of balance, but a molecule called L-dopa (the L enantiomer of dopa) can alleviate the symptoms. The L-dopa must be free of D-dopa, as the latter has many unpleasant side-effects. Chemists now make pure L-dopa for use by patients, and it does not change to D-dopa in the body.

When chemists are seeking a new pharmaceutical they may start from a *natural product* molecule (a molecule synthesised by a plant or other organism) that is known to have a positive effect. Students will probably be familiar with the story of aspirin, a chemical derivative of the salicylic acid present in willow bark and leaves. A more recent example of a natural product being used medicinally is the anti-cancer drug Taxol®. Taxol® is found in yew trees. It acts by binding to protein molecules in the cell and preventing the cell from dividing. However, only small amounts of Taxol® can be isolated from yew trees, so it became important to find ways of synthesising Taxol®. As shown in Figure 3.3, it is an incredibly complex molecule with many different functional groups.



Taxol®

Figure 3.3 – structure of Taxol®

SAQ 1. *See how many different functional groups you can identify in the molecule of Taxol®.*

Synthesising a molecule as large and complex as Taxol® is a major challenge and requires a sound knowledge of many different kinds of reactions, many of which students will have encountered in the organic chemistry part of the course. The race to synthesise Taxol® from basic starting materials was taken on by many groups of chemists worldwide. To do this requires chemists firstly to know the structure of the Taxol®, and then work out ways of making the drug. The structures of such molecules can be worked out by the NMR and X-ray techniques looked at in the previous chapter. Chemists may also use techniques such as computer modelling in the process to plan the best route to take.

The first groups achieved their goal in 1994, but research continues today as chemists are keen to minimise the number of reactions needed to make the synthesis more efficient in terms of cost and resources. They are also keen to make similar molecules with slightly different shapes and functional groups in order to try and find a molecule that is even more effective and with fewer side-effects.

3.2b Delivering drugs

As already shown, it is a major challenge to make drug molecules that interact with specific targets in the cell, but there is a second part of the challenge of finding effective therapies, and that is to get the drug molecule to that target site. This process is termed “drug delivery”.

We have already looked at the different functional groups of Taxol®. Now think about the different environments and barriers that a drug might encounter on its journey to the target cell. If it is in the form of a pill that is swallowed, it will encounter the acid pH of the stomach and the enzymes that are in the intestines to break down food. It will then have to pass from the stomach or the intestines into the bloodstream, which has a different pH and more enzymes. When it reaches the right cell it will have to get through the cell membrane. And how does it know which cells to target? Clearly there are many problems with delivering drugs, one of which is how to get the drug to its target in one piece and avoid it being broken down by enzymes or degraded in extremes of pH.

One successful method of delivery involves the use of liposomes. These are artificial microscopic vesicles consisting of an aqueous core enclosed in one or more phospholipid layers. A phospholipid is a molecule that is hydrophilic (water-loving) at one end and hydrophobic (water-hating) at the other end. Hence, in water-based solutions such as blood, lipids group together to form double layers with their hydrophilic groups on the outside, forming polar interactions with the water, and their hydrophobic groups on the inside of the layer, away from the water. These bilayers can wrap into spherical vesicles.

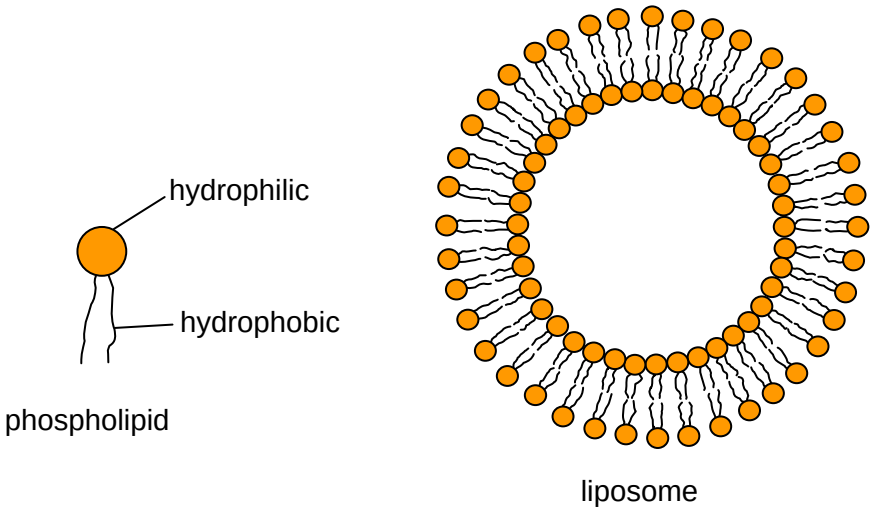


Figure 3.4 – a phospholipid and a cross-section of a liposome

Liposomes are biodegradable and non-toxic and can be used to carry vaccines, drugs, enzymes, or other substances to target cells or organs. They can carry both *hydrophilic* molecules (polar molecules that form hydrogen bonds with water and hence dissolve) and *hydrophobic* molecules (non-polar molecules that do not dissolve in water).

SAQ 2. *Indicate on the diagram where you think the liposome will carry hydrophobic molecules, and where it will carry hydrophilic molecules.*

As liposomes are made from biological molecules they are easily degraded by the body. By modifying the surface of liposomes biochemists have developed long-life liposomes, which do not degrade quickly and have a better chance of reaching their target. Once the liposome reaches its target, the drug can be transferred to the target.

Interested?
If students found liposomes interesting, they should try searching for on the web for virosomes that have diameters of around 150 nm. They are a biological application of nanotechnology (see section **3.3**).

A second method of protecting drugs while they are circulating in the bloodstream is to attach them to polymers. A popular polymer to use is polyethylene glycol, or PEG.

$$HO–(CH_2–CH_2–O)_n–H$$

When the polymer chain is quite short, i.e. when *n* is a relatively small number, PEG is soluble in water. The disadvantage of PEG compared with a liposome is that it can only carry two drug molecules.

SAQ 3. *Draw a PEG molecule with three repeats. Indicate the two places on the molecule where the drug can be attached. Suggest a reaction and conditions by which you could attach a drug molecule with a carboxylic acid group.*

To improve the number of drug molecules that a polymer can deliver, different polymers can be made where some of the monomers have side-chains that can link to the drug molecules.

SAQ 4. *Look back at the organic chemistry part of your course and suggest a pair of functional groups that could be coupled together when one is the side-chain of the polymer and the other is in the drug molecule.*

Summary

- The aim of drug design is to produce drugs that achieve their desired effect with m[illegible] undesirable side-effects.
- Many drug molecules are chiral and it is important that only the enantiomer with the desired effect is used.
- Asymmetric synthesis – the synthesis of just one enantiomer – saves resources.
- Drug molecules act by binding to receptors such as enzymes and must have the correct shape and functional group to interact with the receptor.
- Natural products are often a good starting point for drug design.
- For maximum effect it is important to deliver a drug safely to the point at which it acts.
- Lipsomes and polymers can be used to protect drugs from degradation and deliver them to their target.

3.2	***What the student needs to know***
11.3a	What is meant by *asymmetric synthesis*, The reasons why asymmetric synthesis is desirable when making a new drug molecule. In outline, how a drug molecule interacts with a receptor or enzyme. That many drugs are often derived from natural products.
11.3b	How liposomes and PEG can be used to help deliver drugs to their target organs, and how they work to achieve this.

3.3 – Properties of polymers

By the end of this section, students should be able to:

- *discuss the properties and structures of polymers based on their methods of formation (addition or condensation),*
- *discuss how the presence of side-chains and intermolecular forces affect the properties of polymeric materials (for example spider silk).*

Students have already encountered an application of a polymer in the previous section on drug delivery. In this section they will refresh their knowledge of polymerisation and explore further properties and applications of polymers.

3.3a Addition polymerisation

Check the students' understanding of addition polymerisation with the following SAQ.

SAQ 5. ***(a)*** *Draw the structure of poly(phenylethene).*

(b) *Write an equation to show the addition polymerisation of three poly(phenylethene) molecules.*

(c) *Circle a repeating unit in your diagram of poly(phenylethene).*

(d) *Explain why this polymerisation is called addition polymerisation.*

Polymers made from alkenes contain only carbon and hydrogen atoms. The physical properties of polymers are determined by the intermolecular forces present in the polymer. The electronegativity difference between carbon and hydrogen is small, so the only intermolecular forces present in these polymers are van der Waals' (instantaneous dipole-induced dipole).

The properties of addition polymers can be modified in a number of ways. Addition polymers tend to deform easily and once deformed do not return to their original shape. Students could try stretching a piece of plastic cut from a "soft" plastic bag (not the sort that rustles) placed in contact with their lips. They will find that the plastic film becomes warm as it is stretched. They will also find that the material does not return to its original shape. Generally, the longer the polymer chains, the stronger the van der Waals' forces.

Aside from the length of the polymer chain, another factor is the number of side-chains. Generally, unbranched chains can pack together better than polymers with lots of side-chains. This is the difference between plastic bags that rustle and those that are "soft". The "soft" bags are made from low density poly(ethene) (LDPE), which has lots of side-chains and is relatively weak and easy to deform. The type of bag that rustles is made from high density poly(ethene) (HDPE) which has fewer side-chains. It is less flexible and has a higher melting point. HDPE is also used to make bottles for milk and fruit juices.

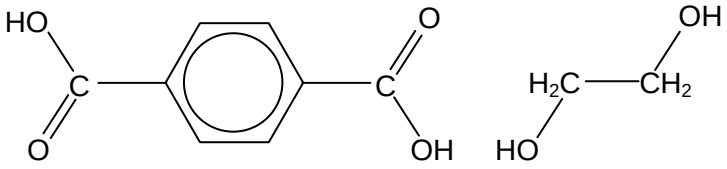
The properties of a polymer are also affected by different kinds of bonding, for example the presence of chlorine atoms in poly(chloroethene) results in some permanent dipole interactions between carbon and chlorine because of the polarity of the carbon-chlorine bond.

3.3b Condensation polymerisation

Unlike addition polymerisation, where an alkene molecule joins to itself, condensation polymerisation usually requires two different molecules that can react together to form a bond such as an ester or amide bond with the elimination of a small molecule such as water. Examples include synthetic polymers such as polyesters and polyamides as well as peptides and proteins.

Polyester is widely used to make a variety of items from drink bottles to clothing and carpeting. It is made by polymerising ethane-1-2-diol with 1,4-benzenedicarboxylic acid (terephthalic acid) with the elimination of water. The product is widely known as *Terylene*. The 1,4 links in this polymer produce a linear polymer that is suitable for spinning into fibres.

The formulae of the two monomers are shown below. Check the students' unde[illegible]f condensation polymerisation using the SAQs.



1,4-dicarboxylic acid ethane-1-2-diol

SAQ 6. ***(a)*** *Write an equation to show the formation of **one** repeating unit of the polyester chain.*

(b) *The amino acid alanine has the following structure – $H_2NCH(CH_3)CO_2H$.*

Draw the structure of the tripeptide formed by three molecules of alanine

(c) *How many water molecules are lost in this condensation reaction to form the tripeptide?*

SAQ 7. *Cellulose is the polymer responsible for the strength of fibres such as cotton. Cellulose is a linear polymer of sugars with many –OH groups. Suggest, in terms of bonding, why cellulose is so strong.*

3.3c Spider Silk

Based on weight, spider silk is five times stronger than steel of the same diameter. There are records where bullets have not broken through a silk handkerchief. George Emery Goodfellow (a doctor at Tombstone, Arizona, US) wrote in the spring of 1881, "I was a few feet from two men quarrelling, they began shooting, two bullets pierced the breast of one man, who staggered, fired his pistol and crumpled onto his back. Despite fatal injuries, not a drop of blood had come from either of the two wounds."

Further investigation located a bullet wrapped within a silk handkerchief. It appears that the bullet had passed through clothes, flesh and bones but not his silk handkerchief. More recently, it has been suggested that a strand of spider silk as thick as a pencil would stop a jumbo jet in flight!

Spider silk is a protein that is in the same group of proteins as hair, nails and ligaments. The Golden Orb-Weaving spider produces a dragline silk (a dragline connects a spider to its web) that is the strongest form of spider silk. The protein in dragline silk is called fibroin. There are actually seven different kinds of spider silk, each being produced by different glands. Each type has a different function. Students could see what they can find out about the other kinds of silk on the web or in the library.

Fibroin has a molecular mass of 200 000 – 300 000 and consists of 42% glycine and 25% alanine, with the remainder coming from just seven other amino acids. The alanine molecules occur in polyalanine regions, where between four and nine alanine molecules are linked in a block. The elasticity of spider silk comes from regions that are rich in glycine. In these regions a sequence of five amino acids is repeated. After each sequence a 180° turn occurs producing a spiral. Ordinary silk, produced by silk moths has a β-pleated sheet structure, held together by hydrogen bonds, see below.

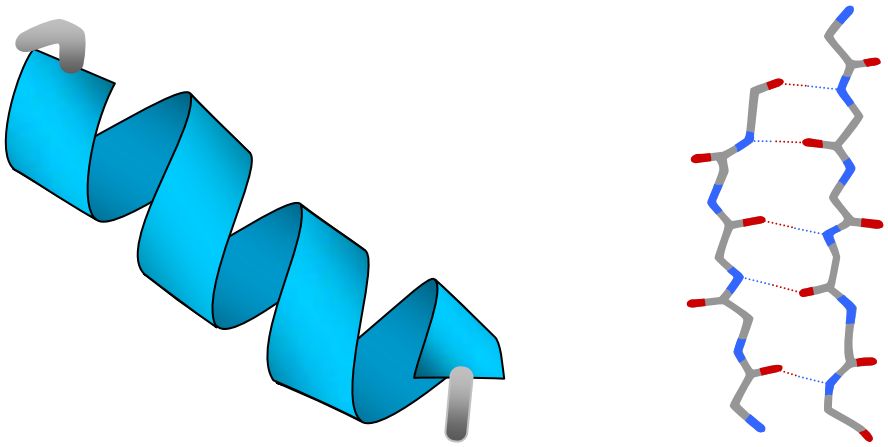


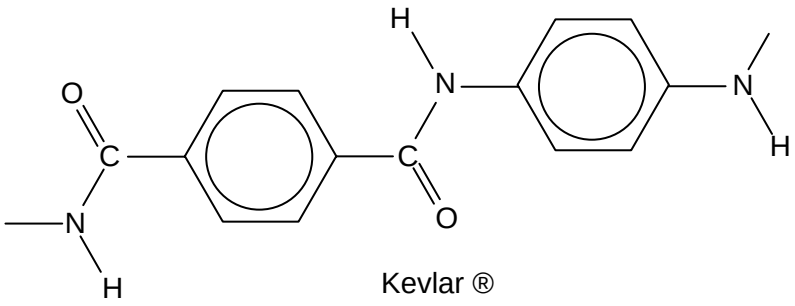
Figure 3.5 – the spiral structure of spider silk and the β-pleated sheet structure of ordinary silk

The most elastic spider silk is 'capture silk' that has about 43 repeats (of the 5-glycine fragment) and can extend to 200% of its length. Dragline silk with only about 9 repeats can only extend by about 30% of its length.

To find out more about the different types and uses of spider silk visit websites given at the end of the chapter.

Clearly the properties of spider silk are very valuable and chemists are researching man-made polymers that have similar properties. Kevlar® is a very tough polymer produced by DuPont™ that is similar in strength to spider silk. Kevlar® is used for bullet-proof vests; re-enforcing Kevlar® with spider silk would make these vests even stronger. Kevlar® has replaced steel in rubber tyres, with the weight reduction leading to a small reduction in fuel consumption.

To find out about Kevlar® production students should visit the websites given at the end of the chapter or see what they can find in the library.



Kevlar ®

SAQ 8. ***(a)*** *When you have carried out this research, write an account of the uses of Kevlar®, together with the reasons for each use.*

(b) *Copy the above structure and then draw another copy alongside your first structure so that N–H aligns with C=O.*

(c) *What type of bonding do you think will occur between these two groups?*

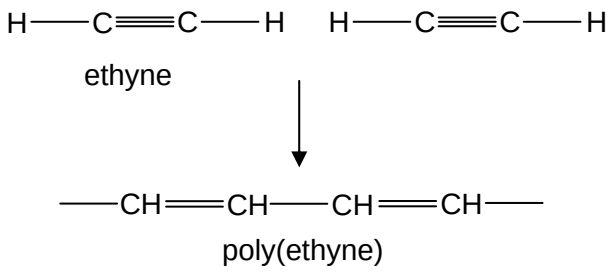
(d) *Explain how this bonding arises.*

(e) *Name another type of polymer that forms spirals with the same type of bonding.*

SAQ 9. *Which type of polymerisation, addition or condensation, would you expect to produce polymers that are biodegradable? Explain your answer.*

3.3d Plastics that conduct electricity or emit light

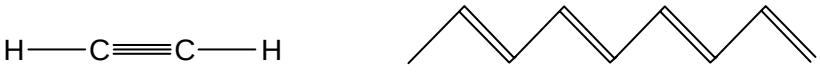
Conducting polymers were discovered by accident by a Japanese student. Polymerisation of ethyne (acetylene) produces poly(ethyne) by addition polymerisation. This material has alternating single and double bonds. Poly(ethyne) has two forms, cis and trans.



SAQ 10. *Draw sections of poly(ethyne) containing three ethyne units to show the cis and the trans isomers.*

The two isomers have different colours, ***trans*-poly(ethyne)** is blue or silver coloured; ***cis*-poly(ethyne)** is red or copper coloured. Molecules that have alternating single and double bonds have "conjugated systems". Benzene is another molecule with a conjugated system of π bonds. The realisation that ***trans*-poly(ethyne)** had conjugated π bonds led to the discovery that this polymer could conduct electricity! Until this discovery, our expectations of polymers were that they do **not** conduct electricity. Three scientists, Alan Heeger, Alan MacDiarmid and Hideki Shirakawa were awarded the Nobel Prize for Chemistry in 2000 in recognition of their work on conducting polymers.

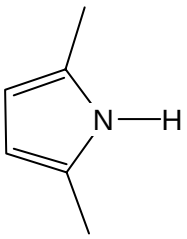
The conjugated system in ***trans*-poly(ethyne)** is shown below.



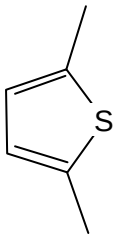
ethyne

trans-poly(ethyne)

Other conducting polymers include compounds such as poly(pyrrole) and poly(thiophene). Technically, these conducting polymers are semi-conductors, unlike the nanotubes in the section on nanotechnology.

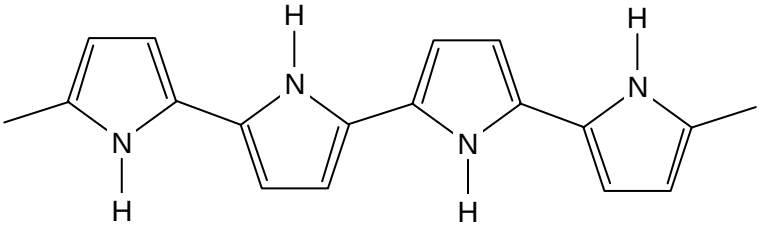


poly(pyrrole) monomer



poly(thiophene) monomer

Poly(pyrrole) has the structure below.



SAQ 11.(a) *Circle the repeating section of poly(pyrrole)*.

(b) *Draw a section of poly(thiophene).*

Uses for conducting polymers are surprisingly few, despite the interest raised by their discovery. For these polymers to conduct, they need to be 'doped', meaning that some electrons are removed (by oxidation) or introduced (by reduction) leaving 'holes' allowing the electrons (or the 'holes') to flow.

Most of the interest in conducting polymers lies in electronics. These polymers are of interest to both physicists and chemists. In the 1980s AGFA engineers had a major problem with their photofilm production. Static discharges were ruining the film, which was very costly to produce. The problem was found to be due to the inorganic salts AGFA traditionally used as an antistatic coating. This coating failed when the humidity dropped below 50%. The parent company Bayer AG used its research team to develop a new antistatic coating. Surprisingly, a conducting polymer poly(thiophene), was found to overcome the problem.

Another use is 'Smart' windows that have been developed to reduce glare from sunlight. [illegible] are coated with a conductive polymer in contact with a layer of black particles. When [illegible] passed through the polymer, these particles align and let light through. When the current is s[illegible] they become disordered and block light.

The traffic lights are changing

Particular interest lies in semi-conducting polymers that have been developed into light-emitting diodes, solar cells, displays for mobile phones and wristwatch size television screens.

Traditionally, traffic lights have been lit with a single bulb that shines through coloured glass. However, we all know that ordinary light bulbs are likely to fail sooner or later, causing confusion at the lights. These new lights, as shown in Figure 3.6, contain an array of OLEDs, organic light emitting diodes. There is no sign of the light bulb that has been replaced by this array. If one of the OLEDs fail, there are still plenty left, so motorists will be able to safely cross on the green.

Figure 3.6 – OLED (organic light emitting diode) traffic lights in London, UK

Light emitting polymers were first discovered in the Chemistry Department at Cambridge University and have now become a major area of research bridging both chemistry and physics. One of the biggest challenges was finding a blue OLED and this has now been achieved.

OLED displays are appearing in a number of applications. For example Kodak have designed a camera with an OLED screen instead of a liquid crystal display screen. The OLED screen has two advantages as it can be viewed even in sunlight as light is being emitted. In addition it has a wider viewing angle. However, current OLEDs, particularly the blue ones, have a shorter lifetime than liquid crystal displays.

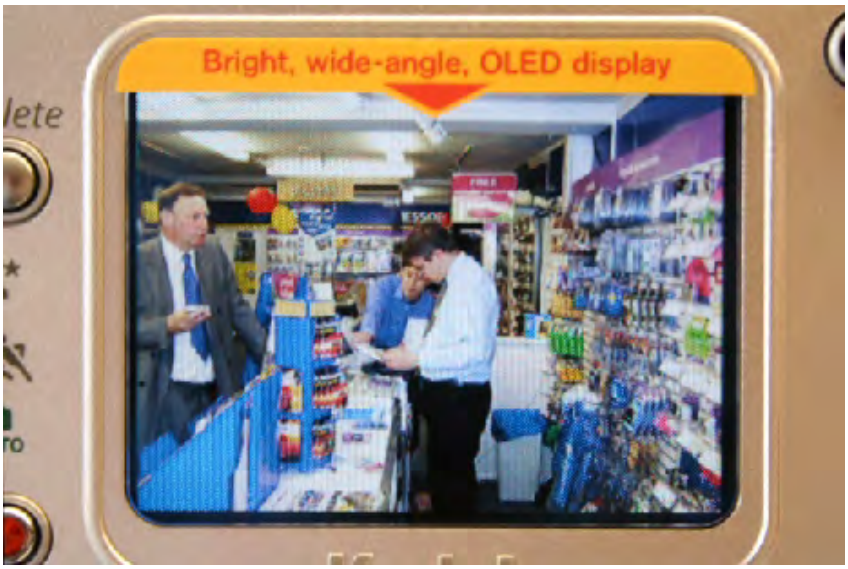
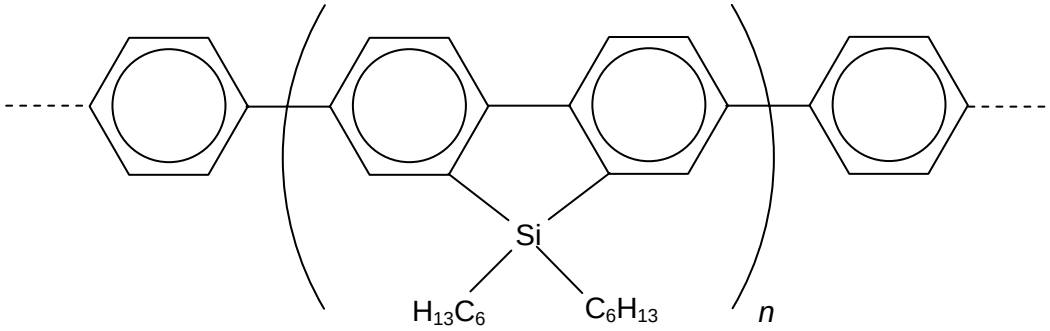


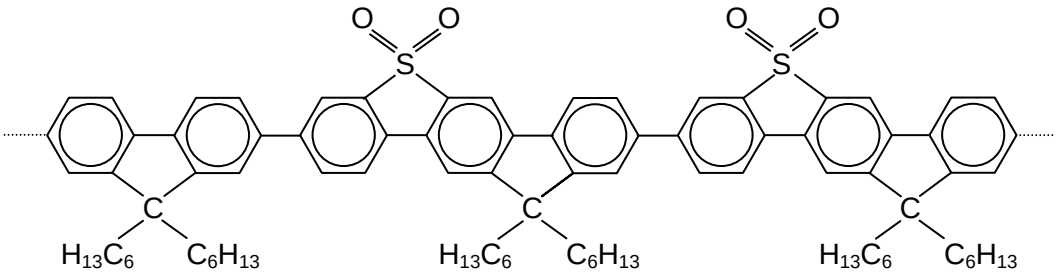
Figure 3.7 – photo of screen on Kodak camera, courtesy of Jessops

Research published in 2005 has found blue-emitting materials that may overcome the problem of the shorter lifespan of blue pixels in OLEDS compared to the red and green-emitting pixels in OLED displays.

A team from Cambridge in the UK created the blue-emitting polymer, shown below.



A second independent team, working between the Donetsk University in the Ukraine and the University of Durham, UK has discovered a similar blue-emitting material. The structure of their polymer is shown below.



SAQ 12. *Identify the repeat unit in this polymer by drawing brackets and adding a label 'n' as in the first polymer above.*

Summary

- Addition polymers are formed from monomers with carbon-carbon double bonds incorporate all of the atoms of the monomer into the polymer.
- Addition polymerisation is used to form simple hydrocarbon polymers.
- Hydrocarbon polymers tend to deform easily. Their properties are primarily dependent on chain length and van der Waals' forces between chains and can be modified by the number and nature of side-chains.
- Conjugated hydrocarbon polymers can conduct electricity or emit light.
- Condensation polymers are formed when two different monomers combine with the loss of a water molecule.
- Condensation polymerisation can result in polymers with a variety of properties, for example elasticity, that depend on their functional groups and side-chains.
- Many natural polymers are condensation polymers.

3.3 What the student needs to know

11.3c+d The characteristic differences between addition and condensation polymerisation. (This is mostly a revision of section 10.8 of the syllabus)
The formation of conducting and semi-conducting polymers – i.e. those that contain long chains of conjugated double bonds.

11.3d The general nature of spider silk (i.e. proteins that use mainly glycine and alanine as their constituent amino acids), and some of its properties.

3.4 – Nanotechnology

By the end of this section, students should be able to:

- *show awareness of nanotechnology and, given information and data, be able to discuss the chemistry involved with reference to the core syllabus.*

3.4a What is nanotechnology?

Nanotechnology has been described as "the science of the very small with big potential". "Very small" in this context means of an order of nanometres – Table 3.1 shows just how small a nanometre is. The technology covers chemistry, physics, biology and related sciences such as materials science. As with other new science developments, the word nanotechnology is slowly becoming familiar to the general public. Much of the news coverage has been very positive, placing the emphasis on the excitement and promise of the science. However, there have also been scare stories predicting overwhelming numbers of nanorobots and grey goo!

So what is nanotechnology and why are scientists so excited about it? Do the public really have cause to be worried, or are we worrying because it is new and we don't understand it? The following section aims to give some background information so that students can start to form their own opinion on this complex issue.

Getting down to nanometres

One nanometre is 0.000000001 m. It can be written as 1 nm or 1×10^{-9} m. Table 3.1 shows the scale of length showing where nanometres fit in.

Table 3.1 – the scale of length

Small	attometre	am	0.000000000000000001 m	1×10^{-18} m
	femtometre	fm	0.000000000000001 m	1×10^{-15} m
	picometre	pm	0.000000000001 m	1×10^{-12} m
	nanometre	**nm**	**0.000000001 m**	**1×10^{-9} m**
	micrometre	μm	0.000001 m	1×10^{-6} m
	millimetre	mm	0.001 m	1×10^{-3} m
	centimetre	cm	0.01 m	1×10^{-2} m
	metre	m	1 m	1×10^{0} m
	decametre	dm	10 m	1×10^{1} m
	hectometre	hm	100 m	1×10^{2} m
	kilometre	km	1000 m	1×10^{3} m
	megametre	Mm	1000000 m	1×10^{6} m
	gigametre	Gm	1000000000 m	1×10^{9} m
Large	terametre	Tm	1000000000000 m	1×10^{12} m

The metre is the standard SI unit of length. Every other unit is stated as a number bigger or smaller than this. The short word put before metre is called a prefix. Many of these are from Greek. The same prefixes are used to change the unit of mass, the kilogram, into smaller and larger units. Atoms and molecules are nano- and picometre sized. Science involving nano-sized particles is called nanoscience.

3.4b Buckyballs – a new allotrope of carbon

A bucky ball, or to give it its full name, buckminsterfullerene, is composed entirely of carbon ato[illegible] is a third allotrope of carbon. A total of 60 carbon atoms are present forming a sphere consisting [illegible] five-carbon and six-carbon atom rings arranged in the same pattern as the seams on a modern soccer ball. It is just less than a nanometre in size.

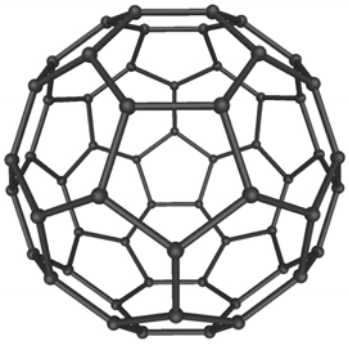


Figure 3.8 – a buckyball, or buckminsterfullerene

SAQ 13.(a) *Explain what is meant by an allotrope.*

(b) *Name the other two allotropes of carbon.*

Buckminsterfullerenes were initially discovered in 1985 during experiments with carbon clusters in supersonic beams. Three scientists shared the Nobel Prize for Chemistry in 1996 for their work on the discovery of fullerenes - Sir Harold Kroto, Robert F Curl Jr and Richard E Smalley. As well as C_{60}, other sized balls have been created. Unlike other forms of carbon, fullerenes may be soluble, as shown in the photograph below. C_{60} is pink and C_{70} is red in solution.

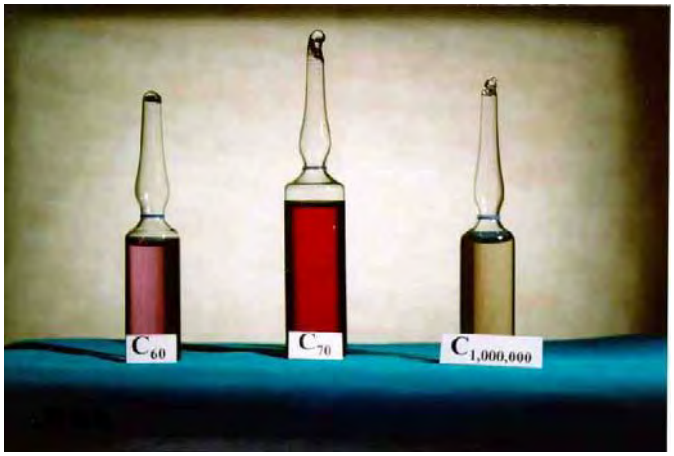


Figure 3.9 – buckyballs in solution

Buckyballs have unusual properties which led to a lot of excitement about their potential. It has been suggested that they

- may be harder than diamond
- may be more slippery than Teflon
- may be insulators or conductors.

These seemingly contradictory suggestions arise through a confusion of the micro and macro properties of buckminsterfullerene. Let us take each of these suggestions in turn, and see how it applies to (a) an isolated molecule, and (b) a bulk sample of the material.

- The bonding within a single molecule of buckminsterfullerene is very strong. The bucky ball is very unlikely to be easily squashed or deformed. So the molecule itself could be described as being “hard”. However, the fact that buckminsterfullerene has a low melting point, and is

soluble in organic solvents, suggests it should have the typical properties of [illegible]d whose intermolecular forces are mainly van der Waals'. So in a bulk sample, we wo[illegible] buckminsterfullerene to be soft, like a polycyclic hydrocarbon such as anthracene (C_{14}[illegible]

- Because they are hard spheres, the molecules of buckminsterfullerene could be looked u[illegible] as nano-sized ball bearings! As students will know from their previous studies of chemistry, graphite is very slippery, because the sheets of hexagonally-arranged carbon atoms can easily slide over one another. Many greases contain graphite. Perhaps in the future "super greases" will contain buckminsterfullerene.
- In buckminsterfullerene the carbon atoms are all joined by delocalised π bonds. So it would be expected that an electric current should easily pass around the molecular sphere. However, transferring the current from one molecule to the next is another matter. Electrons do not find it easy to jump from one molecule to another. In bulk form, samples of the simpler molecules that contain delocalised electrons (such as benzene) are good insulators, although NMR shows that electrons readily move around the aromatic ring within each molecule (see section 2.2b). However, if the delocalisation is extended in space, such as in the long conducting polymers mentioned in section 3.3d, then the energy gap between molecules can be bridged.

These restrictions that apply to the relatively small buckminsterfullerene molecule may be overcome by designing and synthesising larger molecules that contain buckminsterfullerene-type elements. The next section mentions the early work being done in this area.

A buckyball can enclose an atom of another element inside itself. This can be a reactive element or molecule such as a lanthanum atom. The highly reactive atom becomes trapped like a 'tiger in a cage' – while it is protected by the carbon cage it cannot react, but as soon as the cage is removed it can react again. The structure below shows a lanthanum atom in the centre of a buckyball.

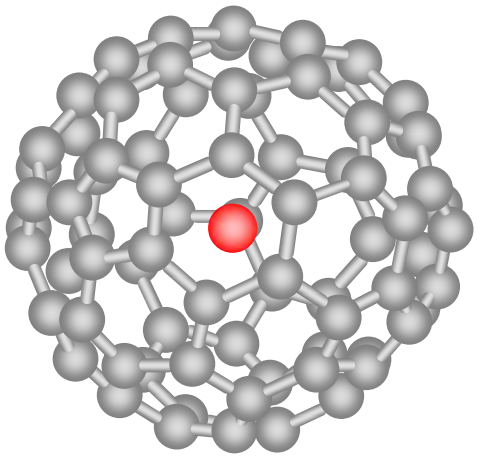


Figure 3.10 – lanthanum atom caged in a buckyball

Carbon nanotubes

The discovery of buckyballs led to the discovery of other forms of carbon that are structurally related, for example, carbon nanotubes. These are cylindrical in structure and also resemble a rolled-up sheet of graphite, with the carbon molecules arranged in repeating hexagons. They have a diameter of a few nanometers and can be open at both ends, sealed at one end or sealed at both ends.

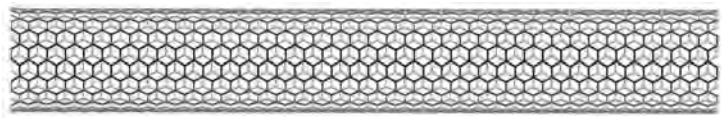


Figure 3.11 – a carbon nanotube open at both ends

Carbon nanotubes have proved to have very useful properties. 'Mini but Mighty', they are [illegible]s stronger than steel. The mechanical (stiffness, strength, toughness), thermal and electrical [illegible] of pure buckytube materials enable a multitude of applications, from batteries and fuel cells t[illegible] and cables to pharmaceuticals and biomedical materials. They are found in the batteries of [illegible] laptop computers.

The world's smallest test-tube has been made from a carbon nanotube and has been accepted for the Guinness Book of World Records. One end of the tube is closed by a fullerene cap that contains both pentagons and hexagons. The tube has a volume of 1×10^{-24} dm^3. The tube has enabled Oxford scientists to use the tube's one-dimensional cavity to provide a template for the synthesis of unbranched polymer chains of $C_{60}O$, fullerene oxide.

Figure 3.12 – a carbon nanotube closed at one end – a nano "test-tube"

This process normally requires expensive catalysts that are sensitive to air and water. Without the one-dimensional cavity the polymer would branch in all directions.

The potential of nanoscale test-tubes for carrying out reactions is being explored. The nanoscale test-tube below has various derivatives attached. These could, for example, be immobilised enzymes enabling fast reactions in the synthesis of new drugs. There are distinct advantages of carrying out these reactions on the nanoscale. In a normal test-tube, the particles have to collide to react and these collisions rely on random movement. In many reactions not all the particles react, or unwanted side-products are produced, as students will have found in the organic chemistry section of the course. These problems result in reduced yield. By contrast, a nanoscale reaction, where individual molecules are brought together, can have an exceedingly high yield.

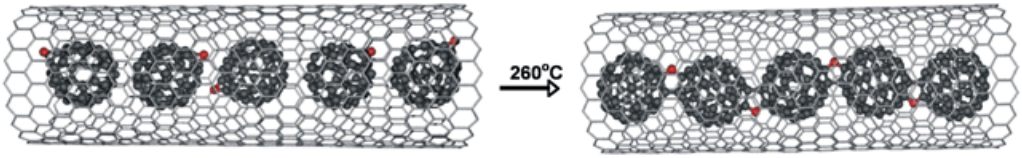


Figure 3.13 – a chemical reaction in a nanotube – polymerisation of $C_{60}O$ to form $(C_{60}O)_n$

Another kind of structure being developed involves buckyball cages containing trapped atoms. These buckyball cages are then entrapped inside a nanotube, rather like peas in a pod. The purpose of this work by scientists in Oxford is to investigate the structure as an information storage display.

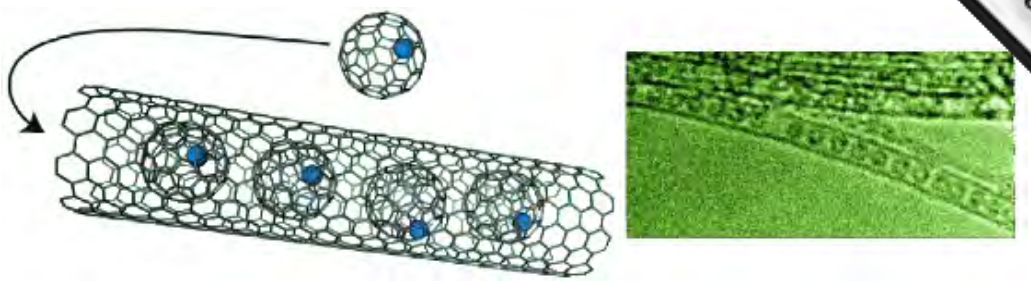


Figure 3.14 – "peas in a pod" – buckyballs containing trapped atoms, themselves inside a carbon nanotube

3.4c Supramolecular chemistry – making super-molecules

In the section on medicinal chemistry we looked at some of the very large drug molecules that chemists make. We also looked at assemblies of phospholipids to form vesicles held together by hydrophobic and hydrophilic effects. Supramolecular literally means "beyond the molecule". The phrase was first coined by the French chemist Jean-Marie Lehn. Together with Americans Donald Cram and Charles Pedersen, Lehn won the Nobel prize in 1987 for his work on molecules that recognise each other. In the same way that biological molecules such as enzymes recognise and bind other molecules, Lehn, Cram and Pedersen created synthetic molecules called crown ethers that recognise and bind (form a complex with) metal ions.

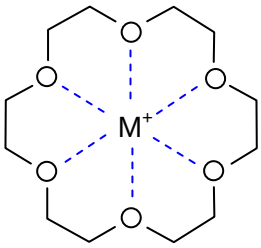


Figure 3.15 – a metal ion bound by a crown ether molecule

SAQ 14. *Suggest the type of bonding holding the metal ion in the centre of the crown ether.*

The field has developed in many directions with chemists synthesising ever more complex and finely-tuned super-molecules from molecules that recognise each other and bind to each other by non-covalent effects, including hydrogen bonding and van der Waals' forces. These super-molecules are being designed to be catalysts, to transport drugs, to transmit electricity and to harvest light, among many other things.

In 1999 a research group in the Netherlands made the first molecular motor. The motor is powered by light and the molecule rotates about a carbon-carbon double bond. The groups either side of the double bond are identical and UV light causes these to undergo cis-trans isomerisation. Because of the large size of the groups, which are chiral, the motor can only rotate in one direction.

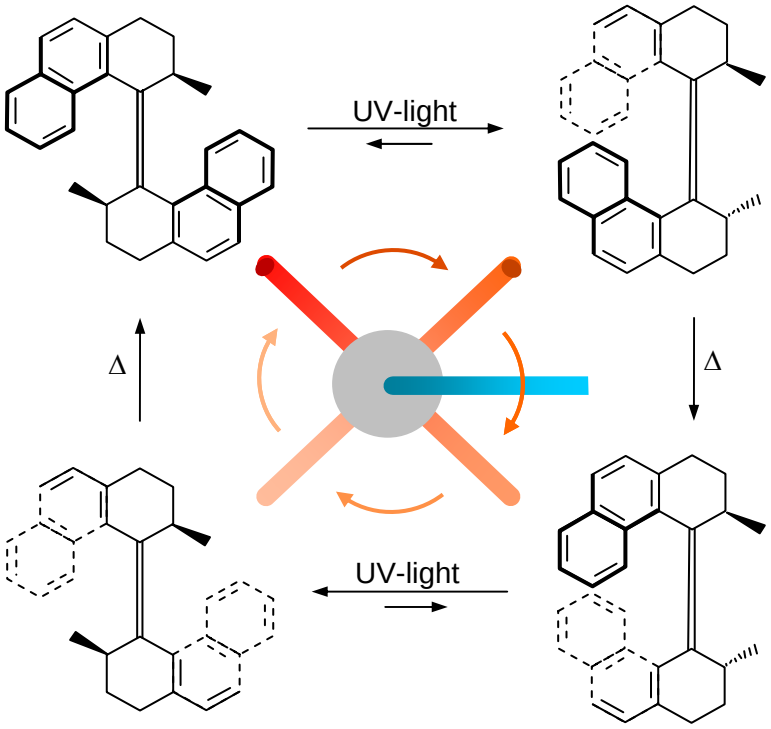


Figure 3.16 – the first light-driven molecular motor

3.4d Quantum dots

Developments in quantum communications and computing could be about to get much easier thanks to the development by Toshiba of a device capable of emitting photons one at a time.

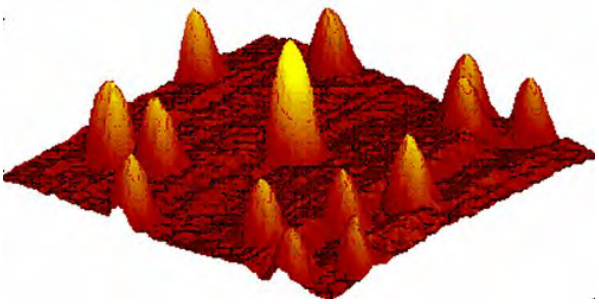


Figure 3.17 – a quantum dot

The device is based on quantum dot technology. Quantum dots are devices that contain trapped electrons. They are made from semiconductor materials such as silicon and have dimensions that are typically between a few nanometres and a few micrometres. The size and shape of these structures, and therefore the number of electrons they contain, can be precisely controlled; a quantum dot can have anything from a single electron to a collection of several thousands.

The remarkable thing about quantum dots is that although they are macromolecules, they behave more like atoms than like the bulk material. When excited, they emit light, and the dots can be tuned to emit light of a specific wavelength, and even to emit single photons. Prior to this development, single photon sources were extremely difficult to make and relied on either using a laser or exciting single atoms. It was hard to prevent multiple photons being emitted. In contrast, Toshiba's quantum-dot emitter reliably generates single photons on demand when excited by short optical pulses.

It is hoped that the quantum devices will be used in optical computers capable of calculations so vast that today's computers could not even finish them.

This is just one example of how differently nano-scale materials can behave from the b[illegible]
For further examples of the unusual properties of nanoparticles, try searching the web for s[illegible]
following examples and for some more.

- Kodak Ultima inkjet paper has a layer of ceramic nanoparticles. A unique blend of additi[illegible] is used to "fix" the dyes in the bottom two layers. The top layer contains ceramic nanoparticles to further stabilise the image. Kodak scientists have achieved a significant leap in longevity and colour reproduction with prints predicted to last for up to 100 years.
- Most modern full spectrum sunscreens contain a variety of compounds that absorb UVA and UVB. Some sunscreens contain particles of titanium dioxide. This compound does not absorb UV but reflects the harmful ultraviolet light away from the individual being protected. Titanium dioxide is often found in products described as 'sunblock'. In the sunscreens the particle size is close to nano-size. At this size it still reflects UV light but doesn't give such a white appearance as larger particles. However, there are some concerns about such small particles being able to penetrate the skin.
- Nanocomposites are materials that have a combination of hardness and toughness that is far greater than the component materials. They are made by grinding down clays or ceramics to a nano-sized powder. The powder is mixed with a polymer to form the material.

By now, students may well be wondering how we can obtain images of some of the devices that we have been discussing. The answer is the Atomic Force Microscope (AFM) together with the Scanning Tunnelling Microscope (STM). These instruments allow us to see individual atoms on surfaces. We can also move atoms around on a surface.

There are worksheets at the end of this chapter on manipulating atoms and pictures that show what can be done.

3.4e Promise and possible problems of nanotechnology

At the beginning of the chapter we mentioned that while scientists are very excited about the potential of nanotechnology, many people have concerns about problems that the new technologies might cause.

The examples in this chapter have demonstrated that nanoparticles can have very different properties from their related bulk materials and this means that they need to be handled differently. Many applications only require very small amounts of nanoparticles, so this reduces risks considerably. However, some uses involve large quantities, for example sunscreens. Large scale manufacture can carry the same risk of explosion as production of other materials which have a small particle size and hence large surface area, for example, milling flour. Like any new chemical products, a full risk assessment is required, both for the production of new materials and their subsequent uses.

Scientists predict that the uses of nanotechnology will go far beyond the current applications, and will be used, for example, in environmental remediation, power transmission and disease diagnosis and treatment. Clearly humans and the environment will be increasingly exposed to nanoparticles, and trials must be carried out to ensure their safety. It is the responsibility of scientists to carry out these trials and assess the risks. New technologies always carry new risks and worries and it is also the responsibility of the scientists to engage in debate with the public to ensure their concerns are addressed and the scientific facts of the technology are communicated.

Summary

- Nanotechnology involves particles whose size can be measured in nanometres.
- A nanometre is 0.000000001 m or 1×10^{-9} m.
- Nano-sized particles have many useful properties that can be different from bulk materials, due to, for example, different surface area to size ratios.
- Super-molecules have the potential to act as catalysts, to transport drugs, to transmit electricity, to harvest light, and many other applications.
- As with all new technologies that have the potential for widespread application, any risk must be adequately assessed.

3.4	***What the student needs to know***
11.3e	The usual sizes of molecules range from 0.1 nm to 10 nm. The structure of buckminsterfullerene and some of its properties. The structures of graphene and nano test-tubes. Some of the uses of nanotubes.

3.5 – Environment and energy

By the end of this section, students should be able to:

- *discuss how a knowledge of chemistry can be used to overcome environmental problems (for example, ground water contamination, oil spillage, CFCs),*
- *discuss how a knowledge of chemistry can be used to extend the life of existing resources and to improve the efficiency of energy production and use*.

3.5a Chemistry to overcome environmental problems

Oil slicks

A major environmental problem is the formation of oil slicks when oil is spilled from oil tankers whilst at sea. There have been many instances where tankers have been sunk, either by collision with other ships or in rough seas. The spilled oil floats on the surface of the water and washes ashore with both immediate and long-term effects on bird and fish populations. It is crucial to the survival of these species, and to local tourist or fishing industries, that such slicks are prevented from spreading and cleaned up as quickly as possible.

A new method for preventing slicks from spreading is to make booms that can be used to soak up oil spills. Porous materials called *sorbents* exist that will soak up oil. However, these materials also soak up water. This makes them much less effective as they soon sink.

In an effort to overcome this problem, chemists at Lawrence Livermore National Laboratory in California have recently increased, by a factor of over 100%, the ability of these sorbents to absorb oil. This discovery promises the development of much more effective sorbent booms.

This major improvement has been achieved by making a porous fibreglass sorbent boom that repels water and allows oil to be absorbed. The secret is to trap fluorinated molecules in the structure of the fibreglass sorbent. The fluorine has hydrophobic properties and so repels water but allows oil in.

To make the new sorbent the researchers dipped fibreglass into a slurry of silica and then into a chemical containing fluorine. The resulting material is dried and then cut into discs. When tested by shaking with a mixture of crude oil and salt water, as present following a tanker spillage, the discs absorbed over 200 times their weight in oil, and they did not sink. The untreated fibreglass absorbed both water and oil and sank. Scientists now need to develop sorbent booms from this new approach to cleaning up oil slicks.

Remediation of contaminated soils

Soil remediation involves the improvement of contaminated land to achieve soil that is the same as it was before the contamination. There are two main approaches to cleaning soil. One is to “wash” the contaminants out of the soil, the other is to chemically degrade the pollutants. The first relies on knowledge of the physical properties of the pollutant, and the second on knowledge of the chemical properties.

Soil remediation using physical properties

Pollutants can move through the soil by diffusion or convection. Diffusion occurs where there are concentration differences in molecules, spreading out the contamination. Convection occurs when

molecules are driven by a fluid such as rain or wind. Water picks up particles as it moves [illegible] soil, carrying them further from the initial spill. These mechanisms that spread the spillage [illegible] the process of cleaning up the contaminated soil. In order to wash pollutants from the gr[illegible] suitable solvent must be chosen that will dissolve the contaminant. To wash out oil-b[illegible] contaminants such as diesel fuels, a trick can be borrowed from the oil industry. Before an oil wel[illegible] completely dry the oil needs to be forced out and a viscose polymer solution is used. To flush contaminants from the ground, a similar viscous fluid or foam can be pumped through the soil from a hole on one side of the region and pumped out of a hole on the other side, with the contaminants dissolved in the foam.

Similarly, gases can be used instead of liquids to transport the pollutant through the soil. Some contaminant chemicals are fairly volatile (they form a gas easily) and can be flushed out of the soil simply by pumping air through. Other contaminants can be made to vapourise by heating the soil with warm air and heating coils.

Another property of contaminants that can be exploited is electrical charge. By inserting electrodes into soil, any charged or polar contaminants can be made to move towards one electrode, depending on whether they are positively or negatively charged. This is the same principle as the electrophoresis experiments discussed in section 2.3c. A disadvantage is that other molecules in the soil will also be charged and will also move towards the electrodes. When this method is used, the soil near the electrode is removed and treated further before being returned.

Soil remediation using chemical properties

Some soil remediation will happen of its own accord as bacteria in the soil break down complex organic chemicals into CO_2. This process is fairly slow and while it can be sped up by adding oxygen or nutrients for the bacteria, it can still take one or two years. Another way of breaking down the contaminants is to add chemicals to the soil.

Polyaromatic hydrocarbons (PAHs) are a major pollutant of contaminated soils. They result from incomplete burning of carbon-containing materials and have structures based on benzene rings. PAHs are frequently *carcinogenic* (that is they may cause cancer). Ozone has been shown to break down PAHs and can be pumped into the ground to degrade the pollutants as it passes through the soil.

SAQ 15. *Using your knowledge of oxidation processes, charge and solubility of transition metals, suggest how heavy metal contamination could be dealt with.*

Remediation of contaminated ground water – plants to the rescue

There are a number of plants that are capable of cleaning up contaminated ground water. *Ground water* is the water that is present below ground. It may be present in the soil or it may be held in porous rock such as chalk. Chalk is particularly good at filtering and purifying water supplies and is used a lot in the UK, France and Namibia. Much mineral water comes from chalk aquifers.

Often, ground water is of vital importance to residents living above it. In Bangladesh, many new wells were sunk to provide an adequate water supply. However, the ground water was not properly tested before being made available to residents. People became ill from drinking this water and they were found to have arsenic in their bodies. The arsenic had come from the rocks surrounding the ground water. Long term consequences of arsenic poisoning include skin cancer, damage to the nervous system and miscarriages. This arsenic contamination is a very serious threat to millions of people.

Chemists at the De Montfort University in the UK decided to try the effect of a notorious water plant known as water hyacinth. This plant causes a problem by clogging waterways in tropical and subtropical areas, including Bangladesh. Previous research at Bangabandhu Sheikh Mujib Medical University in Bangladesh had shown that the whole plant was able to remove arsenic from water and the recent research exploits the affinity of the plant for arsenic. The chemists used powdered dried roots of the water hyacinth to treat arsenic-contaminated water and found that the powder removed arsenic from the water, reducing the level to below the World Health Organisation's guideline value of

0.01 mg dm^{-3}. They believe that powdered root will be easier to use than the whole plant [illegible] effective than other methods of removal.

3.5b Trouble in the stratosphere – replacing CFCs

CFCs were introduced as a refrigerator compressor liquid by the engineer Thomas Midgley in the 1920s. When they were introduced it was in order to replace existing liquids that were toxic, such as ammonia or sulfur dioxide, with something safer.

Unfortunately CFCs were found to cause a huge environmental problem. CFCs are particularly stable and do not break down until they get to the stratosphere. In the stratosphere CFCs absorb ultraviolet light that causes a photo-dissociation of carbon-chlorine bonds as shown below for CF_2Cl_2.

$$CF_2Cl_2(g) \rightarrow CF_2Cl\bullet(g) + Cl\bullet(g)$$

These radicals catalyse the breakdown of ozone to oxygen. In recent years the ozone layer has been thinned significantly over Antarctica, and to a lesser extent over Northern Europe. Each year that passes produces a thinner layer of ozone.

Why should this trouble us? Primarily because the ozone layer provides us with an important sunscreen that prevents the harmful ultraviolet light from damaging our skin and causing skin cancers. The incidence of skin cancers has increased since the use of CFCs.

SAQ 16.(a) *Explain what is meant by the term CFC and suggest an example that is different from the one above.*

(b) *CFCs are not the only compounds that can damage the ozone layer. Search the web for another compound that contributes to ozone loss.*

The first step in addressing this problem was for chemists to understand how CFCs catalyse the breakdown of ozone. Students will have previously learned about the relative strengths of carbon-halogen bonds. As the carbon-chlorine bond is very reactive, it photodissociates in UV light, as shown above.

SAQ 17. *Explain why the carbon-chlorine bond photodissociates and the carbon-fluorine bond does not.*

Chemists have now provided some alternatives to CFCs. In general these are hydrofluorocarbons, for example CH_2FCF_3. Alkanes may also be used. The presence of the C–H bonds is important because this enables the compound to break down before it reaches the stratosphere. And if it does reach the stratosphere, it can't produce the damaging chlorine free-radicals.

CFCs have a second adverse effect on the environment – they contribute to global warming. The new replacements for CFCs may be better in terms of the ozone layer but they are still greenhouse gases. However, they are present in the atmosphere in much smaller quantities than other greenhouse gases such as carbon dioxide.

There are ongoing programmes to monitor ozone levels and there is some evidence that the hole has started to reduce in size. Scientists hope that the 21st century will see the recovery of the layer.

3.5c Green chemistry and sustainability

Increasing awareness of environmental issues has led to much innovation in chemical research. The challenge for chemists is to develop products and processes that are sustainable i.e. they do not impact on the environment in terms of pollution or depletion of resources. To help chemists work towards this aim, twelve principles of green chemistry have been drawn up.

The twelve principles of green chemistry

1. Prevention of waste is cheaper than cleaning it up once it is formed.
2. The synthesis of a new chemical product is designed to ensure that maximum use of materials takes place.

3. New chemicals are produced with no significant toxicity to humans or to [illegible] environment.
4. New chemicals are designed to achieve their use, whilst keeping toxicity low,
5. Use of solvents should be avoided.
6. Energy efficiency should be maximised, ideally reactions should be carried out at roo[illegible] temperature and pressure.
7. Feedstocks should be sustainable.
8. Chemists often need to use temporary modifications to a compound in a synthetic route to a target compound. Such modifications should be minimised.
9. It is better for a reaction to use a catalyst than a reaction that is not catalysed.
10. Chemical products should be designed to break down naturally.
11. Analytical monitoring of reaction processes enables prevention of production of hazardous materials.
12. Safer chemistry reduces the risk of accidents.

SAQ 18.(a) *Using L-dopa as an example, explain principle 2.*

(b) *Using your knowledge of catalysis, explain principle 9. How might use of a catalyst also be good economically? Suggest an instance where use of a catalyst might also be viewed as environmentally non-friendly.*

The following are examples where chemists have employed these principals to design a green process or technology.

Use of supercritical CO_2 as a solvent

A supercritical fluid is a gas that is compressed and heated so that it shows properties of a liquid and a gas at the same time. Carbon dioxide becomes a supercritical fluid at a pressure 7290 kPa and a temperature of 31 °C. Other gases that fairly easily form supercritical liquids include xenon and ethane. The diagram below shows the supercritical region for a gas.

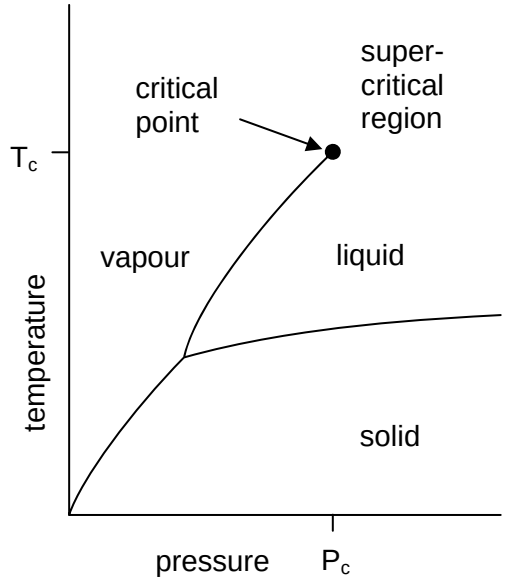
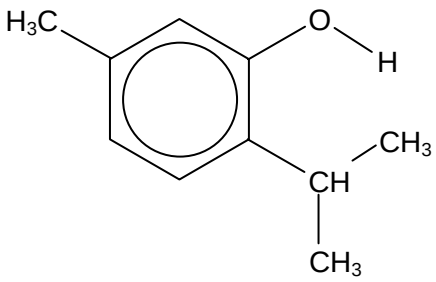


Figure 3.18 – diagram showing the supercritical region

Thymol has the structure shown below. It is used as a disinfectant, a starting material for perfume manufacture, and it also has medicinal properties. Thymol is also used to make menthol.

SAQ 19.(a) *Search the web for the structure of menthol.*

(b) *Suggest reagents that could be used to convert thymol to menthol.*



thymol

A collaboration between researchers at Nottingham University and two companies, Thomas Swan and Co. in England and Schenectady Pratteln in Switzerland, resulted in a process for synthesising thymol using supercritical CO_2 as a solvent and a more environmentally friendly catalyst, anhydrous aluminium oxide.

The use of supercritical CO_2 as a solvent avoids the need for organic solvents, many of which are volatile, flammable and may pose a risk to health and the environment. At the end of the reaction it is easy to separate the product from the solvent – the conditions can be changed so the CO_2 turns to gas. The CO_2 used is a by-product of fermentation, so a waste product is used, and it can be re-used again and again. What's more, by controlling the temperature and pressure, its properties can be fine-tuned to minimise the production of by-products in the reaction and to increase the yield. The process makes chemical reactions possible that were previously too polluting or inefficient.

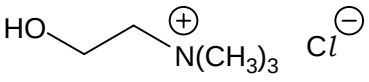
SAQ 20.(a) *Which of the 12 principles does this process fulfil?*

(b) *Search the web for other chemicals produced using supercritical CO_2.*

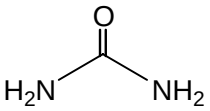
(c) Suggest two hazards of using supercritical CO_2.

Ionic liquids

Ionic liquids are organic salts that have melting points below 100 °C – some are even liquid at R.T.P. Ionic liquids have attracted much interest as solvents for chemical reactions as they can dissolve a wide variety of inorganic and organic compounds. For example, the ionic liquid *Reline 203* has a melting point of 12 °C. It is a eutectic (low melting) mixture of choline (melting point 303 °C) and urea (melting point 13 °C). It readily dissolves ionic compounds such as LiC*l* (its saturated solution contains 2.5 mol dm^{-3}), and even CuO (0.12 mol dm^{-3}) which is very **in**soluble in water. It also dissolves organic compounds such as benzoic acid (0.82 mol dm^{-3}) and even cellulose!



choline chloride



urea

There is, however, much debate as to how green they are. Ionic liquids, unlike organic solvents, have no vapour pressure and can be re-used. However, little is known about their toxicity or their effects on the environment.

Chromium plating is a highly hazardous process that makes use of chromic acid, a highly toxic and cancer forming compound. A company called Scionix has produced an ionic liquid from chromium(III) chloride and choline chloride. This liquid has been shown to reduce the risks to chromium plating personnel, by replacing the previously-used poisonous Cr(VI) compounds with Cr(III) compounds. It also increases the overall current efficiency of the process from 15% with chromic acid to 90% with the ionic liquid.

In this example significant improvements have been made, even though the ionic liquid cannot be classed as green. The only thing green about chromium(III) chloride is its colour!

This process fulfils around eight of the twelve green chemistry principles.

SAQ 21. *For this new chromium plating method, suggest at least four of the twelve green chemistry principles that fulfil these principles.*

Another use of ionic liquids is in electro-polishing of metals such as iron or aluminium, where ionic liquids have replaced sulfuric and phosphoric acids.

Rock-munching bacteria

When the Romans invaded Britain, they were interested in obtaining metals such as copper, tin and lead. In Wales, records from that time tell of copper being made in pits. It is believed that these pits contained bacteria that literally 'eat rock', dissolving copper in the process. The copper was displaced from solution by throwing in worn out iron articles such as swords.

The bacteria involved, *Thiobacillus ferro-oxidans and Thiobacillus thio-oxidans*, are now routinely used in copper mining in the US such as at the huge Bingham Canyon mine.

Figure 3.19 – copper mine

Copper mining produces huge piles of waste that still contain copper. However, it is only recently that bacteria began to be used to extract this remaining copper. The process is termed "bacterial leaching". The bacteria create an acidic solution, dissolving the copper from what was waste ore. The copper is displaced by adding scrap iron, just as the Romans did. 10% of US copper now comes from bacterial leaching.

As their names suggest, the bacteria actually use Fe^{2+} and S^{2-} ions in their metabolic processes, not the copper. They gain their energy to live from oxidation reactions such as the following.

$$Fe^{2+} \rightarrow Fe^{3+} + e^{-}$$

$$S^{2-} + 2O_2 \rightarrow SO_4{}^{2-}$$

The bacteria do not actually feed off the copper ore; the ions do not enter their cells. Overall the bacteria produce a solution of iron(III) sulfate and copper(II) sulfate. The process of bacterial leaching is cheaper, more energy efficient, quieter and less polluting that other methods of extraction.

In Australia, a different problem exists. Gold is traditionally extracted, from ore mixed with oxide or sulfide ores of other metals, using cyanide. However, the cyanide only extracts 10% of the gold and is very toxic to all life. Using bacterial leaching can raise the percentage of gold extracted to 85-100%. However, *Thiobacillus thio-oxidans* prefers a temperature of about 30 °C for optimum performance, much lower than the region where the gold mining takes place.

Thiobacillus thio-oxidans uses its unusual metabolism to turn pyrites (iron sulfide) and arseno-pyrites ores into iron oxides. Any gold in the material can then be removed.

SAQ 22. *Discuss the benefits of using bacterial leaching over the use of cyanide.*

3.5d Chemistry for energy

Biofuels: diesel and ethanol fuels

The use of ethanol as a fuel has received much attention in recent years as a potential alternative to oil-based fuels. It makes use of local resources and is less polluting. For many years Brazil and Columbia have powered their cars with ethanol made from sugar cane. Many countries are overproducing grain crops and are researching ways of using such crops for the production of fuels. In South Africa, chemical plants are being constructed to produce ethanol from maize, to be blended with petrol in proportions ranging from 10% to 85%. In the US, the sugar in maize is also being used to produce ethanol and in Europe, sugar beet is being used to produce ethanol, and oilseed rape is being used to provide a fuel suitable for diesel-powered engines. As well as being used to produce ethanol, biomass can be used as fuel in other ways, for example, in Mauritius, "bagasse", the solid waste from sugar production, has been used to power electricity stations, allowing far more efficient use of biomass.

SAQ 23.(a) *Find three other materials, other than sugar cane, sugar beet and maize, that can be used to produce bioethanol.*

(b) *Bioethanol is widely thought to be less polluting than burning fossil fuels, but some people disagree. Carry out some research to find arguments for and against the use of bioethanol as a clean-burning fuel.*

But is bioethanol really green? Perhaps the biggest surprise has happened in the US, where the incentive to use bioethanol is to help mid-west farmers make use of the grain overproduction. Approximately 40 million tonnes of maize are being converted to bioethanol and the market is growing. The fuel is being sold as E85, where 85% is ethanol, the remainder petrol. However, the energy available from this fuel only just exceeds the energy used in the farming to produce the crop. The crop results in 1.2 units of energy so there is a small net gain of 0.2 units of energy. The process is not as efficient as it could be.

Another question that arises is whether we can make better use of the bioethanol, rather than using it all for transport, for example, by using it to generate electricity.

There are other alternatives to oil-based fuels that are being researched and used. Biodiesel is one of these. The term *biodiesel* refers to fuels that can be used in place of diesel and that are usually made from vegetable oils or animal fats.

SAQ 24. *Find out whether your country has an initiative to produce bioethanol or biodiesel.*

Batteries and fuel cells

A fuel cell is an electrochemical device similar to a battery but unlike a battery, which runs down over time as the reactants it contains are used up, a fuel cell is designed to use a continuous external supply of reactants. A typical fuel cell is a hydrogen cell and the reactants are hydrogen and oxygen. The attraction of fuel cells is that, with the reactants being hydrogen and oxygen, the only product is water. They are therefore emission-free.

SAQ 25. *View the video on the following website, or do your own web or library search for information on hydrogen fuel cells. Discuss the electrochemistry taking place in a hydrogen fuel cell. Include equations for reactions that take place.*

http://www.utcpower.com/fs/com/bin/fs_com_Page/0,5433,03540,00.html

Both alkanes and hydrogen are used in fuel cells to propel vehicles. A prototype Audi A2 has been built that runs on hydrogen, achieving 94 mpg, but with a limited range. Interestingly, the same car, with a diesel engine achieves the same fuel consumption.

Norwegian road construction giant Mesta aims to cut their CO_2 emissions from their vehicles in half by use of hydrogen powered vehicles. They took delivery of their first vehicle in March 2005, hydrogen is stored in three 115 dm^3 bottles at a pressure of 200 bar, giving the vehicle a 120 km range.

The Toyota Prius hybrid petrol/electric car has an electric motor that is driven by a ba[illegible] metal hydride batteries that recharge when the car is coasting down hill and switches to pe[illegible] acceleration is required.

There are some obvious difficulties with running cars powered by hydrogen. Where does hydrogen come from? How can it be stored safely?

Overcoming the hydrogen storage problem

Hydrogen gas must be compressed in order for it to be stored at a small enough volume on a vehicle. This carries a significant risk of explosion.

An alternative approach is to find a solid material to absorb the hydrogen. The hydrogen is bonded to the material and is released by a chemical reaction when needed. The material itself can be recycled. So that the storage material doesn't make the fuel cell too heavy for the vehicle, the elements it is composed of should be in the first two rows of the Periodic Table. A team from the US Department of Energy and the New Jersey Institute of Technology are investigating sodium aluminium hydride $NaAlH_4$, a solid known to reversibly absorb hydrogen molecules. High temperatures are needed and this hydride (and related hydrides) can only store about 5% by weight of hydrogen. The leader of the above team, Graetz, found that doping this compound with titanium made the absorption and release of hydrogen much more efficient. It appears that, unexpectedly, the titanium acts as a catalyst, forming a compound called titanium aluminide on the surface of the $NaAlH_4$.

Another method is to use porous materials which have many molecular sized holes that can absorb hydrogen and release it when needed. Other researchers are investigating organic polymers that can form porous materials.

Sourcing hydrogen

The issue of the hydrogen source for fuel cells is perhaps the biggest issue with this technology. It is attractive to think of the cars on our streets being replaced by cars that only emit water, but the processes used to produce hydrogen are not necessarily environmentally friendly. One option is electrolysis of water, but electricity is required for this process, and in most countries that electricity still comes primarily from power stations that burn fossil fuels. The electrolysis of water usually requires 50% more energy than is stored in the hydrogen produced. Another source of hydrogen is from methane in natural gas, but this process uses up natural resources and generates greenhouse gases. Further research is needed both into sources of hydrogen and into alternative fuel cells before this technology can truly be classed as green.

SAQ 26. *Carry out some research on the web to find out about environmentally sustainable methods of producing hydrogen.*

The future of nuclear power, potentially the greenest of fuels?

The Australian Federal Education and Science Minister, Brendan Nelson, has broken from his party's policy by arguing that use of nuclear energy is a way to cut down on greenhouse gases. Australia is a country with uranium mines.

Meanwhile the UK is again looking at introducing a new generation of nuclear power stations and Iran is looking at nuclear power. The French generate 80% of their electricity from nuclear power and other countries have shown interest in nuclear power.

Nuclear power has always been a controversial issue and there are many arguments both for and against. Some of the key issues for use of nuclear power are as follows.

- Nuclear fuel can be viewed as a clean source of energy with zero emissions of greenhouse gases. In contrast, coal or gas fired power stations may release carbon dioxide, sulfur dioxide and oxides of nitrogen. Only sulfur dioxide is being removed at present (producing a useful by-product of gypsum for plasterboard). Biomass has the same problem, it burns to release CO_2.

- Current use of fossil fuels amounts to 20 million 'fossil fuel years' per year. A te[illegible] (1 × 10^9 kWh) of electricity produces about 20 tonnes of nuclear waste compared [illegible] 20 million tonnes of CO_2 from fossil fuels.
- Nuclear power stations generate radioactive waste products and we must be sure that t[illegible] can be stored safely well beyond our lifetimes. These can be contained most easily as so[illegible] waste and stored underground. The work of chemists has much to contribute in terms of cleaning up nuclear waste.
- There is concern that an accident at or a terrorist attack on a nuclear plant would have much more devastating and long-term consequences than an accident at or attack on a gas or oil-fired power station or pipeline. Unlike Chernobyl, modern nuclear power plants are designed to be fail-safe but opponents of nuclear power argue that the risks are still too great.

Clearly there are arguments, both environmental and economic, for and against nuclear power. Some people believe that nuclear power is a cheap source of energy, others believe that the cost of cleaning up and securing nuclear waste is prohibitive. James Lovelock, a world famous ecologist, says that nuclear power is the environmentally friendly option. Other environmentalists take the opposing position. Chemists have a role both in optimising the process and in investigating waste disposal.

SAQ 27. *Draw a table to compare the benefits and disadvantages to the environment of the technologies you have learned about. Include a column to list areas that require further research by chemists.*

Summary

- Knowledge of chemical and physical properties of pollutants can inform the development of solutions to environmental problems.
- Increasing awareness of environmental issues had lead to increased chemical research into sustainable products and processes.
- The twelve principles of green chemistry summarise the aims for sustainability.
- Chemistry has a large role to play in the development of alternative energy sources.

3.5	***What the student needs to know***
11.3f	The following examples of how some environmental problems can be solved through a knowledge of chemistry: soil remediation, water remediation, the CFC problem and its solution.
11.3g	At least some (!) of the principles of green chemistry. The use of CO_2 as a solvent. The use of ionic liquids as solvents. The use of bacteria to recover metals from low grade ores. The production of bioethanol and biodiesel. The use of fuel cells, using hydrogen or alcohols as fuels. Some of the arguments for and against nuclear power.

3.6 – Key definitions

allotrope: an allotrope of an element is a different crystalline or molecular form of the same ele[illegible]

asymmetric synthesis: making one enantiomer of a chiral compound

biodiesel: diesel made from oilseed rape or other vegetable oil

bioethanol: fuel made by fermentation of sugar sources such as sugar cane or sugar beet

buckminsterfullerene (or buckyballs): the third allotrope of carbon

graphene: a single sheet of graphite

'green chemistry': the development of chemical processes of manufacture that do not harm the environment but provide sustainable routes for the synthesis of chemicals

ground water: water that is below ground: such water provides a source of drinking water by drilling wells

hydrophilic: polar molecules that form hydrogen bonds with water and hence dissolve

hydrophobic: non-polar molecules that do not dissolve in water

liposome: an artificial microscopic vesicle consisting of an aqueous core enclosed in one or more phospholipid layers, used to convey vaccines, drugs, enzymes or other substances to target cells or organs

nanotechnology: science of the very small (with huge potential)

nanotube: a single sheet of graphite may be rolled to form a single walled nanotube (SWNT)

scanning tunnelling microscope: enables chemists to look at single atoms

sorbent: a floating porous material that is designed to absorb oil, but not water from a spillage on the sea

targeting: used in two different ways by chemists: (i) a compound is identified for synthesis; (ii) when using a liposome (or virosome) to carry a vaccine or a drug to cells or organs in the body

remediation (of contaminated soils): the improvement of disturbed land to achieve land that is the same as before the land was disturbed

remediation (of ground water): removal of contaminants that are toxic

3.7 – Resources

Liposomes

http://www.answers.com/
http://www.ingentaconnect.com/content/ben/cpd/

Virosomes

http://*www.bernabiotech.com/rd/platforms/virosomes/*

Spider silk

http://www.chm.bris.ac.uk/motm/motm.htm
http://www.xs4all.nl/~ednieuw/Spiders/Info/spindraad.

Kevlar

http://www.dupont.com/afs/
http://www.lbl.gov/MicroWorlds/Kevlar/index.html

Conducting polymers

http://nobelprize.org/chemistry/educational/poster/2000/index.html

Nanotechnology

http://www.aip.org/pt/vol-57/iss-6/p30.html
http://www.kodak.com/eknec/

Fullerenes

http://www.lbl.gov/Science-Articles/Archive/fullerenes.html
http:// www.vega.org.uk/schools/download/index.php
http://www.nanotech-now.com/nanotube-buckyball-sites.htm
http://www.lps.u-sud.fr/Collectif/gr_23/themes/fullnano/en_AC60.htm

Manipulating atoms

http://www.che.utoledo.edu/nadarajah/webpages/whatsafm.html

Mining

http://www.ias.ac.in/resonance/Aug2004/Aug2004p27-34.htm

Green chemistry

http://www.uyseg.org/greener_industry/index.htm
http://www.uyseg.org/sustain-ed/index.htm
http://www.uyseg.org/sustain-ed/pages/Process/ProcessFrameset.htm

Fuel cells

http://www.utcpower.com/fs/com/bin/fs_com_Page/0,5433,03540,00.html
http://www.rsc.org/chemistryworld/Issues/2006/March/HydrogenOnBoard.asp
http://www.rsc.org/chemistryworld/Issues/2006/March/FuelCells.asp

General

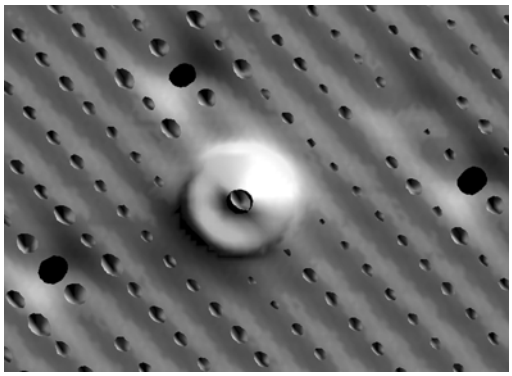
http://www.rsc.org/chemistryworld/
http://www.chemsoc.org/networks/learnnet/index.htm
http://pubs.acs.org/cen/science/science.html
http://www.chemistry.org/portal/a/c/s/1/enthusiasts.html

Nanotechnology worksheet

In this activity you will look at pictures of atoms. See how scientists can move atoms around to [illegible] pictures. Look at the pictures of atoms and read their descriptions. Make as many observations as you can and answer the questions about each picture.

Seeing atoms

Seeing is believing

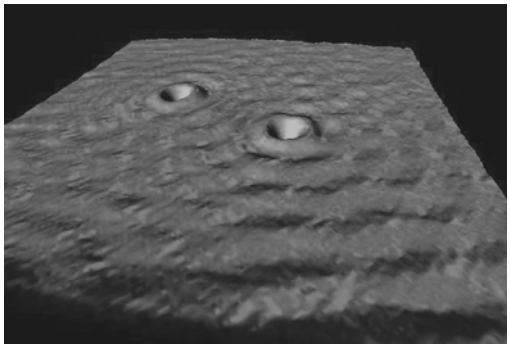


This picture is made from two images laid one on top of the other. This can be done using a computer. The lower picture is of nickel atoms. The big bump is a xenon atom. The tip of the pimple is a peek through to the nickel atom underneath. In a coloured picture produced by a computer, nickel is orange and xenon is blue.

The zit
Reproduced with kind permission from Mike Ross, IBM Almaden Research Center, California USA.

Questions

1. What are the symbols, atomic number and relative atomic mass values of the elements: nickel; xenon?
2. Why is the xenon atom much bigger than the nickel atom?
3. If you could see the nickel and xenon atoms, would they be these colours?



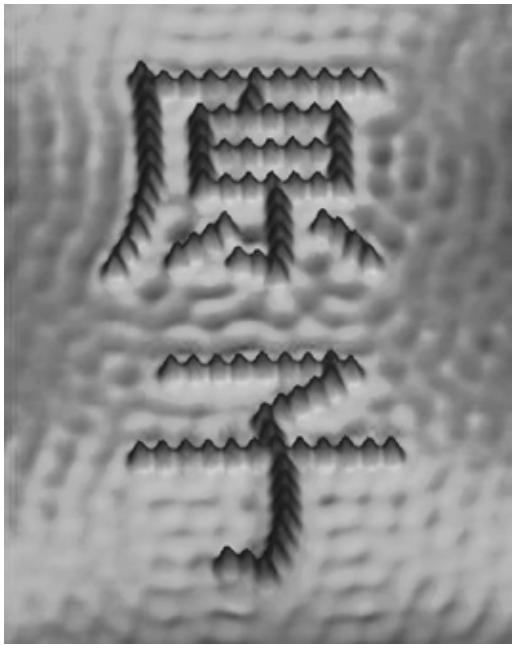
This is a sample of copper which is not perfect – the two dents are probably caused by atoms of another chemical element. The 'foreign' atoms have electron arrangements which are not the same as those of the copper atoms. The copper electrons on the surface are scattered, making patterns.

The dents
Reproduced with kind permission from Mike Ross, IBM Almaden Research Center, California USA.

Questions

4. What do the patterns on the surface look like?
5. In what other substances have you seen this kind of pattern?
6. What does this tell you about how electrons can behave?

Atomic art



Iron atoms have been arranged on a surface of co[illegible] make two Kanji characters which together mean 'atom'. [illegible] their own, the characters mean 'original' and 'child' [illegible] Japanese and Chinese, giving the title for the picture.

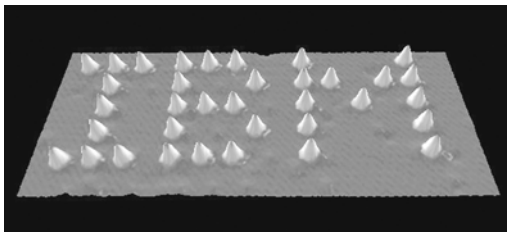
This is how to pronounce the characters:

yuan zi ('you-an zee')	Mandarin Chinese
gen shi ('hard' g, like in 'gun')	Japanese
yuen ji ('you-en jee')	Cantonese Chinese

An 'original child'
Reproduced with kind permission from Mike Ross, IBM Almaden Research Center, California USA.

Questions

7. How many iron atoms have been used in the picture?
8. One iron atom has a diameter of 248 pm. What is the distance across the top of the upper character in nanometres?
9. Is 'original child' a good name for 'atom'? Explain your answer.



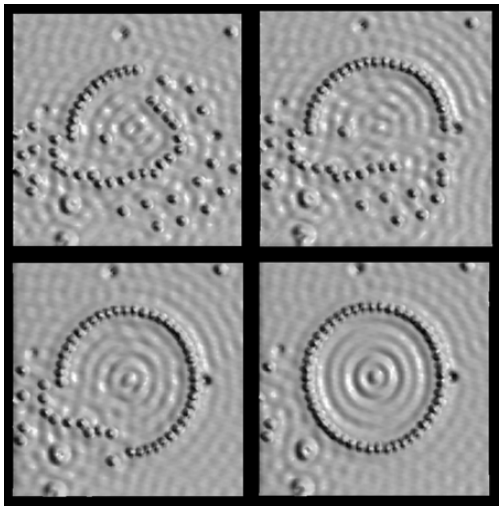
Xenon atoms have been arranged on the surface of nickel atoms to make the letters IBM. This was the first time that individual atoms were moved in a controlled way to make a new arrangement. The picture was taken in 1989.

In the beginning…
Reproduced with kind permission from Mike Ross, IBM Almaden Research Center, California USA.

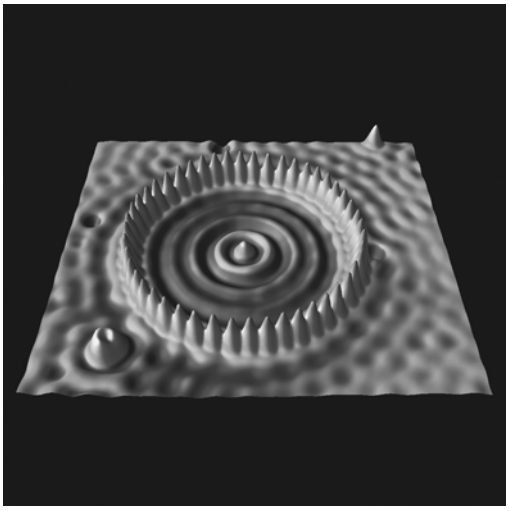
Questions

10. Why would the computer company IBM (International Business Machines) invest in nanotechnology?
11. Explain how this picture is a good advert for IBM and for nanoscience.

Making rings



Ironing the prefect circle



The birthday cake

Reproduced with kind permission from Mike Ross, IBM Almaden Research Center, California USA.

The four pictures show 48 iron atoms being moved into a circle 7.13 nm in diameter on the surface of copper. The tip of an STM is used to move the atoms. The finished ring looks like a birthday cake with candles. A pattern is seen in the centre of the finished ring.

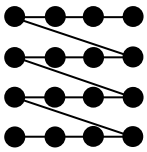
Questions

12. What is the shape of the iron atoms? Is this the shape that might be expected?
13. What might the central pattern be caused by?
14. Why is it useful to be able to move atoms individually?

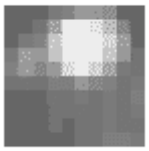
How STM pictures are made

The computer connected to the STM makes a grid of the surface of the substance. The tip of the STM scans the surface at points on the grid. The tip records a tiny electrical current, called the 'tunnelling current', at each point. The current changes as the tip moves up and down depending on the atoms present on the surface. A system inside the STM adjusts the tip movement, keeping it at a constant height. The adjustments are recorded and processed by the computer into an image, showing the changes in the surface of the substance.

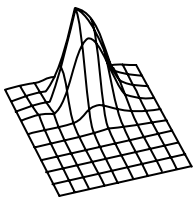
The pictures show the process: (1) scanning; (2) the original image showing light areas as small adjustments; (3) the processed image changing light areas into heights; (4) the processed image coloured in shades of grey.



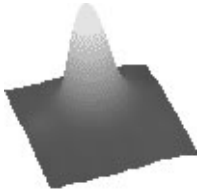
(1)



(2)



(3)



(4)

Producing an STM picture

Reproduced with kind permission from Mike Ross, IBM Almaden Research Center, California USA.

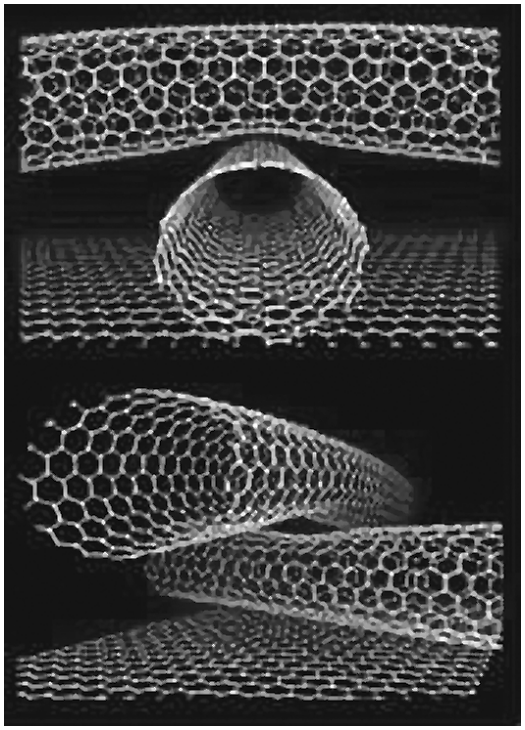
Then the image is coloured to show different features of the atoms.

Carbon and its various forms

Diamond and graphite are two well-known forms of carbon. In 1985 scientists discovered a third form of carbon based on 60 atoms bonded in a football-like structure. Scientists called it 'buckminsterfullerene', or 'buckyball'. This started a search for other carbon structures.

In 1991 a Japanese scientist called Sumio Iijima found carbon nanotubes. These are about 10 000 times thinner than a human hair, made from carbon atoms bonded in sheets and rolled into tubes. A carbon nanotube is about 1 nm in diameter and 1–10 μm long. The tubes are often capped at each end with a half-buckyball structure. Scientists are working to find out more about carbon nanotubes and what they could be used for.

What do nanotubes look like?



Piled high
Reproduced with kind permission from Mike Ross, IBM Almaden Research Center, California USA.

Questions

15. Describe the appearance of the nanotubes in the figure Piled high. What do they remind you of?
16. Name the chemical element from which nanotubes are made.
17. Name another form of this element which has a structure similar to nanotubes. How are nanotubes different from this substance?
18. Explain in terms of chemical bonding why this chemical element exists in several different forms, each with different properties.

4 – SPECIMEN EXAMINATION QUESTIONS

The Chemistry of Life

1. (a) (i) Name **two** secondary structures in proteins.

I ..... **II** .....

(ii) Draw simplified diagrams of each structure, showing the bonding by which each is stabilised.

I **II**

[4]

(b) ATP is an important molecule associated with the supply of energy in cells.
Write a simple equation to represent the hydrolysis of ATP.

..... [1]

(c) The following results were obtained at different concentrations of ATP for the rate of hydrolysis in the presence of the enzyme myosin.

[ATP] /mmol dm^{-3}	0.025	0.050	0.075	0.100	0.150	0.200	0.300
rate /mmol dm^{-3} s^{-1}	0.0065	0.0115	0.0140	0.0158	0.0176	0.0187	0.0196

(i) Use graph paper to plot these results, plotting [ATP] on the *x*-axis, and rate on the *y*-axis.

(ii) ADP is an inhibitor for this reaction, suggest what type of inhibitor it is.

.....

(iii) What would be the effect of ADP on the rate of the reaction at low [ATP]?

.....

.....

(iv) What would be the effect of ADP on the rate of the reaction at high [ATP]?

.....

.....

[6]

Applications of Analytical Chemistry

2. (a) (i) State what is meant by the term *partition coefficient*.

(ii) A solution of 5.00 g of an organic compound X in 50 cm^3 of water was shaken with 100 cm^3 of ether. After separation, the aqueous solution was found to contain 0.80 g of X.

Determine the partition coefficient of X between ether and water.

[4]

(b) Explain briefly how the separation of different components in mixtures is achieved in each of the following chromatographic techniques.

(i) paper chromatography

.....

.....

.....

(ii) thin layer chromatography

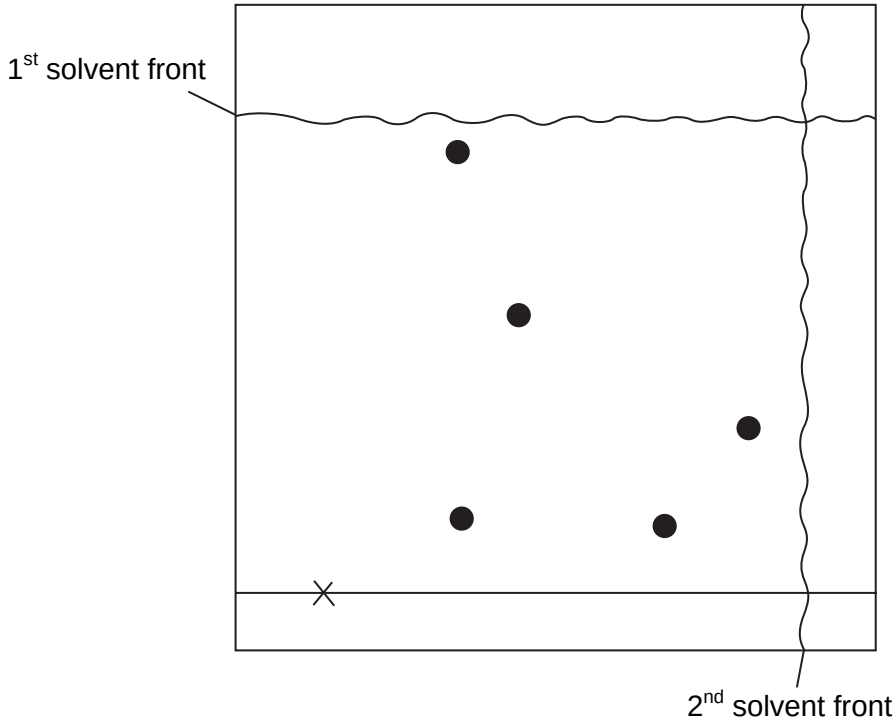
.....

.....

.....

[4]

(c) The diagram shows a two-way chromatogram carried out on the products of the [illegible] of a polypeptide.



X = start point

(i) How many different amino acids were present in the sample?

.....

(ii) What could have been use to make the spots visible?

.....

(iii) Sketch the chromatogram you might expect if only the first solvent had been used.

[4]

Materials and Design

3. **(a)** State what you understand by the term *nanotechnology*.

.....

.....

.....

[2]

(b) 'Nanotubes' are rolled up cylinders of graphite with diameters of about 1 nanometre, and lengths in micrometres.

(i) If a nanotube is 5 micrometres long, how many diameters does this represent?

(ii) These tubes are believed to be stronger than steel. Suggest a possible use for nanotubes.

.....

(iii) One problem in the synthesis of nanotubes is that a mixture of tubes of different length and orientations is produced. Suggest why this is a problem.

.....

.....

.....

[5]

(c) The delivery of cancer-destroying drugs has, in the past, been by injection of the relevant drug into the bloodstream, allowing it to be carried around the body to the tumour. New techniques have been developed which rely on binding the relevant drug molecule to an enzyme.

Suggest the advantages of this new technique, both in economic terms and in terms of the effect on the patient.

.....

.....

.....

.....

.....

.....

[4]