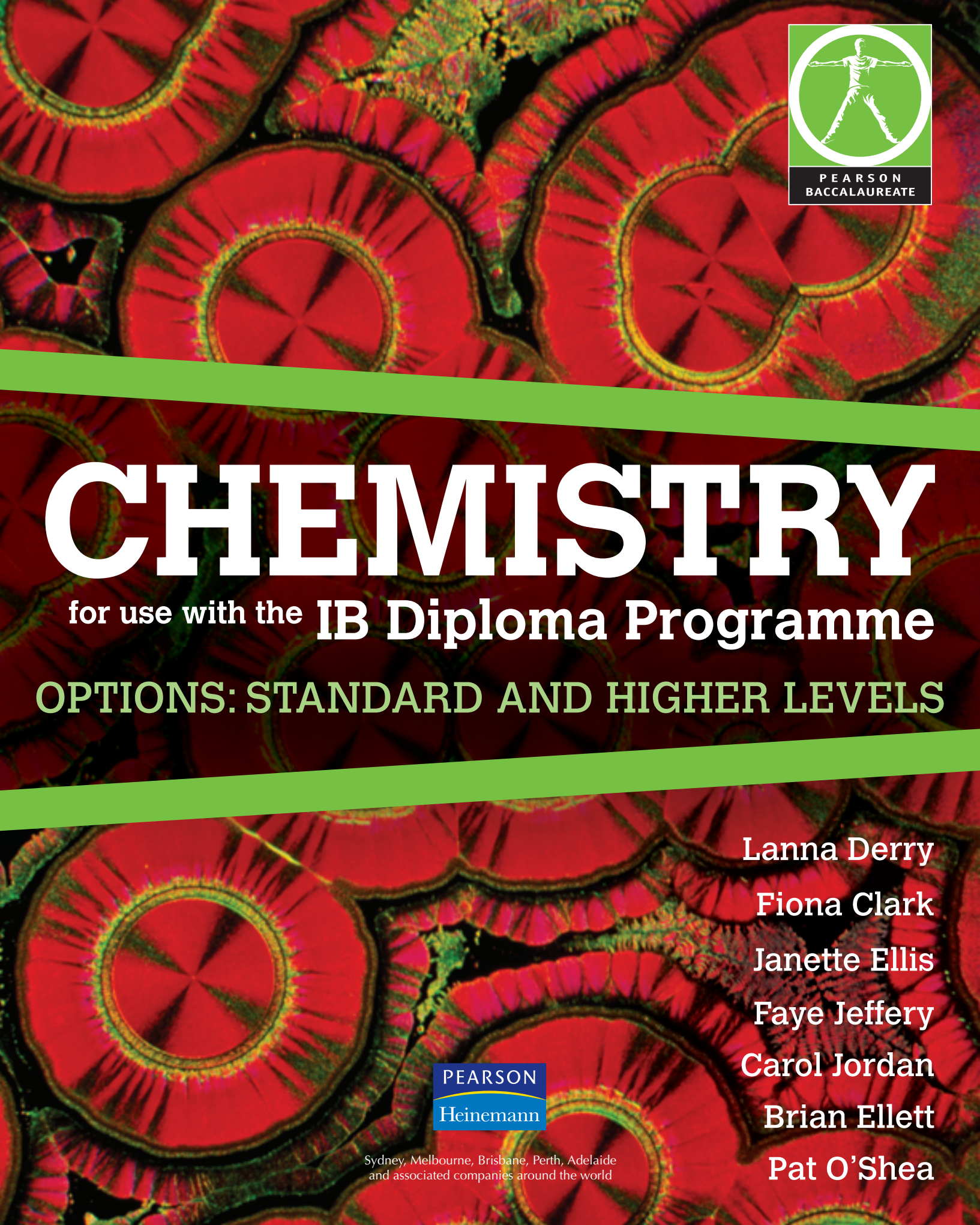


CHEMISTRY

for use with the IB Diploma Programme

OPTIONS: STANDARD AND HIGHER LEVELS

Lanna Derry
Fiona Clark
Janette Ellis
Faye Jeffery
Carol Jordan
Brian Ellett
Pat O'Shea

PEARSON
Heinemann

Sydney, Melbourne, Brisbane, Perth, Adelaide
and associated companies around the world

Pearson Heinemann

An imprint of Pearson Education Australia
(a division of Pearson Australia Group Pty Ltd)
20 Thackray Road, Port Melbourne, Victoria 3207
PO Box 460, Port Melbourne, Victoria 3207
www.pearson.com.au

First published 2009 by Pearson Education Australia
2012 2011 2010 2009
10 9 8 7 6 5 4 3 2 1

Publisher:	Catriona McKenzie
Editor:	Marta Veroni
Text Design:	Meaghan Barbuto
Copyright & Pictures Editor:	Michelle Jellett
Project Editor:	Naomi Murphy
Production Controller:	Jem Wolfenden
Cover Design:	Glen McClay
Cover image description:	Water soluble organic compound with antioxidant properties that are essential in maintaining healthy connective tissues and integrity of cell walls. Found in citrus fruits and vegetables. Deficiency leads to scurvy. Light micrograph.
Cover image:	Getty Images
Illustrator/s:	Wendy Gorton, Nikki M Group
Printed in China	

National Library of Australia Cataloguing-in-Publication entry

Author:	Derry, Lanna.
Title:	Chemistry for use with the IB diploma programme options : standard and higher levels / Lanna Derry … [et al.]
ISBN:	9780733993787 (pbk.)
Notes:	Includes index.
Target Audience:	For secondary school age.
Subjects:	Chemistry--Textbooks. Chemistry--Examinations--Study guides.
Other Authors/Contributors:	Derry, Lana.
Dewey Number:	540

ISBN 978 0 7339 9378 7
Pearson Australia Group Pty Ltd ABN 40 004 245 943

Disclaimer
The selection of Internet addresses (URLs) provided for this book were valid at the time of publication and chosen as being appropriate for use as a secondary education research tool. However, due to the dynamic nature of the Internet, some addresses may have changed, may have ceased to exist since publication, or may inadvertently link to sites with content that could be considered offensive or inappropriate. While the authors and publisher regret any inconvenience this may cause readers, no responsibility for any such changes or unforeseeable errors can be accepted by either the authors or the publisher.

Some of the images used in *CHEMISTRY: For use with the IB Diploma Programme* might have associations with deceased Indigenous Australians. Please be aware that these images might cause sadness or distress in Aboriginal or Torres Strait Islander communities.

CONTENTS

CHEMISTRY

for use with the IB Diploma Programme

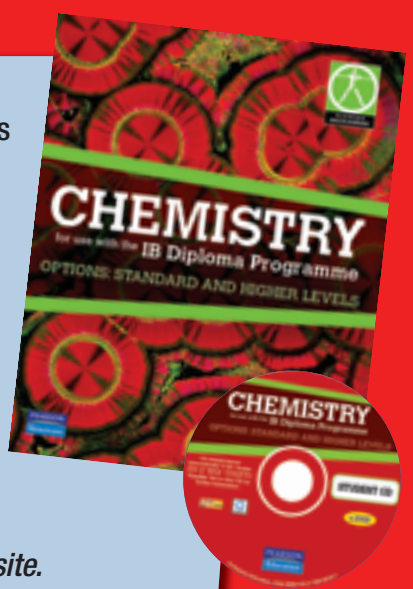
OPTIONS: Standard and Higher Levels

The complete chemistry package

CHEMISTRY: For use with the IB Diploma Programme Options Standard and Higher Levels is the most comprehensive text specifically written for the IB Diploma Programme Options topics.

The content is easy to follow and provides regular opportunities for revision and consolidation. All assessment statements of the IB Diploma Programme Chemistry syllabus are covered in highly structured and meaningful ways.

Coursebook includes Student CD



Each chapter in the coursebook includes:

- focus on the IB Additional Higher Level (AHL) Diploma Programme Chemistry syllabus, topics 12 to 20
- Syllabus Assessment Statements given beside the relevant theory
- stimulating photos and full colour illustrations to support learning of chemical concepts
- theory broken into manageable chunks for ease of learning
- comprehensive exercises for ongoing review and consolidation
- **Chem Complement** boxes, which engage students with interesting extension material and applications to Aims 8 and 9 for Experimental Sciences
- **Theory of Knowledge** boxes, which allow easy integration of this requirement of the syllabus
- ICT activities, which address Aim 7 for Experimental Sciences and are available on the *Companion Website*
- **Chapter summary**, which includes chapter glossary and key points.

Student CD contains:

- an electronic version of the coursebook
- review questions and solutions
- chapter tests of examination questions with solutions
- fully worked solutions to coursebook questions
- a link to the live *Companion Website.*

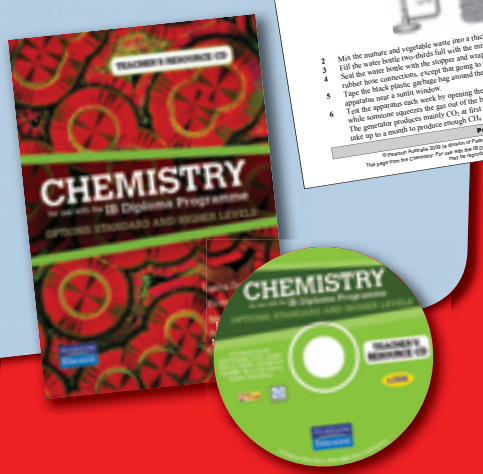
Teacher's Resource CD

The *Teacher's Resource CD* provides a wealth of teacher support material, including:

- **fully worked solutions** to coursebook questions
- **worksheets** for practising skills and consolidating theory; answers are also included
- **teacher demonstrations** to engage students and enhance understanding of concepts
- **practical investigations** to enhance the learning of chemical concepts and for use in meeting the mandated time allocation for practical work
- **practical notes** for the teacher/lab technician
- **risk assessments** for practical activities.

This time-saving resource contains documents available as:

- Microsoft Word documents that can be edited, allowing you to modify and adapt any resources to meet your needs
- PDFs to make printing easy.



Companion Website

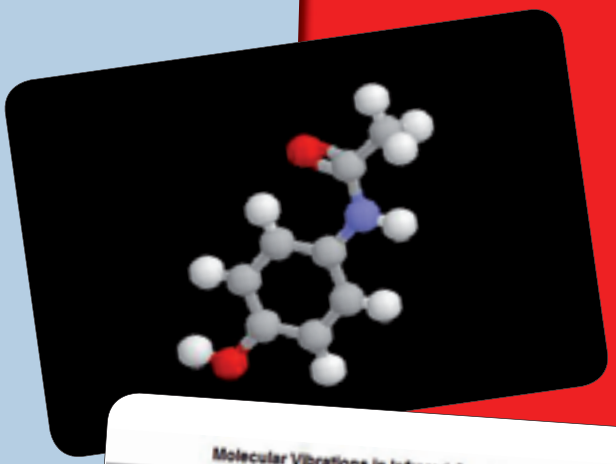
www.pearson.com.au/schools

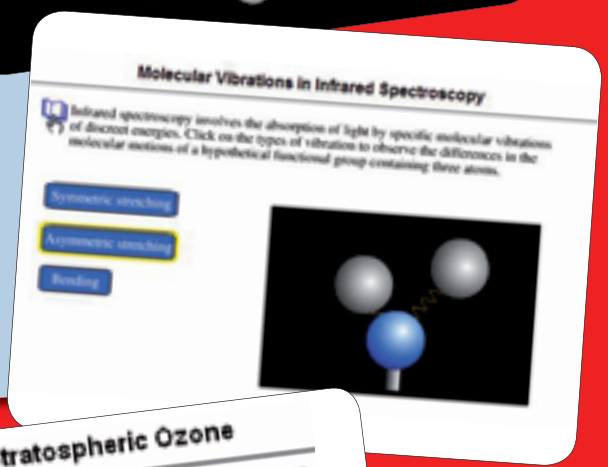
The *Companion Website* addresses Aim 7 for Experimental Sciences by providing easy integration of technology into the classroom. It contains a wealth of support material for students and teachers to enhance teaching and learning in chemistry.

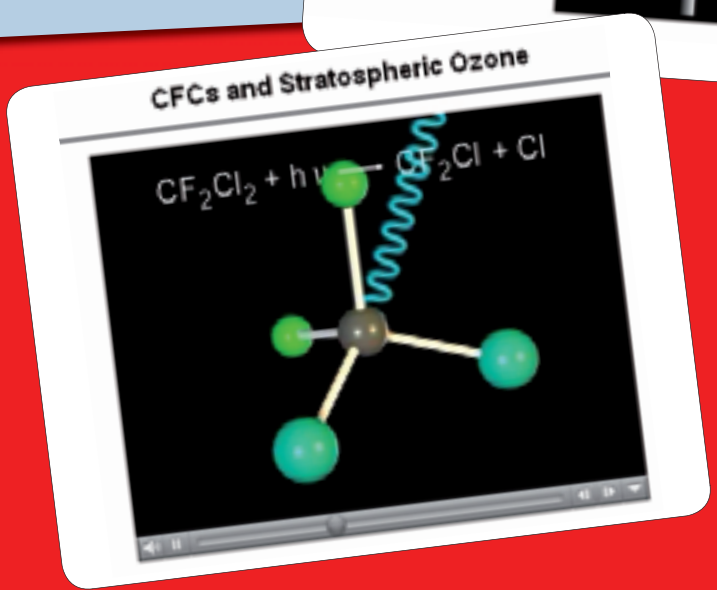
The interactive material on the *Companion Website* allows students to review their work and revise fundamental concepts, as well as providing an opportunity for accelerated learning.

The *Companion Website* contains:

- **Review Questions**—auto-correcting multiple-choice questions for exam revision
- **Interactive Animations**—to engage students in exploring concepts
- **QuickTime Videos**—to explore chemical concepts in a visually stimulating way
- **3D Molecules Gallery**—for interactive viewing and manipulating of molecular structures
- **Web Destinations**—a list of reviewed websites that support further investigation and revision.







For more information on *CHEMISTRY: For use with the IB Diploma Programme* visit www.pearson.com.au/schools

MEET THE AUTHORS

Lanna Derry

Lanna Derry, the lead author of the *CHEMISTRY: For use with the IB Diploma Programme* series, is a highly experienced teacher of IB Chemistry. She has taught senior Chemistry in independent schools for more than twenty years and has authored revision guides for Year 12 Chemistry. Lanna is currently teaching IB Chemistry at Tintern Girls Grammar School, Ringwood East, Victoria, Australia.

Fiona Clark

Fiona Clark is an experienced IB Chemistry teacher and examiner. Fiona is a workshop leader and has been an Assistant Examiner for IB Chemistry examinations. She was a member of the IB Chemistry Subject Review that was responsible for developing the new IB Diploma Programme Chemistry Guide. She is currently Head of Science at Park View Education Centre, Bridgewater, Nova Scotia, Canada.

Janette Ellis

Janette Ellis has experience teaching both IB Chemistry and senior Chemistry. After teaching in Victoria for many years, she is now at Kambala, Rose Bay, New South Wales, Australia.

Faye Jeffery

Faye Jeffery is currently teaching at Melbourne Centre for Adult Education. She has taught Chemistry and Biology for more than twenty years. Faye has written a number of texts for Chemistry and Science.

Carol Jordan

Carol Jordan is currently teaching at the Shanghai American School, Shanghai, China. She is an experienced teacher of IB Chemistry, IB Environmental Systems and Theory of Knowledge. She has been an assistant examiner and senior moderator for internal assessment for IB Chemistry. Carol is a workshop leader and was part of the team responsible for developing the new IB Diploma Programme Chemistry Guide.

Brian Ellett has taught senior Chemistry for more than twenty years and has written a number of texts for Chemistry. He is currently Head of Science at Salesian College, Chadstone, Victoria, Australia.

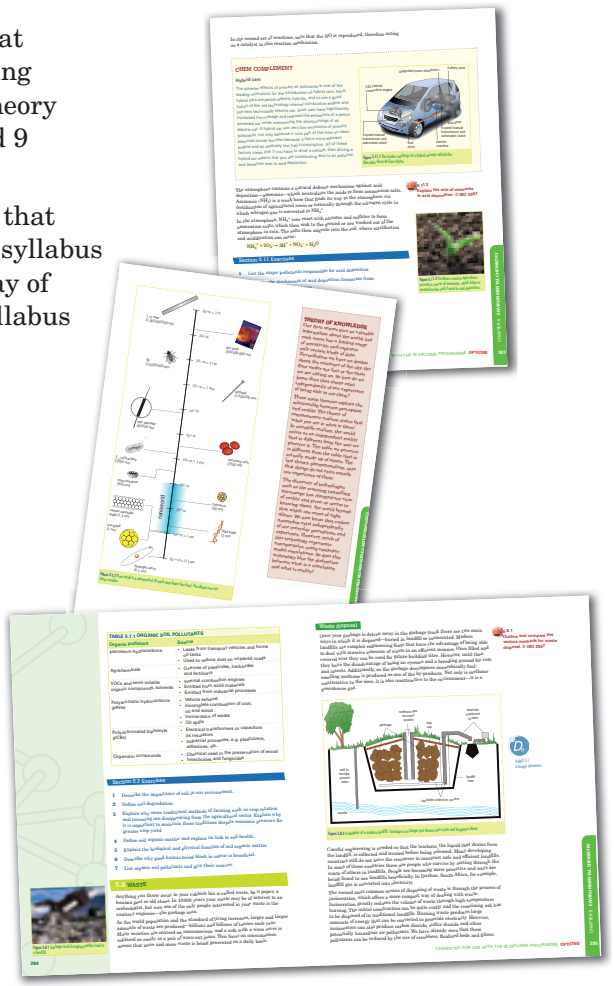
Pat O'Shea is a highly experienced teacher of Chemistry. He is currently Deputy Principal at Loreto College, Ballarat, Victoria, Australia. Pat has presented at many workshops for senior Chemistry teachers.

HOW TO USE THIS BOOK

Our aim has been to present chemistry as exciting, accessible and relevant. The content is carefully structured with regular opportunities for revision and consolidation to prepare students for the IB Diploma Programme Higher Level Chemistry examinations. (Paper 3).

Major features

- *Chapter opening pages* that include a stimulating photo and a simple, student-friendly syllabus-related list of what students should be able to do by the end of the chapter
- **Chem Complement** boxes that engage students with interesting extensions of the Chemistry theory and applications to Aims 8 and 9 for Experimental Sciences
- **Theory of Knowledge** boxes that address the links between the syllabus and aspects of the scientific way of knowing as required by the syllabus
- *ICT activities* that address Aim 7 for Experimental Sciences and are available on the Companion Website
- *Comprehensive exercises* that encourage students to consolidate their learning in a systematic manner while familiarising students with the IB command terms
- *A summary of concepts* at the end of each chapter



Icons in the coursebook

- **Assessment Statement** icons denote Assessment Statements from the IB Diploma Programme Standard Level (Core) Chemistry syllabus.
- **Worksheet** icons denote when a worksheet giving extra practice on a key part of the topic is available. These can be found on the **Teacher's Resource CD**.
- **Prac** icons denote when a practical investigation is available. These can be found on the **Teacher's Resource CD**.
- **Demo** icons denote when a teacher demonstration is available. These can be found in the **Teacher's Resource CD**.
- **Companion Website—Interactive Animation** icons denote when links to an animation are available to support a topic in the coursebook. These can be accessed on the Companion Website.
- **Companion Website—QuickTime Video** icons denote when links to a QuickTime video are available to support a topic in the coursebook. These can be accessed on the Companion Website.

Other features

- **Worked examples** of calculations and chemical structures to aid mastery of difficult concepts
- **Periodic table** with relative atomic masses included on the inside front cover to provide a quick and easy reference

Student CD

This interactive resource contains:

- an electronic version of the coursebook
- fully worked solutions (including diagrams) to all coursebook questions
- review questions and chapter tests and solutions
- a link to the live Companion Website (Internet access required) to provide access to course-related Web links.

Other components

- Companion Website www.pearson.com.au/schools
- Teacher's Resource CD

Other books in the series

CHEMISTRY: For use with the IB Diploma Programme Standard Level

CHEMISTRY: For use with the IB Diploma Programme Higher Level

ACKNOWLEDGEMENTS

We would like to thank the following for permission to reproduce photographs, texts and illustrations. The following abbreviations are used in this list: t = top, b = bottom, c = centre, l = left, r = right.

© ISRIC: FAO & UNEP 1993: p. 366t.

Alamy Limited: p. 108.

Australian Associated Press Pty Ltd (AAP): pp. 253, 359.

Corbis Australia Pty Ltd: pp. 271b, 275tl.

DK Images: p. 85.

Dreamstime: p. 337.

International Nuclear Safety Center: p. 372b.

iStockphoto: pp. 183(pants), 235, 260b, 269, 287, 304, 309, 336, 339, 341 (strawberries), 348, 355l, 362, 367, 369.

Jupiterimages Corporation © 2009: pp. 65br, 294, 305, 293b, 263br, 264b, 259, 260t, 263cr, 263t, 271t, 275tr, 280l, 280r, 282, 283b, 292, 298, 307, 357.

Lanna Derry: p. 285.

NASA: pp. 262r, 283t, 302.

Pearson Australia/Alice McBroom: pp. 335, 341(ham), 366(chocolate), 375.

Pearson Learning Photo Studio: p. 183(wax).

Photolibrary Pty Ltd: pp. 1, 2b, 2t, 3l, 3r, 28, 31b, 31t, 35, 41-43, 48, 50, 52, 65l, 68, 75, 76, 98, 99, 103, 106bl, 106br, 106t, 109, 127, 139, 140-142, 152r, 173, 178, 181tr, 183 (drug patch), 185, 207, 208, 212, 217, 221, 223, 224, 228, 229, 231b, 231t, 236, 239, 244, 249b, 249t, 250, 264t, 270t, 272, 288, 299, 321, 341 (milk), 344, 351, 364, 366 (foam), 387.

Shutterstock: pp. 39, 152l, 176, 183 (golf), 183 (tennis), 261, 278, 293t, 311, 350, 355r, 356, 366 (milk).

Trustees Museum of Applied Arts & Sciences/Powerhouse Museum: p. 147.

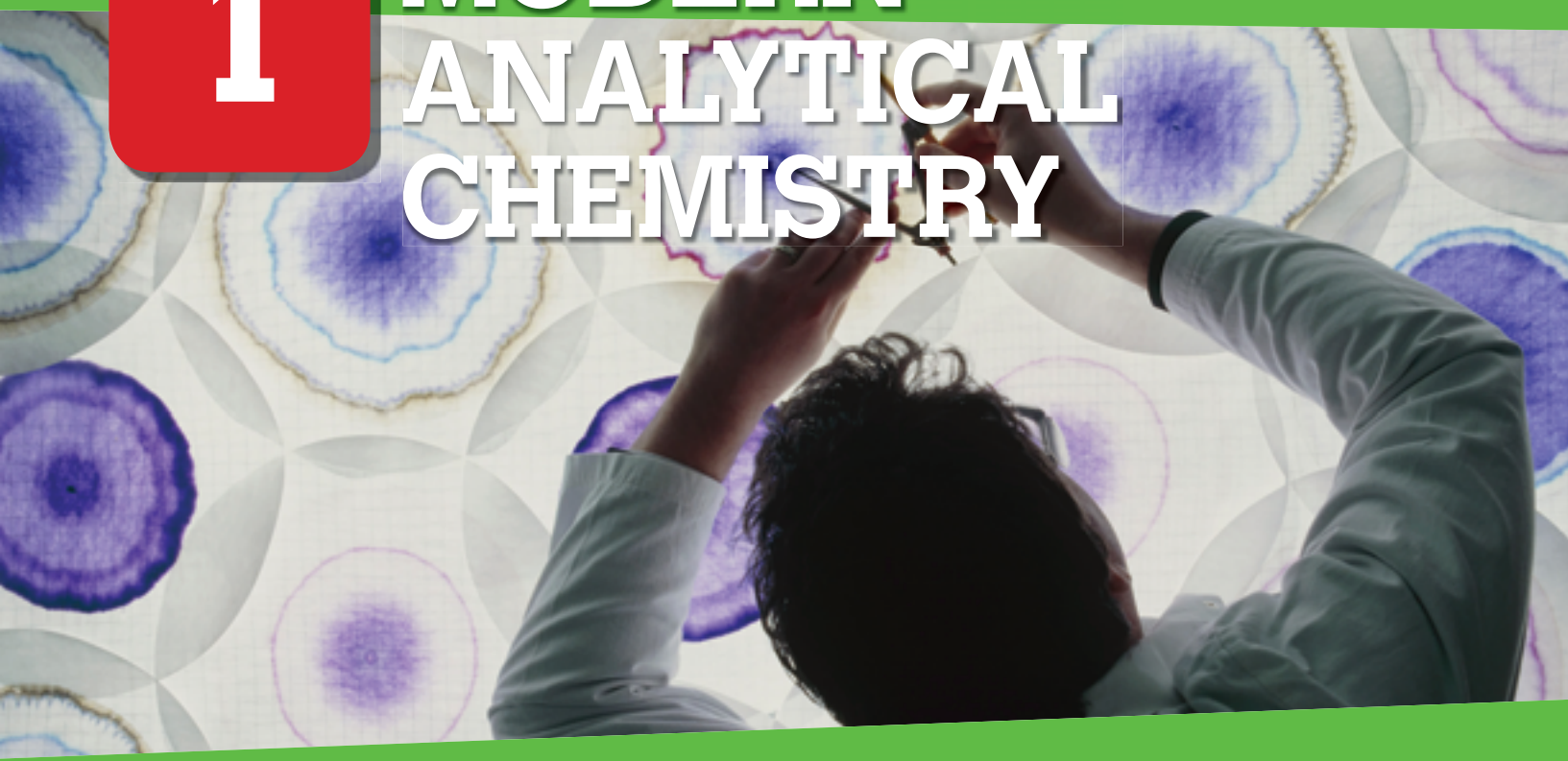
Wikipedia: p. 183 (blackberry); Chris 73: p. 342; Terence Ong: p. 149.

Thanks to the **International Baccalaureate Organization** (IB Organization) for permission to reproduce IB intellectual property. This material has been developed independently of the International Baccalaureate Organization (IB Organization) which in no way endorses it.

Every effort has been made to trace and acknowledge copyright. However, should any infringement have occurred, the publishers tender their apologies and invite copyright owners to contact them.

The Publishers wish to thank Maria Connor, Fiona Clark and Michael McCann for reviewing the text.

1 MODERN ANALYTICAL CHEMISTRY



Chapter overview

This chapter covers the IB Chemistry syllabus Option A: Modern Analytical Chemistry.

By the end of this chapter, you should be able to:

- describe the equipment and procedures used for the analysis of substances using atomic absorption (AA) spectrometry, infrared (IR) spectroscopy, mass spectrometry and nuclear magnetic resonance (NMR) spectroscopy
- describe how the structure of compounds may be deduced using information obtained from various spectroscopic techniques
- analyse IR, mass and NMR spectra in order to gain information about molecular structure
- understand that all chromatographic techniques require a stationary phase and a mobile phase
- describe how chromatography is used to identify and quantify the components of a mixture
- use a calibration curve in quantitative chromatographic and spectroscopic analysis
- calculate R_f values and describe the significance of R_f and R_t values in chromatography
- describe the equipment and procedures used for the analysis of substances by paper, thin-layer and column chromatography
- recognize that chromatographic techniques can be used to separate components of a mixture prior to analysis by other techniques such as mass spectrometry and IR spectroscopy

- describe the equipment and procedures used for the analysis of substances using gas and high performance liquid chromatography HL
- describe the equipment and procedures used for the analysis of substances using UV–visible spectrometry
- analyse UV–visible spectra in order to gain information about molecular structure
- interpret splitting patterns and chemical shifts in NMR spectra to gain information about molecular structure.

IBO Assessment statements A.1.1 to A.10.2

Figure 1.0.1 A roadside breath test determines whether the blood alcohol content of a driver is over the legal limit.

News reports may tell us that a driver has been arrested for driving with a blood alcohol content of 0.17%, an athlete has been tested for drugs and found to have consumed amphetamines, traces of water vapour have been found in the atmosphere of a new planet and that low fat foods have a high sugar content. How do scientists determine such information? It certainly is not just with the laboratory equipment that we have access to daily. Modern methods of chemical analysis involve the use of instruments that routinely analyse samples rapidly and to high levels of accuracy. Although the instruments may be sophisticated and complex, they all use the basic principles of spectroscopy and chromatography.

1.1 INTRODUCTION TO ANALYTICAL CHEMISTRY

AS **A.1.1 State the reasons for using analytical techniques.**

Modern analytical chemistry offers a wide range of problem-solving techniques. Chemists use analytical techniques to identify and confirm the structure of a substance, to analyse its composition and to determine its purity. The quantity and nature of substances present in a mixture may be determined or the progress of a reaction may be followed by analysis. Instrumental analytical techniques are often more accurate and quicker than techniques involving experimental techniques such as those used in acid–base titrations in volumetric analysis or the precipitation of a solid for gravimetric analysis.

Quality control of consumer products requires that the composition of foodstuffs, building materials and even the materials from which our clothes are made be known accurately. Such knowledge can be life saving, or just a guide for people wanting to know the fat content of the food they eat.

Figure 1.1.1 An instrumental analytical technique known as high performance liquid chromatography (HPLC) is used to test the purity of an experimental AIDS vaccine.

In the design and manufacture of drugs, chemical analysis is of the utmost importance. The structure of a newly synthesized compound can be confirmed by spectroscopic methods and its purity measured by chromatographic methods.

Although there may be many analytical instruments with varied components, they are essentially of only three types: those based on the interaction of electromagnetic radiation with the sample (spectroscopic techniques), those based on the interaction of magnetic and/or electric fields with the sample (mass spectroscopy and nuclear magnetic resonance) and those involving chromatography. Instrumental methods have several advantages over the more traditional laboratory methods (volumetric and gravimetric). They are faster, more sensitive, more accurate and less prone to human error. Their main disadvantage is cost!

Analytical techniques

In an attempt to prevent cheating in sport, the urine of competitors is regularly tested for the presence of performance-enhancing drugs. In many schools, students suspected of illicit drug use may also find themselves requested to undergo such tests. Chromatography is the basis of the

drug-testing cards that are used in these situations. In such 'dipstick' tests, a simple colour change in the chemically treated panel of the card indicates the presence of the drug for which the urine is being analysed.

Chemical analysis is widely used in medicine for the detection of unwanted substances in the blood or urine. For example, the initial test for diabetes involves measuring the amount of glucose in the blood and ketones in the urine. Home pregnancy kits are another example of a urine 'dipstick' test. These kits contain a treated panel that will change colour in the presence of a hormone called human chorionic gonadotrophin (hCG).

Although chemical dipsticks are used for identification of a substance in the urine or blood, the exact amount is often determined by instrumental methods. The concentration and identity of a compound in blood or urine can be determined using chromatography (in particular gas–liquid chromatography) and spectroscopy.

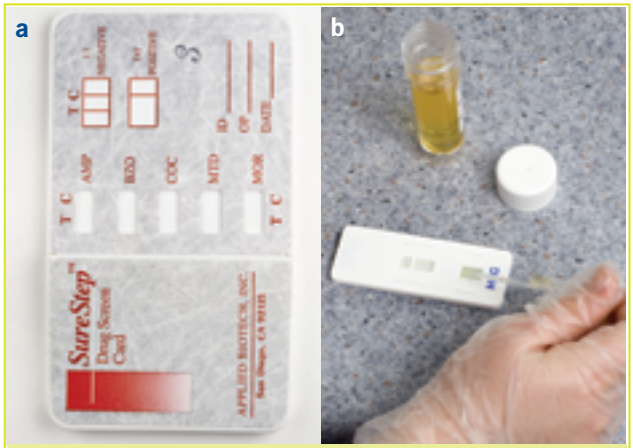


Figure 1.1.2 (a) A drug test card is a simple way to test for the presence of drugs in urine. (b) A pregnancy can be confirmed by a simple urine test for the presence of a particular hormone (hCG).

Modern chemical analysis can detect the presence of tiny, yet toxic, quantities of elements in our water supply and in our air. The technique of atomic absorption spectroscopy can detect up to 70 different metal elements in concentrations as low as parts per billion (ppb = μg per kg) of sample. Environmental protection authorities can use this method to analyse samples of water or air for pollution.

Forensic science makes extensive use of modern chemical analytical methods. The presence of toxic chemicals in tissue samples or hair can be determined using high performance liquid chromatography. This may well reveal the cause of death of a murder victim.

Spectroscopy and chromatography both provide the opportunity for samples to be analysed qualitatively and quantitatively. In **qualitative analysis**, the *identity* of a sample, or its components is determined, whereas **quantitative analysis** is used to determine the *amount* of the sample or its components. The outputs from spectroscopic and chromatographic instruments have several common features. Spectra and chromatograms both show a series of peaks that can be used both qualitatively and quantitatively. The position of a peak on the horizontal axis provides information that identifies the substance causing the peak, while the peak area provides information that can be used to determine the amount of that substance.

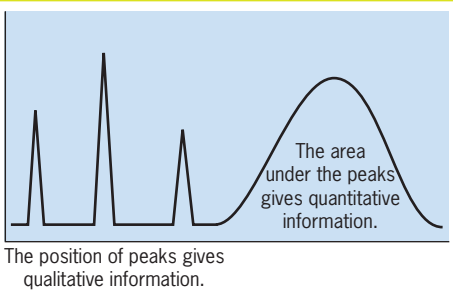


Figure 1.1.3 Spectra and chromatograms have several features in common.

Faced with this array of instrumental techniques, how does a chemist choose the most appropriate technique for analysing a particular sample? Such decisions will depend on a number of factors:

- Is qualitative or quantitative information required?
- How much material is available for analysis?
- How low is the concentration of the component being analysed likely to be?
- What is the chemical nature of the sample?
- What is the level of expertise of the chemist?
- What accuracy is required?

and, of course, the ever practical question:

- How much money is available for the analysis?

CHEM COMPLEMENT

Chemical analysis—the natural way

While many analytical instruments are able to accurately detect and measure minute quantities of substances, few come close to the sensitivity shown by natural systems. Organisms have built-in detection devices that form part of the complex control systems used to regulate the levels of critical chemicals such as hormones and neurotransmitters. Chemists are now making use of these supersensitive biodetectors. For example, hairs from Hawaiian red swimmer crabs can be attached to electrical analysis equipment to detect hormones at concentrations as low as 10^{-12} mol dm^{-3}!

AS **A.1.2**
State that the structure of a compound can be determined by using information from a variety of analytical techniques singularly or in combination.

AS **A.2.1**
Describe the electromagnetic spectrum

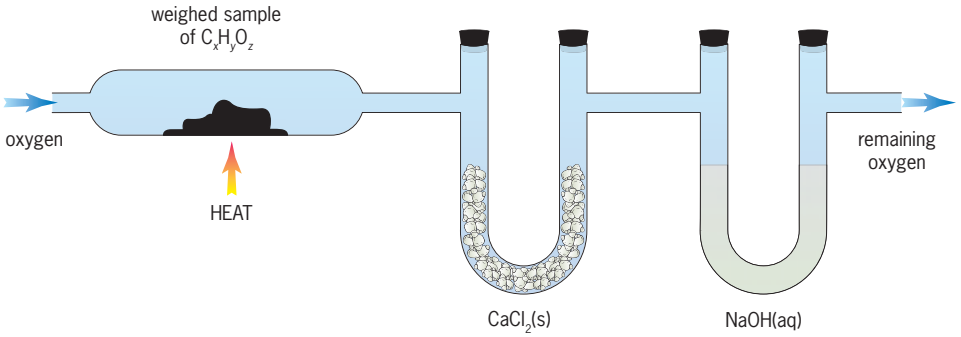


Figure 1.1.4 Determining the empirical formula of an unknown compound.

Consider, for example, the analysis of an unknown organic compound to determine its molecular and structural formula. Classically, a sample of the compound would be burnt in air to produce water, carbon dioxide and any other products. From measurements of the masses of the compound and the products, an empirical formula could be determined. Molar mass, and hence molecular formula, could be determined using measurements of gaseous samples. Finally, chemical tests for functional groups would be conducted. Much of this 'classic' chemistry has been replaced by modern instrumental methods. When such chemical methods are used in conjunction with instrumental methods, chemists can quickly and accurately determine the structural formula of a compound.

A mixture of hydrocarbons may first be separated into its components by high performance liquid chromatography. Then each component can be passed through a mass spectrometer to find its molecular mass and to determine its likely structure. This information may be supported by infrared spectroscopy, which is used to determine the nature of any functional groups. Information from only one technique is usually insufficient to determine or confirm a structure; the results from many different instrumental techniques are used in combination. One major advantage of using instrumental techniques is that great detail can be obtained from a very small amount of substance.

Principles of spectroscopy

The basis of spectroscopic analysis is the effect of electromagnetic radiation on matter. The sample for analysis is exposed to electromagnetic radiation and the effects of the interaction monitored. The study of the radiation absorbed or emitted by matter is called **spectroscopy**. The measurement of the amounts of light absorbed or emitted is called **spectrometry**. Instruments used for viewing the results of interactions between the sample and the radiation are called **spectrometers**. More elaborate instruments that measure amounts of radiation are called **spectrophotometers**. A wide variety of spectroscopic instruments are used in industry. Such instruments include atomic absorption and UV–visible spectrophotometers, and infrared spectrometers.

To understand spectroscopy, we need to review our basic understanding of the electromagnetic spectrum. Light consists of electromagnetic waves. The **wavelength** (distance between successive crests) of visible light ranges from about 8×10^{-7} metres for red light to about 4×10^{-7} metres for violet light. The **frequency** (the number of waves passing a given point each second) of light ranges from about 4×10^{14} for red light to about 7×10^{14} for violet light.

The wavelength and energy of light are related by the equation:

$$E = h\nu = \frac{hc}{\lambda}$$

where ν is the frequency
E is the energy
λ is the wavelength
c is the speed of light (3×10^8 m s^{-1})
h is a constant
(Planck's constant = 6.63×10^{-34} J s)

The energy of light increases as the wavelength decreases, showing an inverse relationship.

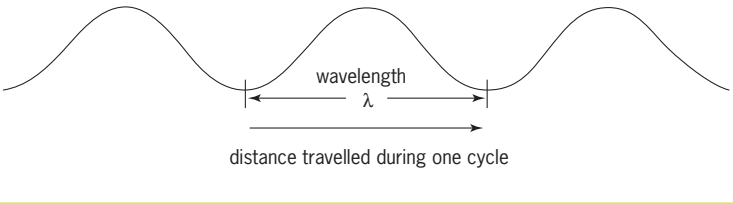


Figure 1.1.5 The wavelength of an electromagnetic wave is the distance travelled by the wave during one cycle. Its frequency is the number of waves that pass a particular point every second.

Of all the colours of visible light, violet light has the highest energy (and shortest wavelength) while red light has the lowest energy (and longest wavelength). Light is part of the broader electromagnetic spectrum, which also includes gamma rays, X-rays, ultraviolet (UV) rays, infrared (IR) waves, microwaves and radio waves.

Wavenumber is another wave property, and is equal to the inverse of wavelength ($\frac{1}{\lambda}$). Wavenumber corresponds to the number of cycles the wave produces in a centimetre and has the units cm^{-1}. The wavenumber has traditionally been the most common method of specifying IR absorption (see section 1.2).

Interactive CW Animation
Electromagnetic spectrum

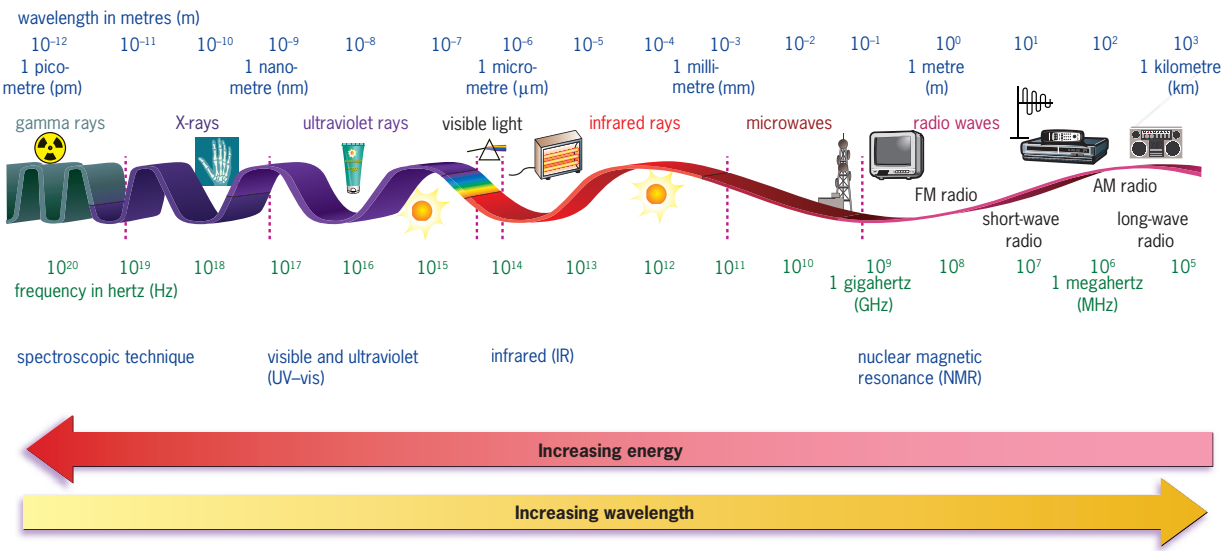


Figure 1.1.6 The electromagnetic spectrum. Different spectroscopic techniques use radiation from different parts of the spectrum.

When sunlight (which contains all wavelengths of visible light) passes through a prism, the different wavelengths are bent (or refracted) through different angles so that the light is broken into its components, producing a continuous spectrum of colours.

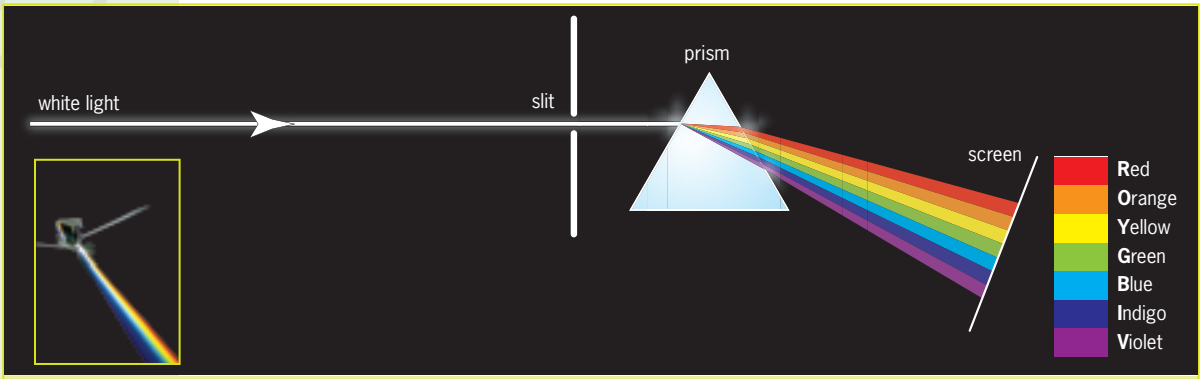


Figure 1.1.7 White light produces a continuous spectrum (ROYGBIV) when passed through a prism.

AS A.2.2 Distinguish between *absorption* and *emission* spectra and how each is produced.© IBO 2007

In chapter 4 of *Chemistry: For use with the IB Diploma Programme Standard Level*, we discussed atomic emission spectra (line spectra). These occur when energy is given to an atom, in the form of heat, electricity or light. An electron can move between energy levels within the atom if it absorbs energy that corresponds to the difference between two energy levels. When an electron moves to a higher energy level, the atom is said to be in an **excited state**. Electrons in higher than usual energy levels are unstable. They quickly return to their lower energy level, their **ground state**, by emitting the energy previously absorbed. This energy is released in the form of light.

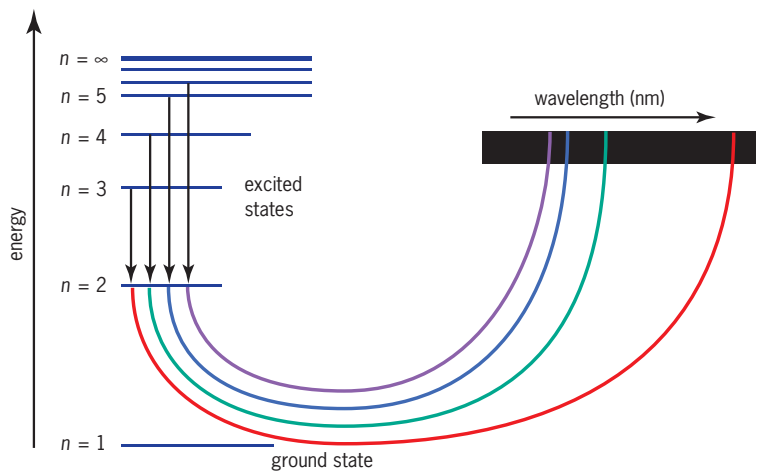


Figure 1.1.8 When electrons move from higher to lower energy levels, they emit light energy.

The energy of the emitted light is equal to the difference in energy between the higher and lower energy levels. Transitions between different energy levels release light of differing wavelengths according to the relationship $E = \frac{hc}{\lambda}$. As several electron transitions are possible, several colours of light could be emitted (although only the most intense colours might be detected by the unaided eye).

Atoms of different elements have different numbers of protons and therefore electrons. They also have different energy levels. Each type of atom will therefore emit light that has a unique set of energy values. This is the basis for **atomic emission spectroscopy**. In atomic emission spectroscopy a very hot flame is used to excite a wide range of metals and the emitted light is passed through a prism, producing an emission spectrum. Light emitted from excited atoms contains only certain characteristic wavelengths of light, which have been produced by the transition of electrons from one energy level to another. When this light is passed through a prism, the spectrum produced is a series of coloured lines separated by 'black' gaps. Such a spectrum is called an **emission** or **line spectrum**. Each metal atom produces a unique line spectrum that can be compared with the line spectrum produced by an unknown sample. If the lines match in terms of intensity, position and frequency of occurrence, it can be established that the unknown sample contains a specific metal(s).

Pr

PRAC 1.1
Flame tests and emission spectra

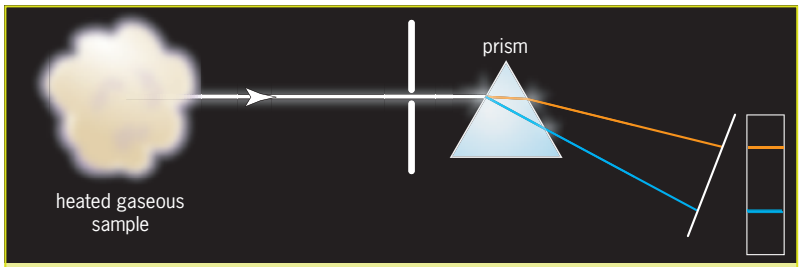


Figure 1.1.9 Light emitted from excited atoms produces an emission spectrum when passed through a prism.

Interactive Animation
Emission spectroscopy

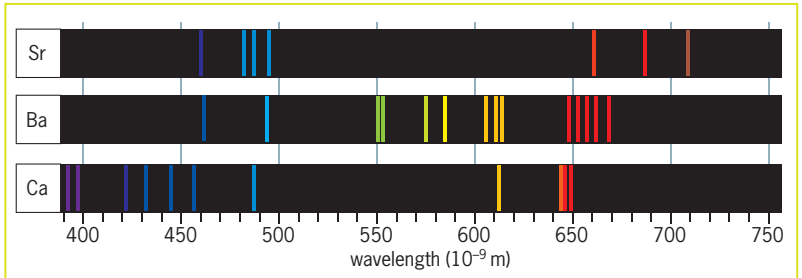


Figure 1.1.10 Emission spectra for a number of elements.

QuickTime Video
Flame tests for metals

If the energy supplied to a cool gaseous element is light, rather than heat, the element can absorb exactly the frequencies of light that are required to excite electrons to higher energy levels. When the remaining light is passed through a prism, 'gaps', black lines on the otherwise continuous spectrum can be seen. This pattern of black lines on the coloured background is an **absorption spectrum**. The black lines on the background of the continuous spectrum represent those energies of light absorbed by electrons moving from the ground state to higher energy levels. This spectrum is also a *line spectrum*. The absorption spectrum will be the complement of the emission spectrum (since the same amounts of energy are involved when electrons absorb and then emit light).

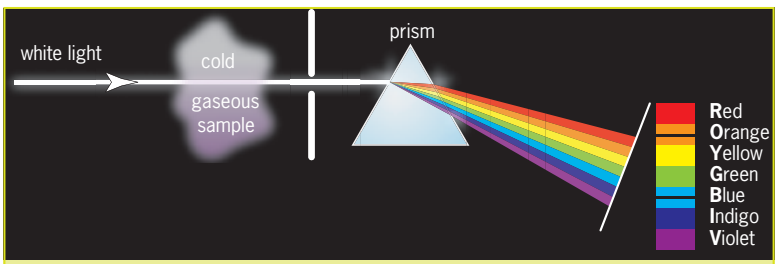


Figure 1.1.11 Light absorbed by atoms produces an absorption spectrum when passed through a prism.

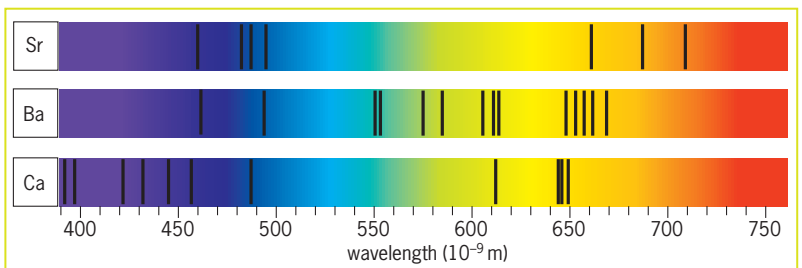


Figure 1.1.12 Absorption spectra for a number of elements.

AS **A.2.3**
Describe the atomic and molecular processes in which absorption of energy takes place. © IBO 2007

As we have seen in the example of absorption spectroscopy, electromagnetic radiation interacts with atoms and molecules. The nature of this interaction depends upon the energy (and hence the wavelength) of the electromagnetic radiation. In the case of absorption spectra, the absorption of energy from visible light by atoms results in electrons being excited and moving to higher energy electron shells. When molecules are involved, electronic transitions are not always the only possibility. As the electromagnetic radiation being absorbed decreases in energy (ultraviolet/visible → infrared → radio waves), the way in which it affects the molecule changes.

If the energy is in the ultraviolet or visible region of the electromagnetic spectrum, enough energy is available for absorption that results in electrons moving up to higher energy levels (electron transition). The spectroscopic technique that makes use of this is called **UV–visible spectroscopy** (see section 1.6). If the electromagnetic radiation is in the infrared region, it is not of sufficiently high energy to move the electrons to higher energy levels. Instead the atoms in a molecule may be made to vibrate in a characteristic manner. The absorption of this electromagnetic radiation results in transitions in vibrational states. These absorptions are not perfectly sharp, they are, in fact, absorption 'bands'. High resolution infrared spectrometers show that for molecules such as HCl, the bands are actually made up of a series of closely spaced absorptions. This occurs because changes in rotation (rotational transitions) often accompany vibrational transitions. This analytical technique is called **infrared spectroscopy** (see section 1.2). At even lower energies, at the frequency of radio waves, the absorption of energy relates to transitions between nuclear-spin energy levels. This is **nuclear magnetic resonance (NMR) spectroscopy** (see section 1.3).

CHEM COMPLEMENT

Investigating space using spectra

Collecting samples of extraterrestrial material is difficult. While material from the Moon and Mars has been analysed directly, most of the information we have about materials in space comes from the light they emit or absorb. In 1802 William Wollaston, an English chemist, noted a number of dark lines in the continuous spectrum of sunlight. These lines were independently rediscovered in 1817 by the German physicist Joseph von Fraunhofer, who studied the lines systematically and measured the wavelengths at which they occurred. He later discovered dark lines in the spectra of stars and noted that some of the lines in these spectra were absent from the Sun's spectrum and vice versa. This clearly indicated that not all of the lines came from the Earth's atmosphere. This led to the proposition that these lines were caused by absorption of the light by elements in the upper layers of the Sun as it passed out towards the Earth. These lines were named Fraunhofer lines in recognition of the careful work done by Joseph von Fraunhofer.

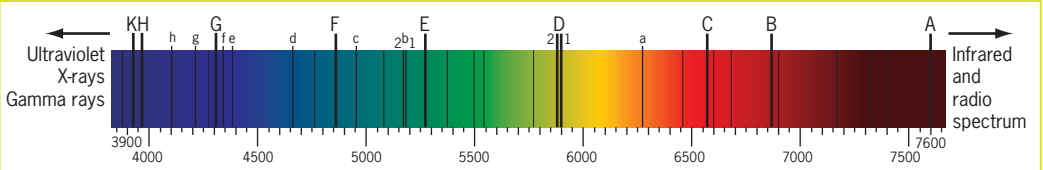


Figure 1.1.13 Fraunhofer lines represent the absorption of light from the Sun by elements in the Sun's outer layers.

Section 1.1 Exercises

1 Describe three general applications of analytical techniques.

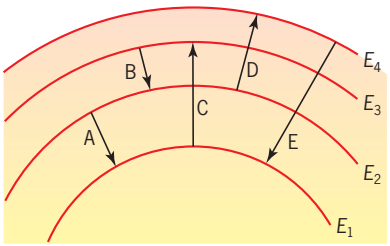
2 Describe how wavelength changes across the electromagnetic spectrum, giving specific examples of regions with long and short wavelengths.

3 Distinguish between wavelength and wavenumber.

4 Which one of the different colours of light in the visible light region of the electromagnetic spectrum has the highest energy? Explain your choice.

5 Explain the basis of all forms of spectroscopic analysis.

6 A representation of the energy levels of an atom is shown. Shown on the diagram are a number of electron transitions, labelled A to E.



a State which transitions would produce lines on an emission spectrum.

b State which transition produce radiation with the longest wavelength.

c Identify the transition that would emit or absorb the largest amount of energy.

d Identify the transitions that involve the absorption of energy.

7 Compare the emission and absorption spectra of hydrogen.

8 Describe what happens to the molecules of a compound when they absorb infrared light.

1.2 INFRARED SPECTROSCOPY

Infrared spectroscopy is a very useful form of spectroscopy, especially in the identification of organic compounds. The infrared region of the electromagnetic spectrum is not of high enough energy to excite electrons to higher states. However, molecules can absorb infrared radiation and changes occur to their bonds. Molecules are not rigid structures; their covalent bonds can be compared to springs that can be stretched or bent. The atoms within a molecule continually move, resulting in the vibrational and rotational motion of the molecule. Vibrations of the molecule equate to the atoms in the molecule changing position as their bonds bend or stretch. You will recall that electrons can only occupy discrete electronic energy levels (shells). Similarly, molecules are only able to occupy discrete vibrational energy levels. Absorption of infrared radiation will give a molecule enough energy to move from one vibrational energy level to a higher (excited) vibrational energy level.

AS **A.3.3 Explain what occurs at a molecular level during the absorption of IR radiation by molecules.** © IBO 2007

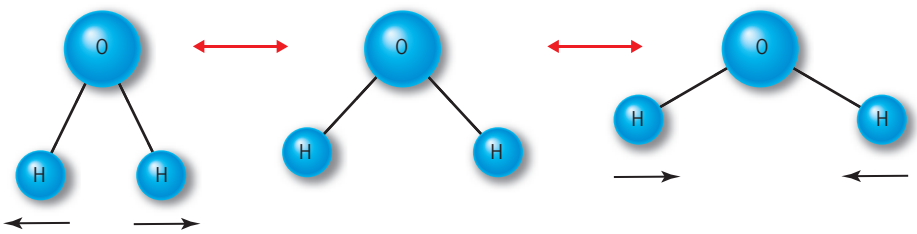


Figure 1.2.1 As a water molecule vibrates, the H–O–H bond angle changes.

Interactive CW Animation
Molecular vibrations

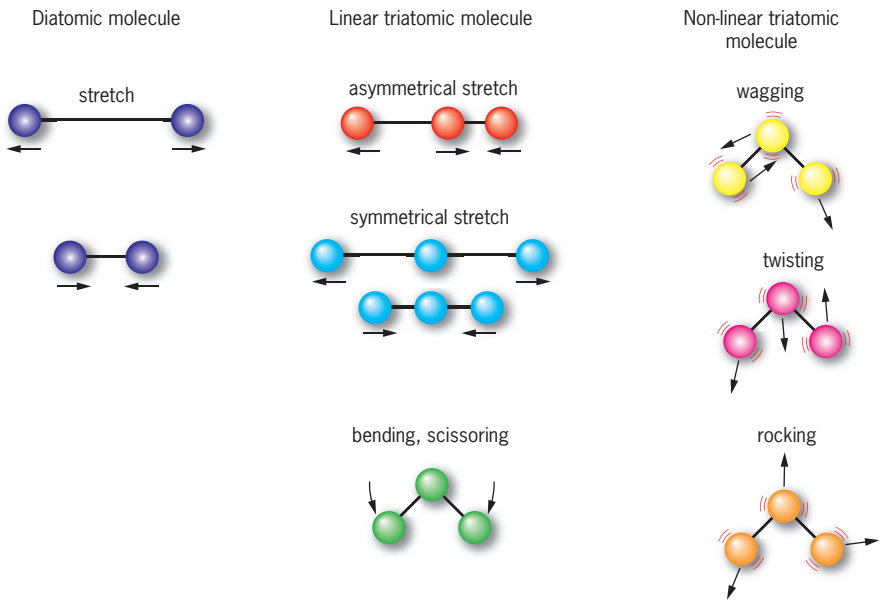


Figure 1.2.2 Diatomic and triatomic molecules can undergo a range of different stretching and bending motions.

When a molecule absorbs infrared radiation, the bending or stretching vibrations change the polarity of the bond and hence the overall dipole moment of the molecule. The dipole moment is a measure of the strength and direction of the charge separation in a molecule. It depends on both the polarities of the individual bonds and the geometry of the molecule. Consider the water molecule in figure 1.2.1. As the hydrogen atoms bend away from each other, the geometry of the molecule changes, so its dipole moment changes. Sulfur dioxide, SO_2, would show a similar change in dipole moment as the molecule bends. Carbon dioxide, a linear triatomic molecule (see figure 1.2.2) could undergo symmetrical and asymmetrical stretching, as well as bending, which would also change its overall dipole moment from zero to a non-zero value. A $–CH_2–$ group in a hydrocarbon would change its geometry as it undergoes stretches similar to those of the non-linear triatomic molecule in figure 1.2.2. A small number of molecules, such as O_2 and N_2, do not absorb IR radiation because they are diatomic molecules made up of identical atoms. Neither the covalent bond, nor the molecule overall has a dipole moment.

Double-beam spectrometry

AS **A.3.1 Describe the operating principles of a double-beam IR spectrometer.**

The main features of a double-beam infrared spectrometer are:

- a source of infrared radiation
- a wavelength selector (**monochromator**)
- a beam splitter, which creates two beams that follow parallel paths
- a sample cell and a reference cell or disc made out of solid NaCl or KBr. (Glass and plastic cannot be used, as they absorb IR radiation, so are opaque to IR radiation.)

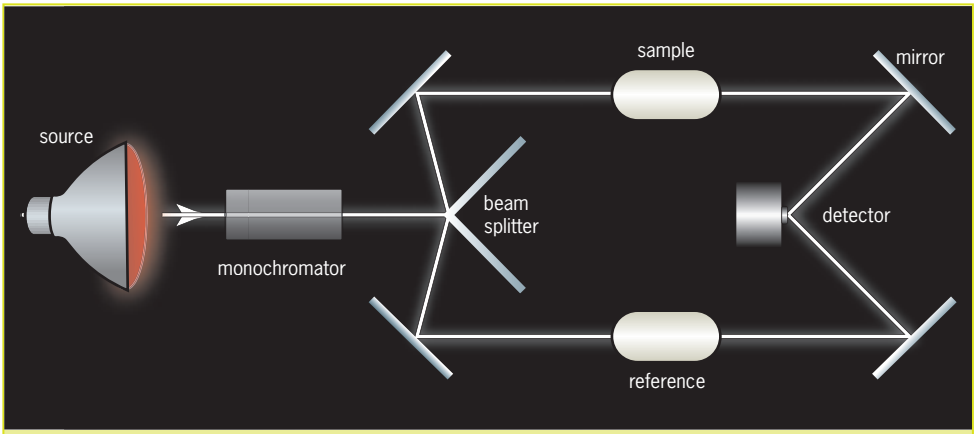


Figure 1.2.3 The double-beam IR spectrometer.

A double-beam infrared spectrometer (or spectrophotometer) allows the radiation passing through the sample to be continually compared with identical radiation that has not passed through the sample. Radiation from the source passes through a monochromator, which only allows radiation of a particular wavelength to pass through. This radiation then strikes a beam splitter and is split into two beams that pass along parallel paths. One beam passes through the cell containing the sample, the other passes through an identical reference cell that contains the same solvent etc. but no sample. The two beams are recombined at the detector. The signals from the sample and reference beams are compared electronically to determine the amount of absorbance due to the sample.

Analysing an IR spectrum

Organic molecules are usually made of several different atoms joined by a variety of covalent bonds. The absorbance of infrared radiation affects these bonds differently. In particular, infrared spectroscopy is important for the information we can obtain about the **functional groups** in the molecule. These groups differ in their composition (the atoms involved in the group) and in the strength of the bonds between the atoms (double or triple bonds). Generally, the range of energies absorbed depends on the strength of the bonds. The stronger the bond (higher bond enthalpy), the higher the value of energy that it absorbs.

AS A.3.2
Describe how information from an IR spectrum can be used to identify bonds.

TABLE 1.2.1 BOND STRENGTH (BOND ENTHALPY) AND ENERGY NEEDED TO EXCITE A MOLECULE TO A HIGHER VIBRATIONAL ENERGY LEVEL

Bond	Average bond enthalpy (kJ mol^{-1})	Infrared absorption range due to stretching vibrations (cm^{-1})
C=C	612	1610–1680
C≡C	837	2100–2260
C–O	360	1050–1410
C=O	743	1700–1750

The mass of the atoms attached by the bond also affects the energy of the infrared radiation absorbed. The higher the mass, the lower the value of the energy of the radiation absorbed. This can be seen in the absorbance of the halogenoalkanes.

TABLE 1.2.2 THE EFFECT OF CHANGING THE MASS OF THE HALOGEN ON THE INFRARED ABSORPTION

Bond	Infrared absorption range due to stretching vibrations (cm^{-1})
C–I	490–620
C–Br	500–600
C–Cl	600–800
C–F	1000–1400

Functional groups in different molecules absorb IR radiation at slightly different values, so an absorption range, rather than a particular value, is assigned to each functional group. The position of an infrared band in an IR spectrum is defined by a wavenumber (cm^{-1}). Radiation with a short wavelength will have a high wavenumber, since a large number of cycles will fit into a centimetre. Conversely, as wavenumber is the reciprocal of wavelength, radiation with a high wavenumber will have a short wavelength. The intensity of a peak on an IR spectrum is measured as % transmittance. A low value of % transmittance indicates that there has been strong absorption of the radiation with that wavenumber.

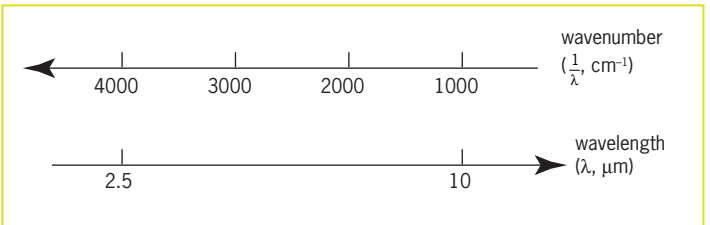


Figure 1.2.4 Wavenumber is the reciprocal of wavelength, so radiation with a high wavenumber has a short wavelength.

AS A.3.4 Analyse IR spectra of organic compounds. © IBO 2007

The result of passing infrared radiation through an organic compound is an infrared spectrum (see figure 1.2.5) showing the % transmittance (or simply transmittance) at different wavenumbers.

As illustrated in figure 1.2.5, each trough (upside-down peak) is called a *band*. Recall that the strength of the covalent bond will determine the energy of radiation absorbed. Each type of bond absorbs infrared radiation over a typical narrow range of wavenumbers. Like a spring, the bond is stretched by the absorption of energy. For example, radiation of higher energy (and hence higher wavenumber) is needed to stretch an O–H bond in an alcohol than a C–H bond, and much less energy is needed to stretch a C–O bond in the same alcohol (see figure 1.2.6).

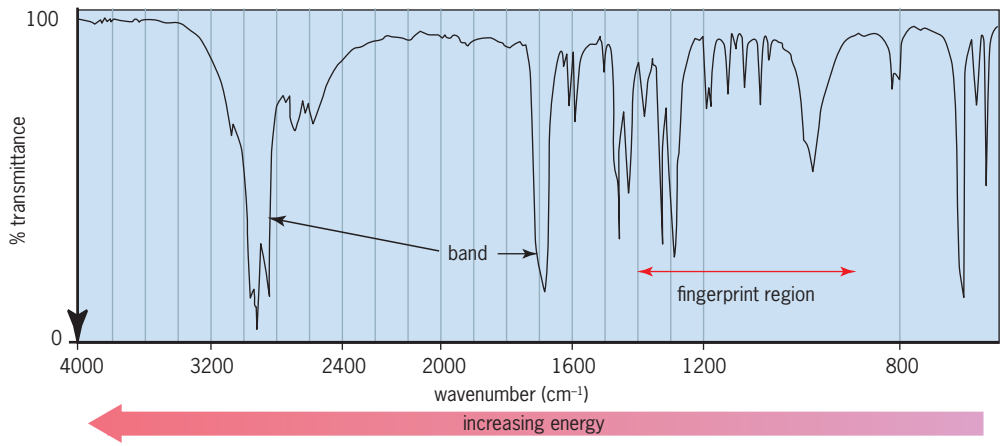


Figure 1.2.5 An infrared spectrum shows a complex series of bands.

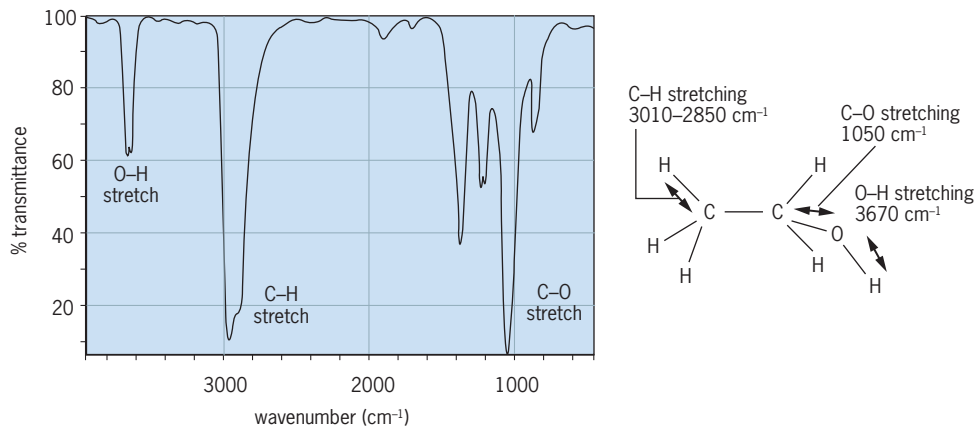


Figure 1.2.6 The infrared spectrum of ethanol shows bands corresponding to the stretching of O–H, C–H and C–O bonds.

Because the functional groups in the organic compound being analysed absorb at characteristic wavenumbers, these characteristic spectral bands can be used to identify specific functional groups within a molecule. Table 1.2.3 shows some of the common functional groups and their characteristic wavenumbers. Some bands, such as the C=O (carbonyl) band, are very useful because they are specific to a limited number of compounds (carboxylic acids, ketones, aldehydes and esters). Others, such as the C–H band, are of limited use because there are so many of them in all organic compounds.

TABLE 1.2.3 CHARACTERISTIC RANGES FOR INFRARED ABSORPTION DUE TO STRETCHING VIBRATIONS OF VARIOUS FUNCTIONAL GROUPS
(These values can also be found in table 17 of the IB Data Booklet. © IBO 2007)

Functional group	Organic molecules in which it is found	Wavenumber (cm^{-1})
C–I	Iodoalkanes	490–620
C–Br	Bromoalkanes	500–600
C–Cl	Chloroalkanes	600–800
C–F	Fluoroalkanes	1000–1400
C–O	Alcohols, esters, ethers	1050–1410
C=C	Alkenes	1610–1680
C=O	Aldehydes, ketones, carboxylic acids, esters	1700–1750
C≡C	Alkynes	2100–2260
O–H	Carboxylic acids	2500–3300
C–H	Alkanes, alkenes, arenes	2850–3100
O–H	Alcohols, phenols	3200–3600
N–H	Primary amines	3300–3500

Interpreting an infrared spectrum requires considerable skill and practice. Bands may overlap, and various neighbouring groups in the molecule may shift the absorption bands of a particular group. Figure 1.2.7 shows the infrared spectrum for 2-methylpropanoic acid, with the characteristic absorption peaks labelled.

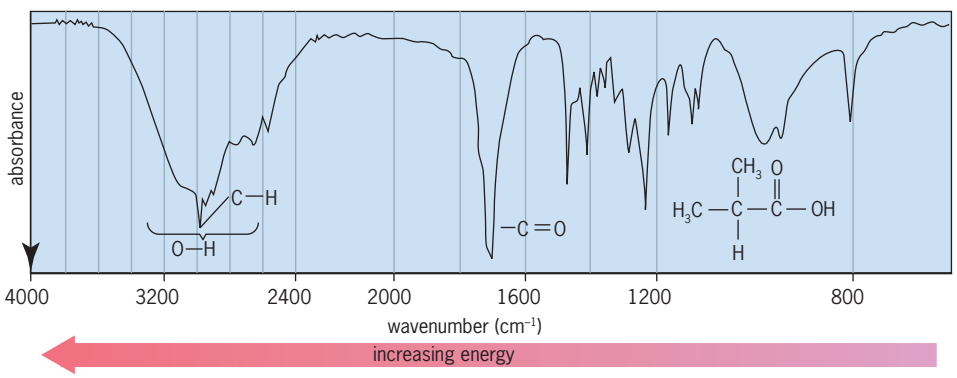


Figure 1.2.7 Infrared spectrum of 2-methylpropanoic acid ($CH_3CH(CH_3)COOH$).

Note that functional group isomers may have very similar infrared spectra. In such cases it may only be possible to identify the bonds present in a molecule, rather than the functional groups.

These difficulties aside, we can look at infrared spectra and recognize some of the more striking bands, and correlate the information with that provided by other types of spectral data. An experienced interpreter would, of course, be expected to gain far more information.

In addition to providing information about functional groups, the infrared spectrum of a molecule can be used as its 'fingerprint'. The region between 900 cm^{-1} and 1400 cm^{-1} of the absorption spectrum is complex and is known as the *fingerprint region*. If this region of the spectrum of an unknown compound is the same as that of a known compound, an identification can be made. Note that the two spectra in figure 1.2.8 have quite different fingerprint regions, even though they both show absorption peaks that correspond to the stretching of C–H bonds and C=O bonds.

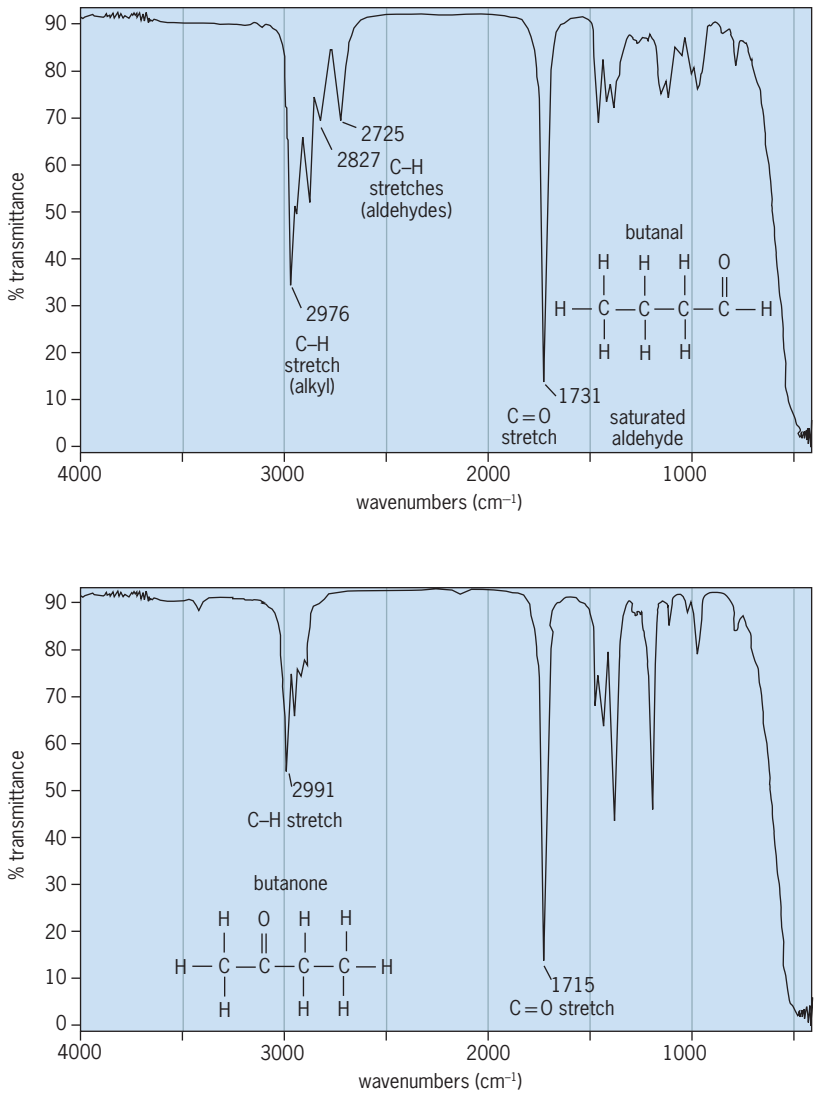


Figure 1.2.8 The infrared spectra of butanal and butanone are very similar.

Worked example 1

The structural formula and infrared spectrum of ethanol is shown in figure 1.2.9. Use table 1.2.3 to identify the bond types corresponding to peaks A, B and C

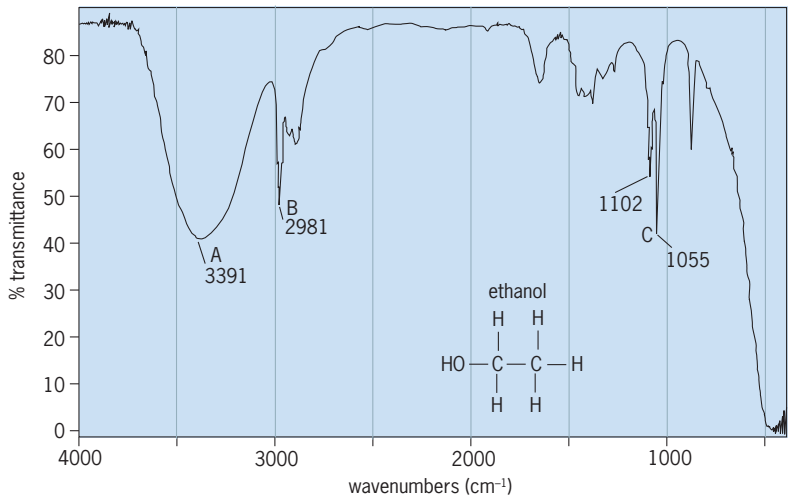


Figure 1.2.9 The infrared spectrum of ethanol.

Solution

Peak	Wavenumber (cm^{-1})	Wavenumber is in the range of absorption for
A	3391	O–H bond in alcohols
B	2981	C–H bond
C	1102 and 1055	C–O bond in alcohols

The spectra of even simple molecules with only one functional group will have so many absorption bands that it is not feasible to try to assign every band in the IR spectrum. Instead, look for obvious bands in the region from 4000 to 1300 cm^{-1}. This will help you to determine the presence of specific bonds that indicate particular functional groups. The flow chart in figure 1.2.10 may be helpful in developing habits for analysing infrared spectra.

Infrared spectra may be complex, but modern infrared spectrophotometers assist identification by providing computerized analysis and presentation of data. Characteristic spectra for thousands of compounds have been produced and stored in databases. The spectrum of an unknown compound can quickly be compared to this reference library of spectra. Even if the unknown compound is a new, previously unrecorded one, comparing its infrared spectrum with those in libraries of characteristic absorption peaks can provide information about possible functional groups and sections of the molecule.

Infrared spectroscopy is a very powerful tool that can be applied to a large range of samples. It can be used to identify the products of organic synthesis reactions. Consider the oxidation of an alcohol to a ketone. If this reaction has

CHEM COMPLEMENT

Increasing degradability

Plastic packaging generates a huge amount of waste. One approach to the problem of the disposal of plastic packaging is to make plastics that are biodegradable or photodegradable. Photodegradable plastics require the inclusion of light-sensitive additives. The carbonyl group (C=O) strongly absorbs infrared radiation (1680–1750 cm^{-1}), so if this group is incorporated into polymer chains used to produce plastic packaging, the items will absorb sunlight and degrade relatively quickly.

been successful, the product will show a carbonyl (C=O) band but no hydroxyl (O–H) band. If no carbonyl band is present, the experiment was not successful. If both carbonyl and hydroxyl bands are present, the oxidation may have progressed further to a carboxylic acid, or it is possible that the product is not pure and there is still alcohol present.

Infrared spectroscopy is often used in conjunction with other techniques such as UV spectroscopy, chromatography, mass spectroscopy and nuclear magnetic resonance spectroscopy to determine the structure of a substance. Applications of infrared spectroscopy include identifying narcotics, identifying compounds in forensic testing, analysing fuels for octane ratings, and determining the structure of naturally occurring substances that may be of use in the pharmaceutical industry.

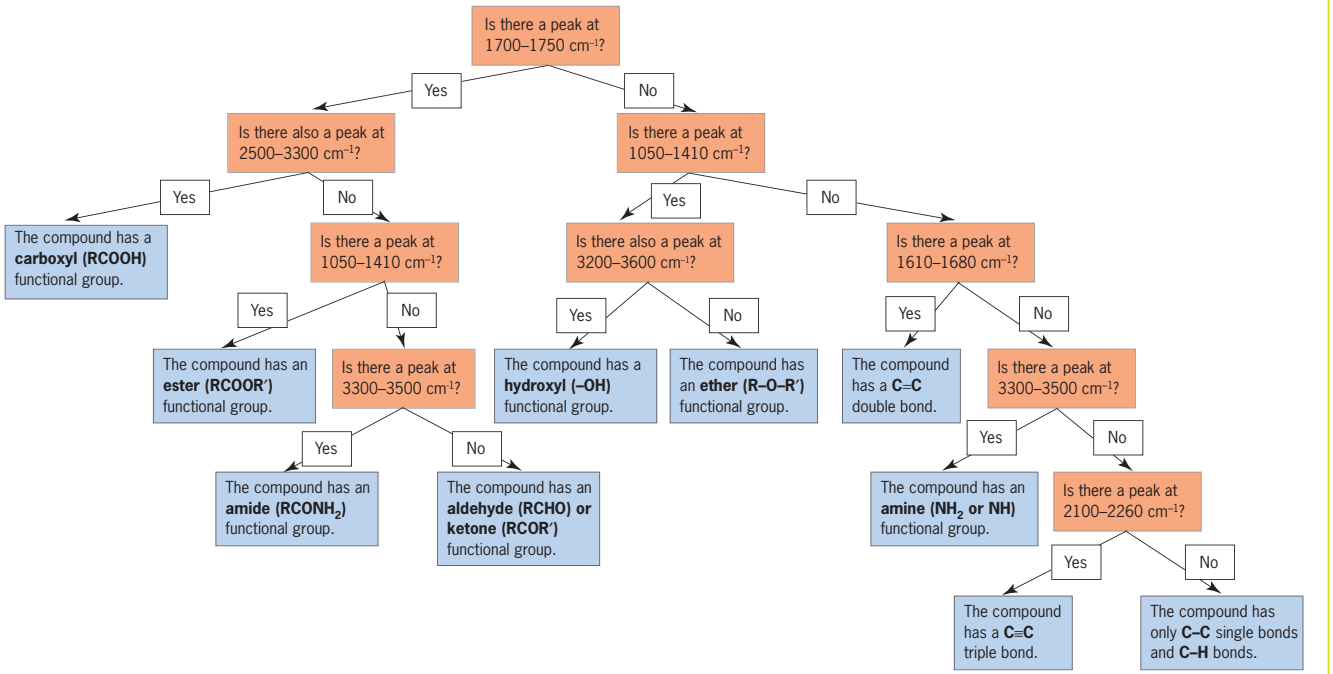


Figure 1.2.10 Flow chart for analysing an infrared spectrum.

Section 1.2 Exercises

1 An infrared spectrometer is used for the analysis of organic compounds. Describe the effect that the passage of infrared radiation has on an organic compound.

2 In several stages, draw an SO_2 molecule as it undergoes:

a a symmetrical stretch

b an asymmetrical stretch

c a bending vibration.

3 Calculate the wavenumber of infrared radiation with the following wavelengths. (Hint: convert the wavelength to cm first.)

a 2×10^{-5} m

b 8×10^{-6} m

c 2.5×10^{-6} m

4 Describe the operating principles of a double-beam infrared spectrometer.

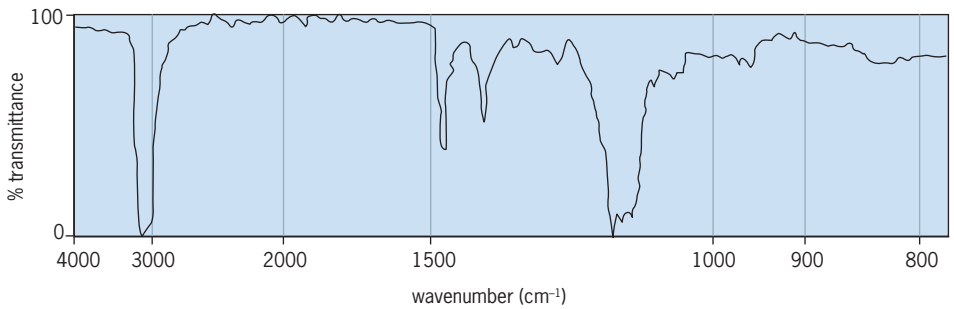
5 Explain why for infrared analysis the sample is placed in a cell, or between discs, of solid sodium chloride, rather than glass or plastic.

6 Describe the location (in terms of the wavenumber) of two peaks that you would expect to find on the infrared spectrum of a sample of propanol.

7 There is one peak that you would expect to occur in all hydrocarbons, that corresponds to C–H. Stating wavenumbers, identify the region in which you would find this peak.

8 The infrared absorption spectrum of an organic compound is shown below.

Use the information in table 1.2.3 to answer these questions.



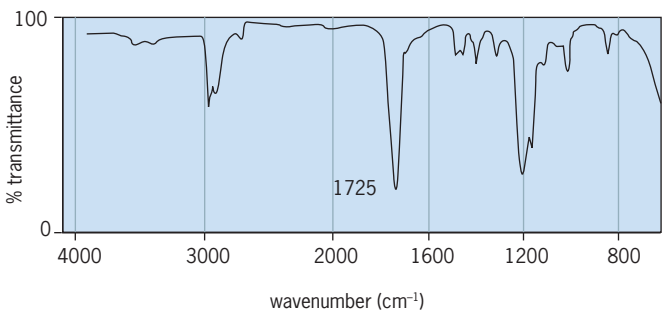
a **i** Suggest one functional group that is present in the compound.

ii Identify one functional group that is not present in the compound.

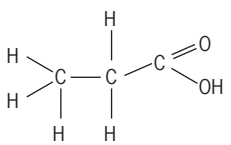
b **i** Stating wavenumbers, identify the region of the spectrum known as the fingerprint region.

ii Describe how this region is used in compound identification.

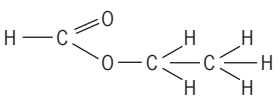
9 The infrared absorption spectrum of an organic compound and two possible structures for the compound are shown below.



A

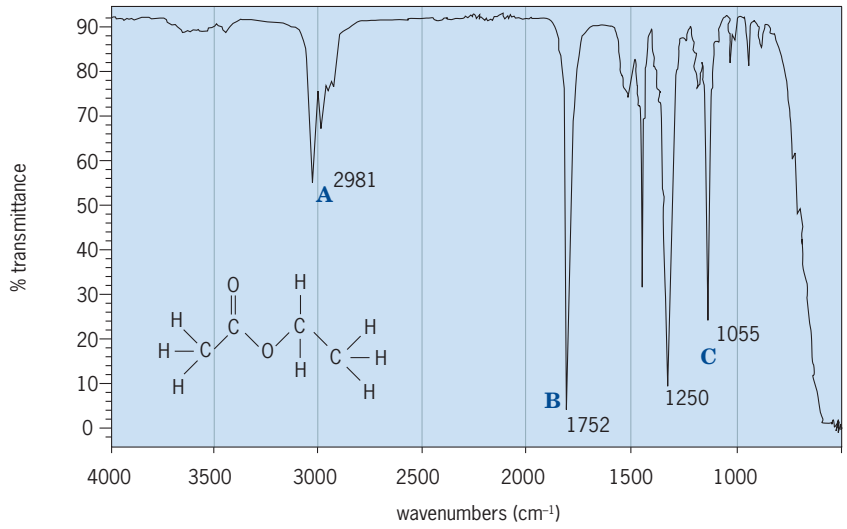


B



The compound is known to be either A, propanoic acid (CH_3CH_2COOH), or B, ethyl methanoate ($HCOOCH_2CH_3$). Using information from the spectrum and table 1.2.3, identify which compound it is. Explain your choice.

10 The figure below shows the structural formula and the infrared spectrum of ethyl ethanoate. Use table 1.2.3 to identify peaks A, B and C.



1.3 MASS SPECTROMETRY

Mass spectrometry was introduced in chapter 1 of *Chemistry: For use with the IB Diploma Programme Standard Level*. As early as the 1920s, mass spectrometry was used to identify the isotopes of elements, along with their relative isotopic masses (I_r) and abundances. Armed with this information, accurate relative atomic masses for elements could be determined. Today, the mass spectrometer is a powerful and accurate tool used extensively for the detection and identification of a wide range of compounds.

A schematic diagram of the mass spectrometer is shown in figure 1.3.1. Inside the mass spectrometer, the vaporized sample is bombarded with a high-energy electron beam to produce positive ions by knocking electrons off the sample atoms or molecules. The positive ions are accelerated by an electric field and then passed through a magnetic field. The positive ions moving in a magnetic field are deflected to varying degrees, depending on their mass-to-charge (m/z) ratio. Assuming that only one electron is knocked off the molecule or atom (this is almost always the case), we find that molecules or atoms with a greater mass are deflected less than those with a smaller mass. Thus, ions are separated by mass and collected, and the mass of each ion is determined. The intensity of the ion beam provides a measure of the abundance of the ion. Results are usually presented as a mass spectrum—a plot of relative mass versus absorbance.

Most modern spectrometers employ a double-focusing feature. This allows resolution of masses to a much greater accuracy than the simpler mass spectrometers. Essentially, the difference lies in the inclusion of an electrostatic analyser. This sorts the ions on the basis of their kinetic energy before they enter the

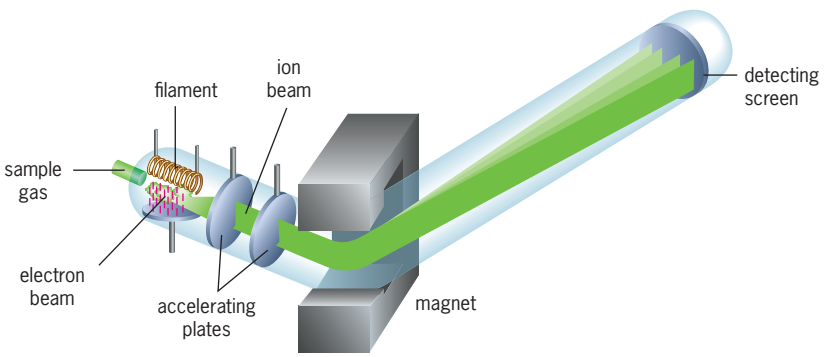


Figure 1.3.1 Schematic diagram showing the major components of a mass spectrometer.

magnetic field. Many instruments incorporate both simple- and double-focusing features, as it is not always necessary to work at very high resolutions.

Fragmentation

AS **A.4.1 Determine the molecular mass of a compound from the molecular ion peak.**

Mass spectrometry (also known as electron impact ionization mass spectrometry, EIMS) is used largely as an analytical tool for molecules, rather than for determination of relative isotopic masses. When a molecular substance is placed in the mass spectrometer, the bombarding high energy electrons break the molecule apart, or *fragment* it. The mass spectrum provides a pattern of peaks (a **fragmentation pattern**) representing the different fragments that have been formed. The masses of the fragments are measured in order to determine the original structure of the molecule.

On a mass spectrum, the horizontal scale is the mass to charge ratio (m/z) of the ions and the vertical scale is percentage abundance with respect to the most intense peak in the spectrum, the **base peak**. The peak with the greatest m/z ratio is the peak corresponding to the **molecular ion**, so this allows the molecular mass to be determined (although the presence of isotopes of chlorine or bromine and very tiny amounts of ^{13}C may have an influence on this—see below). The molecular ion is the ion formed when a molecule has only had an electron knocked off it. No covalent bonds have been broken, so the molecular ion has the same mass as the neutral molecule. In modern instruments the mass of this ion can be found with such precision that two molecules of similar mass can be separated. Thus, carbon monoxide (mass 27.995) can be distinguished from ethene (mass 28.032). Note that the molecular ion is not always the base peak. Having the greatest m/z ratio does not necessarily make this peak the most intense.

The high energy electron beam will initially knock just one electron from molecule M to form the molecular ion, a positive ion M^+:

$$M(g) + e^- \rightarrow M^+(g) + 2e^-$$

If the bombarding electrons in the mass spectrometer have sufficiently high energy they will also cause molecules to fragment, giving rise to a number of lower molecular mass ions.

$M^+(g)$	+	e^-	→	$(M-A)^+$	+	$A\cdot$
molecular ion				fragment ion		uncharged radical

In each fragmentation event there will be a fragment ion (which is accelerated and deflected in the mass spectrometer) and an uncharged free radical (which is neither accelerated nor detected, but its existence can be inferred by the difference in mass between the molecular ion and the observed fragment). The mass spectrum of these fragments is recorded in the same way as a mass spectrum of isotopes. The ions are accelerated by an electric field and are then deflected by a magnetic field. Then they are detected, and the signal received is converted into a mass spectrum. The molecular ion can break into almost every type of fragment, right down to individual atoms; however, only the positive ions reach the detector. The uncharged radicals are evacuated from the mass spectrometer.

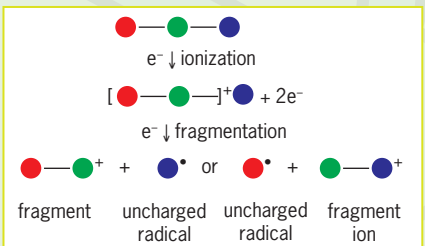


Figure 1.3.2 In the mass spectrometer the bombarding electrons ionize the molecule and may also create fragments of the molecule.

Analysing fragmentation patterns

The masses of the fragments in the mass spectrum give clues as to the structure of the molecule. For example, a fragment of an organic molecule with a mass of 15 is most likely to correspond to a CH_3 fragment ($12 + 3 \times 1 = 15$), whereas one with a mass of 17 would correspond to an OH fragment. If the masses of two fragments of an organic molecule differ by 14, this is most likely to be the result of the loss of $–CH_2$ ($12 + 2 \times 1 = 14$). Some common mass fragments formed by organic molecules are shown in table 1.3.1.

AS **A.4.2 Analyse fragmentation patterns in a mass spectrum to find the structure of a compound.**

TABLE 1.3.1 COMMON MASS SPECTRUM FRAGMENTS FORMED BY ORGANIC MOLECULES

Relative mass	Fragment	Calculation of mass of fragment
M	Molecular ion	*M*(compound)
15	CH_3	$12 + 3 \times 1 = 15$
M – 15	*M* – CH_3	
17	OH	$16 + 1 = 17$
M – 17	*M* – OH	
29	CH_2CH_3 and CHO	$2 \times 12 + 5 \times 1 = 29$ (CH_2CH_3) or $12 + 1 + 16 = 29$ (CHO)
M – 29	*M* – CH_2CH_3 and *M* – CHO	
31	CH_3O	$12 + 3 \times 1 + 16 = 31$
M – 31	*M* – CH_3O	
45	COOH	$12 + 2 \times 16 + 1 = 45$
M – 45	*M* – COOH	

Figure 1.3.3 shows the mass spectrum of 1-propanol ($CH_3CH_2CH_2OH$), with several fragment peaks labelled.

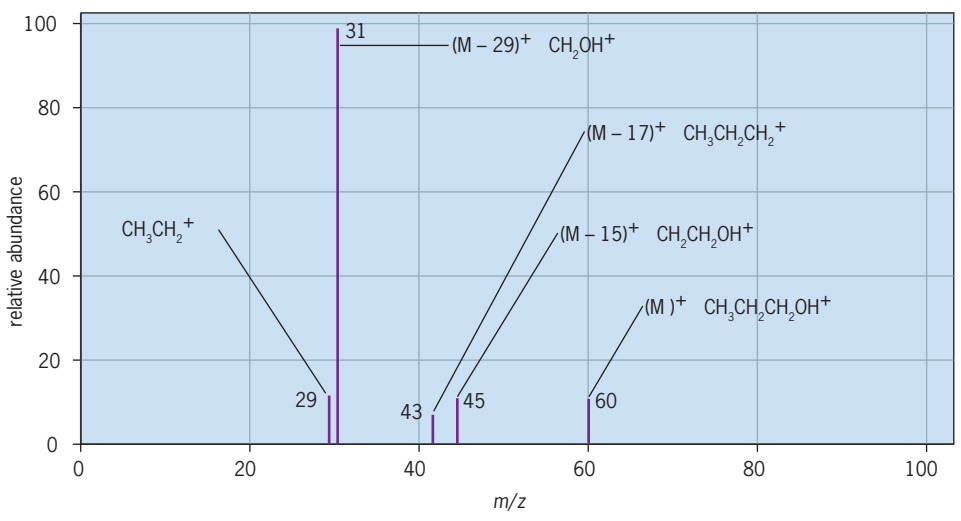


Figure 1.3.3 Mass spectrum of 1-propanol ($CH_3CH_2CH_2OH$).

Ws WORKSHEET 1.1
Analysing mass spectra

Isotopes in halogenoalkanes

Chlorine and bromine have isotopes that occur in sufficiently high abundance that they cause distinct peaks in the mass spectrum of halogenoalkanes. The two isotopes of chlorine are ^{35}Cl (75.5% abundance) and ^{37}Cl (25.5% abundance). If chlorine is present in the compound being analysed, there will be twin peaks for any fragment containing chlorine. For example, the molecular ion for CH_3Cl will have two peaks: one at $m/z = 50$ ($CH_3{}^{35}Cl^+$) and the other at $m/z = 52$ ($CH_3{}^{37}Cl^+$) with their heights in a 3 : 1 ratio. Bromine also has two isotopes: ^{79}Br (50.5% abundance) and ^{81}Br (49.5% abundance). The two peaks corresponding to the bromine isotopes in CH_3Br will have equal intensity. In a dihalogenoalkane such as CH_2Cl_2, the molecular ion will have three peaks since there are three possible combinations of the isotopes: $CH_2{}^{35}Cl_2$ ($m/z = 84$), $CH_2{}^{37}Cl_2$ ($m/z = 88$ and $CH_2{}^{35}Cl{}^{37}Cl$ ($m/z = 86$).

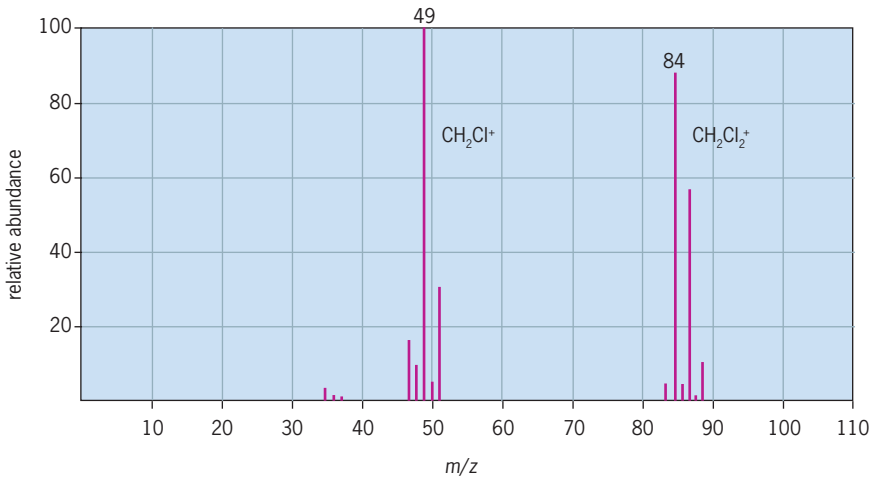


Figure 1.3.4 The mass spectrum of dichloromethane.

The mass spectrum is unique to a compound, so it can be used to identify a substance. When asked to find the structure of a compound from the fragmentation pattern, the best place to start is with the molecular ion peak. This will have the largest m/z ratio in the mass spectrum and, unless the compound is a halogenoalkane, it will be alone at the far right of the spectrum. The mass of the molecular ion peak will be equal to the molar mass of the compound. However, the more stable an ion is, the more likely it is to form, o resulting in a higher peak height. In some cases, such as for alkanes and alcohols, the molecular ion peak may be small, but it should not be ignored.

The next peaks to look for are the common ones listed in table 1.3.1. Remember if a peak occurs at $m/z = 15$, there may also be a peak at $M - 15$, thus confirming the molar mass. In figure 1.3.5, the mass spectrum of ethanoic acid, there is a peak at 15 and also one at 45. The molecular ion appears at $m/z = 60$.

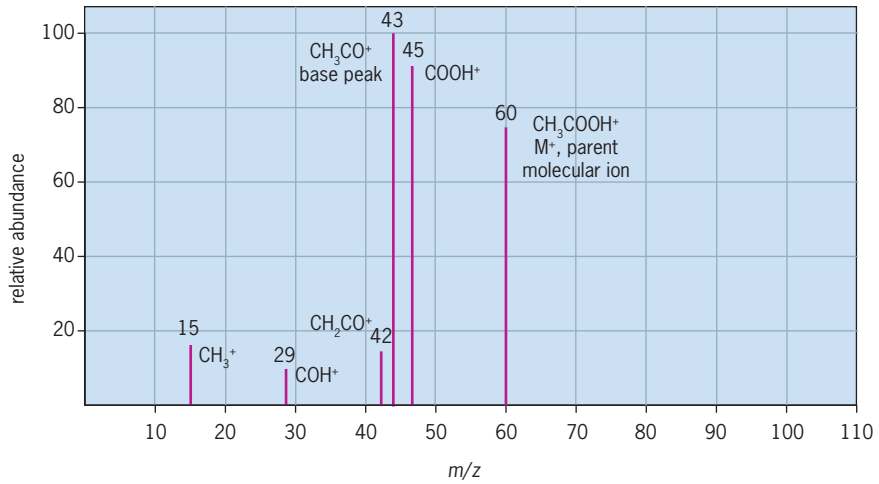


Figure 1.3.5 This mass spectrum of ethanoic acid shows the major peaks that occur.

Worked example 1

Carefully observe the following mass spectrum and deduce the structural formula of the compound to which it belongs.

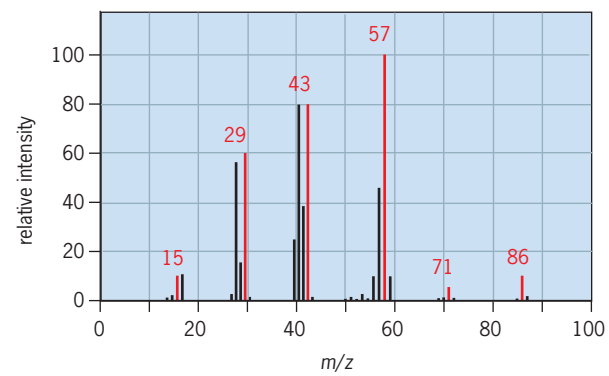


Figure 1.3.6 The mass spectrum of an unknown compound.

Solution

- The molecular ion peak has *m/z* ratio of 86, so the molecular mass is equal to 86.
- There is a peak at 15 (CH_3) and one at 71 ($M - 15$).
- There is a peak at 29 (C_2H_5).
- The fragmentation pattern contains clusters of peaks 14 mass units (CH_2) apart.
- The compound appears to be an alkane with molecular mass 86.
- Hexane, C_6H_{14}, has a molecular mass of 86.

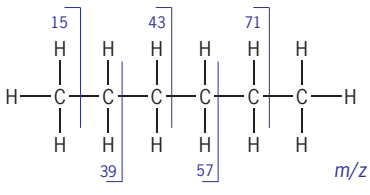


Figure 1.3.7 The fragmentation pattern of hexane.

Worked example 2

The mass spectrum of compound A is shown below. Deduce the structural formula of compound A.

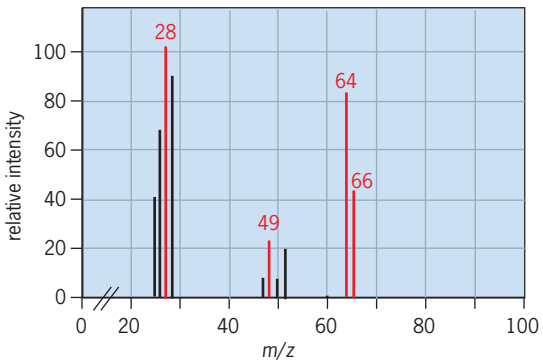


Figure 1.3.8 The mass spectrum of compound A.

Solution

- There are two molecular peaks at $m/z = 64$ and $m/z = 66$, with the second much smaller than the first, suggesting a chlorine atom.
- Molecular ion (M^+) peak has m/z ratio of 64: $M - Cl = 29$
- The base peak is at $m/z = 28$, with $m/z = 29$ (CH_3CH_2) beside it.
- The scale does not go down to $m/z = 15$; however, there is a peak at $m/z = 49$, which is equivalent to $M - 15$, so there must be a CH_3 fragment.
- The compound is most likely to be C_2H_5Cl.

Comparison of the mass spectrum of a substance with computer files of the mass spectra of many organic compounds allows for rapid identification of unknown substances. This technique is very useful in identifying the presence of performance-enhancing drugs in urine and blood samples. Mass spectra are also useful in determining the possible structure of newly synthesized or newly isolated compounds. Coupled with data from infrared spectra, a mass spectrum provides a powerful tool for identifying compounds.

THEORY OF KNOWLEDGE

If someone told you that finding new ways to solve some of society's greatest threats to human health involved using technology that has existed since the beginning of the 20th century, you would probably give them a strange look. One of the recent challenges of mass spectrometry has been its application to medical research. Traditional mass spectrometry is useful for analysing small to medium-sized molecules; but biological molecules, particularly macromolecules such as proteins, that are large, fragile to ionize and sensitive to temperature are not compatible with the high temperatures and high-energy electrons found in the average mass spectrometer.

Mass spectrometry has been part of the chemist's toolkit for many years, but being able to make accurate measurements of the molecular masses of large proteins was a dream. With many considering it impossible, two chemists using a lot of imagination and creativity caused a revolutionary breakthrough in our understanding of protein structure. In 1988 John Fenn from the USA published a paper on an electrospray mass spectrometry method that involved a simple modification to a typical mass spectrometer. Instead of the sample being vaporized and then ionized, the process was reversed, and the sample was first ionized in solution and then vaporized. At the same time, in Japan Koichi Tanaka, using a soft laser desorption method, found a way to use an intense laser pulse to ionize macromolecules. Since then these new technologies have become the standard methods for the structural analysis of biological molecules. Society has benefited from these advances in knowledge. In the field of drug development many more molecules can now be analysed, new mechanisms for the spread of malaria have been discovered, the early diagnosis of some cancers is now possible, and hazardous molecules formed during food preparation have been identified.

This story shows how, using technology, chemists can bring information, insights and analytical skills to bear on matters of public concern. It also highlights how long periods of continuity and major shifts in thinking are also persistent features of the way in which new knowledge is acquired in science. Like the advances in mass spectrometry, these changes often involve looking at existing knowledge in completely different ways.

- How do beliefs about what is valued in society influence the pursuit of which analytical technology tools will be developed and how they are used? Consider the social, political and economic forces.
- Compare the ways in which a technological invention such as the mass spectrometer might be comparable to imagining and creating a work of art.

Section 1.3 Exercises

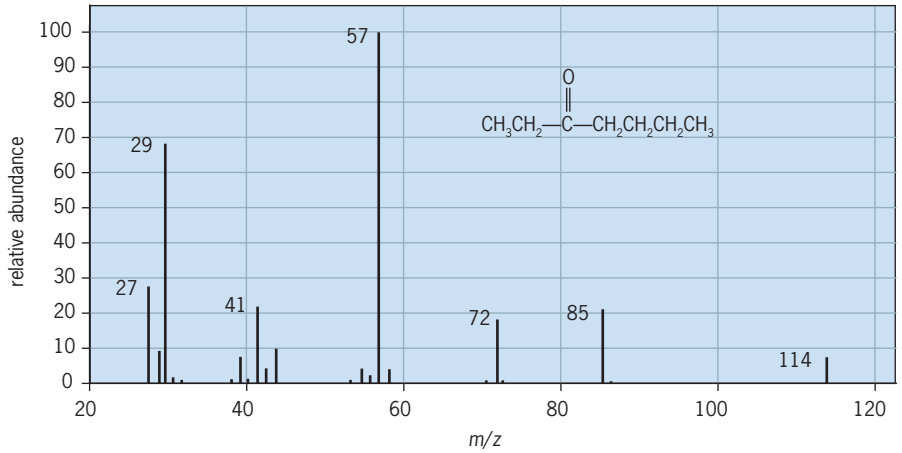
1 Describe how the molecules become positive ions in the mass spectrometer.

2 Complete the table below stating the masses of the following fragments.

Fragment ion	*m*/*z* ratio
CH_3^+	
$C_2H_5^+$	
OH^+	
CH_2Cl^+	
CH_3CO^+	
CH_3O^+	

3 Including the molecular ion, state the molecular formula and the *m*/*z* ratio of three fragments you would expect to find in the mass spectrum of methanol.

4 The mass spectrum of 3-heptanone is shown below.



Suggest the probable formula for the ions producing peaks at mass number:

a 114

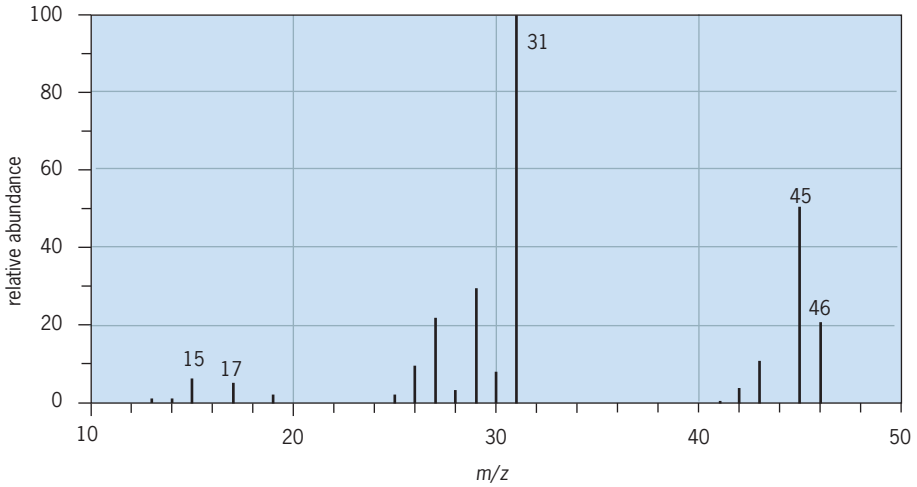
b 85

c 57

d 29

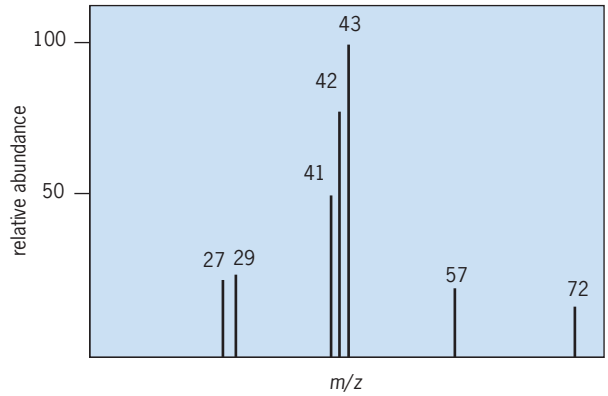
5 The mass spectrum of compound B contains major peaks at *m/z* = 88, 73, 45, 43 and 29. Confirm that this fragmentation pattern is consistent with compound B being ethyl ethanoate.

6 A student believes that the mass spectrum of compound C (below) is that of ethene. Explain why the student is mistaken, using at least two peaks on the spectrum to support your answer.



7 Referring again to the spectrum of compound C in question **6**, deduce the identity of compound C, giving reasons for your deduction.

8 The mass spectrum of compound D is shown below. Deduce the identity of compound D and explain why you made that deduction.



1.4 NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY

Nuclear magnetic resonance (NMR) spectroscopy makes use of the interaction between atoms and both a magnetic field and electromagnetic radiation. NMR spectroscopy is one of the most powerful tools for investigating the structure of organic molecules and it has important medical applications. The technique is based on the fact that the nuclei of some atoms, when placed in a strong magnetic field, absorb radiation in the radio wave region of the electromagnetic spectrum.

As early as 1930, it was found that certain atomic nuclei have a property called **spin**, and that, in spinning, these nuclei create a magnetic field and so behave as if they were tiny bar magnets. Nuclei that exhibit this property include ^{1}H, ^{13}C, ^{19}N and ^{31}P. They all have an odd number of nucleons (protons and neutrons). Of these, the most commonly used for analysis in NMR is the ^{1}H nucleus because it is so common in organic matter. When placed in a strong magnetic field, the ^{1}H nucleus has two possible orientations of its magnetic field: in the same direction as the magnetic field with spin $= +\frac{1}{2}$, or in the opposite direction with spin $= -\frac{1}{2}$. In figure 1.4.1 it can be seen that if radiation of the correct wavelength (and so of the correct energy) is applied to the nucleus, the energy is absorbed and the spinning nuclear magnet flips and becomes aligned at the higher energy state (spin $= -\frac{1}{2}$). The appropriate wavelength that causes this behaviour is in the radio wave region of the electromagnetic spectrum. When the flip occurs, the nucleus is said to be in *resonance* with the applied radiation, hence the name nuclear magnetic resonance.

A schematic diagram of the NMR spectrometer is shown in figure 1.4.2. A sample is dissolved in a suitable solvent and placed in a thin glass tube. The tube is placed between the

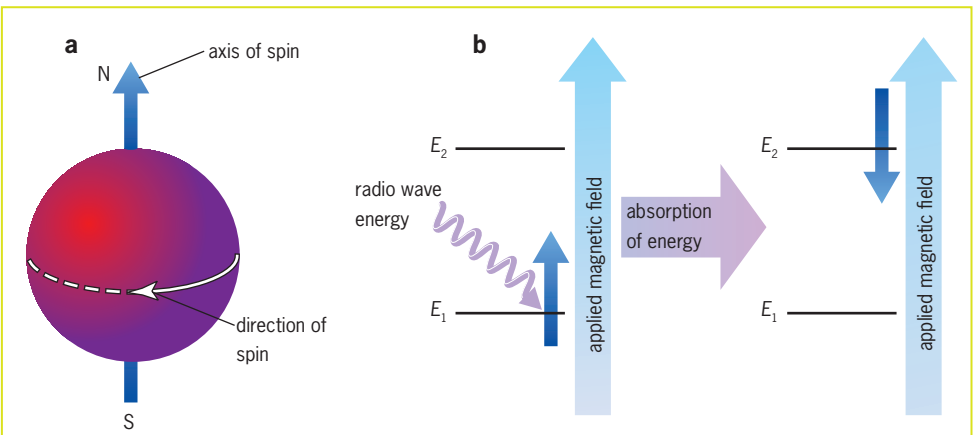


Figure 1.4.1 (a) A spinning nucleus creates a magnetic field and so acts like a tiny bar magnet. (b) Spinning nuclei have two energy levels (E_1 and E_2) in an applied magnetic field. Absorption of energy causes transitions between these energy levels.

poles of a magnet. A strong magnetic field is applied and the sample is irradiated with radio waves from an antenna coil (green coil in figure 1.4.2). The sample absorbs energy and hydrogen nuclei are excited. The energy is then emitted and is picked up by the radio receiver coil (blue in the figure 1.4.2) and the signal is passed on to the computer.

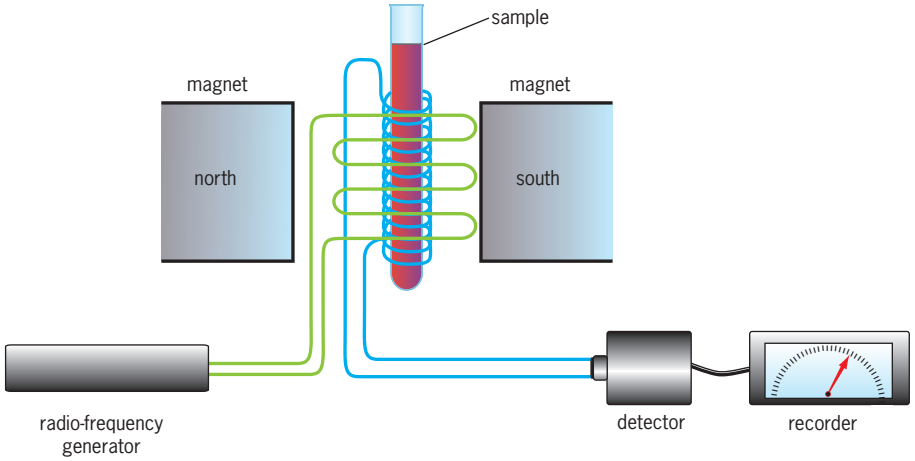


Figure 1.4.2 Schematic diagram of the NMR spectrometer.

Depending on the applied magnetic field, a different frequency of radio waves will be needed to excite the nuclei. As the strength of the applied magnetic field increases, the frequency (and hence energy) of the radio waves needed to excite the nuclei from spin $= +\frac{1}{2}$ to spin $= -\frac{1}{2}$ will increase.

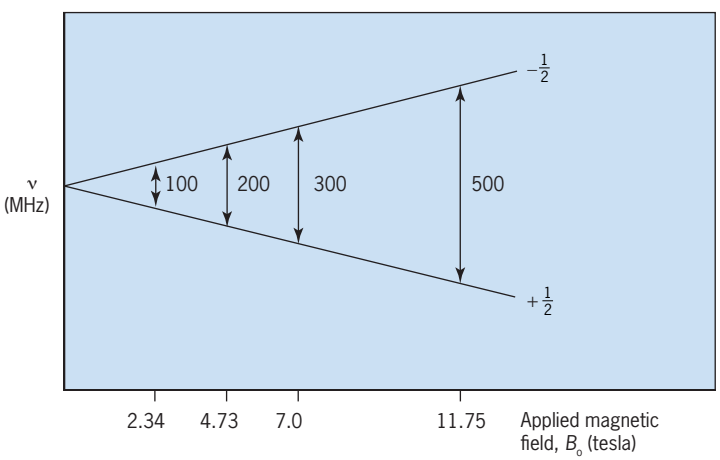


Figure 1.4.3 As the strength of the applied magnetic field increases, the frequency (and hence energy) of the radio waves needed to excite the nuclei will increase.

In practice, the wavelength is kept constant and the nuclei are exposed to a range of applied magnetic field strengths. As the field strength increases, ^{1}H nuclei absorb and produce an NMR signal. The environment of the ^{1}H nuclei influences the strength of the magnetic field that it experiences. The atom to which the hydrogen is bonded and the presence of other atoms bonded to that atom change

CHEM COMPLEMENT

Spinning nuclei

The basic research that culminated in NMR spectroscopy dates to the 1930s, when two physicists, Otto Stern and Isidor Rabi, determined that certain nuclei have spin. For their research into this phenomenon, Stern and Rabi were awarded the Nobel Prize in Physics in 1943 and 1944 respectively. In the 1940s, Edward Purcell and Felix Bloch developed the technique of NMR to work out complex molecular structures. Purcell and Bloch shared the Nobel Prize in Physics in 1952 for this research. In the 1960s, scientists used the principles of NMR to develop the diagnostic tool of magnetic resonance imaging (MRI). The initial basic research into nuclear spin ultimately led to the development of this important medical tool. In 2003 the Nobel Prize for Physiology and Medicine was awarded to Paul C. Lauterbur and Sir Peter Mansfield for their work on magnetic resonance imaging.

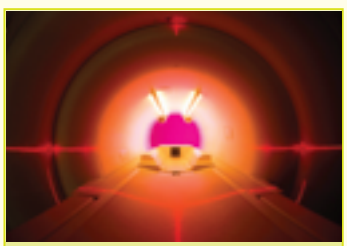


Figure 1.4.4 A view of the inside of an MRI (magnetic resonance imaging) scanner.

the way in which the hydrogen nucleus responds to a given magnetic field. Electrons around each nucleus are spinning and so have an associated magnetic field that shields the nucleus from the applied magnetic field. In response, a greater amount of energy is needed for the nucleus to be excited.

An actual NMR spectrometer is a very expensive instrument that is highly dependent on modern computer technology. In the past 20 years this analytical tool has advanced significantly due to advances in superconducting magnets and computers. The magnetic fields that are used are very great indeed, making the NMR measurements more sensitive than with smaller magnetic fields. If we compare the magnetic field used in an NMR spectrometer to the approximate value of the Earth's magnetic field, we find that the field in the NMR spectrometer is approximately 10 000 times that of the Earth's magnetic field. However, even with such high magnetic fields, the change in energy between the two spin states of the ^{1}H nuclei is very small: 10^{-5} times that of the energy transitions undergone in infrared spectroscopy and 10^{-7} times that of electronic transitions in UV–visible spectroscopy.

Deducing structures using ^{1}H NMR spectra

AS **A.5.1 Deduce the structure of a compound given information from its ^{1}H NMR spectrum.**

The exact point of absorption of energy by the ^{1}H nucleus also depends on its immediate electronic environment. Electrons close to the ^{1}H nucleus shield it slightly from the magnetic field and so alter the position at which energy is absorbed. Thus the energy absorbed by a ^{1}H nucleus bonded to an oxygen atom (–OH) will be different from that absorbed by a ^{1}H nucleus bonded to a carbon atom (for example, in a CH_3 group of an organic compound). Chemically equivalent nuclei (those in the same environment) all absorb at the same position on the NMR spectrum, while those in different environments absorb at different positions. The position of the peak along the horizontal axis of an NMR spectrum is known as the **chemical shift**. Hydrogen atoms that are bonded to more electronegative atoms have a greater shift than those in more electron-dense environments. Chemical shift is discussed in more detail in the higher level part of this option.

Worked example 1

The structural formula of butan-1-ol is given in figure 1.4.5. In this molecule the hydrogen atoms exist in four different environments.

1 The three green hydrogen atoms are in the CH_3 group at the end of the chain.

2 The two pink hydrogen atoms are in a CH_2 group adjacent to the CH_3 and CH_2OH groups.

3 The two blue hydrogen atoms are in a CH_2 group that is bonded to a carbon atom that also has an –OH bonded to it.

4 The one red hydrogen atom is bonded to the oxygen atom.

When an NMR spectrum for this compound is examined, we find four peaks (from lowest chemical shift to highest) with a peak area ratio 3:2:2:1.

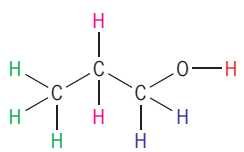


Figure 1.4.5 The structural formula of propan-1-ol showing the four different environments of the hydrogen atoms.

The intensity of a signal on an NMR spectrum gives an indication of the number of nuclei causing the signal. An NMR spectrum can therefore provide information on the 'types' of ^{1}H nuclei in molecules, as well as (from the relative intensities of the signal) the number of nuclei of each type.

The area under each signal peak is measured electronically (integrated) and the results presented in a 'stair-step' fashion. The height of each step is

proportional to the number of nuclei causing the peak. Therefore, to compare peak areas we simply need to measure the step heights in the integrated spectrum. WORKThese integrated spectra are extremely useful in interpreting ^{1}H NMR spectra.

WORKSHEET 1.2
Chemical detectives

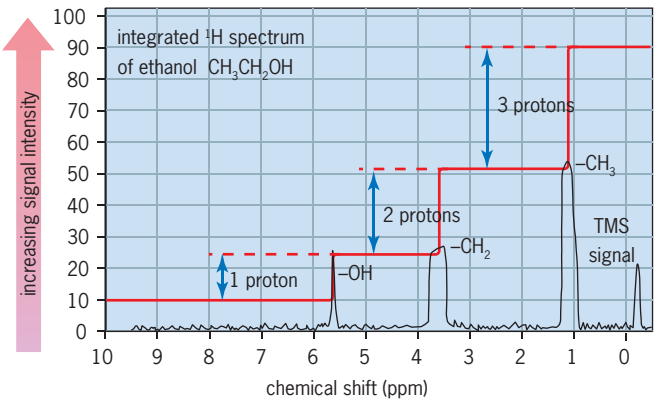


Figure 1.4.6 An integrated NMR spectrum allows determination of the relative numbers of each type of nucleus.

Worked example 2

The NMR spectra and structural formulas of propanone and of dimethoxymethane are given in figure 1.4.7.

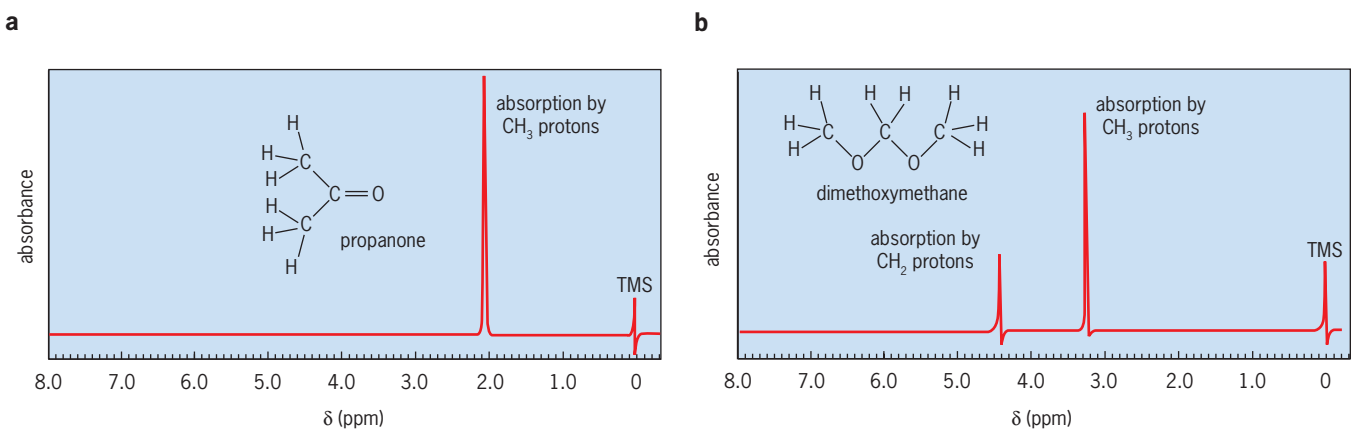


Figure 1.4.7 The NMR spectra and structural formulas of (a) propanone and (b) dimethoxymethane. The NMR spectra of these compounds are very straightforward due to the carbonyl and ether functional groups.

a Propanone (CH_3COCH_3): There is only one environment for the hydrogen atoms in this molecule. The six hydrogen atoms are identically placed in two $–CH_3$ groups bonded to a carbonyl group. As a result there is only one peak in the NMR spectrum. (Note that the peak marked TMS is a reference peak and not due to the propanone sample.)

b Dimethoxymethane ($CH_3OCH_2OCH_3$): This ether has hydrogen atoms in two different environments. Six of the eight hydrogen atoms are part of $–CH_3$ groups that are bonded to the ether functional group (–O–). The other two hydrogen atoms are in a $–CH_2–$ group in the middle of the molecule (bonded to two ether functional groups). The NMR spectrum has two peaks with intensities in the ratio 3:1, resulting from the 6:2 ratio of hydrogen atoms.

Magnetic resonance imaging

A.5.2 Outline how NMR is used in body scanners. © IBO 2007

Since the 1960s, nuclear magnetic resonance has also been applied to scans of the human body, used in medical diagnosis. This application, known as **magnetic resonance imaging (MRI)**, allows the visualization of internal organs. An MRI scanner works in the same way as a NMR spectrometer except that the patient takes the place of the sample.

Protons in water molecules within human cells are detected in an MRI scan. For a particular magnetic field strength, water absorbs radiation at a particular frequency. Different parts of the body have different tissue water contents and therefore absorb radiation to different extents. As the MRI scanner is moved around the body, absorption data is accumulated. Cells in normal tissue respond differently from cells in a tumour. The MRI scans the body systematically, producing pictures that are 'slices' through the body. The final product is a three-dimensional image of various organs, depending on the requirements of the scan. Other nuclei, ^{31}P for example, may also be scanned for and used in diagnosis. Nuclear magnetic resonance is especially useful in medicine because the energy of the radio waves is harmless and there are no known side-effects. Patients are conscious during the MRI scan and their responses to stimuli can be measured. For example, the way the brain works when a patient is shown images, hears sounds and carries out other mental processes can be measured using MRI. The non-invasive procedure of MRI has replaced many invasive, exploratory surgical procedures. The images produced are used to diagnose and monitor conditions such as cancer and tissue damage that results from a heart attack.

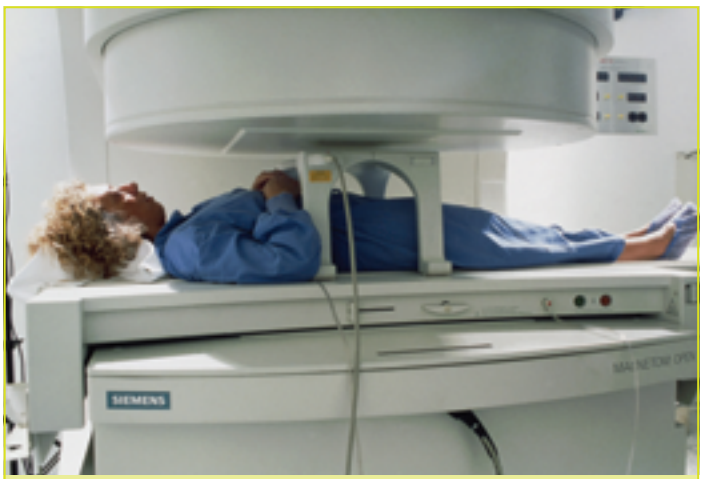


Figure 1.4.8 A patient undergoing a scan in a MRI scanner.

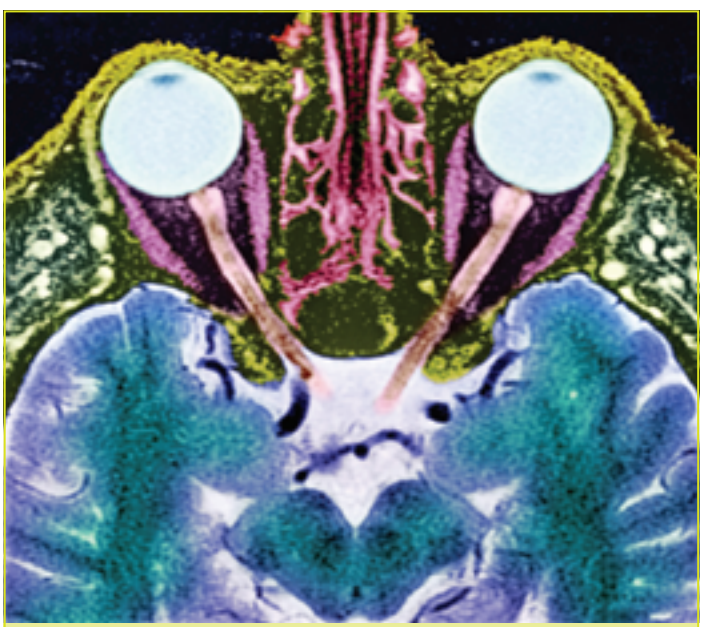


Figure 1.4.9 An MRI scan of a cross-section through the head showing the eyes, optic nerves and brain.

THEORY OF KNOWLEDGE

One important way of understanding the nature of matter at a molecular level is to develop technology tools to 'see' with. Advances in mass spectrometry have enabled us to fill in the gaps in what we know about protein structure, but NMR technology has added to this body of knowledge by providing three-dimensional pictures of protien molecules. In the 1980s, while Fenn and Tanaka were expanding the applications of mass spectrometry to biological molecules, Swiss chemist Kurt Wüthrich was trying to use NMR spectroscopy to determine the structures of proteins and nucleic acids. Facing the challenge of analysing the structures of these large molecules with very complicated spectra and many peaks, Wüthrich found a way to determine which peaks corresponded with which atoms. What was ingenious about his discovery was that he was able to carry out the analysis in an aqueous solution, similar to the molecules' natural environment in a cell. This technique is now widely used to study the structure of proteins, and has made it possible for other scientists to extend their research using NMR spectroscopy in new and entirely different ways. Wüthrich shared the 2002 Nobel Prize in Chemistry with Fenn and Tanaka.

- Comment on the claim that 'Technology affects society more directly than science because technology solves practical problems and serves human needs.'
- Is scientific knowledge valued more for its own sake or for the technology that makes it possible?
- The field of modern analytical chemistry today depends entirely upon technology for its existence. What are the implications for the advancement of our knowledge of molecular structure?

Section 1.4 Exercises

1 NMR spectroscopy involves the interaction of atoms with two influences. Describe the two influences on atoms that are used in NMR spectroscopy.

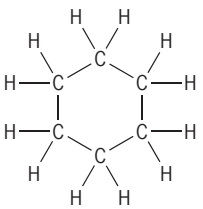
2 Describe the property of nuclei that allows them to behave like tiny bar magnets.

3 Describe the operating principles of an NMR spectrometer.

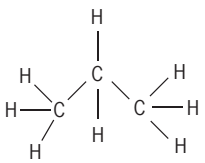
4 Identify the factors that influence the number and the relative sizes of the peaks on a ^{1}H NMR spectrum.

5 Describe the ^{1}H NMR spectrum of each of the following compounds in terms of the number of peaks and their relative sizes.

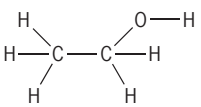
a Cyclohexane



b Propane

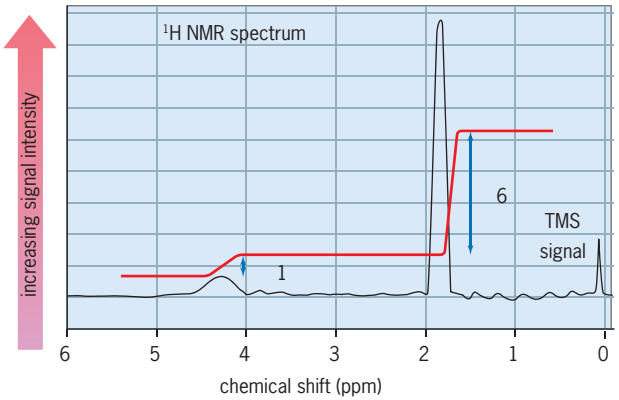


c Ethanol



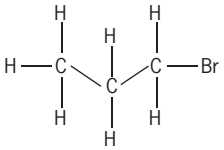
6 A ^{1}H NMR spectrum consists of two peaks, one with a relative peak area of 3 and the other at a higher chemical shift and a relative peak area of 2. Explain whether it is possible that this ^{1}H NMR spectrum could correspond to butane, C_4H_{10}.

7 The ^1H NMR spectrum of a compound of molecular formula C_3H_7Br is shown below.

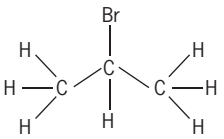


possible structures

X



Y

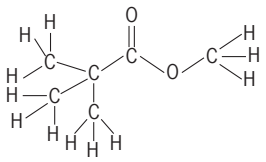


a State how many types of ^1H nuclei the molecule contains.

b State the ratio of numbers of these types of nuclei.

c Two possible structures for the molecule are also shown. Deduce which one is correct: X or Y?

8 On the ^1H NMR spectrum of the compound whose molecular structure is shown:



a How many peaks would you expect to see?

b What is the expected ratio of the areas under these peaks?

9 Describe how magnetic resonance imaging (MRI) can be used to locate a tumour on a human brain.

1.5 ATOMIC ABSORPTION SPECTROSCOPY

AS A.6.2 Describe the principles of atomic absorption. © IBO 2007

Studying a substance by analysing the absorption of electromagnetic radiation is called **atomic absorption spectroscopy**. Consider again what happens when electrons in atoms become excited—they absorb a specific amount of energy to move from a low energy level to a higher one. If this energy for excitation is provided by light (rather than heat), an absorption spectrum can be created (see figure 1.1.12, p. 7). This spectrum consists of a coloured background (the incoming white light spectrum) with a series of black lines that correspond to those energies of light absorbed by the electrons moving between levels. This absorption spectrum will be the complement of the emission spectrum, since the same amounts of energy are involved when electrons absorb and then emit light. (See figure 1.1.10, p. 7)

The use of atomic absorption (AA) spectroscopy was pioneered by Australian chemist Sir Alan Walsh (1916–1998) in the 1950s. The modern AA spectrophotometer is used worldwide to analyse for up to 70 metals in concentrations as low as parts per billion (ppb). It can be used for analyses as wide ranging as the detection of polluting heavy metals in waterways, and the presence of sodium ions in mineral water. A schematic diagram of the AA spectrophotometer (figure 1.5.1) allows us to identify the three key stages of analysis.

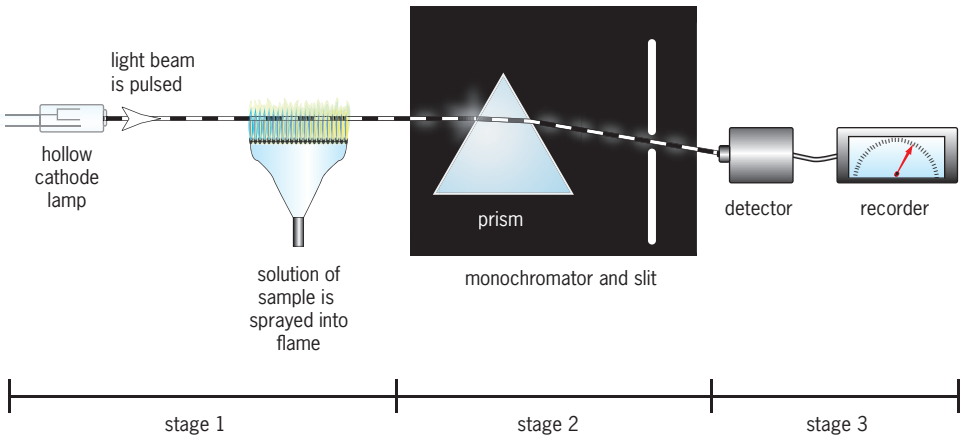


Figure 1.5.1 Schematic diagram of the atomic absorption spectrophotometer.

A.6.3
Describe the use of each of the following components of the AA spectrophotometer: fuel, atomizer, monochromatic light source, monochromatic detector and readout.

In stage one, a sample of the substance being analysed is sprayed (usually in acid solution) into a flame. The fuel that is usually used for this flame is acetylene mixed with air, or a nitrous oxide–acetylene mixture is used for a hotter flame. A slot-type burner is used to increase the path length, and therefore to increase the total absorbance. Sample solutions are usually aspirated (sucked up) with the gas flow into a nebulizing/mixing chamber (the atomizer) to form small droplets before entering the flame. The purpose of such a hot flame is to provide enough energy to convert the sample to atoms. This atomization is a key part of the instrument. The flame used is chosen so that only a small fraction of the atoms undergo excitation by heating. More recent AA spectrophotometers use a graphite furnace atomizer. The graphite furnace has several advantages over a flame. It is a much more efficient atomizer than a flame and it can directly accept very small absolute quantities of sample. It also provides a reducing environment for easily oxidized elements. The main intention is for the atoms to be excited by light from the selected light source.

The light source is chosen so that it produces the exact wavelength of light required by atoms in the sample for excitation (a **monochromatic light source**). This is achieved by using a hollow cathode lamp made of the same metal as the one being analysed. In this lamp, high voltages are used to excite metal atoms that are present as a low pressure vapour in the lamp. As these excited atoms return to their ground state, they emit their characteristic wavelengths of light. It is these wavelengths that will be absorbed by the metal atoms in the flame.

In stage two, the monochromator is used to select a wavelength to be passed to the detector. The stronger the wavelength of light that is selected, the more accurate the absorption data will be. The detector is set to detect the same wavelength of light as that selected by the monochromator.

In stage three, the intensity of the transmitted light is compared to the intensity of the incident light, and the absorbance is calculated. The greater the concentration of the sample, the greater will be the absorbance.

One problem should be evident in the design. The light coming from the atomized sample will include both the transmitted (unabsorbed) light and the light emitted as excited electrons return to their ground state. To overcome this problem, the incident light is pulsed. The detector is set to record only this pulsed light, and not the continuous light emitted by the excited atoms.

In all forms of spectroscopy, the relationship between the amount of a substance and the intensity or amount of light must be established in a process known as *calibration*. Known concentrations of the substance under investigation are analysed. A plot is constructed of the amount of substance versus the absorbance of light. This plot is known as a calibration curve. From the calibration curve, the concentration of an unknown sample is determined by **interpolation**.

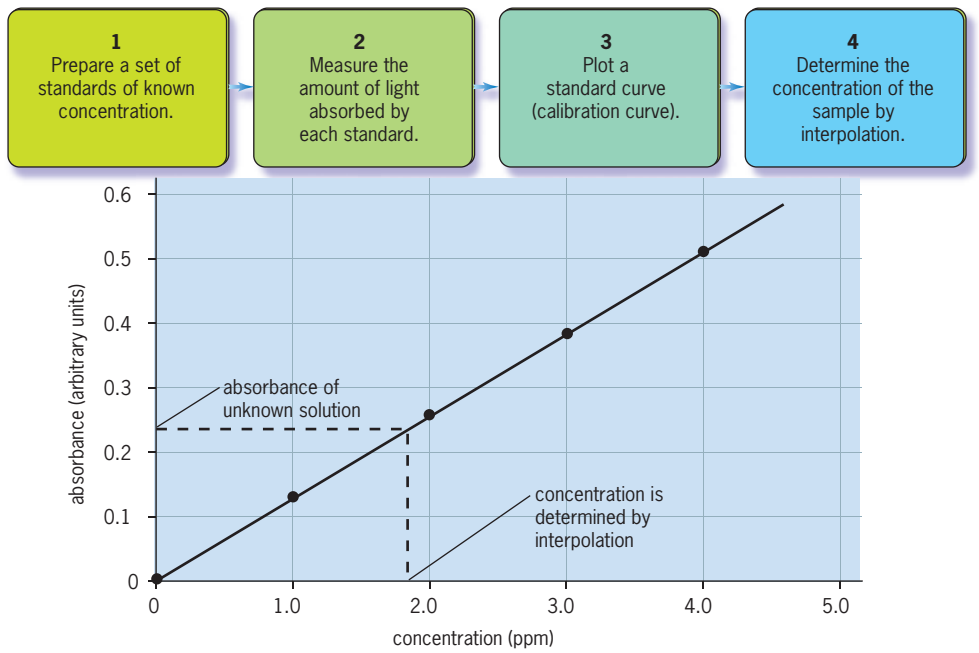


Figure 1.5.2 Spectrometry relies on the construction of a calibration curve to relate the amount of light absorbed to the amount of substance present.

Atomic absorption spectroscopy is a rapid, sensitive and selective technique with many and varied uses. The selectivity of the technique allows metals to be analysed without having to separate them from other components. Examples of the use of AA spectroscopy include:

- monitoring metals in the quality control of steel and other alloys
- analysing trace amounts of metals in foods
- determining metal concentrations in iron tablets, vitamins and other nutrient supplements
- testing air, water and soil for heavy metal contamination
- testing the grades of ores used in the mining industry
- testing for excess or deficiency of metals in body fluids such as blood and urine
- determining levels of trace elements such as Mn in soils.

A.6.1 State the uses of AA spectroscopy. © IBO 2007

Figure 1.5.3 An analytical chemist using an atomic absorption spectrometer.

Worked example 1

Many sports drinks contain high levels of sodium and potassium ions. To determine the potassium content in a particular sports drink, a 5.0 cm^3 sample was diluted to 50.0 cm^3 and the absorbance of the diluted solution and of several standard solutions was measured using AA spectroscopy. The results are shown in the table.

Concentration of solution (ppm)	Absorbance
0.00	0.010
1.00	0.080
2.00	0.150
3.00	0.220
4.00	0.290
Diluted sports drink	0.185

AS A.6.4 Determine the concentration of a solution from a calibration curve. © IBO 2007

a Plot the calibration curve for this experiment.

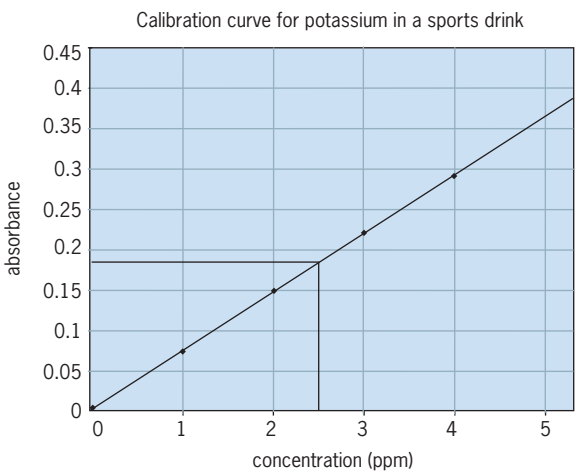
b Calculate the concentration of potassium ions in the sports drink in:

i ppm **ii** g dm^{-3}

c If there is 600 cm^3 of sports drink in an average bottle, calculate the mass of potassium ions that a player would consume if he consumed the whole bottle of drink after a football match.

Solution

a Calibration curve.



b **i** By interpolating the calibration curve, the concentration of potassium ions in the diluted sports drink is 2.5 ppm.

5.0 cm^3 of sports drink was diluted to 50.0 cm^3

$2.5 \times \frac{50.0}{5.0} = 25$

The concentration of potassium ions in the undiluted sports drink was 25 ppm.

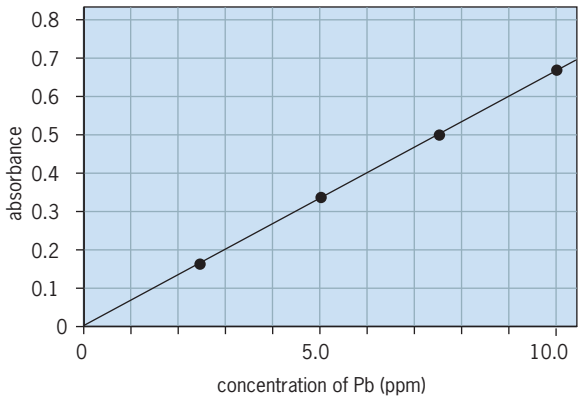
ii 1 ppm = 1 mg dm^{-3} so the concentration of potassium ions in the sports drink is 25 mg dm^{-3} = 0.025 g dm^{-3}

c $25 \times \frac{600}{1000} = 15$

If a football player consumed 600 cm^3 of the sports drink, he would consume 15 mg of potassium ions.

Worked example 2

The lead level in a sample of contaminated soil was determined by AA spectrometry. A 2.0 g sample was dissolved in acid and then diluted to a total volume of 100.0 cm^3. The absorbance of this solution, determined in a spectrometer set at a wavelength of 218 nm, was found to be 0.60. Several standard solutions of lead were tested under the same conditions and the calibration curve shown at right was generated. Determine the mass of lead in the 2.0 g sample of soil.



Solution

From the graph, an absorbance of 0.60 corresponds to a concentration of 9.0 ppm (mg dm^{-3}).

We have 9.0 mg of Pb in 1 dm^3.

$9.0 \times \frac{100}{1000}$ mg of Pb in 100 cm^3 (containing 2.0 g soil sample)

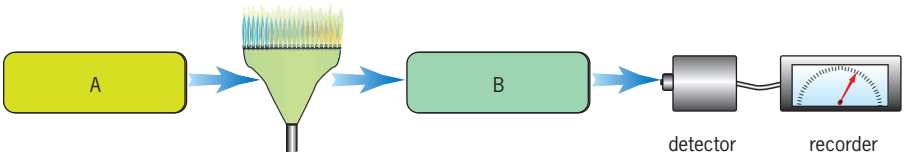
0.90 mg of Pb was present in 2.0 g of soil.

Section 1.5 Exercises

1 Describe the effect on metal atoms of the absorption of light energy.

2 Explain why a fuel such as acetylene is required for the flame in an AA spectrophotometer.

3 The diagram below shows the main components of an AA spectrophotometer in schematic form.



a In what state is the sample when analysed using AA spectroscopy?

b Explain the function of component B.

c Describe how component A is chosen.

4 Explain the purpose of pulsing the incident monochromatic light source.

5 Explain the purpose of a calibration graph in AA spectrometry.

6 The sodium content of a sports drink was analysed using AA spectrometry. A 10.0 cm^3 sample was diluted to 100.0 cm^3. The absorbance of the diluted solution and of several standard solutions was measured using AA spectroscopy. The results are shown in the following table.

Concentration of solution (ppm)	0.00	1.00	2.00	3.00	4.00	5.00	6.00	Diluted sports drink
Absorbance	0.000	0.070	0.140	0.210	0.280	0.350	0.420	0.330

a Plot the calibration curve for this experiment.

b Calculate the concentration of sodium ions in the sports drink in:

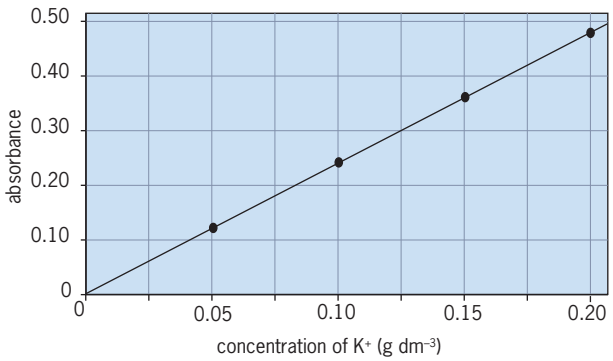
i mg 100 cm^{-3} **ii** mol dm^{-3}

c If there is 600 cm^3 of sports drink in an average bottle, calculate the mass of sodium ions that a player would consume if he consumed the whole bottle of drink after a football match.

7 The potassium content of an apple was measured using AA spectroscopy. A 6.0 g sample of apple was treated with nitric acid and the resulting solution made up to 100.0 cm^3. Using several potassium standard solutions, a calibration curve was constructed. The absorbance of the apple solution was then determined to be 0.3.

a Calculate the mass (in gram) of potassium in the sample of apple.

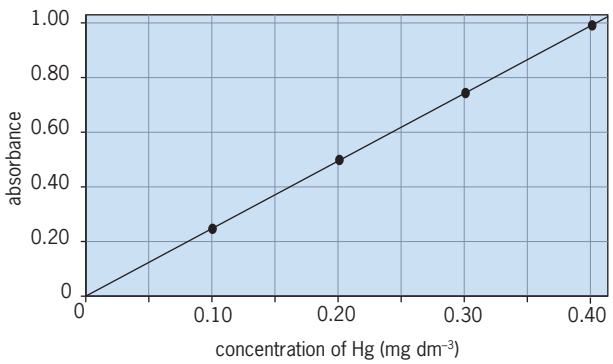
b Apples contain metal ions other than K^+. Explain why these other ions do not interfere with the analysis performed to determine the K^+ content of the apple.



8 The mercury content of a fish flesh sample was analysed by AA spectrometry. A 2.00 g sample of fish was treated with acid to dissolve the mercury. The solution was then made up to a volume of 100.0 cm^3. The absorbance of this solution at 254 nm was found to be 0.650. The absorbances of a series of standard mercury solutions were recorded at 254 nm and used to produce the calibration graph shown.

a Calculate the mass (in milligram) of mercury in the fish sample.

b Explain why the wavelength of 254 nm was chosen for this analysis.



1.6 CHROMATOGRAPHY

The samples provided to an analytical chemist are often complex mixtures. Consider, for example, biochemical samples such as blood and urine. How many components do these contain? Separating the components of such mixtures is often a crucial part of an analysis. Chromatography is particularly useful for separating complex mixtures. It is not only used as a preparative technique for separating mixtures, but it is also a means of identifying and quantifying the components of a mixture. It is used extensively in the areas of quality control, in testing for the presence of banned substances in urine of athletes, and in forensic testing.

Figure 1.6.1 The changes in leaf colours during autumn suggested to Mikhail Tsvet that leaves must contain at least two different pigments.

During autumn many green leaves turn various shades of red. In 1906, Russian botanist Mikhail Tsvet (1872–1919) reasoned that leaves must therefore contain at least two different pigments: red and green. Tsvet conducted an experiment to test this idea. Leaves were crushed and their pigments extracted with petroleum ether. This extract was added to the top of a column containing fine powdered chalk and washed through with the solvent. Several bands of various shades of green and yellow developed along the column, confirming Tsvet's idea that more than one pigment existed. He named the process chromatography (from the Greek *khroma*, meaning 'colour', and *graphe*, meaning 'writing').

The term *chromatography* is now used to describe a range of techniques used to separate, identify and quantify the components of a mixture. Most chromatography is used for analytical purposes; that is, to answer the 'What is it?' (*qualitative*) and 'How much is there?' (*quantitative*) questions for samples of unknown composition. Some chromatography is preparative; that is, it is used to obtain pure samples of the components of a mixture.

AS **A.7.1**
State the reasons for using chromatography. © IBO 2007

Many mixtures do not appear to be mixtures at all. Consider black ink. It flows from the pen onto the page and shows no sign of separating, yet in chromatography several different colours are found to be present in that black ink (see figure 1.6.4). In this case, the chromatography experiment is a qualitative technique. For quantitative experiments, column chromatography is much more appropriate than paper and thin-layer chromatography. It allows each component of the mixture to be collected and the individual amounts measured.

AS **A.7.2**
Explain that all chromatographic techniques involve adsorption on a stationary phase and partition between a stationary phase and a mobile phase. © IBO 2007

All chromatographic techniques involve two phases: a **stationary phase** and a **mobile phase**. In Tsvet's experiment these phases were the solid (powdered chalk) packed into the column and the solvent (petroleum ether) respectively. The mobile phase moves over or through the stationary phase. The mixture to be separated is dissolved in the same solvent as the mobile phase and applied to the stationary phase. Separation of the components of a mixture occurs because the components adsorb (form bonds) to a solid stationary phase with different strengths. The stronger the **adsorption**, the more slowly the component moves as the mobile phase sweeps over or through the stationary phase. Components undergo continuous adsorption and **desorption** (breaking of the bonds) to differing degrees. The rate of movement of each component therefore depends on the strength of its adsorption to the stationary phase and its tendency to desorb into the mobile phase. Examples of separation by adsorption include column chromatography (see pages 42–43), thin-layer chromatography (see pages 41–42) and high performance liquid chromatography (see pages 66–68).

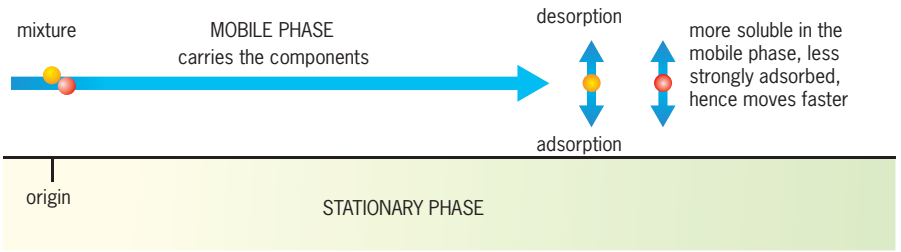


Figure 1.6.2 All types of chromatography involve two phases: one stationary and one mobile.

Other types of chromatography involve a liquid as the stationary phase. This liquid may be part of the structure of the stationary phase, as in the case of paper chromatography, or it may be a very viscous liquid coated onto solid beads (as in gas–liquid chromatography, see pages 64–65). In this case, separation occurs due to **partition**. This is a term used to describe the relative solubility of a substance in two different liquids. In chromatography, partition determines the degree of separation of the components of the mixture in the liquid stationary phase and in the liquid mobile phase. If the mobile phase is gaseous, then the volatility of the components will contribute also. The more soluble (or more volatile) the component is in the mobile phase than in the liquid stationary phase, the faster it will travel along the chromatography paper or column.

The choice of stationary and mobile phases is vital if good separation of components is to occur. This choice is often the result of trial and error. Chromatographic techniques differ according to the nature of the two phases, but the essential processes of adsorption and desorption leading to separation remain the same. The length of stationary phase over which the chromatogram is run is also important. The longer the stationary phase, the better the separation of components is likely to be. Chromatography is sometimes referred to as separation science. Often, as its name suggests, chromatography is used to separate coloured substances, but it can also be used to separate, identify and quantify a vast range of colourless substances.

TABLE 1.6.1 SUMMARY OF ADSORPTION AND PARTITION IN CHROMATOGRAPHY

	Stationary phase	Mobile phase	Types of chromatography that employ this technique
Adsorption	Solid	Liquid	Column chromatography Thin-layer chromatography High performance liquid chromatography
Partition	Liquid	Liquid/gas	Paper chromatography Gas–liquid chromatography Thin-layer chromatography (in some cases)

Paper and thin-layer chromatography

The simplest form of chromatography involves the use of strips of absorbent paper, e.g. filter paper, with water bound up in its structure, as the stationary phase. In a technique known as *paper chromatography*, a solvent (the mobile phase) moves the components of a mixture up the strips of paper. At the beginning of the experiment a spot of the mixture is placed about 1 cm from the bottom of the paper and the end of the paper is dipped into the solvent, ensuring that the spot of the mixture is not submerged directly in the solvent. The solvent moves up the paper by capillary action, moving past the mixture spot and carrying the components with it. Partition is achieved because some components adsorb strongly to the water that is bound to the fibres of the paper and make slow progress along the chromatogram, while others desorb readily into the solvent and progress quickly. It is important to remove the paper from the solvent and to dry it quickly before the *solvent front* (the visible line of the solvent moving up the paper) reaches the top of the paper. If the solvent goes all the way to the top of the paper, calculations cannot be made of the ratio of distance travelled by the component to distance travelled by the solvent (see R_f values p. 42).

Figure 1.6.3 shows a typical arrangement for paper chromatography. The stationary phase is, in fact, the water held in the absorbent pores of the paper. Polar components in the mixture adsorb strongly to this polar stationary phase. The solvent may be somewhat non-polar, perhaps even a mixture of solvents, so it carries the non-polar components readily. Once the solvent reaches a point close to the top of the paper, the strip is dried to produce the chromatogram.

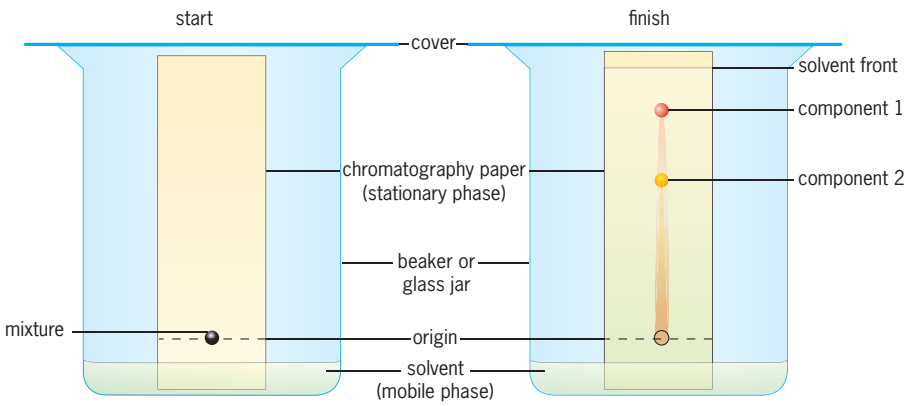


Figure 1.6.3 During paper chromatography the solvent moves up the paper (stationary phase) to produce a chromatogram.

Paper chromatography is an inexpensive, convenient and simple procedure used to separate the coloured components of dyes, inks and food colourings. It was once routinely used to separate amino acid mixtures. Since amino acids are colourless, the chromatogram was treated with a chemical, usually ninhydrin, to make the component amino acids visible. Now the technique of electrophoresis is more commonly used to separate amino acid mixtures. In other cases, chromatograms in which non-coloured components such as sugars have been separated are developed using iodine vapours or other chemicals that are visible under ultraviolet light.

To identify the separated components, a comparison is made between the new chromatogram and chromatograms of known substances produced under the same conditions. Identification involves the calculation of a **retention** or

AS A.7.3
Outline the use of paper chromatography, thin-layer chromatography (TLC) and column chromatography.

QuickTime CW Video
Paper chromatography of inks

Pr
PRAC 1.2
Paper chromatography of food dyes

Figure 1.6.4 Coloured components of inks can easily be separated by paper chromatography. The ink on the far left was black but can be seen to contain yellow, purple and blue components.

retardation factor (R_f) for each component and known substance. The R_f is determined by comparing the distances moved by each substance and by the solvent. Note that R_f values must be less than 1, since the solvent moves furthest from the origin.

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

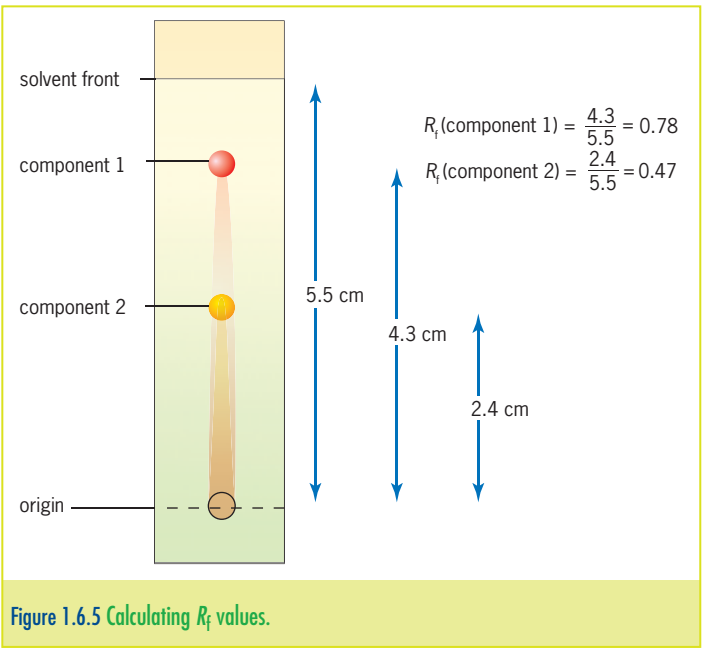


Figure 1.6.5 Calculating R_f values.

PRAC 1.3
Thin-layer chromatography of leaf pigments

Figure 1.6.6 Thin-layer chromatography provides good separation of the pigments in green plant material such as leaves.

Thin-layer chromatography (TLC) is similar to paper chromatography, but the paper is replaced by a plate made of glass or plastic and coated with a thin layer of fine powder such as aluminium oxide. Development of a chromatogram is achieved in the same way as for paper chromatography. Component separation depends on the degree of adsorption to the surface layer of oxide (the stationary phase), and desorption into the solvent (the mobile phase) moving over the plate's surface. The R_f values are calculated and compared as for paper chromatography.

Thin-layer chromatography is faster and more sensitive than paper chromatography, and is capable of greater resolution (clearer separation) of mixture components. It allows for the use of a wider range of solvents, some of which would be corrosive to paper.

Paper and thin-layer chromatography are essentially qualitative techniques, used for separation and identification. Thin-layer chromatography can be used quantitatively if the components are scraped off the thin layer, extracted and then analysed for the amount of substance. However, this technique has very limited use in quantitative analysis.

Column chromatography

Column chromatography uses the same principles as paper and thin-layer chromatography. In simple column chromatography, the stationary phase is a solid—usually aluminium oxide or silicon dioxide—packed into a glass or plastic column. The sample is applied at the top of the column, and solvent (the mobile phase), under the influence of gravity, slowly drips down the column and out the bottom. As the solvent carries the mixture through the column, the various components separate.

WORKSHEET 1.3
Amino acid analysis

Separation occurs on the basis of the different degrees of adsorption to the stationary phase and desorption into the mobile phase. The less strongly adsorbed components move more quickly through the column. Components are collected as they **elute** (pass out) from the column. These collected components can then be identified and quantified.

Figure 1.6.7 Components of a mixture can be separated and collected by column chromatography.

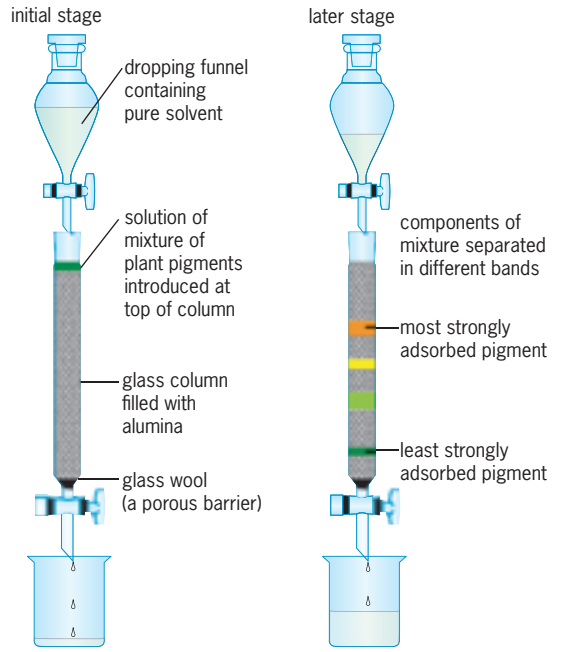


Figure 1.6.8 Separation of plant pigments by column chromatography.

Identification involves comparison with known substances passed through the column under the same conditions. The R_f values cannot be used, since all components travel the same distance (the length of the column). Instead, the **retention time (R_t)** for each substance is calculated. The retention time is the time taken for the substance to pass through the column. The identity of each component is determined by analysis of the eluted pure sample using a variety of techniques. For example, a coloured component may be analysed using UV–visible spectroscopy.

Section 1.6 Exercises

1 Explain what is meant by the terms *stationary phase* and *mobile phase* in chromatography.

2 When separation of the components of a mixture occurs, describe how the rate of adsorption and desorption of the component closest to the top of the chromatogram compares to that closest to the bottom of the chromatogram.

3 Explain how an R_f value is calculated.

4 For each of the following types of chromatography, identify the:

i stationary phase **ii** mobile phase

a Paper chromatography

b Thin-layer chromatography

c Column chromatography

5 Describe how the mobile phase moves through the stationary phase in:

a paper chromatography

b thin-layer chromatography

c column chromatography.

6 For paper chromatography, explain the following:

a The level of the solvent must be lower than the origin where sample spots are placed.

b The origin line must be marked with a graphite pencil rather than a pen.

c R_f values are always less than one.

7 A student wishes to use paper chromatography to separate the coloured dyes used in a packet of candy-covered chocolate sweets. A blue sweet and a red sweet were moistened with water and the food dyes from the sweets were spotted onto a strip of chromatography paper. The paper was then dipped into a solvent mixture of butanol, ethanol and ammonia and allowed to develop over 2 hours. The paper was dried and the distance moved by the solvent and each coloured component of the dyes was measured. The results were as follows:

The solvent front had moved 18.3 cm.
Blue sweet—yellow component: 3.2 cm; blue component: 6.7 cm
Red sweet—orange component: 4.5 cm; red component: 5.8 cm

a State the number of different dyes that were used to produce the blue sweets and the red sweets.

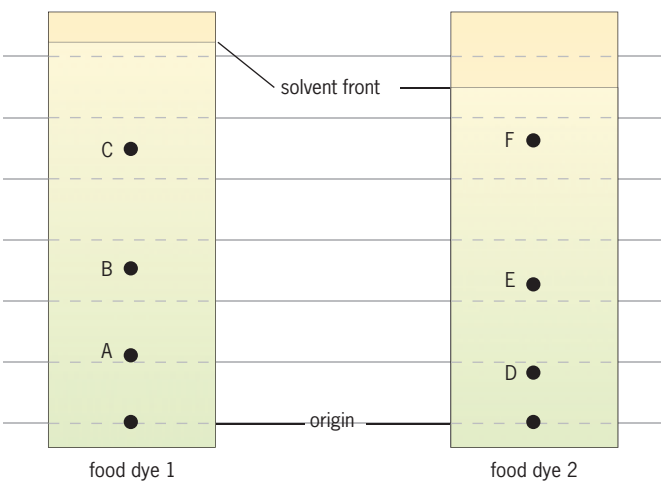
b Calculate the R_f value for each component.

c Draw the chromatogram.

8 **a** Compare paper and thin-layer chromatography and describe two similarities between the techniques.

b Describe two advantages of thin-layer chromatography over paper chromatography.

9 Thin-layer chromatography was used to investigate the colourings in two food dyes. Two chromatograms were obtained under identical conditions. These are shown below.



a Explain why the colourings in the food dyes separated during the chromatography procedure.

b Which components (A to F) are found in both food dyes?

10 **a** Explain why column chromatography rather than thin-layer chromatography would be used if samples of the components are required for further analysis.

b Explain why paper or thin-layer chromatography, rather than column chromatography, might be used for the initial investigation of a mixture.

1.7 VISIBLE AND ULTRAVIOLET SPECTROSCOPY

HL

A UV–visible spectrophotometer is used to analyse samples that absorb in the ultraviolet (200–400 nm) and visible (400–800 nm) regions of the electromagnetic spectrum. This analysis is applicable to coloured substances, which absorb visible light. In addition, many colourless substances, especially organic compounds, absorb radiation of higher energy than visible light (ultraviolet light), as their electrons are excited between electronic energy levels and so may be analysed by UV–visible spectroscopy.

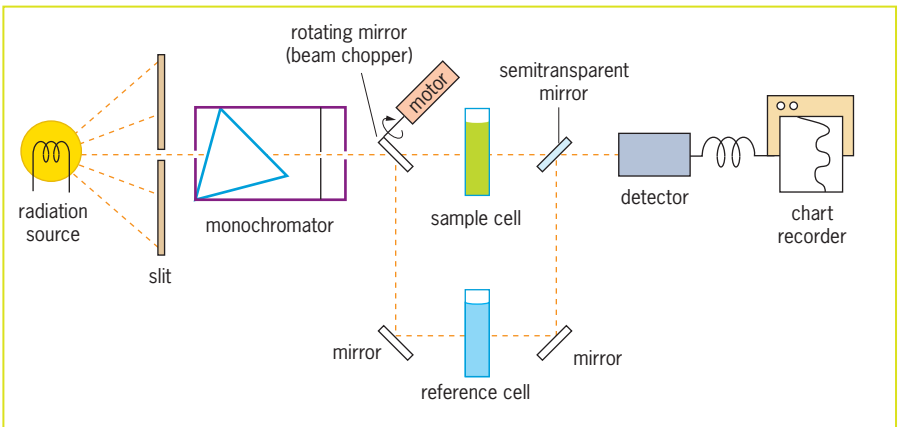


Figure 1.7.1 Schematic diagram of the UV–visible spectrophotometer.

The schematic diagram in figure 1.7.1 allows us to identify the key stages in UV–visible spectroscopy. As in infrared spectroscopy, a double beam spectrophotometer is used for greater accuracy. The sample in this case is in solution. Light to be passed through the sample is generated by a light source that provides radiation in both the visible and ultraviolet ranges. The selection of wavelength for analysis occurs via a monochromator, which selects and focuses the chosen wavelength. The light beam is split into two beams and pulsed. One beam passes through the sample while the other beam passes through a reference cell.

The detector converts the light signal to an electrical signal, comparing the sample beam and the reference beam to remove any absorbance due to the solvent. The recorder uses the electrical signal to produce a recording of the absorbance of the sample.

The sample and reference cells are usually square-shaped and made from a special type of clear glass that is able to transmit UV radiation. Care must be taken in handling these cells that no fingerprints distort the optical clarity of the cell.

UV–visible spectroscopy plays a significant role in analysis, particularly in biochemical and environmental applications. Examples include the analysis of hemoglobin in blood and the analysis of nitrates and phosphates in water samples. UV–visible spectrometry has two main uses: qualitatively it is used to identify unknown compounds in solution by comparing known absorption spectra with the spectrum of the unknown, and quantitatively it may be used to determine the concentration of a known compound in solution. This technique is quite sensitive; it is capable of measuring concentrations as low as 10^{-5} mol dm^{-3} with around 98% accuracy.

When used qualitatively, recordings are made of the absorbance of the sample over a range of wavelengths, producing an absorbance spectrum. Comparison of this spectrum with the spectra of known substances allows identification of the unknown.

For quantitative use, the instrument is set to one wavelength. This wavelength is chosen by examining the absorption spectrum. Usually, the wavelength corresponding to the maximum absorbance will be used, provided that other substances present in the sample do not absorb this wavelength. Any interference by other substances absorbing can be a source of error in UV–visible spectrophotometry. Using the selected wavelength, the absorbance of a set of solutions of known concentration is recorded and a calibration curve constructed. The unknown concentration is found by interpolating this curve once the absorbance of the unknown is determined.

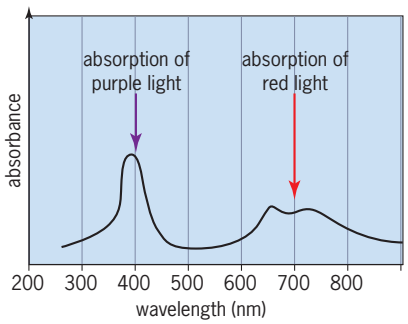


Figure 1.7.2 An absorbance spectrum shows the absorbance of light by the sample over a range of wavelengths. This is the spectrum of a green Ni^{2+} solution.

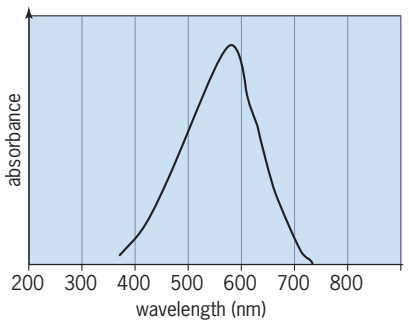


Figure 1.7.3 Which wavelength would be chosen for use in the quantitative analysis of a compound with this absorption spectrum?

A.8.6 Determine the concentration of a solution from a calibration curve using the Beer–Lambert law.

As in AA spectroscopy, the greater the concentration of the solution being analysed, the greater the absorbance of the solution. Similarly the path length, the distance through the solution that the light must travel, also influences the absorbance, as does the molar absorptivity of the substance. A substance with a high molar absorptivity is very effective at absorbing light. An example of a substance with a very high molar absorptivity is β-carotene, the strongly coloured organic substance that is responsible for the orange colour of carrots.

The Beer–Lambert law relates absorbance, A, to the molar absorptivity, ε (in dm^3 mol^{-1} cm^{-1}), the path length, l (in cm), and the concentration of the solution, c (in mol dm^{-3}):

$$A = \varepsilon l c$$

Absorbance is directly proportional to ε, l and c.

A calibration curve can be plotted for absorbance against path length, or molar absorptivity, but it would most commonly be plotted for absorbance against concentration in order to determine the concentration of an unknown solution.

Worked example 1

The concentration of iron in a sample of lake water was determined by UV–visible spectroscopy. Iron, present as Fe^{2+} ions, was reacted to form an orange-yellow complex. The absorbance of a series of standard solutions and a sample of the lake water were measured at 480 nm.

Determine the concentration of iron in the lake water in ppm.

Concentration of Fe^{2+} (ppm)	Absorbance
0	0
5.0	0.17
10.0	0.33
15.0	0.49
20.0	0.65
Sample	0.35

Solution

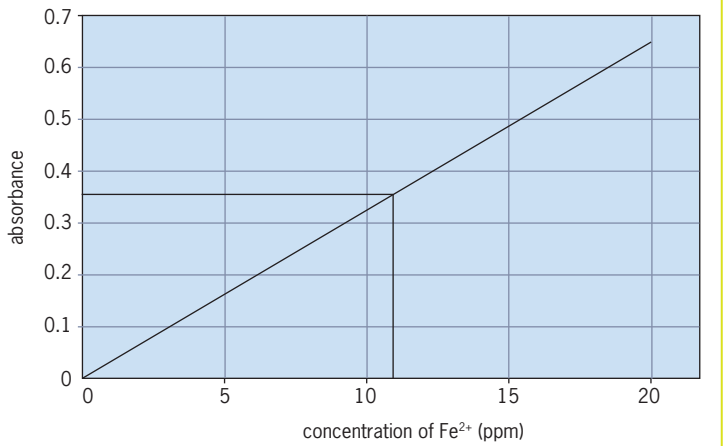


Figure 1.7.4 The calibration curve for the absorbance of iron.

From the calibration curve, the concentration of iron in the lake water can be seen to be 10.9 ppm.

CHEM COMPLEMENT

Colorimetry

Colorimetry is a much simpler, less expensive, but less accurate form of absorption spectrometry. Its design is very similar to UV–visible spectrometry, but the wavelength used is selected by a filter rather than a monochromator. In addition, the light source used provides only visible light, not ultraviolet radiation. Comparison of the schematic diagrams in figures 1.7.1 and 1.7.5 highlights the similarities and the differences.

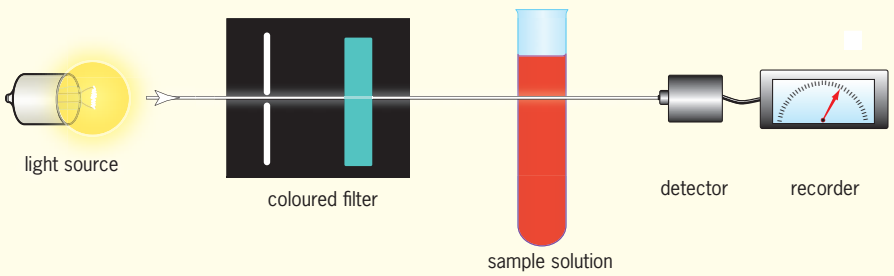


Figure 1.7.5 Schematic diagram of a colorimeter.

Selection of the wavelength to be used for analysis in colorimetry involves the use of coloured filters. Recall the concept of complementary colours from *Chemistry: For use with the IB Diploma Programme Higher Level*, chapter 3. When we view an object, our eyes receive light reflected by the object. This reflected light is composed of the wavelengths not absorbed by the object. We see the complementary colour of the colour absorbed. For example, copper(II) solutions absorb wavelengths in the orange region, and reflect those in the blue-green region. The more concentrated a solution, the more of a certain colour that has been absorbed, making the colour appear darker. In colorimetry, the filter used is the complementary colour to that of the solution being analysed.

While it is unlikely that a school laboratory would have a UV–visible spectrophotometer, a colorimeter is a much more affordable and compact instrument and can be used for a range of practical investigations.

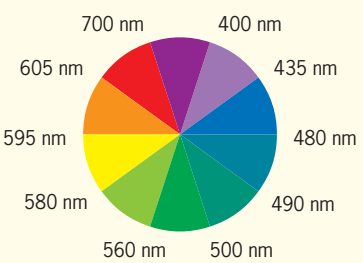


Figure 1.7.6 Complementary colours are opposite one another on a colour wheel.

Figure 1.7.7 Solutions of known concentrations are used to calibrate a colorimeter.

Pr

PRAC 1.4
Determination of phosphate in washing powder

Colourful transition metal compounds

In UV–visible spectrometry, the substance being analysed is often coloured. Many non-coloured substances, however, can be reacted with reagents to produce a coloured complex that can be analysed spectroscopically. In this way, a wider range of elements and compounds can be analysed using UV–visible spectrometry.

The chemistry of transition metal complexes is studied in detail in *Chemistry: For use with the IB Diploma Programme Higher Level,* chapter 3; however, it is useful to consider here the factors that affect the colour of transition metal complexes.

AS A.8.1
Describe the effect of different ligands on the splitting of the d orbitals in transition metal complexes. © IBO 2007

Transition metal ions have one or more unpaired electrons in their d subshells. The ions are also highly charged and quite small. A ligand is a small molecule with an electron pair that can be donated to a transition metal ion in a coordinate (dative) bond. Typically two to six ligands may surround a transition metal ion in a complex. Common ligands include water, H_2O; ammonia, NH_3; the chloride ion, Cl^-, and the cyanide ion, CN^-.

AS A.8.2
Describe the factors that affect the colour of transition metal complexes. © IBO 2007

The d orbitals of a transition metal ion that has formed coordinate bonds with ligands in a complex will be affected by the ligands surrounding the ion. Repulsion occurs between the non-bonding electrons of the ligand and the electrons in the d orbitals of the metal ion. Instead of being all the same energy, as they would be when the ion is isolated, the orbitals are split into two groups. The closer a ligand can get to the metal ion the further apart the d orbitals will split. The amount of splitting depends on four factors:

1. *Nature of the transition metal ion*. Different transition metals have different ionic radii, so even if the charges on the ions are the same, the attraction of the ion for the electrons of the ligand will vary from one transition metal to another.
2. *Charge on the transition metal ion*. An ion with a greater charge will attract the non-bonding pair of electrons on the ligand more strongly than one with a smaller charge, so the ligand will be pulled closer to the ion with the greater charge. This increases the repulsion between the non-bonding electrons of the ligand and the d orbital electrons, increasing the splitting.
3. *Size of the ligand*. Smaller ligands can get closer to the metal ion, increasing the repulsion between the non-bonding electrons of the ligand and d orbital electrons and increasing the splitting. The order of increasing amount of splitting for some common ligands is:

   $Cl^- < H_2O < NH_3 < CN^-$
4. *The stereochemistry of the complex*. The splitting for tetrahedral and for square planar complexes will be different from that for an octahedral complex.

For example the chromium(III) ion, Cr^{3+}, has five 3d orbitals with the same energy. However, when the ion is surrounded by six water ligands, the five d orbitals are split into three orbitals of lower energy and two orbitals of higher energy.

When white light falls on the complex, energy is absorbed and electrons in the lower energy d orbitals are excited to a higher energy d orbital. The wavelength of energy absorbed, depends on the difference in the energy of the split orbitals. The remaining wavelengths of the light combine to give the colour that is observed (the complementary colour).

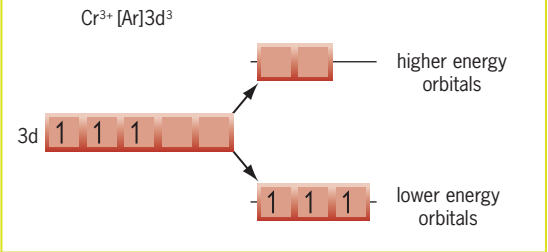


Figure 1.7.8 When the Cr^{3+} ion is surrounded by six water ligands, the five d orbitals are split into three orbitals of lower energy and two of higher energy.

For example, if a complex ion absorbs light of a wavelength of 450 nm in the blue range of visible light spectrum, then the ion appears yellow, its complementary colour in the opposite segment of the colour wheel (see figure 1.7.6) . Similarly, for the hexaaquacopper(II) ion to appear blue it must absorb light with a wavelength between 580 and 595 nm.

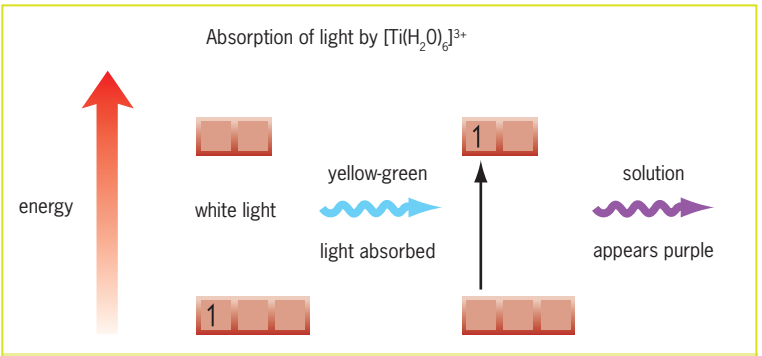


Figure 1.7.9 When white light is shone on $[Ti(H_2O)_6]^{3+}$, it absorbs strongly in the yellow-green region, causing the solution to appear purple.

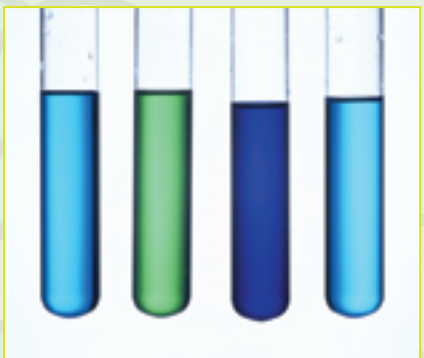


Figure 1.7.10 The colour of the copper ion solution depends on the ligands surrounding the copper ion. These four test tubes contain $[Cu(H_2O)_6]^{2+}$, $[CuCl_4]^{2-}$, $[Cu(NH_3)_4(H_2O)_2]^{2+}$ and $[Cu(EDTA)]^{2-}$.

Because the energies of the d orbitals are determined by the presence of ligands, it follows that different ligands will produce compounds of different colours. One of the most well-known examples of this is the copper tetrammine complex $[Cu(NH_3)_4(H_2O)_2]^{2+}$, which is royal blue in colour whereas the hexaaquo complex $[Cu(H_2O)_6]^{2+}$ is a sky blue in colour. The reaction in which the water ligands are displaced by ammonia ligands is:

$$[Cu(H_2O)_6]^{2+}(aq) + 4NH_3(aq) \rightarrow [Cu(NH_3)_4(H_2O)_2]^{2+}(aq) + 4H_2O(l)$$

hexaaquacopper(II) ion, sky blue → tetraamminediaquacopper(II) ion, royal blue

The presence of ammonia ligands changes the energy of the d orbitals sufficiently to produce the more 'purplish' royal blue colour.

Another common example of the replacing of ligands to form a different complex involves the addition of 1.0 mol dm^{-3} ammonia to the $[Ni(H_2O)_6]^{2+}$ complex. The ammonia ligands replace all six water ligands:

$$[Ni(H_2O)_6]^{2+}(aq) + 6NH_3(aq) \rightarrow [Ni(NH_3)_6]^{2+}(aq) + 6H_2O(l)$$

hexaaquanickel(II) ion, green → hexaamminenickel(II) ion, blue

Absorption of UV and visible light by organic molecules

AS **A.8.3** State that organic molecules containing a double bond absorb UV radiation. © IBO 2007

AS **A.8.4** Describe the effect of the conjugation of double bonds in organic molecules on the wavelength of the absorbed light. © IBO 2007

While many organic compounds are colourless, there are substantial numbers of coloured organic compounds. Indeed organic dyes such as indigo, saffron and the exotic Tyrian purple have been used for centuries. The bright orange of β-carotene (found in carrots and other orange vegetables) is familiar to most people, as is the green colour of chlorophyll, the light-capturing molecule found in leaves. When observing the molecular structures of some of these compounds in figures 1.7.11 and 1.7.12, the most obvious feature that they have in common is the large number of carbon–carbon double bonds. These are often **conjugated** (alternating with single bonds) and the pi electrons are delocalized within the structure. It is this part of the structure that allows the molecule to absorb ultraviolet and visible light.

Chlorophyll is a substance with a vital role in the food chain. In this role the absorption of visible light is fundamental (see figure 1.7.11). There are two forms of chlorophyll which differ by a single functional group. In chlorophyll *a* the R group is a methyl group, $–CH_3$, while in chlorophyll *b*, the R group is an aldehyde functional group, –CHO. The two forms of chlorophyll have maximum absorption at slightly different wavelengths, adding greater light-capturing power to chloroplasts in which these pigments are found. Plants can obtain all their energy requirements from the blue and red parts of the spectrum. In the large region between 500 nm and 600 nm very little light is absorbed. This is the green region of the spectrum, and the reflection of this green light makes the chlorophyll (and consequently leaves) appear green.

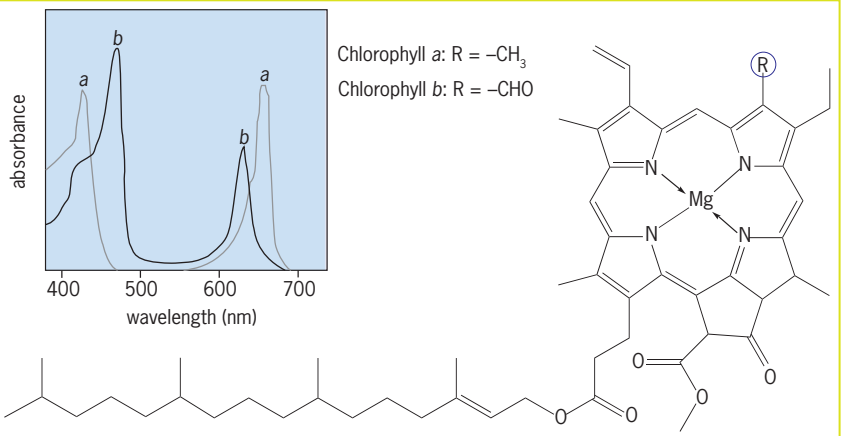


Figure 1.7.11 Visible spectrum and structure of chlorophyll. This molecule absorbs strongly at 420 nm (violet) and 660 nm (red) giving it a green colour.

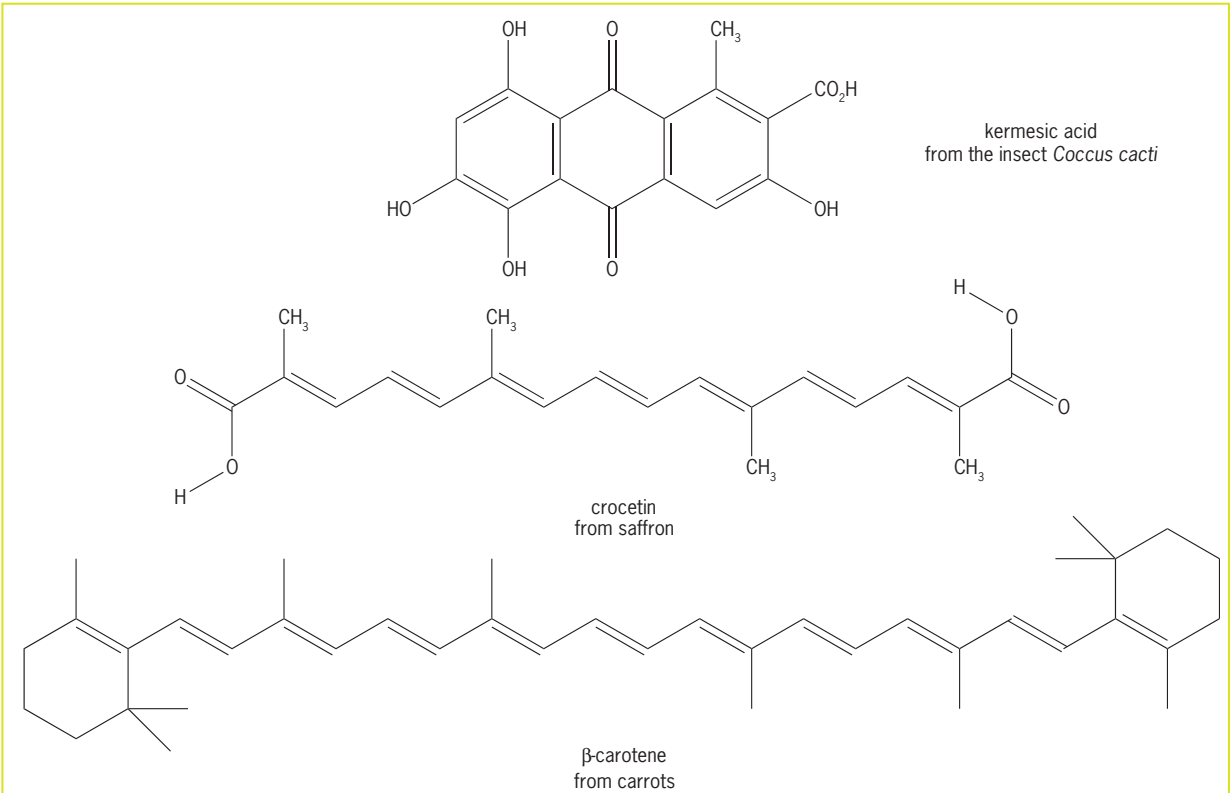


Figure 1.7.12 The conjugation of double bonds in these molecules allows them to absorb light in the UV and visible regions of the electromagnetic spectrum. This feature makes these compounds strongly coloured: kermesic acid is crimson, crocetin is yellow and β-carotene is orange.

CHEM COMPLEMENT

The royal colour

The colour purple has been a symbol of royalty for centuries. As early as 1600 BC in the city of Tyre, the ancient Phoenicians used a purple-red dye known as Tyrian purple to colour cloth. This dye was extracted from the hypobranchial gland of the *Murex brandaris* mollusc. About 12 000 shellfish needed to be crushed to extract 1.5 g of the pure dye. As a result the purple dye was so expensive that only royalty could afford clothes dyed with it.

Although Tyrian purple is often translated as 'scarlet', it appears that the method of extraction could have caused variation in the colour. Another dye extracted from a related sea snail *Hexaplex trunculus* could produce a purple-blue colour when processed in the shade, or a sky-blue indigo colour when processed in sunlight. Sea snails in other parts of the world such as the tropical eastern Pacific have been found to produce a similar substance that makes a purple dye in the sunlight. The snails are predatory and use the secretion as part of this behaviour.

The structure of the dye indigo is very similar to Tyrian purple, the only difference being the replacement of two hydrogen atoms with bromine atoms.

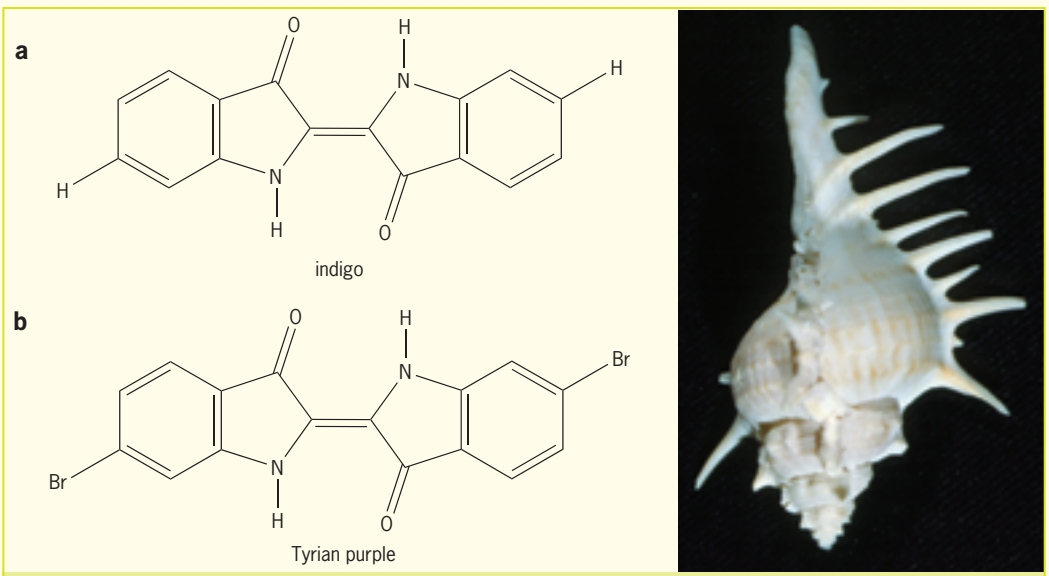


Figure 1.7.13 The shell of a *Murex* sea snail with the molecular structures of (a) indigo and (b) Tyrian purple.

Absorbance of high energy UV light causes molecular electrons to be excited to higher energy molecular orbitals. Although molecular orbital theory is beyond this course, it is important to realize that absorption of this energy promotes electrons from the highest *occupied* molecular energy level (**HOMO**) to the lowest *unoccupied* molecular energy level (**LUMO**). When a double bond or a triple bond occurs in a molecule, the only radiation with enough energy to achieve this effect is the highest energy UV radiation. A group such as a carbon–carbon double bond that is able to absorb UV or visible radiation is known as a **chromophore**. The wavelength of the light absorbed by just one carbon–carbon double or triple bond in a molecule is below 200 nm. For example, ethene absorbs UV radiation at 171 nm and 1-hexyne (with one triple bond) absorbs at 180 nm. This absorption is below the range of UV–visible spectroscopy; however, when the double or triple bonds are conjugated (alternated with single bonds) there is a shift in the wavelength of the

absorbed light (a **bathochromic shift**) to a longer wavelength and UV light of wavelength greater than 200 nm is absorbed. This is because the difference in energy between the HOMO and LUMO molecular energy levels decreases as the degree of conjugation increases.

Phenolphthalein is part of another group of organic compounds that are valued for their colour: acid–base indicators. These are weak acids that are two different colours in their acid and conjugate base form. The structure of phenolphthalein is shown in figure 1.7.15 in its acid and base forms. Notice that the acid form of phenolphthalein has less conjugation than the base form. This change in structure when the H^+ ion is donated is enough to increase the wavelength of light that is absorbed by phenolphthalein into the visible region and to make the compound coloured. Retinol (vitamin A) is another well-known coloured organic compound.

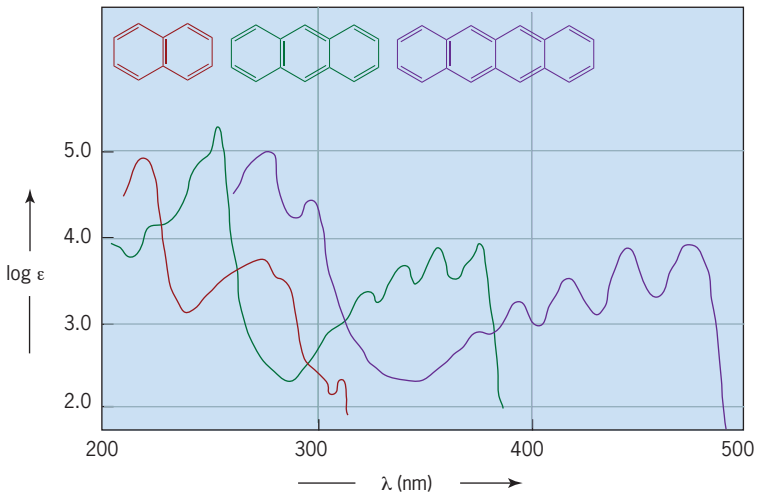


Figure 1.7.14 The UV–visible absorption spectra of naphthalene (2 rings), anthracene (3 rings) and tetracene (4 rings). As the degree of conjugation increases the wavelength of absorbed light increases. While naphthalene and anthracene are colourless, tetracene is orange.

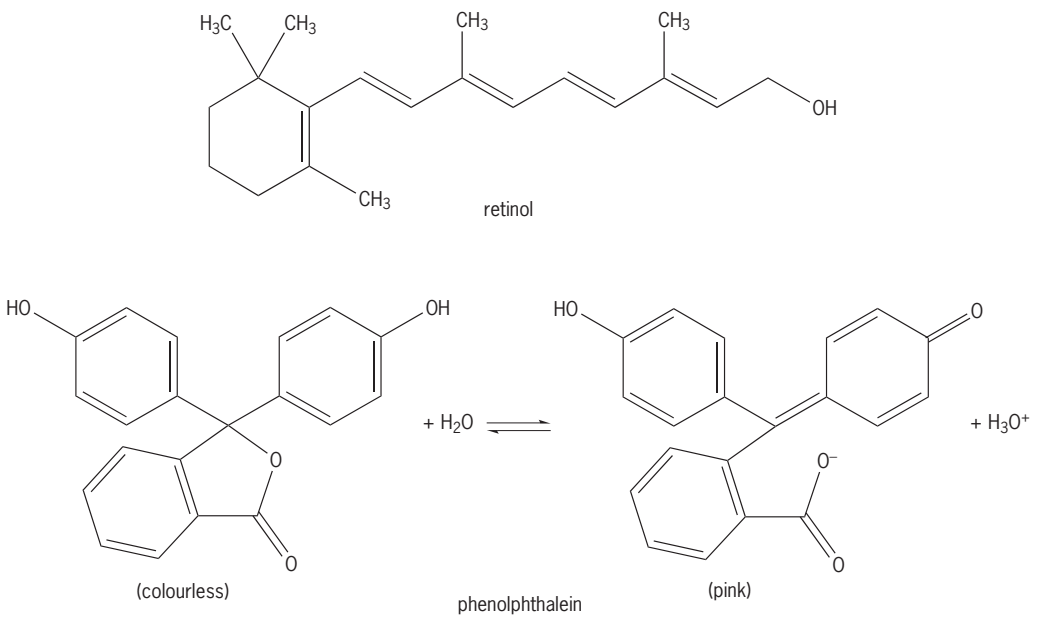
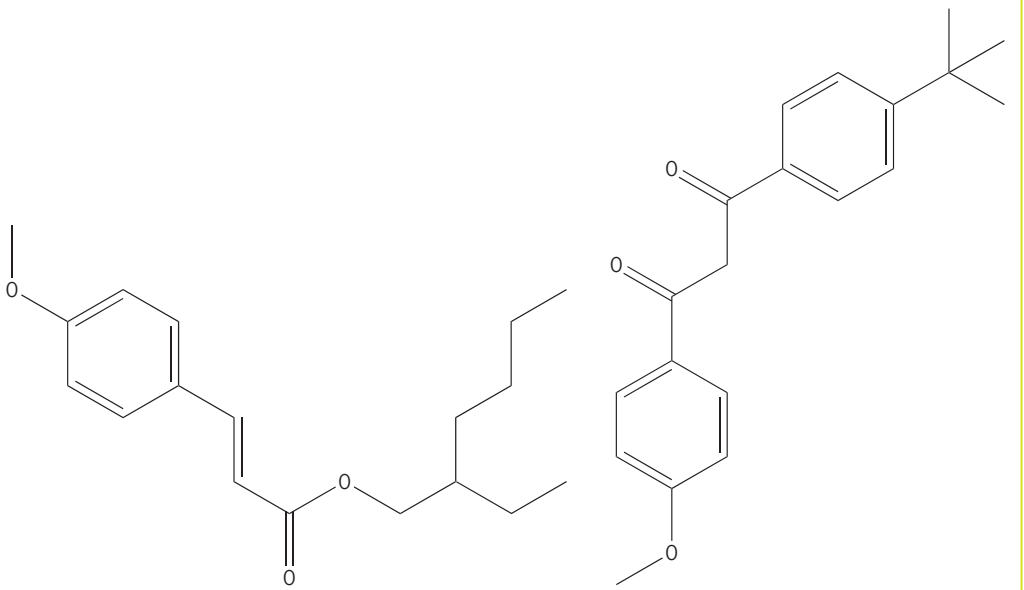


Figure 1.7.15 Retinol is yellow while phenolphthalein is colourless in its acid form, but pink in its basic form.

A group of pharmaceutical preparations that have become vital to our existence, especially in the southern hemisphere where the hole in the ozone layer increases the risk of sunburn, is sunscreens. The function of sunscreens is to prevent UV radiation from penetrating our skin. UV radiation is harmful to skin cells and can cause mutations that lead to skin cancer. Sunscreens can act in two ways: they may absorb most of the UV radiation or they may reflect most of UV radiation. Those that reflect the UV rays contain TiO_2 or ZnO,

which are both white compounds and may physically block the passage of UV radiation into the skin. The compounds that are used to absorb the UV radiation contain conjugated carbon–carbon double bonds. Octyl methoxycinnamate is a compound that absorbs UV radiation of wavelengths 290–320 nm (UVB radiation), while butyl methoxydibenzoylmethane ('parsol') absorbs UV radiation of wavelengths 320–400 nm (UVA radiation).



octyl methoxycinnamate butyl methoxydibenzoylmethane

Figure 1.7.16 These two molecules are used to absorb UV radiation in sunscreens.

AS A.8.5 Predict whether or not a particular molecule will absorb UV or visible radiation. © IBO 2007

The structure of a molecule gives a good indication of whether it will absorb UV or visible radiation. The larger the number of conjugated double bonds, or delocalized bonds, such as in a benzene ring, the greater the chance that visible light will be absorbed and the molecule will be coloured. The presence of just one double bond will result in UV absorption, most likely below 200 nm. If two double bonds are present, but they are isolated (i.e. not conjugated) then the absorption will be at a higher wavelength, but still not in the visible region. Compare the absorbance of 1-pentene with that of isoprene: 1-pentene with one carbon–carbon double bond absorbs at just 178 nm, isoprene has two conjugated carbon–carbon double bonds, and its absorbance is at the short end of the UV range (just within the range of a UV–visible spectrometer). The compounds in figure 1.7.12 all contain considerable conjugation of double bonds and delocalization of electrons. These compounds absorb visible light and appear coloured. Our uses for them over the centuries have been based on these colours.

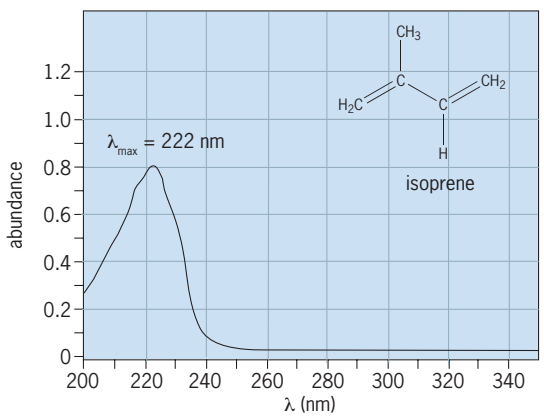


Figure 1.7.17 Isoprene is a colourless liquid that is used to make synthetic rubber. The UV–visible absorption spectrum shows that it absorbs strongly at a wavelength of 222 nm.

Section 1.7 Exercises

1 Compare the wavelengths of UV and visible radiation.

2 Explain why electron transitions occur when UV light is absorbed, whereas when IR light is absorbed only molecular vibrations occur.

3 Consider the visible spectrum of chlorophylls *a* and *b* given in figure 1.7.11.

State the four wavelengths of light at which chlorophyll *a* and chlorophyll *b* absorb light most strongly.

4 Explain why transition metal complexes are coloured and so absorb visible light.

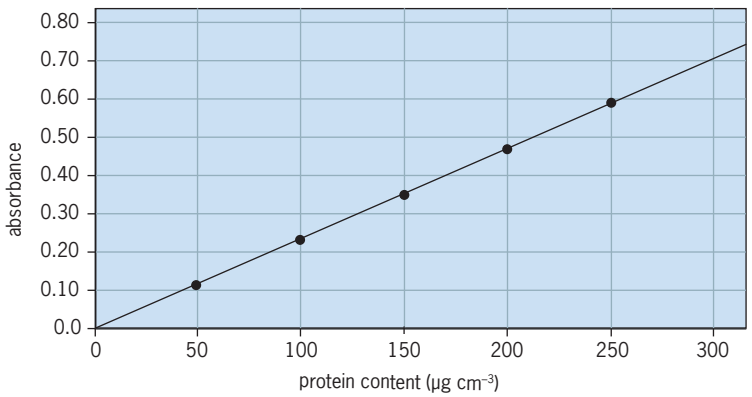
5 List the factors that affect the colour of transition metal complexes.

6 Explain how UV–visible spectrometry may be used for:

a qualitative analysis

b quantitative analysis.

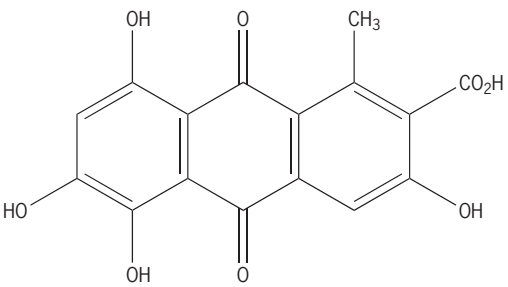
7 The protein content of a biological sample may be determined using a reagent to convert the protein into a coloured complex. The absorbance of the complex is then measured using a UV–visible spectrophotometer set at an appropriate wavelength. Using standard protein solutions, the calibration curve shown at right was obtained. A sample containing protein was diluted 100-fold. The diluted sample was found to have an absorbance of 0.40.



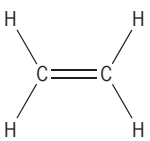
a Determine the protein content of the sample in μg cm^{-3}.

b Briefly describe how an 'appropriate wavelength' for this analysis would be determined.

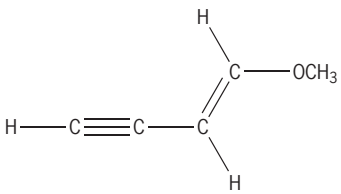
8 Consider the following molecules.



I



II



III

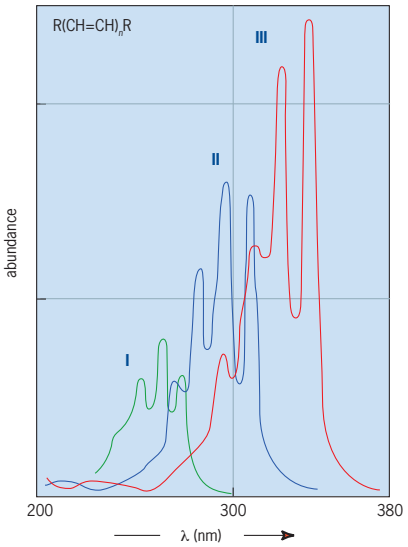
a Order these molecules from shortest maximum wavelength of light absorbed to longest maximum wavelength of light absorbed.

b Explain whether you would expect any of these molecules to be coloured.

9 The UV–visible absorption spectra for three compounds is shown in the figure to the right. All of the compounds have the general formula $R(CH{=}CH)_nR$, where n is the number of carbon–carbon double bonds in the molecule.

a State which compound has the greatest value for n.

b Explain the relationship between the number of double bonds and the maximum wavelength at which an organic compound absorbs radiation in the UV–visible region.



1.8 NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY HL

Recall from section 1.4 that ^1H NMR spectroscopy involves the application of a strong magnetic field to a sample and irradiation with radio waves. The sample absorbs energy and hydrogen nuclei are excited. As these hydrogen nuclei relax from their excited state they emit radio frequency energy that is picked up by a radio receiver and passed on to be recorded as a spectrum.

In section 1.4, the environment of different hydrogen nuclei was discussed as influencing the signal produced by those nuclei. The location of these NMR signals needs to be compared to a reference signal, just as IR and UV–visible absorption data needs to be compared to the absorption of the solvent.

AS **A.9.1 Explain the use of tetramethylsilane (TMS) as the reference standard.**

The reference standard used in NMR is a compound called tetramethylsilane (TMS), $(CH_3)_4Si$. It has been chosen as the reference because of the following properties:

- All 12 hydrogen atoms in TMS are equivalent.
- It is chemically unreactive.
- It is easily removed from the sample after the measurements have been taken.
- It produces a single, sharp NMR signal that does not interfere with other resonances.

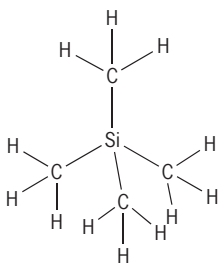
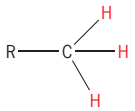
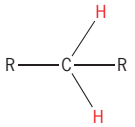
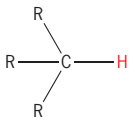
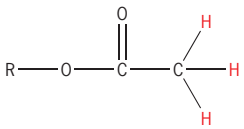
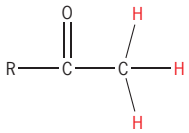
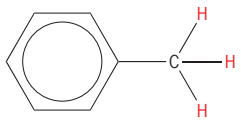
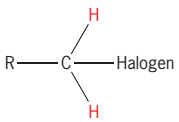
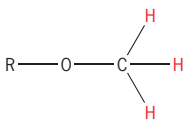
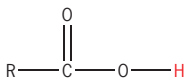
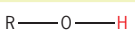
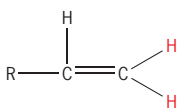
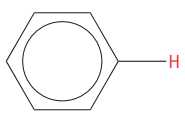
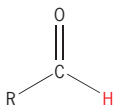


Figure 1.8.1 Tetramethylsilane.

It is important to realize that the magnets used in different NMR spectrophotometers are not identical. So resonance frequencies of identical protons may vary from one instrument to another. An NMR spectrophotometer is calibrated by reporting other signals in terms of how far they have shifted relative to the reference signal. In doing this, the scale is divided by the frequency of the radiation used (it may be 100 MHz or even 500 MHz, depending on the spectrometer) and this value is then multiplied by 10^6 (otherwise the value is too small). This produces the quantity known as *chemical shift*, which has the symbol δ and units ppm (parts per million). Note that NMR spectrophotometers are very sensitive and very precise. The proton resonances that they record all fall within a 12 ppm range of each other. The chemical shifts of several types of ^1H nuclei are shown in table 1.8.1.

TABLE 1.8.1 CHARACTERISTIC ^{1}H NUCLEI NMR CHEMICAL SHIFTS

(These values can also be found in table 18 of the IB Data Booklet. © IBO 2007)

Type of ^{1}H nucleus (proton)	Chemical shift (ppm)
	0.9–1.0
	1.3–1.4
	1.4–1.6
	2.0–2.5
	2.2–2.7
	2.5–3.5
	3.5–4.4
	3.3–3.7
	9.0–13.0
	4.0–12.0
	4.5–6.0
	6.9–9.0
	9.4–10.0

The different electronegativities of the surrounding atoms influence the chemical shift of ^1H nuclei. In the case of tetramethylsilane, TMS, silicon is *less* electronegative than carbon, so the electron density (and hence shielding) around the methyl hydrogens is greater than in the carbon derivative, 2,2-dimethylpropane. This causes a downward shift for the resonance of those ^1H nuclei and this resonance is defined as zero. Elements that are more electronegative than carbon reduce the electron density around the carbon atom, they deshield the hydrogen atoms that are bonded to that carbon atom. The proton shifts for a number of methyl derivatives is given in table 1.8.2 and the deshielding effect is shown in figure 1.8.2.

Another example of the effect of increasingly electronegative neighbours can be seen in the comparison of halogenoalkanes. The more electronegative the neighbouring halogen, the greater the chemical shift, and the more halogen atoms that are bonded to the carbon, the greater the chemical shift.

TABLE 1.8.2 THE EFFECT OF ELECTRONEGATIVITY ON CHEMICAL SHIFT FOR METHYL PROTONS

Increasing electronegativity					
Compound	$(CH_3)_4C$	$(CH_3)_3N$	$(CH_3)_2O$	CH_3F	Decreasing electronegativity
Chemical shift (δ)	0.9	2.1	3.2	4.1	
Compound	$(CH_3)_4Si$	$(CH_3)_3P$	$(CH_3)_2S$	CH_3Cl	
Chemical shift (δ)	0.0	0.9	2.1	3.0	

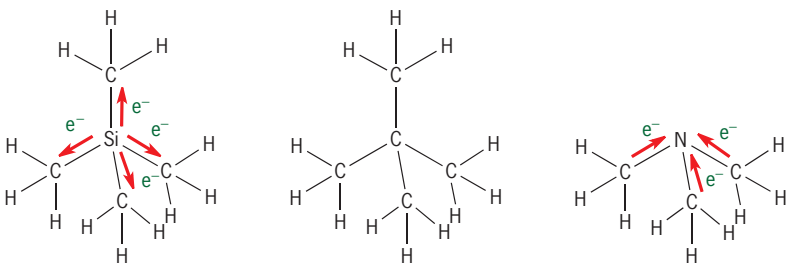


Figure 1.8.2 As the electronegativity of the atom bonded to the methyl group increases, the shielding of the methyl protons decreases and their chemical shift increases.

TABLE 1.8.3 THE EFFECT OF ELECTRONEGATIVITY ON CHEMICAL SHIFTS FOR METHYL PROTONS IN HALOGENOALKANES

Increasing number of halogen atoms				
	X = Cl	X = Br	X = I	Decreasing electronegativity
$\delta(CH_3X)$	3.0	2.7	2.1	
$\delta(CH_2X_2)$	5.3	5.0	3.9	
$\delta(CHX_3)$	7.3	6.8	4.9	

Analysing ^{1}H spectra

AS A.9.2
Analyse ^{1}H spectra. © IBO 2007

In analysing a ^{1}H NMR spectrum many parts of the spectrum provide useful information.

- The number of peaks gives an indication of the number of different environments of ^{1}H atoms in the molecule (see p. 29).
- The chemical shift of the different peaks indicates the nature of the neighbouring atoms.
- The integrated ratio of the peak sizes indicates the ratio of the numbers of hydrogen atoms with the different environments.
- The extra detail given by splitting patterns helps to locate the hydrogens by identifying how many hydrogen atoms are bonded to neighbouring atoms.

Closer inspection of absorption peaks reveals that they are split into a number of closely spaced peaks, called doublets, triplets and so on. This splitting is the result of the influence of neighbouring nuclei on the nucleus to which the peak corresponds. Neighbouring nuclei are also tiny bar magnets and so they influence the effect of the applied magnetic field on the nucleus under investigation. Valuable information about these neighbouring groups can be obtained by analysing the splitting patterns of each signal in the NMR spectrum.

Splitting patterns can be complex, but one useful feature is seen in the spectrum in figure 1.8.3. The peak at 3.8 ppm, due to the CH_2 protons, is split into four, indicating that there are three neighbouring protons (the CH_3 protons). Similarly, the peak at 1.8 ppm is a triplet. These protons have two neighbouring protons (the CH_2 group). This illustrates the rule that a proton that has n equivalent neighbouring protons gives a signal peak that is split into a multiple of $n + 1$ peaks.

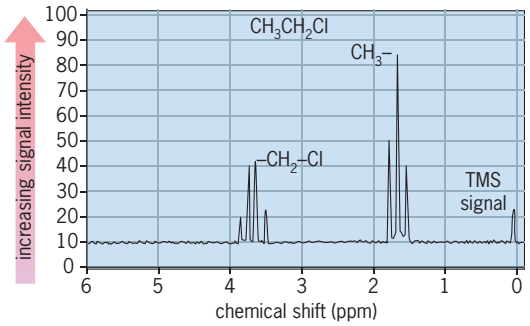
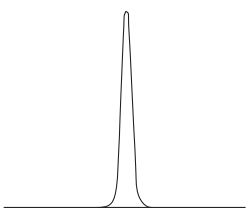
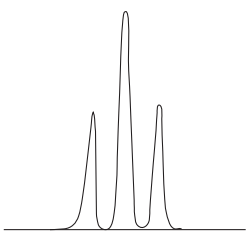
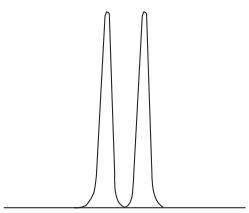
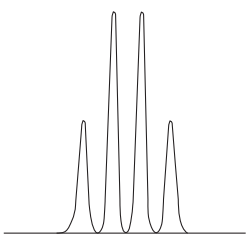


Figure 1.8.3 The NMR spectrum of chloroethane. The peak at 1.8 ppm is a triplet and that at 3.8 ppm is a quartet.

TABLE 1.8.4 SUMMARY OF SPLITTING PATTERNS

Splitting pattern	Looks like this	Gives the following information	Splitting pattern	Looks like this	Gives the following information
Singlet	 singlet peak	The neighbouring functional group is most likely a hydroxyl group.	Triplet	 triplet peak	The neighbouring carbon atom has two hydrogen atoms bonded to it.
Doublet	 doublet peak	The neighbouring carbon atom has one hydrogen atom bonded to it.	Quartet	 quartet peak	The neighbouring carbon atom has three hydrogen atoms bonded to it.

Worked example 1

Consider the following NMR spectrum of a compound with molecular formula C_3H_8O. Use table 1.8.1 (p. 57) to help determine the structure of the molecule.

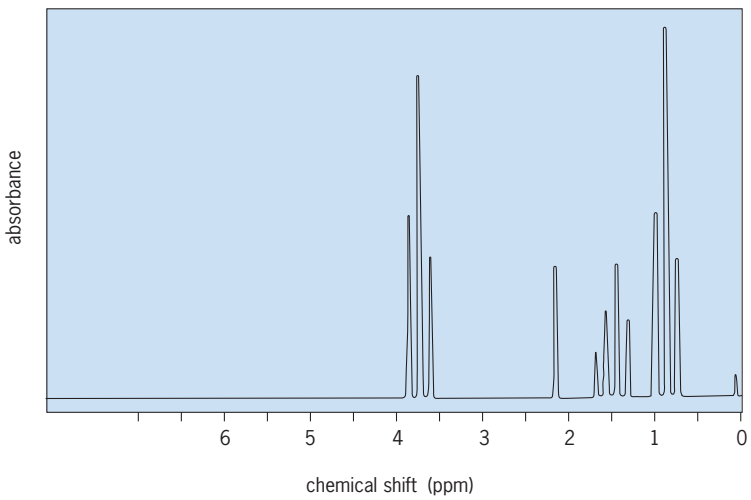


Figure 1.8.4 The NMR spectrum of C_3H_8O.

Solution

- There are four sets of peaks in the spectrum, indicating four different environments for the ^1H nuclei.
- Note that the formula is C_3H_8O, so if there are four different environments, there must be a hydroxyl, –OH, group to provide the fourth position for the hydrogen atoms to be bonded.
- The triplet peak centred at $\delta = 3.8$ ppm could correspond to the hydroxyl proton. Since it is a triplet, the adjacent carbon must have two ^1H atoms bonded to it.
- The triplet peak at $\delta = 0.9$ ppm will most likely correspond to methyl protons (chemical shift = 0.9–1.0 ppm). Since it is a triplet, the adjacent carbon must have two ^1H atoms bonded to it.
- The quartet peak at $\delta = 1.4$ corresponds to CH_2 protons that must be adjacent to a methyl group, CH_3, since it is a quartet and the chemical shift corresponds to that of a CH_2 group (chemical shift = 1.3–1.4 ppm).
- Finally, the singlet at $\delta = 2.1$ should correspond to the CH_2 group that is closest to the hydroxyl group. The high electronegativity of the oxygen atom will increase the chemical shift of the CH_2 which would normally be in the range 1.3–1.4 ppm.

This spectrum does not have an integration trace; however, we have enough information to determine the structure. The structure of the molecule is shown in figure 1.8.5.

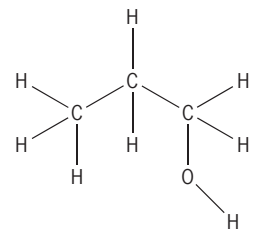


Figure 1.8.5 The compound with molecular formula C_3H_8O is propan-1-ol.

Worked example 2

A compound has the molecular formula C_3H_6O. Consider the following NMR spectrum and use table 1.8.1 (p. 57) to help determine the structure of the molecule.

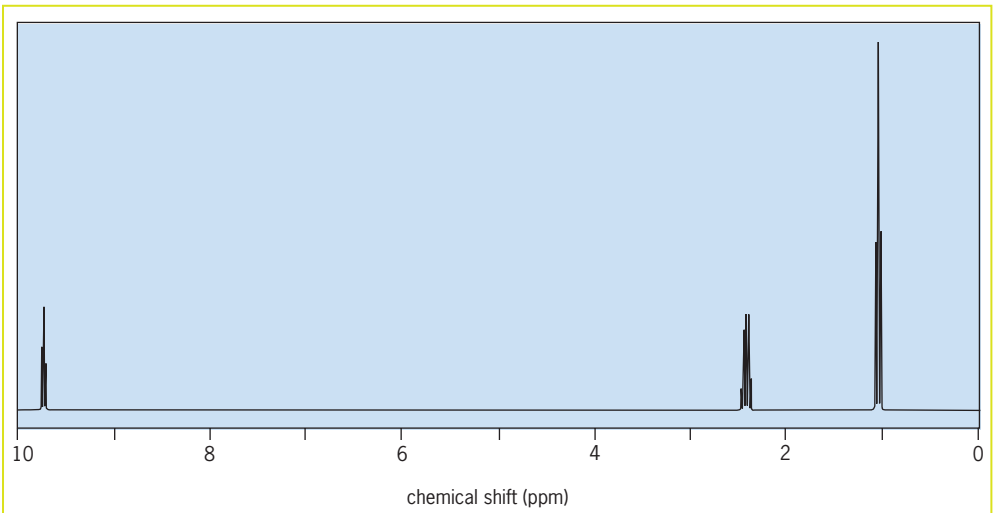


Figure 1.8.6 The NMR spectrum of C_3H_6O.

Solution

- There are three sets of peaks in the spectrum, indicating three different environments for the ^1H nuclei.
- Note that the formula is C_3H_6O, so the three different environments means that hydrogen atoms are bonded only to carbon atoms. There is no hydroxyl group, so the single O will be either an aldehyde or a ketone. With only three carbons and three different environments, a ketone is not possible (would have only one environment for ^1H).
- The triplet peak centred at $\delta = 9.7$ ppm will correspond to the proton in the aldehyde functional group. The very large chemical shift is due to the adjacent carbonyl group, which deshields the proton. Since it is a triplet, the adjacent carbon must have two ^1H atoms bonded to it.
- The multiplet peak at $\delta = 2.4$ must belong to hydrogens in the middle of a chain since the peak has split into more peaks than an adjacent CH_3 group would cause (a quartet) and the chemical shift corresponds to that of a CH_2 group (chemical shift = 1.3–1.4 ppm) that has been deshielded by the adjacent carbonyl group (increasing the chemical shift).
- The triplet peak at just over $\delta = 1.0$ ppm will most likely correspond to methyl protons (chemical shift = 0.9 – 1.0 ppm). As it is a triplet, the adjacent carbon must have two ^1H atoms bonded to it.
- This spectrum does not have an integration trace; however, we have enough information to determine the structure.
- The structure of the molecule is shown in Figure 1.8.7.

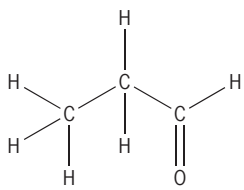


Figure 1.8.7 The compound with molecular formula C_3H_6O is propanal.

Section 1.8 Exercises

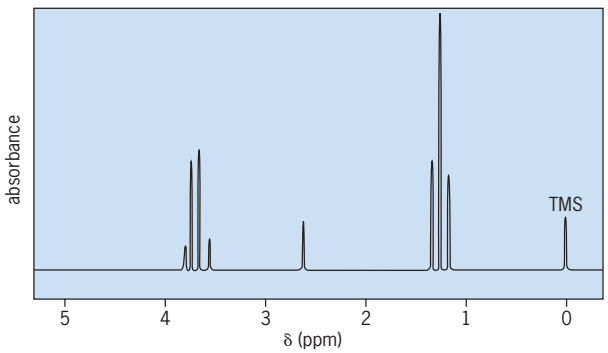
1 **a** State the full name and draw the structure of the compound represented by the letters TMS.

b Explain why TMS is used as a reference standard in ^{1}H NMR spectroscopy.

2 **a** Explain what is meant by the term *chemical shift*.

b Describe how chemical shift is used as a reference standard to identify the environment of the ^{1}H nuclei that correspond to a given peak.

3 Examine the following NMR spectrum and answer the following questions.



a Copy and complete the following table, using table 1.8.1 to help you.

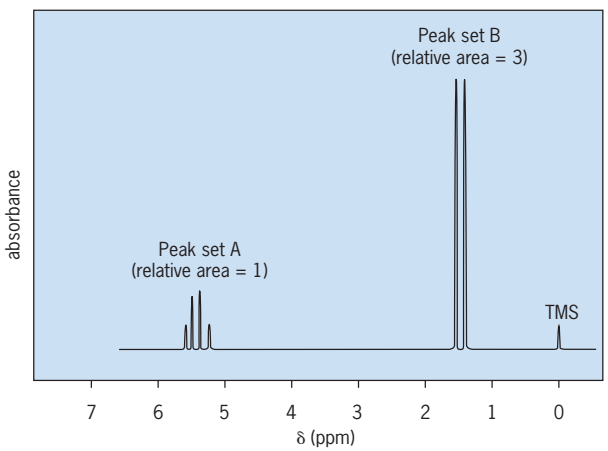
Chemical shift (δ)	Type of ^{1}H nucleus	Splitting pattern	Neighbouring carbon
3.8	$-CH_2-$	Quartet	
2.6		Singlet	
1.3		Triplet	Has 2 Hs

b Using the information you have gathered in the table, draw the structure of this compound, which has the molecular formula C_2H_6O.

4 Explain how NMR spectra would help to distinguish between two isomers such as $H_3C-O-CH_3$ and H_3C-CH_2OH.

5 The NMR spectra of the following compounds all have one feature in common. Describe the feature: C_6H_6, $(CH_3)_4C$, CH_2Cl_2, $HClC{=}CHCl$

6 The ^{1}H NMR spectrum for C_4H_8 is given below. Examine the spectrum and answer the following questions.

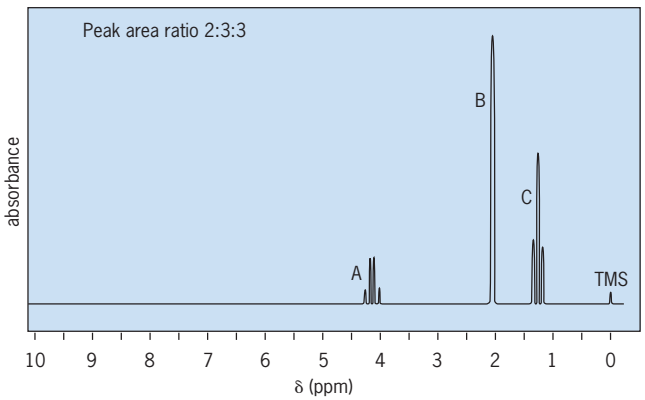


a Copy and complete the following table using table 1.8.1, or the IB Data Booklet © IBO 2007, to help you.

Chemical shift (δ)	Type of ^{1}H nucleus	Splitting pattern	Neighbouring carbon
5.4			
1.5			

b Using the information you have gathered in the table, draw the structure of this compound.

7 Consider this NMR spectrum of a compound with molecular formula $C_4H_8O_2$.



a State how many different ^{1}H environments there are in this molecule.

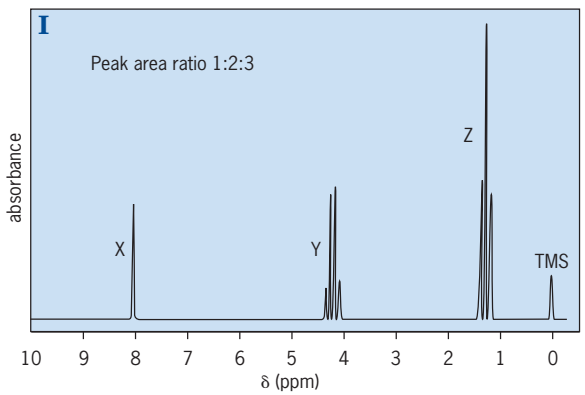
b Comment on the comparison between the number of ^{1}H environments and the number of carbon atoms in the molecular formula.

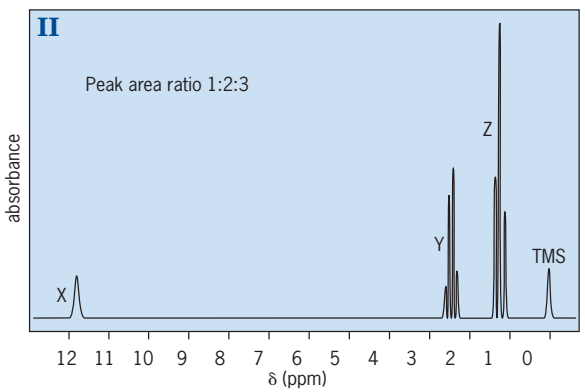
c Peak A is a quartet. Explain the impact of this observation on your determination of the structure of this molecule.

d Assuming that there is not an –OH group in this compound, explain why peak B is a singlet.

e Draw the structure of the molecule that matches this NMR spectrum.

8 The following two NMR spectra belong to two structural isomers, each with the molecular formula $C_3H_6O_2$. One of these spectra corresponds to ethyl methanoate and the other to propanoic acid.





a Draw the structural formula of ethyl methanoate and of propanoic acid.

b Describe the effect that being bonded to the O–C=O functional group will have on the chemical shift of hydrogen atoms.

c Describe one feature that is identical in the two NMR spectra.

d Identify which molecule corresponds to each of the two NMR spectra.

1.9 CHROMATOGRAPHY

HL

Two types of column chromatography—**high performance liquid chromatography (HPLC)** and **gas–liquid chromatography (GLC)**—are used routinely in analytical laboratories. Both use the same principles as the simple column chromatography described in section 1.6, but both involve the use of sophisticated (and expensive) instruments.

AS **A.10.1 Describe the techniques of gas–liquid chromatography (GLC) and high-performance liquid chromatography (HPLC).**

Gas chromatography

A schematic diagram of a gas chromatograph is shown in figure 1.9.1. The stationary phase is packed in a long column, (often several metres long) which is then coiled into a helix. Typically, the column is several millimetres in diameter, although some columns used are even thinner, of capillary size.

The column is filled with fine granules of an inert solid. These granules serve as a support for a high boiling point, viscous liquid. In this technique, known as gas–liquid chromatography (GLC), separation of a mixture depends on the interaction of the components of the mixture with the stationary and gas phases. In GLC, the mobile phase is a gas, called a carrier gas, which is inert and is usually nitrogen, helium or argon.

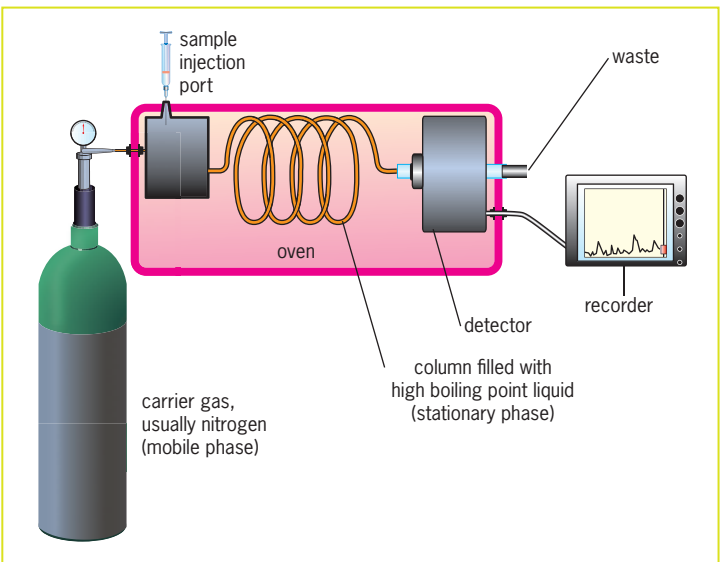


Figure 1.9.1 Schematic diagram of a gas–liquid chromatograph.

Ws WORKSHEET 1.4 GLC

The sample to be separated must be vaporized. This limits the use of gas chromatography to substances that do not decompose when heated, often substances with molar masses of less than 300 g mol^{-1}. The sample is injected, using a syringe, into a steady gas flow at the start of the column. Very small samples are required, around several microlitres (μL). The components of the sample adsorb to the stationary phase and partition occurs as they desorb and are carried along in the gas stream. The rate of flow of the gas is controlled by the temperature, and the entire column is housed in an oven.

Components are detected as they elute from the column, either by their effect on the thermal conductivity of the gas or, more commonly, by their effect on the electrical conductivity of a flame. In a flame ionization detector (FID), burning hydrogen ionizes the eluted components, which produce an electrical current that is then amplified and recorded in the chromatogram.

Results are presented as a graph. Identification of each component is based on the retention times, R_t. The retention time of a component is compared to that of standard samples run through the column under the same conditions. An identical retention time indicates identical components. Quantification is based on the area under each peak, which is calculated automatically by an attached computer. A calibration curve is prepared by measuring peak areas for samples of known concentration.

Gas–liquid chromatography is extremely sensitive, and can detect specific components in masses as small as 10^{-12} g. It can be used to separate complex mixtures such as the hydrocarbons, that make up crude oil. Gas–liquid chromatography is frequently used in combination with mass spectrometry and infrared spectroscopy, and together these techniques make an incredibly powerful tool for identifying and quantifying mixtures. The mixtures are separated using GLC. As the separated components elute from the column, they are passed directly into a mass spectrometer, which is sensitive enough to detect and measure the concentration of minute quantities. A computer connected to the mass spectrometer matches the spectra to a database of spectra of known substances. This technique is used to detect narcotics, anabolic steroids, analgesics and other banned substances in the urine samples of athletes, and to detect toxins in food samples.

Figure 1.9.2 A gas chromatograph (left) connected to a mass spectrometer (right) provides a powerful analytical tool for forensic chemists.

Other uses of GLC in identifying compounds that can vaporize without decomposing include:

- analysing blood alcohol levels
- monitoring pollutants in air and other gas samples, such as gases from underground mines
- analysing foods—for example, checking the fatty acid content of vegetable oils and determining the presence of certain food colourings
- identifying the source of oil spills by fingerprinting crude oil found at different locations.

AS A.10.2
Deduce which chromatographic technique is most appropriate for separating the components in a particular mixture.

CHEM COMPLEMENT

Name that cockroach

Gas chromatography is being used to aid in the identification of cockroach species! Cockroaches pose a major problem in many parts of the world, where they consume large quantities of grain and other food crops. Different species of cockroach present different problems and so require different methods of control, but first the species concerned needs to be identified. Biologists traditionally identify species by comparing features with those listed in guides for identification. This can be time-consuming and, for closely related species, quite difficult. Enter the chemist armed with a gas chromatograph! It has been found that the outer, waxy layer of a cockroach is distinctive for each species. By dissolving this waxy coating and analysing it using gas chromatography, the species is found with some certainty in about 30 minutes.

Figure 1.9.3 Which species of cockroach is this? GLC may be used to efficiently and accurately identify it.

CHEM COMPLEMENT

Over the limit?

Motorists are often stopped for random breath testing. An initial screening test consists of blowing into a white tube connected to a black box. This test, based on an electrochemical method, gives an estimate of the driver's blood alcohol content (BAC). If this estimate indicates a BAC over the legal limit (ranging from 0 in countries such as Pakistan, Saudi Arabia and Hungary, through 0.05% in many countries of the world, to 0.08% in countries such as the US, Malta and Ireland), the driver must be tested using a breathalyser.

A breathalyser is essentially a type of colorimetric analysis. A measured amount of the driver's breath sample is bubbled through an acidified potassium dichromate ($K_2Cr_2O_7$) solution. Any alcohol present reacts according to the equation:

$$3CH_3CH_2OH(aq) + 2Cr_2O_7^{2-}(aq) + 16H^+(aq) \rightarrow 3CH_3COOH(aq) + 4Cr^{3+}(aq) + 11H_2O(l)$$

The extent of the colour change (from chromium(VI), orange, to chromium(III), green) caused by the breath sample is used to determine the amount of ethanol present. This colour is compared with that of a test sample containing only the acidified potassium dichromate solution.

The BAC is most reliably tested by GLC of a blood sample. Nitrogen gas is blown through a measured sample of the driver's blood. The nitrogen removes any alcohol from the blood and carries it, as a vapour, through the GLC column. Any alcohol eluted from the column is detected and measured electronically. By comparing the peak area with those of alcohol samples of known concentration, the BAC can be calculated.

High performance liquid chromatography

A schematic diagram of a high performance liquid chromatograph is shown in figure 1.9.4. The stationary phase is a solid packed in a column that is much shorter than those used in GLC (usually 10–30 cm in length). The nature of this solid varies with the particular analysis being conducted. The solid particles are small and tightly packed, so the mobile phase (the solvent) is pumped through the column to ensure an efficient flow rate. The solvent under pressure gives HPLC its alternative name: high pressure liquid chromatography.

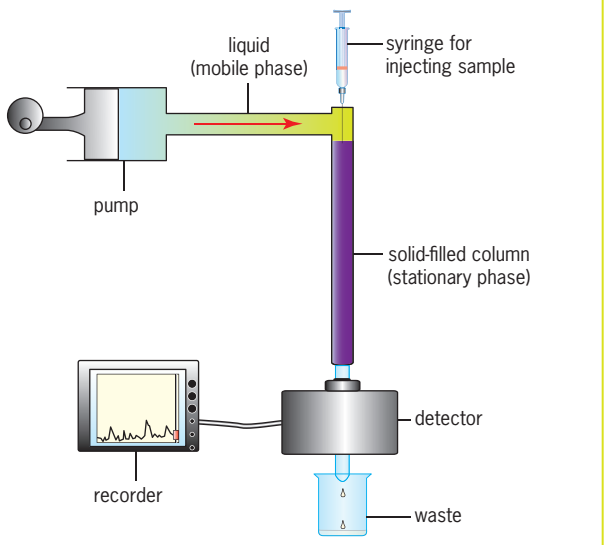


Figure 1.9.4 Schematic diagram of a high performance liquid chromatograph.

Components separate as they pass through the column as a result of their different degrees of adsorption and desorption. As the components elute from the column they are detected in the solvent stream by measuring a variety of physical properties of the eluted components, including refractive index, conductivity, fluorescence and, most commonly, absorption of UV–visible light. The recorder produces a printout, known as a chromatogram.

Figure 1.9.5 shows a chromatogram for a mixture of three hydrocarbons. The horizontal axis shows the retention time. The most strongly adsorbed component (octane, C_8H_{18}) has the longest retention time. The least strongly adsorbed component (butane, C_4H_{10}) moved quickly through the column and so has a shorter retention time. The order of retention times can be explained in terms of the increasing strength of van der Waals' forces with increasing molecular size. The larger molecules, with the stronger van der Waals' forces, adsorb more strongly to the column and so take longer to be eluted. The relative amounts of each hydrocarbon are calculated by reference to the vertical axis. The area under each peak provides a measure of the relative amounts of each hydrocarbon. Here, there is approximately twice the amount of octane as hexane.

Identification of components on the chromatogram is made by comparing retention times with those of known substances obtained under the same conditions. The concentration of each component is determined by comparing peak areas with those of the same substance at known concentrations. The peak areas for samples of known concentration are determined and plotted against concentration. In this way, a calibration curve is constructed to determine unknown concentrations. Calibration is conducted with each test, as results depend on factors such as column type and length, and the flow rate.

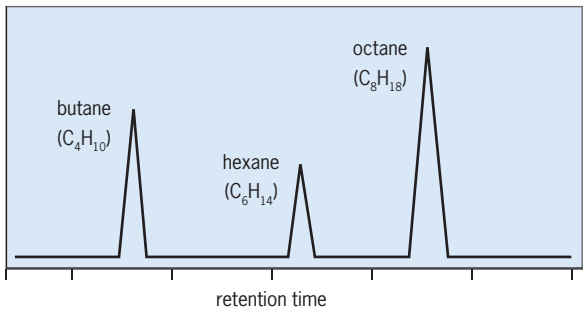


Figure 1.9.5 A chromatogram for a mixture of three saturated hydrocarbons, obtained using HPLC.

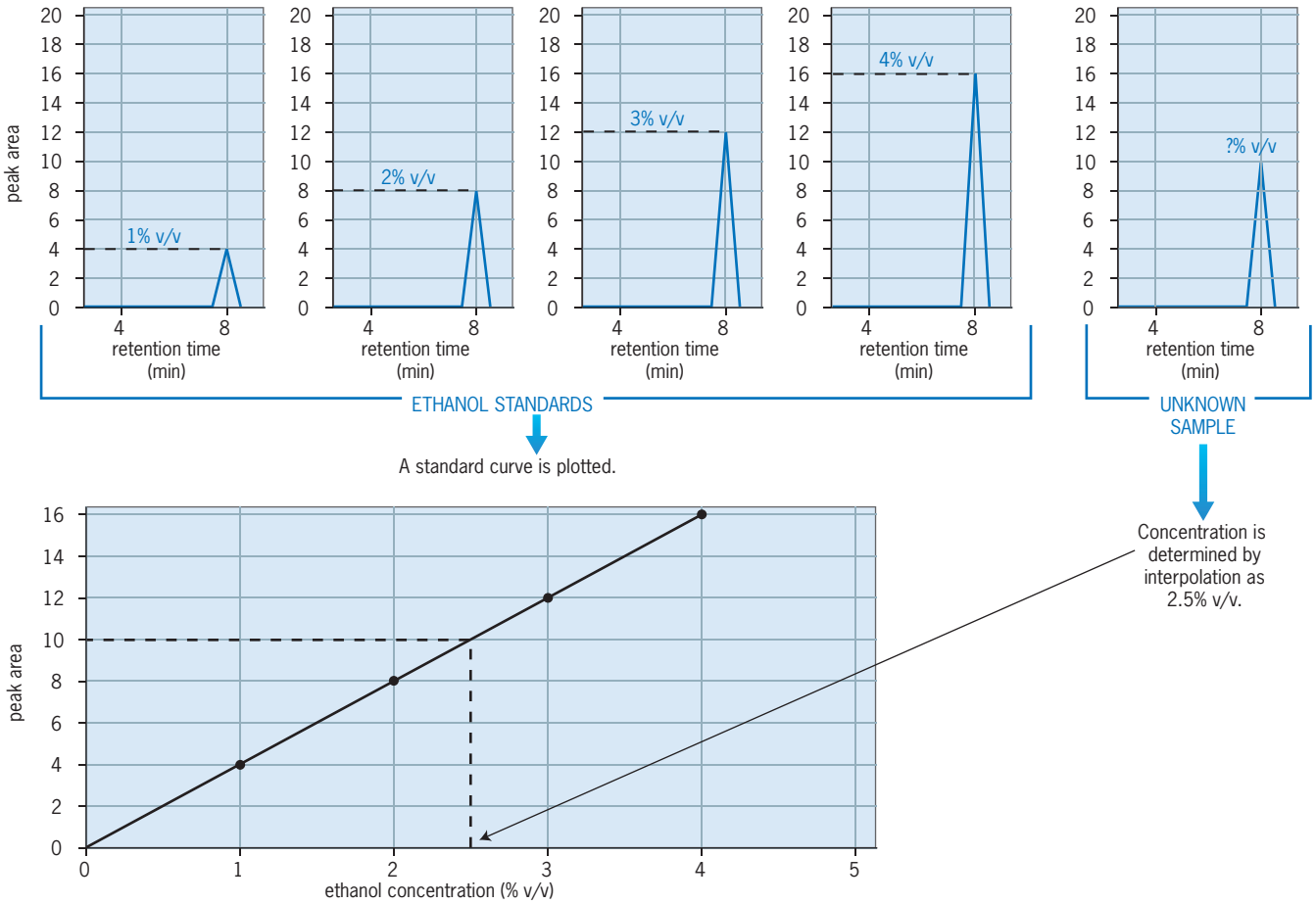


Figure 1.9.6 Determining concentration using standard solutions in HPLC.

HPLC is used for non-volatile molecular substances (with relative molar mass of approximately 300 g mol^{-1} or greater). It is well suited to temperature-sensitive materials that would decompose if used in a gas–liquid chromatograph. It requires only a small sample (approximately 10 μL) and is efficient and relatively fast. Advances made in shortening the column length have meant that the procedure is now faster and less expensive than previously, as shorter column lengths means less solvent is needed and the retention times are shorter.

Figure 1.9.7 An HPLC being used in a chemistry laboratory.

HPLC has a wide range of uses in testing and analysing temperature-sensitive compounds.

These include:

- testing pharmaceutical products
- quality control of insecticides and herbicides
- analysing foods and beverages, such as analysing the additives (antioxidants) in margarine, sugars in fruit juice and vitamins in other foods
- analysis of alcoholic beverages
- polymer analysis
- analysis of oil
- biochemical and biotechnology research
- detecting drugs, such as barbiturates, in blood samples.

Section 1.9 Exercises

1 In all forms of column chromatography, the material packed into the column is usually finely divided to produce a large surface area of solid. Explain why this is important.

2 For each of the following techniques describe:

i the mobile phase

ii the stationary phase.

a Gas–liquid chromatography

b High performance liquid chromatography

3 By considering the types of samples that can be analysed using the two methods, describe an advantage of HPLC over GLC.

4 List two types of samples that could be analysed by HPLC, but not by GLC, and explain why they cannot be analysed by GLC.

5 **a** Describe the general features of a sample that can be analysed using GLC.

b Describe the general features of a sample that can be analysed using HPLC.

c List three uses of GLC.

d List three uses of HPLC.

6 When answering the following questions, select from the following list of terms:

peak heights, peak areas, peak widths, retention time (R_t)

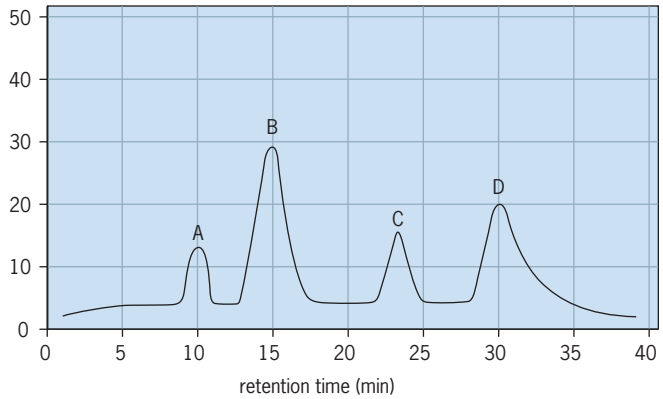
State which feature of a chromatogram obtained during an HPLC analysis is:

a used to identify the components of a mixture

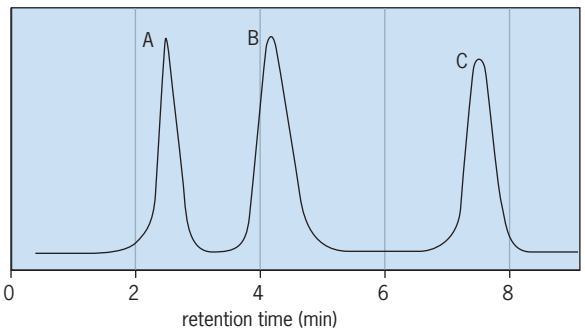
b used to determine the concentration of each component of the mixture

c least affected by a change in concentration of the sample components.

7 A gas chromatogram of a mixture containing the alkanes pentane (C_5H_{12}), butane (C_4H_{10}), octane (C_8H_{18}) and decane ($C_{10}H_{22}$) is shown to the right. Identify the four peaks, A to D, on the chromatogram. Explain your choice.



8 An analytical laboratory used HPLC to analyse samples of soft drink containing the additives vitamin C (A), saccharin (B) and caffeine (C). The chromatogram obtained is shown below.



Explain how this chromatogram can be used to provide information about the soft drink:

a qualitatively

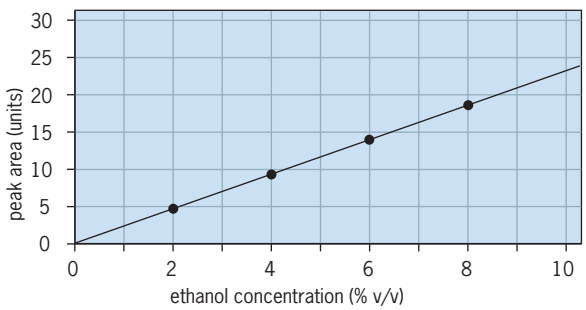
b quantitatively.

9 Compare the design and function of an HPLC and a GLC.

a State three similarities between them.

b State two differences between them.

10 Gas–liquid chromatography was used to test the alcohol content of a new white wine blend. A set of ethanol standards was prepared and, each standard was injected into the GLC column. The peak area for each standard was determined and used to generate the calibration curve shown below. A wine sample was diluted by a factor of 3 and then injected into the column. The peak area for the diluted sample was found to be 12.



a Determine the volume of ethanol in 100 cm^3 of the wine.

b Suggest an alternative method for the determination of the ethanol content of the wine.

Chapter review questions and tests are available on student CD.

Chapter 1 Summary

Terms and definitions

Absorption The taking in (of radiation) which causes an effect (vibration, rotation, electron transition).

Absorption spectrum A pattern of black lines on the background of the continuous spectrum in which the black lines represent those energies of light absorbed by the electrons moving from the ground state to higher energy levels.

Adsorption The formation of weak bonds between the stationary phase and the components of a mixture during chromatography.

Atomic absorption (AA) spectroscopy Study of (usually the amount of) light absorbed by an element.

Atomic emission spectroscopy Study of the light emitted by an element.

Base peak The peak on a mass spectrum with the greatest height.

Bathochromic shift In UV–visible spectroscopy this is the movement towards absorbing longer wavelengths of light that occurs when more double bonds are present in a sample.

Chemical shift (δ) The position of the peak along the horizontal axis of an NMR spectrum when compared to the position of the peak due to the standard, TMS.

Chromophore The double- or triple-bonded part of a molecule that causes it to absorb UV or visible radiation.

Conjugated Alternating multiple single and double (or triple) bonds in a molecule.

Desorption The breaking of weak bonds between the stationary phase and the components of a mixture during chromatography.

Elute To come out of the bottom of the chromatography column.

Emission (line) spectrum Spectrum of bright lines on dark background generated when an element is excited and then releases energy as light.

Excited state The state of an atom when electrons are in higher than usual energy levels.

Fragmentation pattern A pattern of peaks or lines on a mass spectrum caused by the breaking up of a molecule during mass spectrometry.

Frequency Number of waves passing a given point each second.

Functional group An atom or group of atoms bonded to a molecule that affects the chemical reactivity of the molecule.

Gas–liquid chromatography (GLC) A type of chromatography that uses a liquid stationary phase and a gaseous mobile phase or carrier gas.

Ground state The energy level to which unstable electrons return, releasing light as they do so.

High performance liquid chromatography (HPLC) A type of column chromatography that uses a solid stationary phase and a liquid mobile phase under pressure.

HOMO Highest occupied molecular orbital. It is the molecular orbital from which electrons are excited when they absorb UV radiation.

Infrared (IR) spectroscopy Study of the effects of infrared electromagnetic radiation that is absorbed by a substance in solution.

Interpolation The process of obtaining a value from a graph by reading from the line between points on the graph.

LUMO Lowest unoccupied molecular orbital. It is the molecular orbital to which electrons are excited when they absorb UV radiation.

Magnetic resonance imaging (MRI) Medical technique using the principles of NMR spectroscopy to obtain a three-dimensional image of organs and parts of the body.

Mass spectrometry Study of the movement of charged particles in the influence of a magnetic field in order to determine the masses of those particles.

Mobile phase In chromatography, the solvent or gas that passes over the stationary phase, and carries the components of the mixture with it.

Molecular ion The positive ion formed in a mass spectrometer when the molecule has only lost one electron.

Monochromatic light source Source of light of only one wavelength.

Monochromator The part of a spectrophotometer that allows only one wavelength of light to be transmitted.

Nuclear magnetic resonance (NMR) spectroscopy Study of the effects of radio wave electromagnetic radiation on certain nuclei placed in a magnetic field.

Partition The distribution of a mixture between two phases.

Percentage transmission In IR spectroscopy, the proportion (percentage) of light that passes through a sample.

Qualitative analysis Analysis in which the type of compounds or functional groups present in a sample are identified.

Quantitative analysis Analysis in which the amounts of components of a mixture or of specific compounds present in a sample are identified.

Retention (retardation) factor (R_f)

$R_f = \frac{\text{distance moved by component}}{\text{distance moved by solvent}}$, a measure of how far a component of a mixture has travelled in comparison to the solvent during paper or thin-layer chromatography.

Retention time (R_t) The time taken for a solute to move through a column in column chromatography (including GLC and HPLC).

Spectrometer An instrument that is used for viewing the results of the interaction between a sample and the electromagnetic radiation being passed through it.

Spectrometry The measurement of the amount of light absorbed or emitted.

Spectrophotometer A more elaborate instrument that measures amounts of radiation.

Spectroscopy The study of the radiation absorbed or emitted by matter.

Spin The behaviour of an electron as if it were spinning about an axis. The spinning charge generates a magnetic field whose direction depends on the direction of spin.

Stationary phase The solid phase (or liquid coated on a solid) that does not move during chromatography.

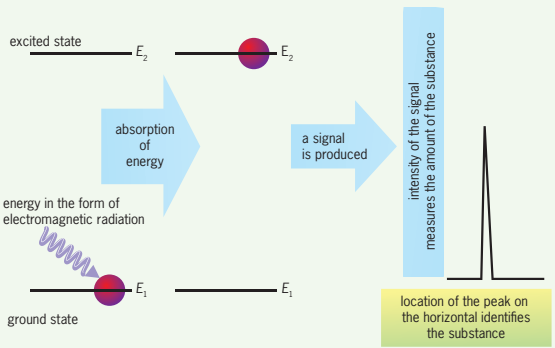
UV–visible spectroscopy Study of electromagnetic radiation in the ultraviolet and visible regions when it is absorbed by a substance in solution.

Wavelength The distance between successive crests of waves.

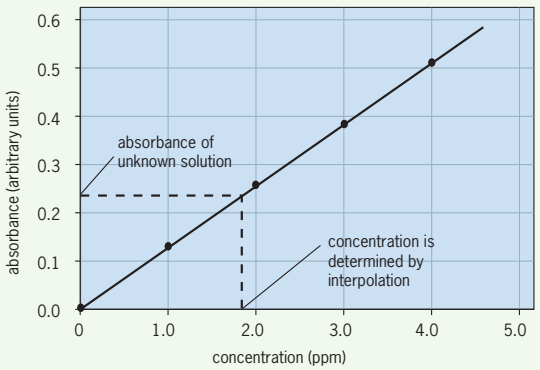
Wavenumber The inverse of wavelength $\left(\frac{1}{\lambda}\right)$, measured in cm^{-1}.

Concepts

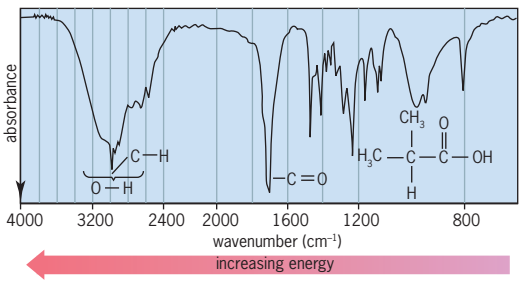
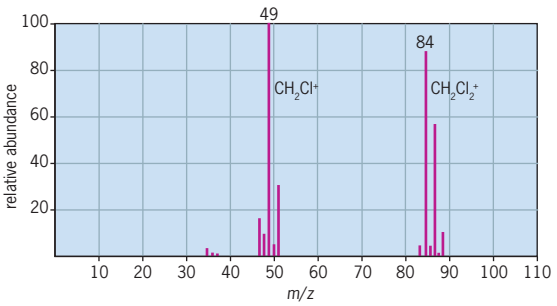
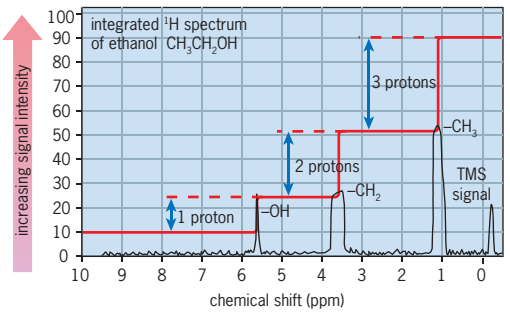
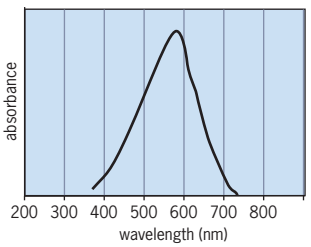
- Many forms of spectroscopy involve the absorption of electromagnetic radiation by matter to produce either line or band spectra. The position of lines or bands on the horizontal axis is used to determine the identity of the substance. The intensity of the line or band is used to determine the concentration of the substance.



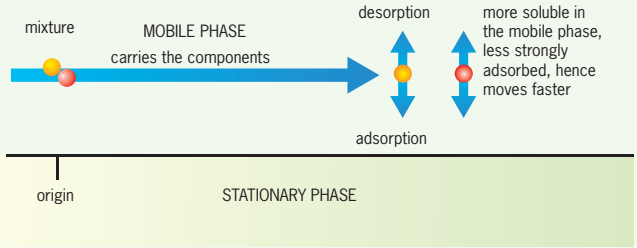
- Atomic absorption (AA) spectroscopy involves the absorption of specific wavelengths of light by the vaporized atoms in a sample solution sprayed into a flame. It is used in the analysis of metals, and can be used because the wavelength of light required to excite an electron is specific to each element, thus identifying the element.
- Spectra may be used quantitatively. This involves measuring the absorbance (at a set wavelength) of solutions of known concentration and the preparation of a calibration curve. The absorbance of the unknown solution (at the set wavelength) is measured, and its concentration determined by interpolation of the calibration curve.



- Infrared spectroscopy involves the absorption of specific wavelengths of infrared radiation by a sample in solution. It is used to identify particular functional groups in organic compounds and can be used because the wavelength of radiation required to stretch and vibrate the covalent bonds in molecules is characteristic of the functional group.
- Mass spectroscopy is widely used to identify the structure of organic molecules. It requires the sample to be vaporized and subjected to a stream of high-energy electrons. The resulting positive ions are accelerated and passed through a magnetic field, which causes their path to alter, depending on the charge–mass ratio of the ions. The fragmentation patterns when the molecules are broken are used to help determine the molecular structure.
- Nuclear magnetic resonance (NMR) spectroscopy is used to determine the precise structure of organic compounds. In NMR spectroscopy, the sample is placed in an extremely strong magnetic field and ^{1}H nuclei in different chemical environments absorb radio waves of different frequencies.

	Infrared spectroscopy	Mass spectrometry
Type of radiation used	Infrared: wavenumbers of 800–4000 cm^{-1}	No radiation; a magnetic field deflects fragments of different sizes
Effect of radiation	Causes molecular vibrations	
Appearance of spectrum		
Used to identify	Functional groups	Molecular mass and fragment masses
	NMR spectroscopy	**UV–visible spectroscopy** HL
Type of radiation used	Radio waves 100–800 MHz in addition to magnetic field	UV (200–400 nm) and visible (400–800 nm)
Effect of radiation	^{1}H nuclei excited to opposite spin state	Molecular electron excited to higher energy orbital (HOMO to LUMO)
Appearance of spectrum		
Used to identify	^{1}H environments and hence the structure of organic compounds	• Conjugation of double bonds • Quantity of sample

- In all forms of chromatography, components of a mixture are separated when a mobile phase moves over a stationary phase. Separation occurs because the components adsorb and desorb to different degrees. The weaker the bonds between a sample component and the stationary phase, the faster the component moves.



- The results of paper and thin-layer chromatography are presented as a chromatogram. The retention factor (R_f) of each component is then found.

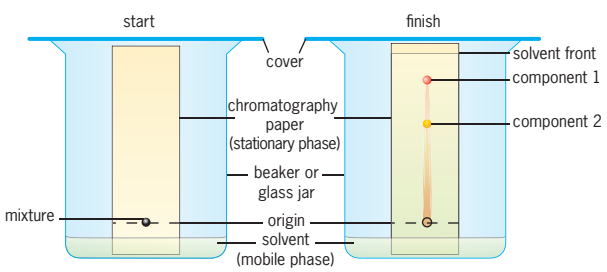
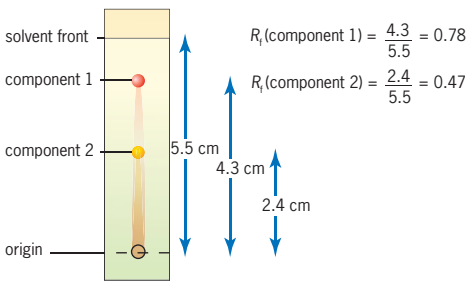
$$R_f = \frac{\text{distance moved by component}}{\text{distance moved by solvent}}$$

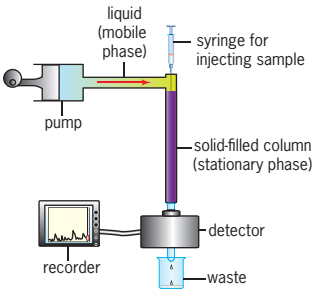
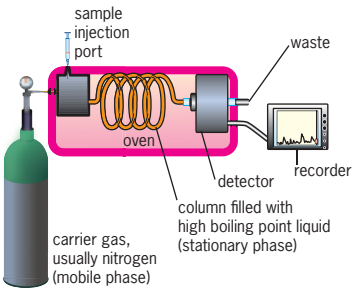
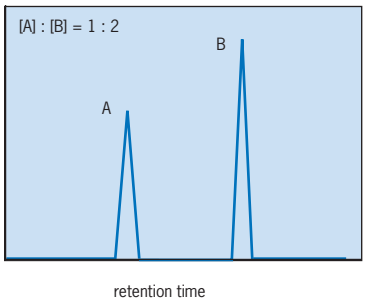
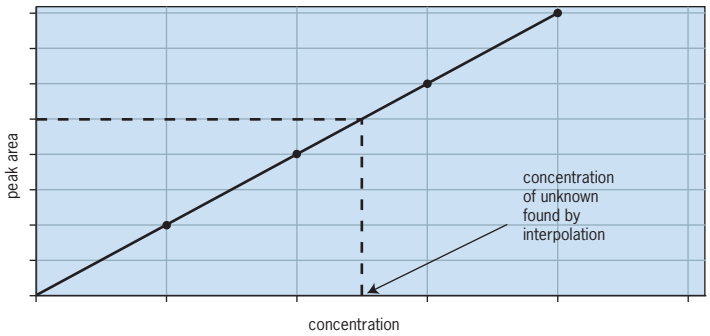
- Identification of components in paper and thin-layer chromatography is made by comparing component R_f values to the R_f values of pure, known samples obtained under identical conditions.

HL

- UV–visible spectroscopy involves the absorption of specific wavelengths of visible light and ultraviolet radiation by a sample in solution. UV–visible spectroscopy is used in the analysis of coloured substances and many organic compounds and can be used because the wavelength of light required to excite an electron in a molecule is characteristic of that compound.
- Splitting patterns and chemical shifts of ^{1}H nuclei in NMR spectra enable the complete determination of the structure of an organic compound, given the molecular formula.
- The results of gas–liquid chromatography and high performance liquid chromatography are presented as a chromatogram, showing a series of peaks at the various times taken for the components to elute from the column.
- Identification of components in column chromatography including gas–liquid chromatography and HPLC is made by comparison of retention times (R_t), where R_t is the time taken for a component to move through the chromatography column.
- Gas chromatography may be used in combination with mass spectrometry or infrared spectroscopy to detect the presence of banned substances in the urine samples of athletes and is also used in a variety of forensic tests.

See the table on the next page.

Features	Paper and thin-layer chromatography
Equipment	
Stationary phase	Paper Plastic or glass plate coated with fine powder
Mobile phase	Liquid solvent or mixture of solvents
Collection of data	
Uses	Coloured compounds, amino acids

Features	High performance liquid chromatography (HPLC) HL	Gas–Liquid chromatography (GLC) HL
Equipment		
Stationary phase	Solid packed in a column	Inert granules coated with high boiling point liquid and packed in a column
Mobile phase	Liquid solvent or mixture of solvents	Carrier gas: N_2, Ar or He
Collection of data		
Uses	Non-volatile organic compounds	Volatile organic compounds

2 HUMAN BIOCHEMISTRY

Chapter overview

This chapter covers the IB Chemistry syllabus Option B: Human Biochemistry.

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

- use enthalpy of combustion data to calculate the energy value of a food
- recognize the important structural features of proteins, carbohydrates and lipids
- describe the condensation reactions involved in the formation of polypeptides, polysaccharides and triglycerides
- describe the primary, secondary, tertiary and quaternary structure of proteins and recognize the importance of these levels of structure to molecular function
- explain how proteins can be analysed by chromatography and electrophoresis
- compare the structural properties of starch and cellulose and understand the importance of dietary fibre
- describe the difference in structure between saturated and unsaturated fatty acids referring to linoleic and linolenic acids
- identify vitamins as water-soluble or fat-soluble, with attention to the structures of vitamins A, D and C
- list the major functions of proteins, carbohydrates, lipids and micronutrients in the human body
- outline the production and function of hormones in the body
- describe how oral contraceptives work

- recall that enzymes are highly efficient and specific biological catalysts HL
- describe the mechanism of enzyme action, as well as inhibition
- determine the kinetic parameters of an enzyme, V_{max} and K_m, graphically
- compare the structural features of nucleic acids and their condensation polymers (DNA and RNA)
- describe the use of and steps involved in DNA profiling
- compare the energy production of aerobic and anaerobic respiration.

IBO Assessment statements B.1.1 to B.9.2

Can you name the world's foremost chemical institution? In what location is the science of chemistry most successfully applied? To answer these questions, you need look no further than your own body, for it is the site of a complex of extraordinary chemical processes. Every second of every day, thousands of chemical reactions occur in your body cells, each carefully controlled and monitored, each contributing to your overall functioning. That you are generally unaware of this relentless industry is further tribute to its efficiency. Biochemistry is the study of the many thousands of reactions occurring in the bodies of all living things, and the study of the biomolecules that participate in these reactions.

AS **B.1.1**
Calculate the energy value of a food from enthalpy of combustion data. © IBO 2007

2.1 ENERGY

We depend on food to supply our bodies with energy to move, keep warm and function in so many other ways. The balance between energy supplied by the food we eat and the energy used by us in our everyday activities is a controversial one in these days of 'stick-like' fashion models and newspaper headlines that tell us that people are growing too fat. In some parts of the world children can only dream of the opportunity to eat as much as their bodies need every day.

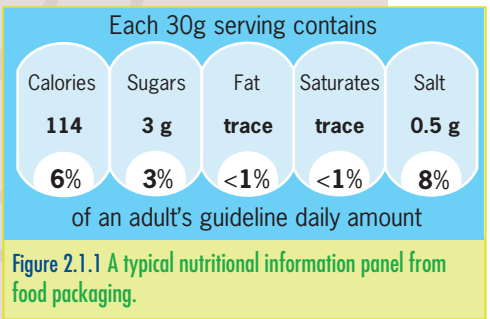


Figure 2.1.1 A typical nutritional information panel from food packaging.

If more energy is taken in from food than is used up, weight gain will follow. Similarly if more energy is used than we supply our body with, weight loss will occur. It is important to know how much energy is available from a particular food, so processed foods are labelled with energy content, as well as percentages of carbohydrates, fats and protein.

Calorific value (energy content) of foods is measured in kJ g^{-1}, kJ 100 g^{-1} or even kJ mol^{-1} if the food is a pure substance, such as glucose. It is measured using calorimetry (see *Chemistry: For use with the IB Diploma Programme Standard Level*, chapter 6). An accurately weighed sample of the food is burnt in oxygen in a **bomb calorimeter** (figure 2.1.2). This is a sealed metal container or 'bomb' containing a crucible in which the food is held. A sparking device is used to start the combustion of the food in the oxygen-rich atmosphere. Heat released during the combustion of the food is conducted through the metal 'bomb' to a measured amount of water in an insulated container surrounding it. The temperature of the water increases and the temperature change is measured with a thermometer.

Pr
PRAC 2.1
Determining the calorific value of a biscuit

The specific heat capacity of water is used together with this change in temperature to determine the amount of energy that has been released by the combustion of the food, using the equation

$$q = m \times c \times \Delta T$$

where q = heat energy change in J
m = mass of substance in g
c = specific heat capacity of the substance in J $°C^{-1}$ g^{-1}
ΔT = change in temperature in °C or K

Interactive CW Animation
Bomb calorimetry

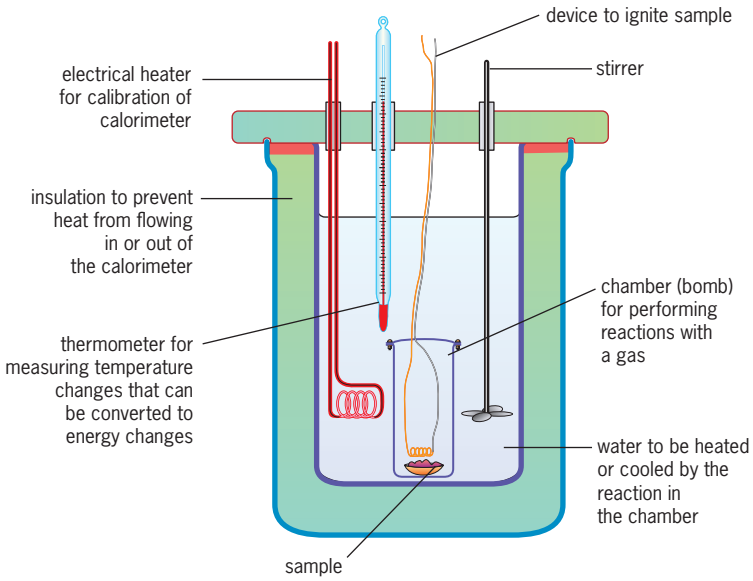


Figure 2.1.2 A bomb calorimeter can be used to measure the energy value of foods.

Worked example 1

A 0.600 g piece of bread was burnt completely in a bomb calorimeter. The heat produced raised the temperature of 500 cm^3 of water in the calorimeter from 16.00°C to 19.30°C. Assuming that the calorimeter absorbs a negligible amount of heat, what was the calorific value of the bread in kJ g^{-1}?

Solution

$\Delta T = 19.30 - 16.00 = 3.30°C$

Note that it is the water whose temperature is rising by 3.30°C, rather than the bread, so the mass of water must be used and also the specific heat capacity of the water.

$m(\text{water}) = 500$ g (since density of water = 1 g cm^{-3})
specific heat capacity of water = 4.18 J g^{-1} $°C^{-1}$

$$q = m \times c \times \Delta T$$
$$= 500 \times 4.18 \times 3.30$$
$$= 6897 \text{ J}$$

So, 0.600 g of bread releases 6897 J of energy upon combustion.

The calorific value of the bread is found by dividing the energy released by the mass of bread that was combusted:

$$E = \frac{6897}{0.600} = 11\,495 \text{ J g}^{-1} = 11.5 \text{ kJ g}^{-1}$$

The calorific value of the bread is 11.5 kJ g^{-1}.

In some cases, the specific heat capacity of the calorimeter itself, usually given as a calorimeter constant, will need to be considered in the calculation.

Worked example 2

Some macaroni pasta weighing 4.00 g was burnt completely in a bomb calorimeter with a calorimeter constant of 660 J °C^{-1}. The heat produced by the combustion of the pasta raised the temperature of 400 cm^3 of water in the calorimeter from 20.00°C to 27.50°C. Remembering to take account of the heat absorbed by the calorimeter, determine the calorific value of the macaroni in kJ g^{-1}.

Solution

Note that the calorimeter constant is equal to the energy required to raise the temperature of the calorimeter by 1°C (the equivalent of $m \times c$).

$$\begin{aligned} q &= (\text{calorimeter constant} \times \Delta T)_{\text{calorimeter}} + (m \times c \times \Delta T)_{\text{water}} \\ &= (660 \times 7.50) + (400 \times 4.18 \times 7.50) \\ &= 17\,490 \text{ J} \\ &= 17.5 \text{ kJ} \end{aligned}$$

$$E = \frac{17.5}{4.00} = 4.37 \text{ kJ g}^{-1}$$

The calorific value of the macaroni is 4.37 kJ g^{-1}.

The calorific value of a food can also be worked out theoretically using enthalpy of combustion data of the various components of the food. Each food group—**carbohydrates**, **fats** and **proteins**—has a particular enthalpy of combustion, although the different varieties of these food groups will have different enthalpies of combustion. For example, the enthalpy of combustion of monosaccharides, the simplest carbohydrates, is 15.7 kJ g^{-1}, while that of polysaccharides is 17.6 kJ g^{-1}. When required to perform a calculation of the energy value of a food, it is always best to use only the data supplied to you.

TABLE 2.1.1 ENTHALPIES OF COMBUSTION OF THE THREE FOOD GROUPS

Nutrient	Enthalpy of combustion, ΔH_c (kJ g^{-1})
Carbohydrate	17
Protein	17
Fat	37

When the percentage composition of a food is known, the calorific value can be calculated using these values. The composition and calorific value of a range of foods is given in table 2.1.2.

TABLE 2.1.2 COMPOSITION AND CALORIFIC VALUE OF SOME FOODS

Food	Carbohydrate (%)	Protein (%)	Fat (%)	Calorific value (kJ g^{-1})
Rice (white)	79	7	Negligible amount	15.2
Bread (wholemeal)	39	11	4	9.7
Roasted peanuts	18	26	50	24.3
Avocado	6	2	17	7.2
Pizza (with cheese)	31	11	8	10.0
Blackberries	13	1	1	2.5
Almonds	19	19	54	25.4

Worked example 3

A sample of unsalted cashews is quoted as having 29% carbohydrate, 18% protein and 46% fat. The remaining 8% is water, which does not supply energy. Find the calorific value of the cashews, in kJ g^{-1}.

Solution

Each enthalpy of combustion (for carbohydrate, protein and fat) is multiplied by the proportion of the food that consists of that food group.

$$E = \frac{(\%\ \text{carbohydrate} \times \Delta H_c\text{carbohydrate}) + (\%\ \text{protein} \times \Delta H_c\text{protein}) + (\%\ \text{fat} \times \Delta H_c\text{fat})}{100}$$

$$= \frac{(29 \times 17) + (18 \times 17) + (46 \times 37)}{100}$$

$$= 25\ \text{kJ g}^{-1}$$

The calorific value of the cashews is 25 kJ g^{-1}.

Section 2.1 Exercises

1 The composition of roasted peanuts and of almonds is given in table 2.1.2.

a State which nut has the greater fat content.

b Using data from table 2.1.1 and table 2.1.2, explain why the calorific value of almonds is greater than that of the roasted peanuts.

2 In an experiment it was found that 100 g of roasted peanuts yielded 2430 kJ of energy. Calculate the number of kJ of energy that was due to the fat in the peanuts.

3 When the sums of the percentages of carbohydrates, proteins and fats in each of the foods in table 2.1.2 are calculated, it is found that they are not equal to 100%. Explain what the rest of the composition of the foods would be.

4 Calculate the calorific value (in kJ g^{-1}) of 250 cm^3 (assuming a mass of 250 g) of full-cream milk, using the data for heat of combustion in table 2.1.1, if it contains 9.7 g fat, 11.4 g of carbohydrate and 8.4 g of protein.

5 A biscuit weighing 5.00 g was combusted in a bomb calorimeter in the presence of excess oxygen. The temperature of the water surrounding the 'bomb' was found to increase from 20.10°C to 24.85°C. If the calorimeter constant was 700 J g^{-1} and there was 500 cm^3 of water in the calorimeter, calculate the heat of combustion of the biscuit, in kJ g^{-1}.

2.2 PROTEINS

Proteins are natural polymers (biopolymers) that are essential to life. All organisms, even the tiniest viruses, contain proteins. Proteins have many different functions in living things. For example, they act as biological catalysts (**enzymes**), they give structure (hair, muscle, feathers and nails), they provide energy and, in some cases, they are hormones.

Amino acids

B.2.1
Draw the general formula of 2-amino acids. © IBO 2007

All proteins contain the elements carbon, hydrogen, oxygen and nitrogen, while many also contain sulfur. Proteins are polymers of **2-amino acids**, the general structure of which is shown in figure 2.2.1. Proteins differ from the man-made polymers discussed in our studies of Organic chemistry in that 20 different amino acid monomers may be used to produce proteins. Each amino acid differs according to its side chain, which is represented in the general structure of the amino acid as 'R' (see table 2.2.1, p. 81). Plants are able to synthesize amino acids from CO_2, H_2O and mineral ions such as NO_3^- and SO_4^{2-}. All 20 of the amino acids are required for protein synthesis, but humans are thought to be able to synthesize only 10 of them. The remaining 10 are called essential amino acids and must be obtained from the diet.

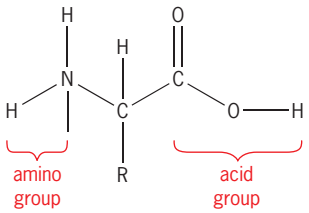


Figure 2.2.1 The general structure of a 2-amino acid.

B.2.2
Describe the characteristic properties of 2-amino acids. © IBO 2007

Notice that in their general structure amino acids contain both an acidic group (–COOH) and a basic group ($–NH_2$). This means that, depending on the surroundings, an amino acid may act as an acid or a base. In an acidic environment, with an accompanying high concentration of H^+ ions, the basic amino group of the amino acid will gain an H^+ ion, forming a positively charged ion. In a basic environment, where the H^+ concentration is low, the carboxyl group of the amino acid will lose an H^+ ion, forming a negatively charged ion. This enables an amino acid solution to exert a buffering action. A buffer solution resists changes in pH when a small amount of an acid or a base is added; however, an amino acid buffer will only occur when the concentrations of the positively charged conjugate base and the original amino acid are approximately equal. That is, the amino acid becomes a buffer once some acid has been added.

In a neutral solution an amino acid may both lose an H^+ ion from the carboxyl group and gain an H^+ on the amino group. Such an ion with both a negative and positive charge on it at the same time is known as a **zwitterion** (from the German, zwei meaning two). The pH at which an amino acid occurs as a zwitterion is called its **isoelectric point**. At this pH the amino acid carries no net electrical charge. One way in which the isoelectric point is put to use is that an amino acid will precipitate out of solution at its isoelectric point. Table 2.2.3 (p. 87) includes the isoelectric points for the 20 different amino acids.

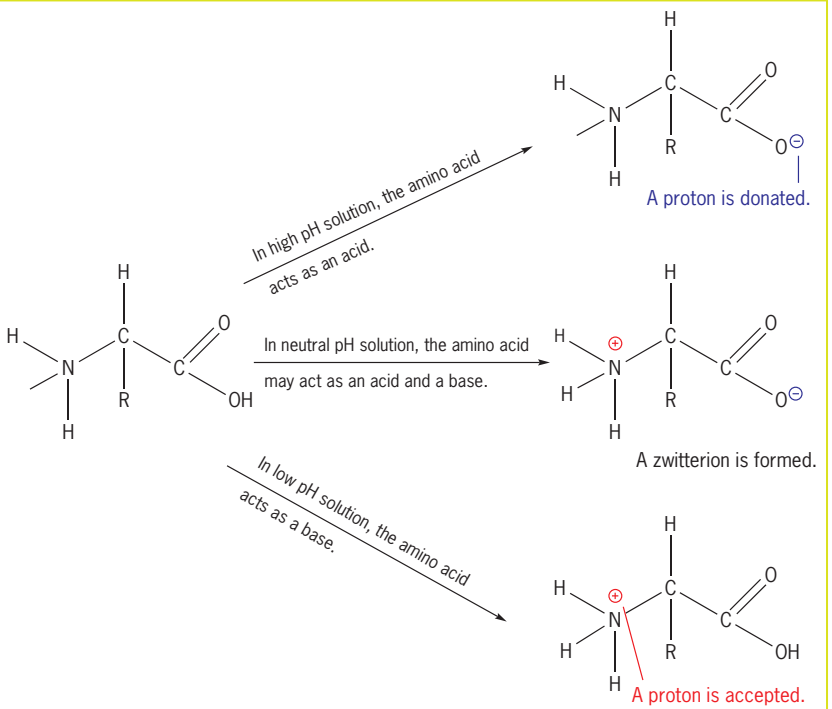
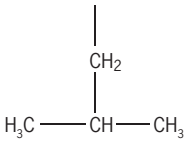
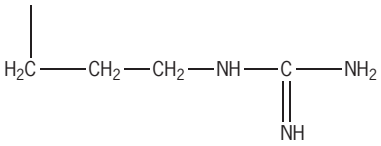
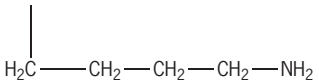
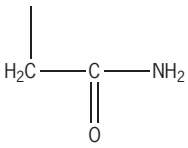
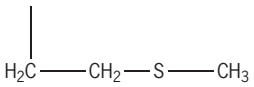
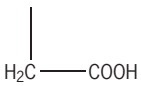
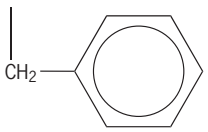
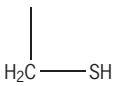
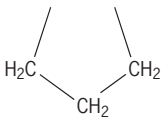
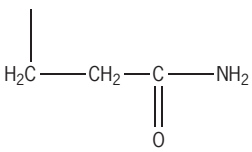
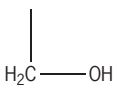
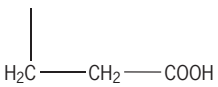
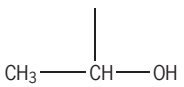
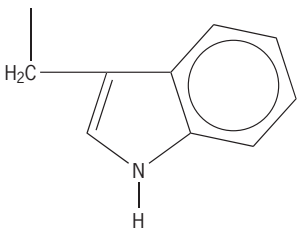
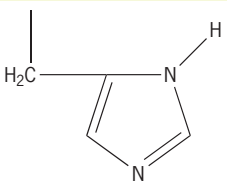
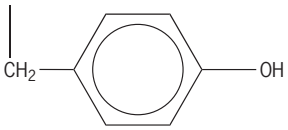
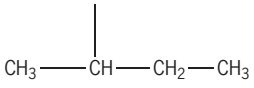
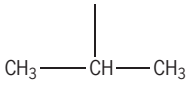


Figure 2.2.2 Amino acids can function as both acids and bases.

TABLE 2.2.1 THE FUNCTIONAL GROUPS THAT DISTINGUISH THE 20 AMINO ACIDS

Common name of amino acid	Symbol	Structural formula of side chain (R group)	Common name of amino acid	Symbol	Structural formula of side chain (R group)
Alanine	Ala		Leucine	Leu	
Arginine	Arg		Lysine	Lys	
Asparagine	Asn		Methionine	Met	
Aspartic acid	Asp		Phenylalanine	Phe	
Cysteine	Cys		Proline	Pro	
Glutamine	Gln		Serine	Ser	
Glutamic acid	Glu		Threonine	Thr	
Glycine	Gly		Tryptophan	Trp	
Histidine	His		Tyrosine	Tyr	
Isoleucine	Ile		Valine	Val	

AS B.2.3
Describe the condensation reaction of 2-amino acids to form polypeptides. © IBO 2007

Amino acids may undergo condensation reactions with other amino acids. A condensation reaction will occur between the $–NH_2$ functional group of one amino acid and the –COOH functional group of another amino acid. A H from the $–NH_2$ and an –OH from the –COOH group join to form a molecule of H_2O and a new link is formed between the two amino acids. One example is shown in figure 2.2.3a. When two amino acids react together, two different products, called **dipeptides**, are possible. The condensation reaction between three amino acids can produce six possible **tripeptides** (see figure 2.2.4). The link formed (–CONH–) is an amide bond, but it may be called a **peptide link** when it joins amino acids. For each amino acid that bonds to another amino acid by a condensation reaction, a water molecule is formed. This means that the molecular mass of a dipeptide is equal to that of the two amino acids minus the molecular mass of water (18.02) and that of a tripeptide is equal to the molecular masses of the three amino acids minus the molecular mass of two water molecules (36.04).

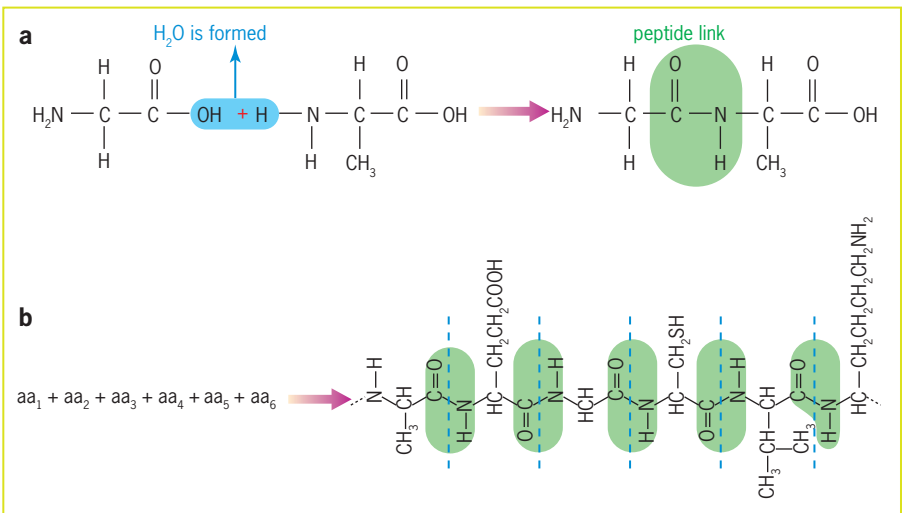


Figure 2.2.3 (a) Two amino acids undergo a condensation reaction to form a dipeptide. (b) Many amino acids undergo condensation reactions to form a polypeptide.

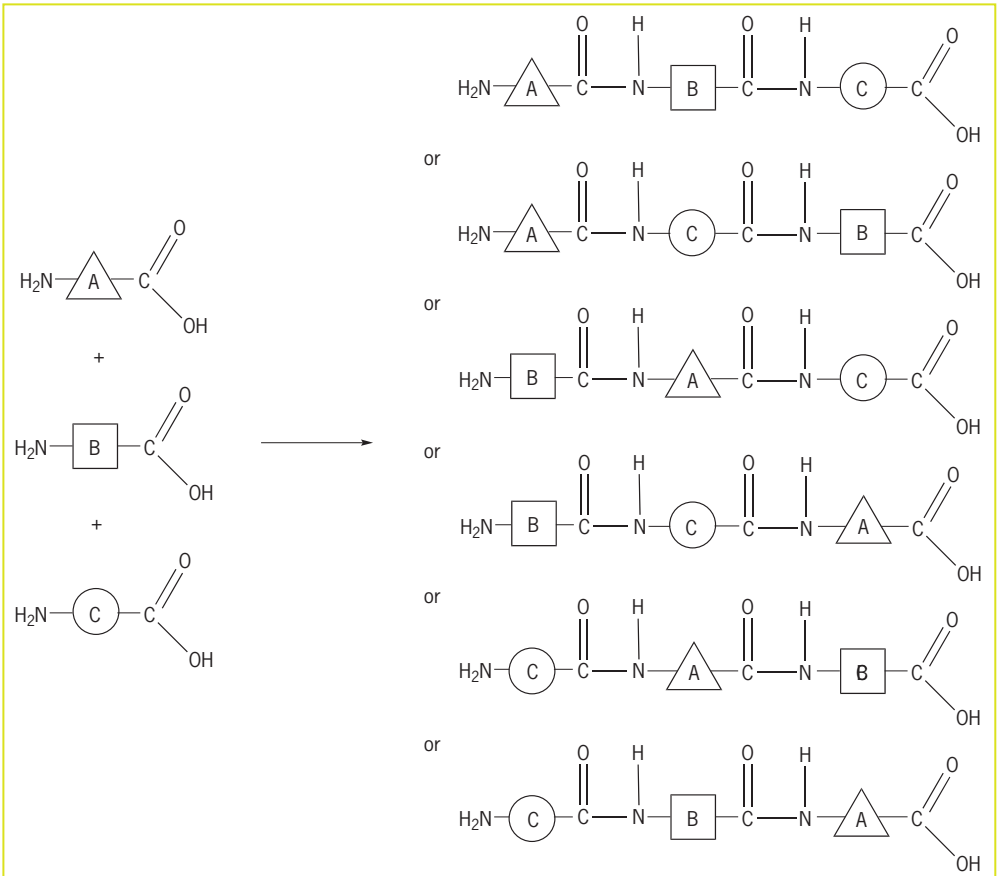


Figure 2.2.4 Three amino acids A, B, and C may produce a possible six different tripeptides.

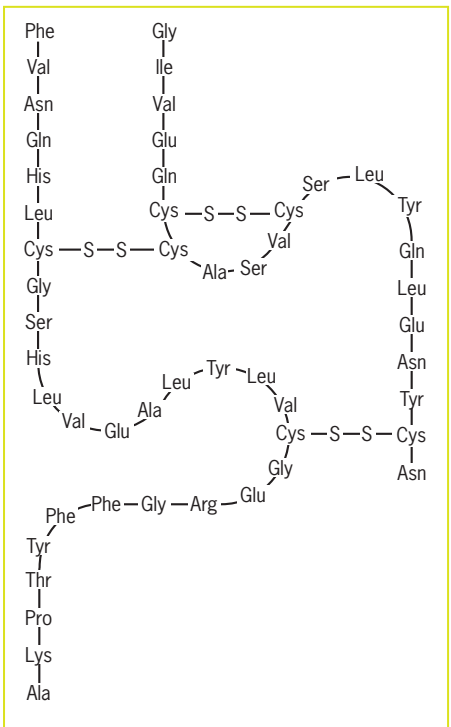


Figure 2.2.5 The amino acid sequence of bovine insulin. This order of particular types and numbers of amino acids represents the primary structure of this protein.

Polypeptides and proteins

The reaction of many amino acids produces a condensation polymer called a **polypeptide**, in which hundreds or even thousands of amino acids are joined. There are many millions of different combinations of sequences of the 20 different amino acids. If there is a large enough number of amino acids involved, the polymer produced is called a *protein* rather than a polypeptide. The actual size distinction is somewhat arbitrary, but the molar mass for a protein is usually above 10 000 g mol^{-1}. Insulin, which is made up of two peptide chains and has a total of 51 amino acids, is considered to be the smallest protein.

Each protein contains a fixed number of amino acids in a particular sequence. This order, number and type of amino acids making up the polymer is called the **primary structure** of the protein. The importance of primary structure can be seen in figure 2.2.4 in which the six possible tripeptides that could be made from just three amino acids are illustrated. The order in which the amino acids bond together makes a significant difference to the final structure of the protein being synthesized.

Interactive CW Animation
Peptide bond

AS **B.2.4**
Describe and explain the primary, secondary (α-helix and β-pleated sheets), tertiary and quaternary structure of proteins. © IBO 2007

CHEM COMPLEMENT

How big is big?

The molecular formula for the protein hemoglobin, which carries oxygen in red blood cells, is $C_{3032}H_{4816}O_{780}N_{780}S_8Fe_4$. This has a molar mass of 64 450 g mol^{-1}. While this may seem large compared with 'ordinary' molecules, the head of a pin could hold about a billion average-sized proteins.

QuickTime CW Video
Proteins and amino acids

Ws
WORKSHEET 2.1
Protein structure and function

The presence of oxygen and nitrogen atoms in a protein results in significant polarity within the polymer chains. Dipole–dipole bonding occurs at regular intervals between parts of the polypeptide chains, resulting in two major structures: α-helices, which are coils of protein, like springs; and β-pleated sheets, which have a corrugated appearance. These helices and sheets are referred to as the **secondary structure** of the protein. Notice that the particular type of dipole–dipole bonding involved in this case is hydrogen bonding.

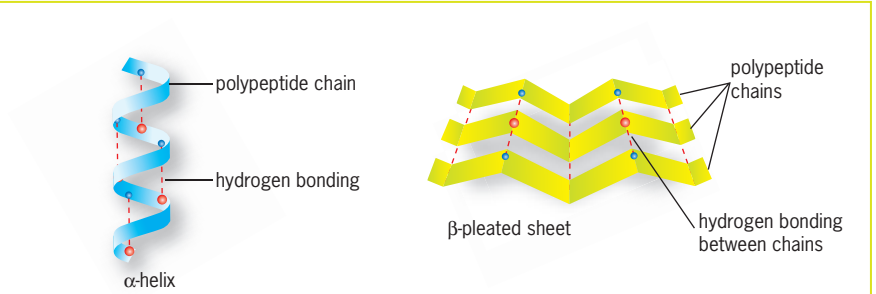


Figure 2.2.6 The secondary structure of a protein may take the form of an α-helix or a β-pleated sheet.

If the coiled protein molecules fold over themselves to form a particular three-dimensional shape, the protein is said to have a **tertiary structure**. The unique, tertiary structure of a protein is responsible for the unique function of that protein. Amino acids from one segment interact with those from another, quite different, section. These interactions may be due to hydrogen bonds, van der Waals' forces between non-polar side groups, ionic attractions between ionized side groups, ion–dipole attractions or covalent bonds (disulfide bridges) that form when sulfur-containing side groups react. An example of disulfide bridges holding a protein in its tertiary structure can be seen in the structure of bovine insulin in figure 2.2.5. Hydrogen bonds may occur between the atoms of two amide groups, the atoms of a peptide bond and an amino acid side chain or even two amino acid side chains. In the folding of a tertiary structure those amino acid residues with hydrophobic (non-polar) side chains (such as glycine, alanine and isoleucine) will be folded into the inner part of the protein so that the hydrophilic (polar or ionic) side chains can form hydrogen bonds with water and enable the protein to be miscible with the mainly aqueous environment in which it operates.

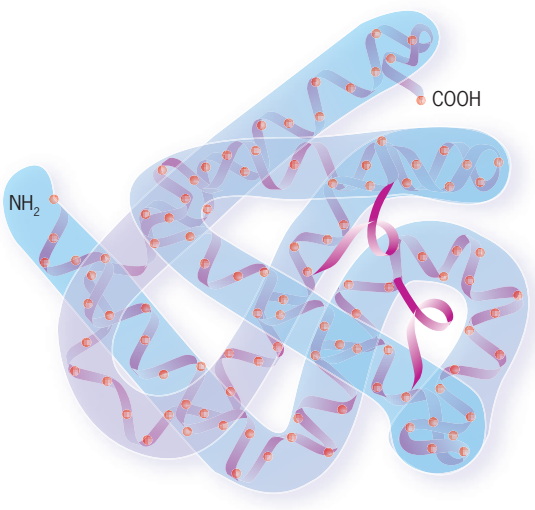


Figure 2.2.7 The tertiary structure of myoglobin. The twisted polypeptide chain is folded into a three-dimensional structure.

Separate polypeptide chains may interact with each other to give further complexity to the structure. This is known as the **quaternary structure**. This is the overall arrangement of the polypeptide chain to form the 'working shape' of the protein. Hemoglobin (figure 2.2.8) has a quaternary structure that involves four polypeptide chains grouped together around four iron ions.

CHEM COMPLEMENT

Hemoglobin—the protein of life

Red blood cells, or erythrocytes, are manufactured by the bone marrow at a rate of about 2 million per second and live for about 3 months in the body. Each red blood cell contains approximately 250 million molecules of hemoglobin. The hemoglobin molecule consists of two parts: the iron-containing pigment called heme, and the protein section called globin. The protein chains are coiled into α-helices and they are folded in the same way as other globin proteins such as myoglobin. Hemoglobin is a tetramer that consists of two of each of two types of protein subunits—with 141 and 146 amino acid residues respectively. The structure of the heme group is complex (see figure 2.9.1, p. 132), but contains four nitrogen atoms that together form a bond with an iron atom. The globin proteins wrap around the heme group and protect it from being oxidized (and destroyed) by the oxygen it is intended to transport. It is to the iron of the heme group that an oxygen molecule is attached, so each hemoglobin molecule is able to carry four oxygen molecules. In this form, the molecule is the characteristic bright red colour of oxygenated blood, and is called oxyhemoglobin.

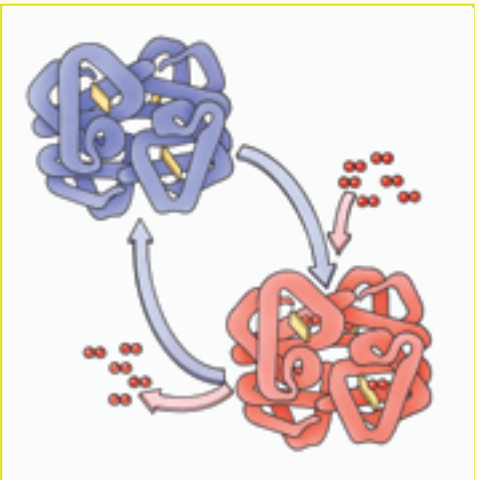


Figure 2.2.8 Hemoglobin (blue) combines with oxygen to form oxyhemoglobin (red).

Physical and chemical agents may destroy these three-dimensional protein shapes, changing the nature of the protein. Such changes are called **denaturation**. For example, heat breaks hydrogen bonds, so heating a protein strongly destroys the helical structure. In some proteins, heat causes the unfolding of the polypeptide chains, followed by a more random re-forming of bonds, resulting in precipitation or coagulation. This is what happens when an egg is boiled. Ironing of silk clothes also breaks the hydrogen bonds in the helical structures, therefore getting rid of the wrinkles in the silk. Similar changes can occur by mechanical treatment; for example, whipping an egg white to make a meringue. Addition of chemicals may also denature proteins. Detergents open hydrophobic (non-polar) regions of the polymer, while specific reducing agents break the disulfide bridges. Acids and bases may affect ionic bonding between side groups, while some metal ions interact with –SH side groups.

Proteins perform a wide range of functions in the body. They are used for structure, movement, transport, catalysis of chemical reactions, energy transformation and storage, protection, control and buffering. Most of the energy supplied to cells is used in making proteins and allowing them to perform their functions. Proteins may be identified according to their structural type, which largely controls their function.

AS **B.2.6 List the major functions of proteins in the body.**

Structural proteins tend to be insoluble in aqueous solution, are long and fibrous, even rope-like and include muscle, hair, tendons and ligaments.

Globular proteins are basically spherical and are water-soluble. These include hormones, enzymes, antibodies and other transport proteins such as hemoglobin.

TABLE 2.2.2 THE RANGE OF FUNCTIONS THAT DIFFERENT PROTEINS PERFORM

Type of protein	Function	Example	Explanation of functions
Structural	Structure	Collagen	• Connective tissue in skin—most abundant protein in the body (25%) • Major component of cartilage, ligaments, tendons, bone and teeth
		Keratin	The main constituent of structures that grow from the skin such as hair, nails and claws and beaks (in animals)
Enzymes	Catalysis	Amylase	Breaks down starch into maltose in saliva
		Catalase	Catalyses the decomposition of hydrogen peroxide to water and oxygen
Hormones	Control	Insulin	A chemical messenger that causes liver and muscle cells to take in glucose and store it in the form of glycogen
Energy transformation proteins	Energy transformation	Rhodopsin	Found in the retina of the eye. Traps light energy and converts it to electrical energy of a nerve impulse
Immunoproteins	Protection	Antibodies	Used by the immune system to identify and neutralize foreign objects, such as bacteria and viruses
Transport proteins	Transport	Hemoglobin	Transport of oxygen in the blood from the lungs to cells
Energy source	Supply of energy in extreme situations	Muscular protein	Will only be used if carbohydrates and fats are in low supply in the body.

Analysis of proteins by electrophoresis and chromatography

AS **B.2.5 Explain how proteins can be analysed by chromatography and electrophoresis.**

The primary structure of a protein relates to the number of amino acids, the identity of those amino acids and the arrangement of the amino acids in the protein chain. Part of this information can be obtained using electrophoresis or chromatography.

Electrophoresis is an analytical technique that separates charged substances according to the different rates at which they move (depending on their size and electrical charge) when subjected to a potential difference. In preparing a protein sample for electrophoresis, the protein is denatured and hydrolysed, when it is heated in a solution of hydrochloric acid for a period of time. This separates the amino acids from each other. In PAGE (**p**oly**a**crylamide **g**el **e**lectrophoresis) a tiny sample of protein solution is placed into a well in the centre of a polyacrylamide gel. The gel is moistened with buffer solution and a voltage is applied. Depending on the pH of the buffer solution, the amino acids separate according to their charge and their mass.

Let us consider a mixture of amino acids being separated by PAGE. If the pH of the buffer solution is equal to the isoelectric point of one of the amino acids in the mixture, that amino acid will not be charged in the buffer solution and it will not move beyond the starting position. Amino acids with isoelectric points that are greater than the pH of the buffer solution will be positively charged and will migrate towards the negative terminal. The greater the mass of the amino acid, the slower will be its progress through the gel. Those amino acids

with isoelectric points that are less than the pH of the buffer solution will have a negative charge and will migrate towards the positive terminal, with the smaller amino acids moving more quickly than the larger ones. The gel may be 'developed' by spraying with **ninhydrin**, an organic dye that changes the colour of the amino acids and the amino acids will appear as bands along the length of the gel. The identity of these bands can be determined by comparison with standards—known amino acids that have been run under identical conditions—or have known isoelectric points such as those listed in table 2.2.3.

TABLE 2.2.3 ISOELECTRIC POINTS AND MOLAR MASSES OF AMINO ACIDS

Common name of amino acid	Symbol	Isoelectric point	Molar mass ($g\ mol^{-1}$)	Common name of amino acid	Symbol	Isoelectric point	Molar mass ($g\ mol^{-1}$)
Alanine	Ala	6.0	89	Leucine	Leu	6.0	131
Arginine	Arg	10.8	174	Lysine	Lys	9.7	146
Asparagine	Asn	5.4	132	Methionine	Met	5.7	149
Aspartic acid	Asp	2.8	133	Phenylalanine	Phe	5.5	165
Cysteine	Cys	5.1	121	Proline	Pro	6.3	115
Glutamine	Gln	5.7	146	Serine	Ser	5.7	105
Glutamic acid	Glu	3.2	147	Threonine	Thr	5.6	119
Glycine	Gly	6.0	75	Tryptophan	Trp	5.9	204
Histidine	His	7.6	155	Tyrosine	Tyr	5.7	181
Isoleucine	Ile	6.0	131	Valine	Val	6.0	117

Worked example 1

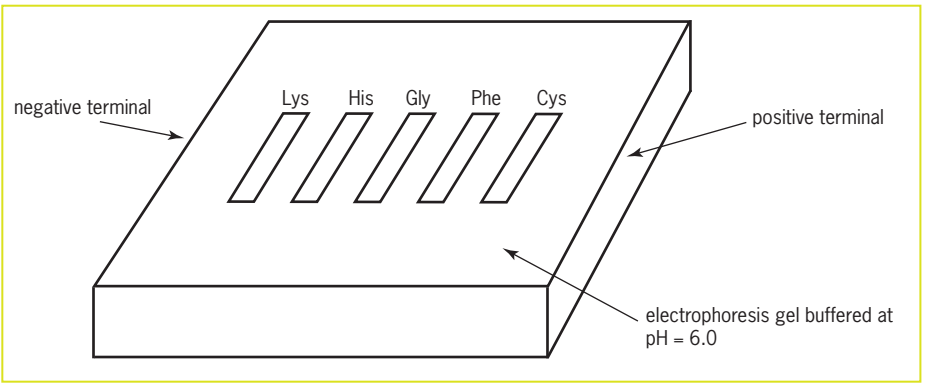
A mixture of five amino acids (glycine, cysteine, lysine, phenylalanine and histidine) is separated by gel electrophoresis in a buffer solution of pH = 6.0. Draw the finished gel after the amino acids have been separated.

Solution

Amino acid	Isoelectric point	Molecular mass	Charge on amino acid in buffer of pH = 6.0
Cysteine	5.1	121	Negative
Glycine	6.0	75	Neutral
Histidine	7.6	155	Positive
Lysine	9.7	146	Positive
Phenylalanine	5.5	165	Negative

- The isoelectric point of glycine is equal to the pH of the buffer solution. Glycine will not be charged and so will not move from the starting position in the centre of the gel.
- Cysteine and phenylalanine will migrate towards the positive terminal. Cysteine has a lower molecular mass than phenylalanine, so it will have migrated at a greater rate than phenylalanine and will be further along the gel.

- With a positive charge, histidine and lysine will migrate towards the negative terminal. Lysine has a lower molecular mass than histidine, so it will have migrated at a greater rate and will be closer to the negative terminal.



Proteins can also be analysed by gel electrophoresis without being hydrolysed into individual amino acids. The technique known as SDS–PAGE involves the denaturing of the proteins in the presence of a detergent called sodium dodecyl sulfate (SDS) that coats the proteins with a negative charge. The protein mixture is placed at one end of the gel and a potential difference is applied across the gel. The proteins migrate towards the positive terminal and separate according to their molecular masses, with the lower molecular mass proteins moving at a greater rate. Marker proteins are run as standards to calibrate the process. These proteins have known molecular masses, so the unknowns can be compared to them. A range of different stains may be used to show the positions of the proteins on the completed gel, with silver staining and a stain called Coomassie blue being found to be accurate for low concentrations of proteins. In addition the composition of the polyacrylamide gel can be changed to give greater resolution for higher or lower molecular mass proteins.

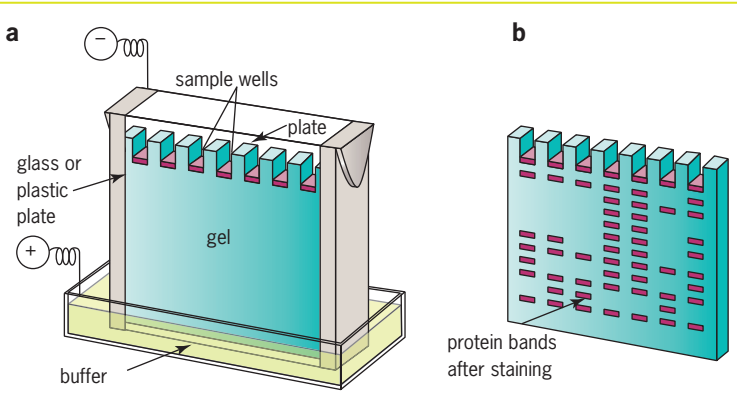


Figure 2.2.9 Proteins that have been separated using gel electrophoresis, produce bands on the electrophoresis gel.

Chromatography separates components of a mixture according to their relative attraction to a stationary phase and a mobile phase. In paper chromatography, the stationary phase is a strip of filter paper. This is particularly suitable for separating amino acids quickly and cheaply, although thin-layer chromatography can also be used. To release the amino acids from the protein chain, the peptide bonds must be hydrolysed. The solution of amino acids is spotted onto the paper which is then placed in a developing tank containing the mobile phase which will be a solvent or a mixture of solvents. When the mobile phase has moved almost the full length of the paper, the different amino acids will have moved different distances due to the relative attraction of their different side groups to the stationary and mobile phases. Since the amino acids are colourless, their positions on the chromatogram cannot be distinguished until the finished chromatogram is sprayed with a solution of ninhydrin. This gives most amino acids a purple colour (proline turns yellow). The amino acids that have separated from the mixture can be identified by comparing their positions on the chromatogram with those of pure samples of amino acids that have been run on a chromatogram under the same conditions. The calculation of a **retention factor (R_f)** for each amino acid will give a more definite identification. The R_f is determined by measuring the distances moved by each amino acid and by the solvent:

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

QuickTime CW Video
Paper chromatography of inks

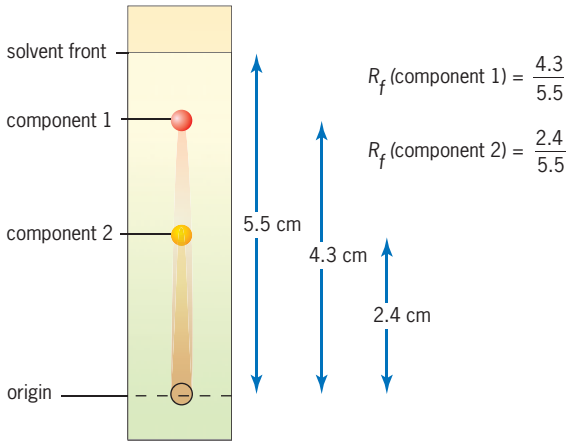


Figure 2.2.10 Calculating R_f values.

A more sophisticated chromatographic method for the separation of proteins, such as the proteins in milk and those found in blood, is **high performance liquid chromatography (HPLC)**. In this technique the mixture in solution is swept along a column containing the stationary phase under high pressure. Separation of the components of the mixture occurs due to their relative attractions to the stationary and mobile phases. The components are identified by a UV detector since these colourless compounds absorb in the ultraviolet region of the electromagnetic spectrum.

Ws
WORKSHEET 2.2
Amino acid analysis

Section 2.2 Exercises

1 Draw the general structure of an amino acid, labelling the functional groups that are common to all amino acids.

2 **a** Describe what is meant by the term *isoelectric point*.

b Distinguish between a zwitterion and an amino acid with no charge.

3 **a** Using table 2.2.1 to identify the correct functional groups, draw a possible structure for the amino acid with molecular formula:

i $C_3H_7O_2N$

ii $C_5H_9O_4N$

iii $C_4H_8O_3N_2$

b Draw the structure of the amino acid alanine ($H_2NCH(CH_3)COOH$) in a solution of:

i pH = 2

ii pH = 11

4 Amino acids combine in condensation reactions to form dipeptides and polypeptides.

a State the name of the functional group formed by the condensation reaction between two amino acids.

b Draw the structure of a dipeptide formed when alanine and glycine react together and circle the functional group that joins the two amino acid residues.

c State the name of the other product formed (other than a dipeptide) when alanine and glycine react together.

5 The side groups of three amino acids are shown in the table below.

Abbreviation	Amino acid	Side group
Ser	Serine	$-CH_2OH$
Val	Valine	$-CH(CH_3)_2$
Cys	Cysteine	$-CH_2SH$

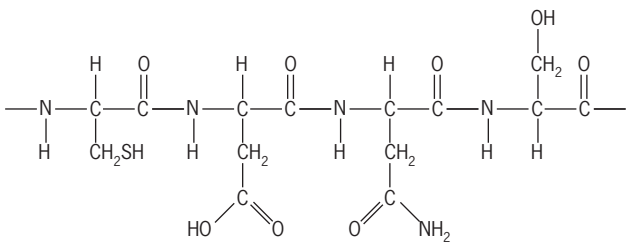
a Draw the molecular structures of two possible products of the condensation reaction between serine and valine.

b The notation SerCysVal represents a tripeptide formed from serine, cysteine and valine.

i Draw a possible molecular structure for CysSerVal.

ii State how many structures are possible for tripeptides formed from serine, cysteine and valine.

6 The diagram below shows a segment of a protein.



a State how many amino acids were used to produce this segment.

b State how many peptide (amide) links are included.

7 Complete the following table, which links the structure of a protein to the bonding involved.

Protein structure	Bonding or intermolecular forces involved in the structure
Primary	
Secondary	
Tertiary	

8 Describe how the amino acids present in a protein can be separated using paper chromatography.

9 a Explain why amino acids move along an electrophoresis gel.

b Explain how the different isoelectric points of amino acids can affect their movement during electrophoresis.

c Explain how the different molar masses of amino acids can affect their movement during electrophoresis.

d A mixture of aspartic acid, glutamic acid, valine, lysine and tyrosine were separated using gel electrophoresis at pH 5.7. Draw the finished electrophoresis gel, showing the final positions of the five amino acids.

10 List the major functions of proteins in the body, giving one example of a protein that performs each function.

2.3 CARBOHYDRATES

The name *carbohydrate* is derived from the observation that many members of this group have the empirical formula $C_x(H_2O)_y$, where x and y are whole numbers. Examples include the sugar found in grapes (glucose, $C_6H_{12}O_6$) and cane sugar (sucrose, $C_{12}H_{22}O_{11}$). Carbohydrates are often called saccharides (from the Latin, *saccharum*, for sugar) because of the sweet taste of these simple members of the group. Carbohydrates are among the most abundant components of living things. They serve several functions: as an energy source and storage (starch, glycogen and glucose), as a structural material (cellulose) and as an essential component of the genetic material (ribose and deoxyribose in nucleic acids).

AS **B.3.1 Describe the structural features of monosaccharides. © IBO 2007**

Glucose is one of the simplest of carbohydrate types, known as **monosaccharides** (simple sugars). All monosaccharides have the empirical formula CH_2O. They contain a carbonyl (C=O) group and have at least two hydroxyl (–OH) groups. Fructose, found in fruits and honey, is also a monosaccharide, as is galactose. All are white, crystalline solids with a sweet taste, and all have the formula $C_6H_{12}O_6$. They are structural isomers of each other and are known as hexoses because they have six carbon atoms in their formula.

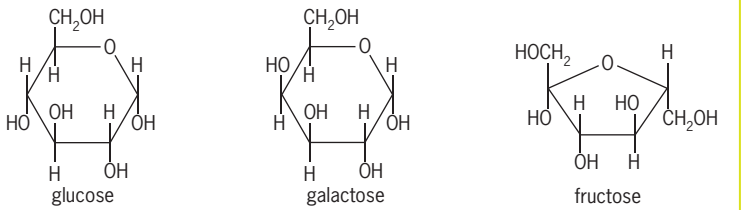


Figure 2.3.1 Structures of the monosaccharides glucose, galactose and fructose.

Note that most of the carbon atoms are not drawn in these simplified structures. In each case, a carbon atom is to be found where four lines intersect; that is, at the vertices of the hexagon or pentagon.

AS **B.3.2 Draw the straight-chain and ring structural formulas of glucose and fructose. © IBO 2007**

In solution, three isomers of monosaccharides are in equilibrium—two with ring structures and a straight-chain molecule.

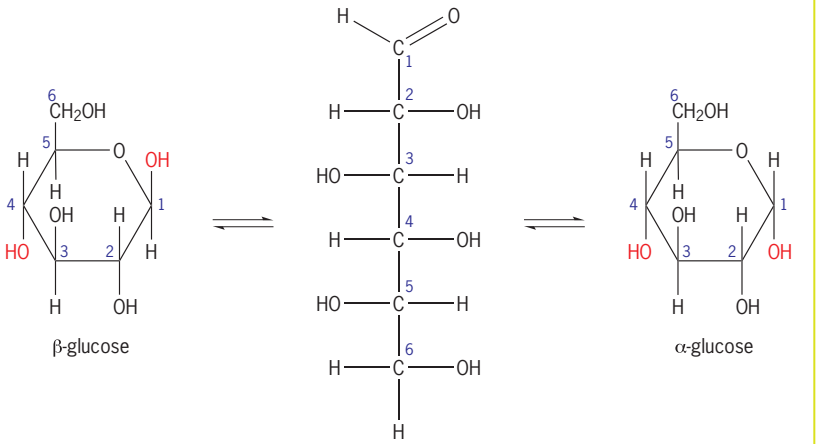


Figure 2.3.2 Glucose exists in solution in equilibrium between three isomers.

PRAC 2.2
Identifying sugars

While the ring structures of glucose appear to be very similar, there is a fundamental difference between the two. When the orientation of the hydroxyl groups on carbons numbered 1 and 4 are compared, the difference between α-glucose and β-glucose becomes apparent. In α-glucose, both hydroxyl groups are pointing 'downwards', whereas in β-glucose the hydroxyl group on carbon 1 is pointing 'upwards' and that on carbon 4 is pointing 'downwards'. This difference is most important when the polymers of glucose—cellulose and starch are made. While free rotation around all bonds except the C=O bond in the linear glucose molecule is possible, the hydroxyl groups are fixed in their positions (pointing up or down) once the ring structure is formed.

AS B.3.3
Describe the condensation of monosaccharides to form disaccharides and polysaccharides. © IBO 2007

Disaccharides (di- meaning two) form by condensation reactions between two monosaccharides. Like the condensation reaction between amino acids, the combination of a H atom from one monosaccharide and an –OH group from the other monosaccharide results in the formation of a water molecule. The linkage formed between the two monomers is a **glycosidic** or **ether linkage**, and the product formula is $C_{12}H_{22}O_{11}$. Sucrose, the most common disaccharide, is added as a sweetener to foods. The condensation reaction between α-glucose and β-fructose produces sucrose. Lactose is present in the milk of mammals, making milk an energy source for the young. It is less sweet than sucrose, possibly a design of nature to protect the taste buds of the young. The condensation reaction between α-glucose and β-galactose produces lactose. Notice the very slight difference in structure between glucose and galactose, in the arrangement of the atoms around carbon number 4. Maltose is found in grains, particularly in barley. Barley is the source of maltose for the brewing of beer. Amylose, a polymer of maltose is hydrolysed to produce maltose which is fermented to produce ethanol and carbon dioxide. The condensation reaction between two α-glucose molecules produces maltose.

Interactive CW Animation
Carbohydrate structures

The condensation reaction between monosaccharides does not necessarily stop with the disaccharide. The reaction may continue on to produce a polymer, a **polysaccharide** containing thousands of glucose units. These polysaccharides are less soluble than the smaller saccharides and do not taste sweet.

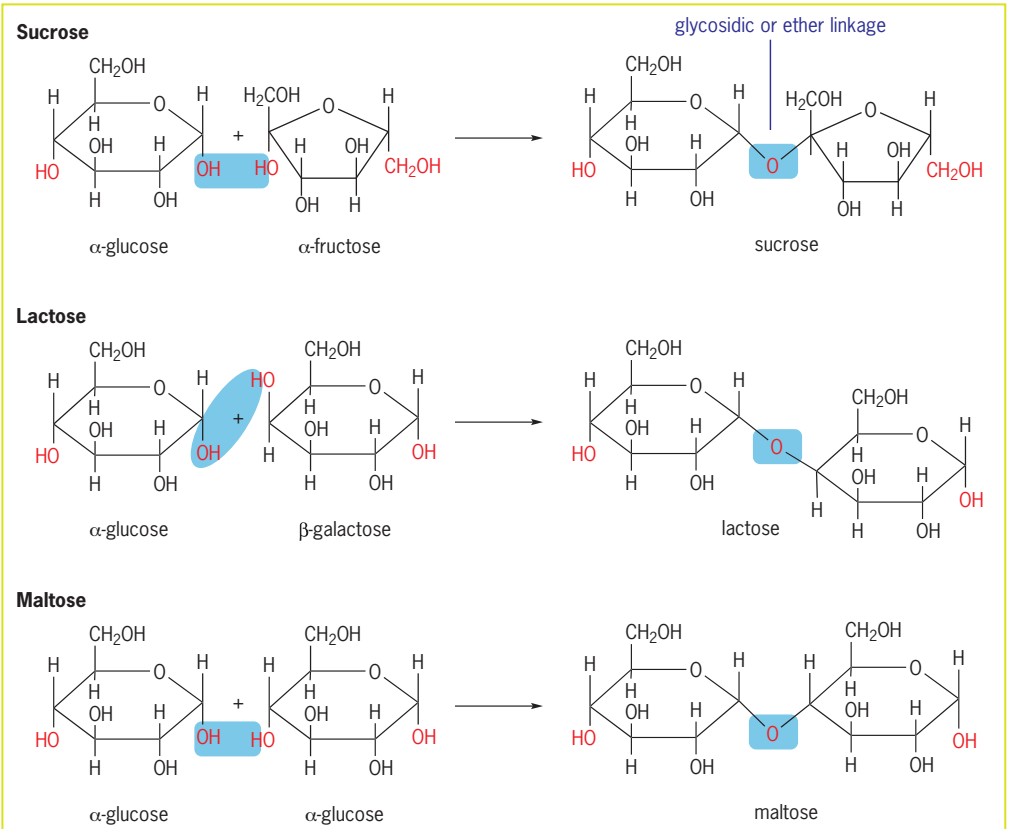


Figure 2.3.3 Condensation reactions form sucrose, lactose and maltose.

Polymerization of glucose in living things can produce three different polysaccharides: starch, cellulose and glycogen. How can the same monomer produce different polymers? The answer lies in the two forms of glucose: α-glucose and β-glucose. The different combinations of these isomers lead to different polymer structures, one of which is **starch**, a food source for animals and a food store for plants. Foods such as potato and sago are well known for their high starch content. When α-glucose is polymerized by plants, the product is starch.

There are two forms of starch: **amylose** and **amylopectin**. Amylose is a straight-chain polymer in which the α-glucose units are linked only at carbons 1 and 4. This is know as an α-1,4 linkage. Amylose chains take on a helical structure. It is these helices that make the common test for starch using iodine possible. The iodine molecules fit neatly inside the helical structure of amylose and bind with it. As a result, the starch–iodine complex absorbs visible light of a different wavelength to iodine alone and a blue-black colour will be observed when amylose is present.

CHEM COMPLEMENT

Polysaccharides in drugs

Polysaccharides are often used as 'fillers' in drugs. While fillers are inactive as far as the drug is concerned, they may affect the rate of drug delivery. Starch, for example, is often added to pills and capsules for oral delivery. When swallowed with water, rapid uptake of water by the starch causes the tablet to swell and fall apart, allowing the drug to be delivered in the stomach and the intestines.

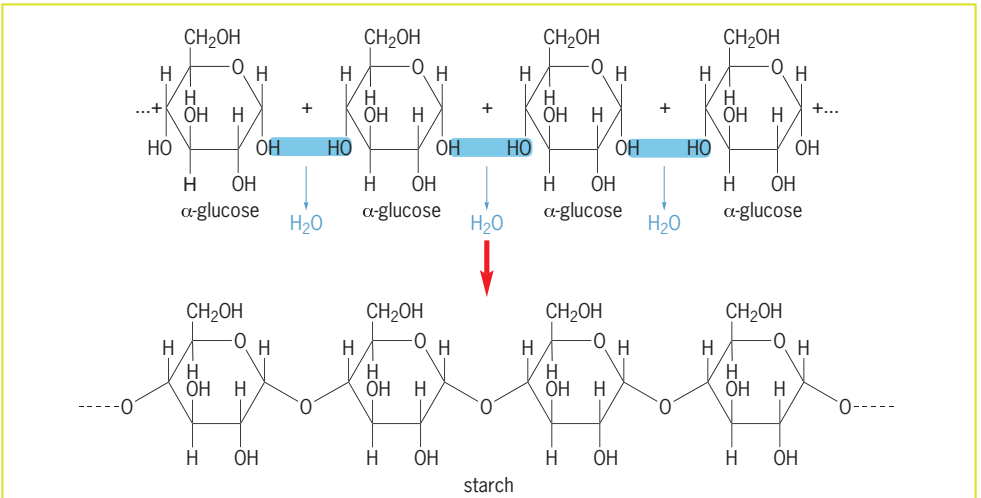


Figure 2.3.4 The condensation polymerization of α-glucose to produce starch (amylose).

The second and more abundant form of starch, amylopectin, is a branched structure with both α-1,4 and α-1,6 linkages. Carbon number 6 is the one carbon that is not part of the ring, and it is a condensation reaction between the hydroxyl group on carbon number 6 and a hydroxyl group on carbon number 1 of another glucose molecule that starts a new branch. Branches in the amylopectin chain are not frequent; they occur every 24–30 glucose units.

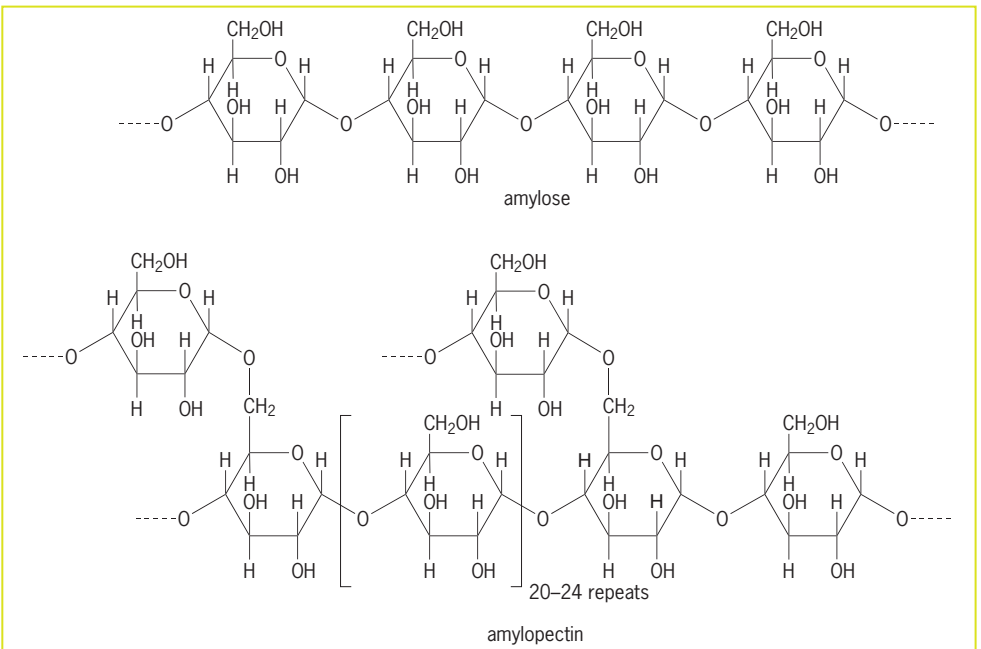


Figure 2.3.5 Two forms of starch: amylose, a straight chain polymer, and amylopectin, a branched polymer.

A second polysaccharide, **cellulose**, is the most abundant molecule in living tissue, making up around 50 per cent of the total organic carbon in the biosphere. It is a structural polysaccharide that is found in the walls of plant cells. Cotton is almost pure cellulose, while wood is about 50 per cent cellulose. The cellulose polymer is a straight-chain molecule, with these molecules acting rather like stiff rods. Extensive hydrogen bonding between the molecules produces a strong material. Cellulose has a β-1,4 linkage; that is, carbon number 1 of a β-glucose forms a glycoside (ether) linkage with the carbon 4 of a β-glucose.

AS B.3.5
Compare the structural properties of starch and cellulose, and explain why humans can digest starch but not cellulose.

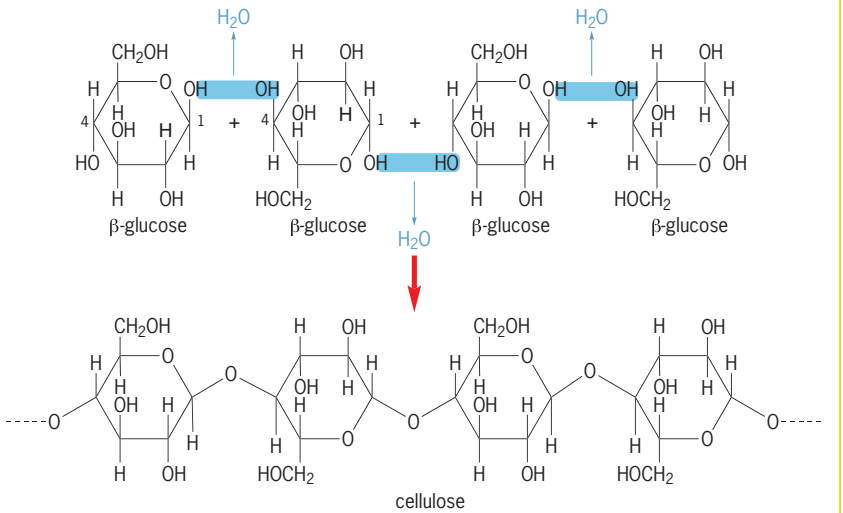


Figure 2.3.6 Condensation polymerization of glucose to produce cellulose. Note that the β-glucose molecules are inverted to show the formation of the glycosidic linkages.

Digestion of polysaccharides involves the **hydrolysis** (adding water) of the bonds between the monosaccharide residues. Enzymes catalyse these reactions in the digestive tracts of animals, including humans; however, humans and most other animals lack the enzyme cellulase, which is required to hydrolyse the β-1,4 linkage, so cellulose cannot be digested. Animals such as sheep, cows and the Australian native koala can digest cellulose because their stomachs and other digestive system parts contain bacteria that provide the necessary enzyme for cellulose breakdown.

Cellulose forms a significant component of human **dietary fibre**—plant material that is not hydrolysed by enzymes secreted by the human digestive tract but may be digested by microflora (bacteria) in the gut. Other examples of dietary fibre include hemicellulose, lignin and pectin.

AS B.3.6
State what is meant by the term *dietary fibre*.

A diet high in cellulose (dietary fibre) provides 'bulk' to aid the passage of food through the digestive system. Foods that are high in dietary fibre include whole grain foods, fruits and vegetables. Such 'bulk' helps prevent constipation and hemorrhoids; reduces the calorific (energy) intake of a diet, thus helping to prevent obesity and diabetes mellitus (type 2 diabetes); and helps reduce the risk of diseases such as colon cancer, diverticulosis, and irritable bowel syndrome.

AS B.3.7
Describe the importance of a diet high in dietary fibre.

The third polysaccharide formed by the polymerization of α-glucose is **glycogen**, a highly branched polymer with a similar structure to amylopectin, but with more branching. Animals polymerize α-glucose to form glycogen, storing some excess energy in this form. Glycogen is stored in muscle and liver, and serves as a ready, short-term store of energy. The branched structure of glycogen makes it a more readily available source of energy than starch, since it can be hydrolysed more rapidly than the long chains of starch.

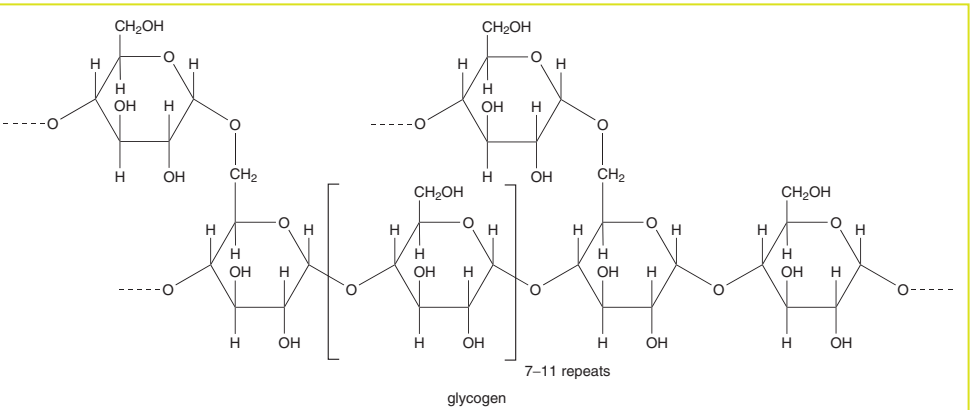


Figure 2.3.7 Glycogen is a polymer of α-glucose which is produced by the liver and to a lesser extent, muscles.

The reactions of carbohydrates are summarized in figure 2.3.8 and table 2.3.1.

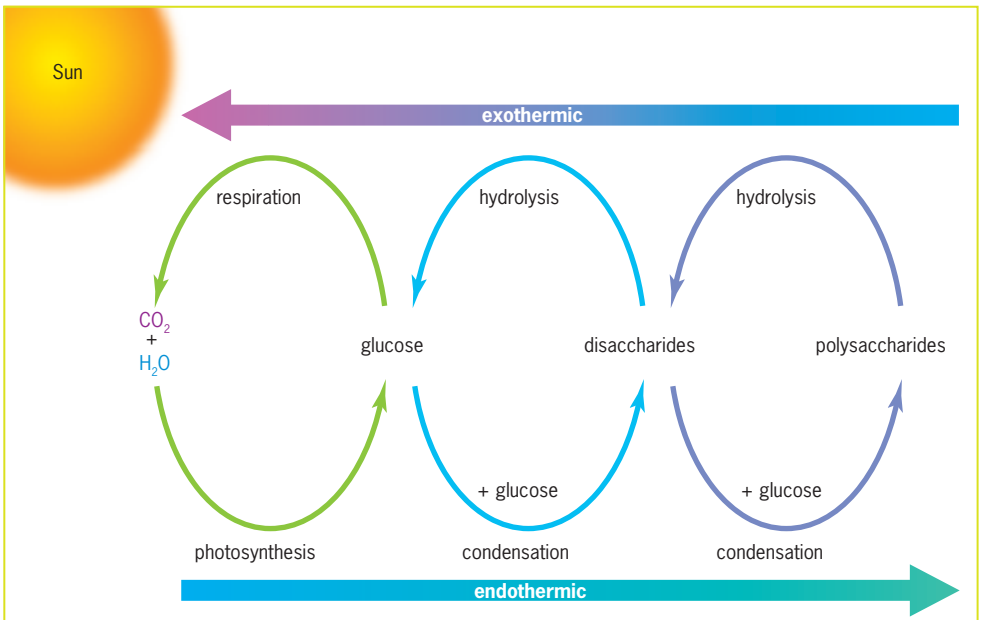


Figure 2.3.8 A summary of reactions involving glucose.

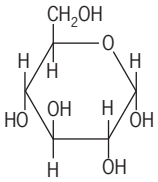
AS B.3.4
List the major functions of carbohydrates in the human body. © IBO 2007

TABLE 2.3.1 MAJOR FUNCTIONS OF CARBOHYDRATES IN THE HUMAN BODY

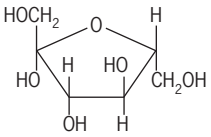
Function	Type of carbohydrate	Source of carbohydrate
Major source of energy for cellular respiration	Glucose	• Starch in plant-based foods • Glycogen in animal-based foods
Formation of energy reserves	Glycogen	Glucose in body not required for immediate energy needs.
Precursors of other biologically important molecules (e.g. heparin, an anticoagulant)	Glucose	• Starch in plant-based foods • Glycogen in animal-based foods

Section 2.3 Exercises

1 Shown below are the molecules of the structural isomers glucose and fructose.



glucose



fructose

a Compare the structures of glucose and fructose, taking particular note of functional groups present in the two molecules.

b State the meaning of the term *structural isomers*.

c Draw the disaccharide formed if these two molecules combine with each other.

d Deduce how the mass of the disaccharide will compare with that of the two separate monosaccharide molecules.

2 Draw the straight-chain structural formulas of glucose and fructose.

3 Describe the difference between the α and β isomers of glucose.

4 **a** State the name of the type of reaction that occurs between two glucose molecules to form the disaccharide maltose.

b State the molecular formula of the other product that is formed in this reaction.

5 **a** State the names of four polymers formed from glucose.

b Describe the different functions that each of these polymers has in living things.

6 Classify each of the substances named below as monosaccharides, disaccharides or polysaccharides.

a Amylopectin
b Cellulose
c Sucrose
d Glucose
e Fructose
f Lactose
g Glycogen
h Amylose

7 Compare amylopectin with amylose.

8 Considering both the structure and the ability of humans to digest these two polysaccharides, compare starch with cellulose.

9 Dietary fibre is considered to be an important part of a healthy diet.

a State what is meant by the term *dietary fibre*.

b Explain the role of dietary fibre in a healthy diet.

c Describe how a lack of dietary fibre may adversely affect the health of a human being.

10 List the major functions of carbohydrates in the human body, giving one example of a carbohydrate that performs each function.

2.4 LIPIDS

AS **B.4.1 Compare the composition of the three types of lipids found in the human body.** © IBO 2007

Oils and fats are part of a class of molecules called **lipids**. Fats and oils are **triglycerides** (see figure 2.4.5), while other members including **phospholipids** and **steroids** have different structures from the triglycerides. Triglycerides are generally very large, non-polar molecules that are insoluble in water. A triglyceride is formed by the condensation reaction between three **fatty acid** molecules (carboxylic acids with very long hydrocarbon chains) and one **glycerol** molecule. It therefore contains three ester linkages (see figure 2.4.5). However, not all lipids are hydrophobic. Phospholipids, together with **glycolipids**, cholesterol and proteins are the major components of all biological membranes.

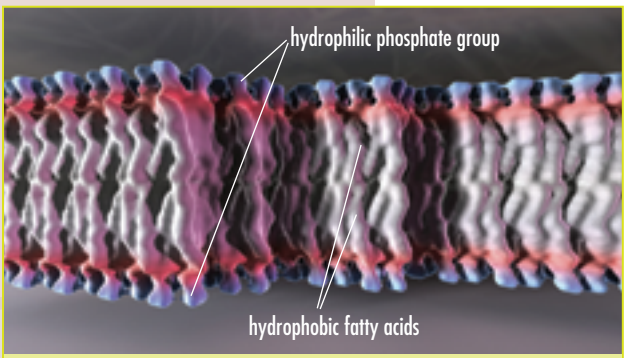


Figure 2.4.1 A phospholipid bilayer such as that found in cell membranes.

A phospholipid molecule has a hydrophilic (polar) head that consists of a phosphate group and hydrophobic (or **lipophilic**) tail made up of two long hydrocarbon chains. In a membrane, twin layers (a bilayer) of phospholipids form, with the non-polar tails lining up against one another, forming a membrane with hydrophilic heads on both sides facing the aqueous surroundings. Lecithin (phosphatidylcholine) is a phospholipid that can be isolated from egg yolk. (The Greek word for egg yolk is *lekithos*.)

Cholesterol is a **sterol**, an alcohol with a fused ring system. This substructure is also found in steroid hormones such as testosterone and progesterone (see figure 2.6.2, p. 113). Cholesterol is classified as an alcohol because it has a hydroxyl group (–OH) in position 3 of the ring system. Cholesterol is produced by the liver and is found in all body tissues, where it helps to control the permeability of cell membranes. Cholesterol derivatives in the skin are converted to vitamin D when the skin is exposed to sunlight. Steroids are primarily made up of carbon and hydrogen with a small amount of oxygen.

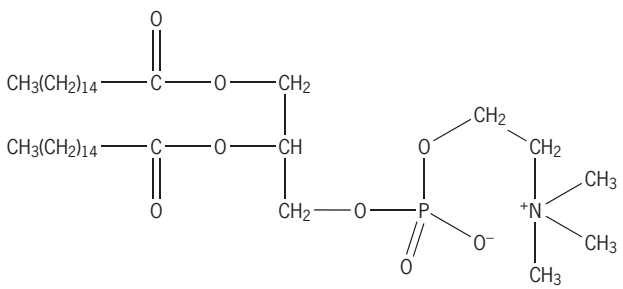


Figure 2.4.2 The structure of lecithin.

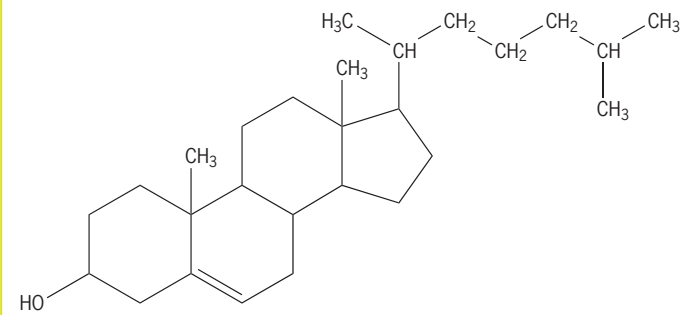


Figure 2.4.3 Cholesterol is made up of four fused six- and five-membered rings with a hydroxyl group on the first ring and a short hydrocarbon chain on the fourth ring.

AS **B.4.2 Outline the difference between HDL and LDL cholesterol and outline its importance.** © IBO 2007

Since cholesterol is insoluble in blood, it is transported in the circulatory system within **lipoproteins**. There is a large range of lipoproteins within blood, of which two are **low density lipoprotein (LDL)** and **high density lipoprotein (HDL)**. There is no difference between the chemical composition of the cholesterol that is carried by the various lipoproteins; however, HDL molecules are smaller and denser than LDL molecules due to the larger proportion of protein (with a higher molecular mass than cholesterol) in HDL. The LDL molecules contain much more cholesterol than HDL molecules.

Higher concentrations of LDL (or inversely, low concentrations of HDL) are strongly associated with cardiovascular disease. LDL promotes the narrowing of arteries (atherosclerosis) by accumulating beneath the inner elastic wall of the artery and the smooth muscle surrounding it. This disease process leads to heart attack, stroke and other diseases caused by the blockage of large peripheral arteries. As a result, cholesterol bound up in LDL is known as 'bad cholesterol'. On the other hand, it is hypothesized that high concentrations of HDL can remove cholesterol from cells and reduce atherosclerosis by removing cholesterol from blockages within arteries and transport it back to the liver for excretion or re-utilization. For this reason HDL-bound cholesterol is sometimes called 'good cholesterol'.

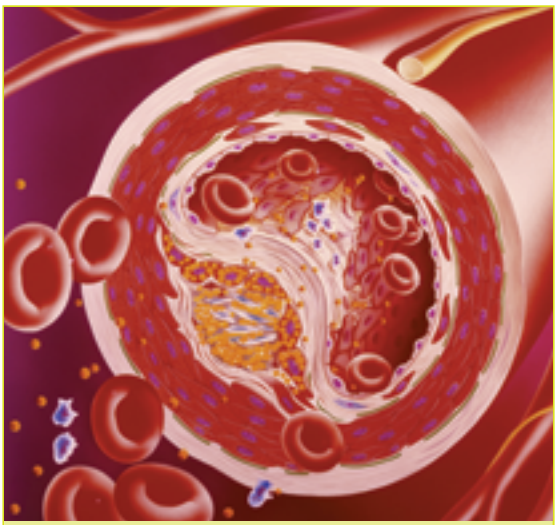


Figure 2.4.4 The accumulation of LDL in the artery wall causes arteries to become narrowed, leading to coronary heart disease.

Triglycerides are formed by the condensation reaction between glycerol, a polyalcohol, and fatty acids, long-chain carboxylic acids. Note that an ester link forms (figure 2.4.5). Triglycerides are usually classified according to the type of fatty acids involved in their formation, although a triglyceride does not have to be composed of three identical fatty acids. We find that most naturally occurring fats contain a mixture of **saturated**, **mono-unsaturated** and **polyunsaturated fatty acids** so they are classified according to the predominant type of unsaturation present.

AS B.4.6 Describe the condensation of glycerol and three fatty acid molecules to make a triglyceride. © IBO 2007

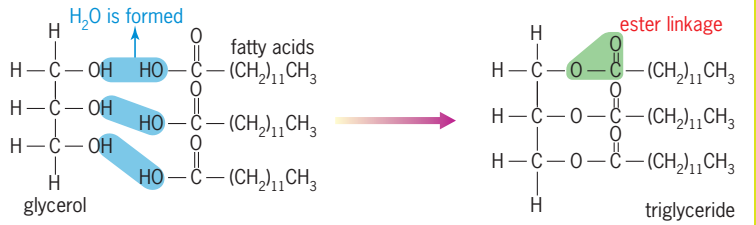


Figure 2.4.5 Triglycerides form by condensation reactions.

Saturated fatty acids, like other saturated hydrocarbons, have a hydrocarbon chain that contains no carbon–carbon double bonds. Mono-unsaturated fatty acids have one carbon–carbon double bond and polyunsaturated fatty acids have more than one carbon–carbon double bond in the hydrocarbon chain. The formula of a fatty acid can be used to determine the number of carbon–carbon double bonds. An unsaturated fatty acid will have the general formula $C_nH_{2n+1}COOH$. For every double bond that is present in a fatty acid (the degree of unsaturation), two hydrogen atoms will be lost from the formula. Consequently, the general formula of a mono-unsaturated fatty acid will be $C_nH_{2n-1}COOH$ and of a polyunsaturated fatty acid with two carbon–carbon double bonds will be $C_nH_{2n-3}COOH$. The pattern continues for other polyunsaturated fatty acids.

AS B.4.3 Describe the difference in structure between saturated and unsaturated fatty acids. © IBO 2007

Unsaturation in fatty acid leads to 'kinks' in the chain. The unsaturated molecules therefore do not pack closely together. This leads to weaker van der Waals' forces between the chains and a lower melting

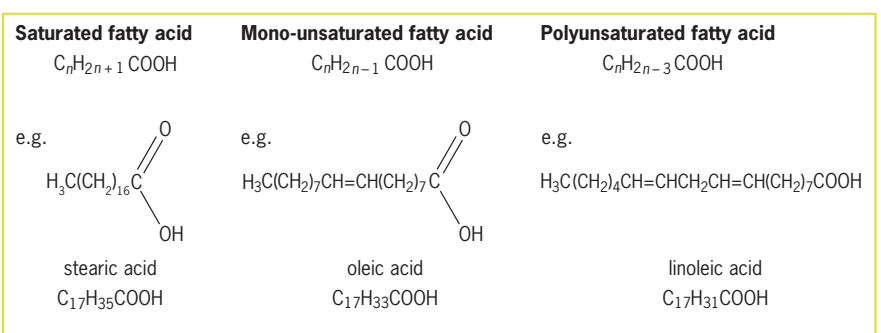


Figure 2.4.6 The formulas of fatty acids match a pattern.

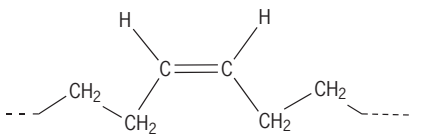
point. Generally we classify fats as triglycerides that are solid at room temperature, and oils as those that are liquids at room temperature. Oils are more likely to contain a large number of carbon–carbon double bonds, so their molecules do not pack together as well as those with no carbon–carbon double bonds, thus explaining their lower melting point.

CHEM COMPLEMENT

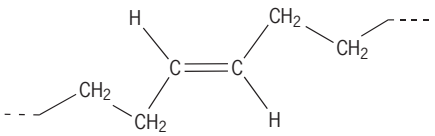
Margarine and the presence of trans fats

Margarine is the man-made equivalent of butter, the saturated fat derived from the cream from cow's milk. With butter not always being as available as consumers would like, the possibility of a mass produced vegetable-based substitute became attractive. The difference between butter and vegetable oils is the degree of saturation. Butter is made up of unsaturated triglycerides, so has a higher melting point than polyunsaturated vegetable oils. Hydrogenation, first developed in the 1890s by Paul Sabatier and further improved by the German chemist Wilhelm Normann in 1901, was found to be the way to decrease the number of double bonds in a triglyceride and consequently increase its melting point.

The process uses metal catalysts such as nickel and palladium; however, it has been shown that trans fatty acids are formed if the unsaturated fat leaves the surface of the catalyst too quickly. These trans fats were suggested as early as 1988 to be the cause of large increases in coronary artery disease and deaths from heart disease. In recent years health organizations have lobbied governments for the removal of trans fats from foods and currently many countries require that trans fats be included on the labelling of foods so that consumers can avoid them since their presence is just as bad for health as the saturated fats that they have replaced.



cis arrangement about a double bond



trans arrangement about a double bond

Figure 2.4.7 The difference between the *cis* and *trans* arrangement of atoms around a double bond.

Another class of unhealthy fatty acids originate from the industrial hydrogenation of plant oils to reduce their levels of unsaturation. These fatty acids are known as *trans* fatty acids because the hydrocarbon chain takes up a *trans* rather than *cis* configuration around the double bonds in the fatty acid chain. With this configuration, the fatty acid chain becomes straight like that of a saturated fatty acid, rather than kinked as in the case of naturally occurring polyunsaturated fatty acids. This lowers the melting point of the lipid making it more convenient to use, but the consumption of *trans* fatty acids is quite detrimental to health. It increases the risk of coronary heart disease by raising the levels of LDL cholesterol in the blood and also lowering the levels of HDL cholesterol.

Two important fatty acids are linoleic acid, $C_{18}H_{32}O_2$, and linolenic acid, $C_{18}H_{30}O_2$. These fatty acids are essential fatty acids; they are essential in the diet of all mammals. Linoleic acid is an omega-6 fatty acid, while linolenic acid is an omega-3 fatty acid. Notice that their structures are very similar, only differing by one carbon–carbon double bond and the consequent reduction in number of hydrogen atoms. These two substances work together in the body to promote health. Linoleic acid is used in the biosynthesis of prostaglandins, while linolenic acid is converted into two other omega-3 fatty acids: eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA) both of which have much more use in the body than linolenic acid itself. Omega-3 fatty acids such as linolenic acid, EPA and DHA help reduce inflammation, while most omega-6 fatty acids including linoleic acid tend to promote inflammation. An inappropriate balance of these essential fatty acids contributes to the development of disease, while a proper balance helps to maintain and even improve health. A healthy diet should consist of roughly two to four times more omega-6 fatty acids than omega-3 fatty acids. Since both linoleic acid and linolenic acid are abundant in many vegetable oils, it is easy to take in enough of these essential fatty acids in the daily diet.

AS B.4.4
Compare the structures of the two essential fatty acids linoleic (omega-6 fatty acid) and linolenic (omega-3 fatty acid) and state their importance.

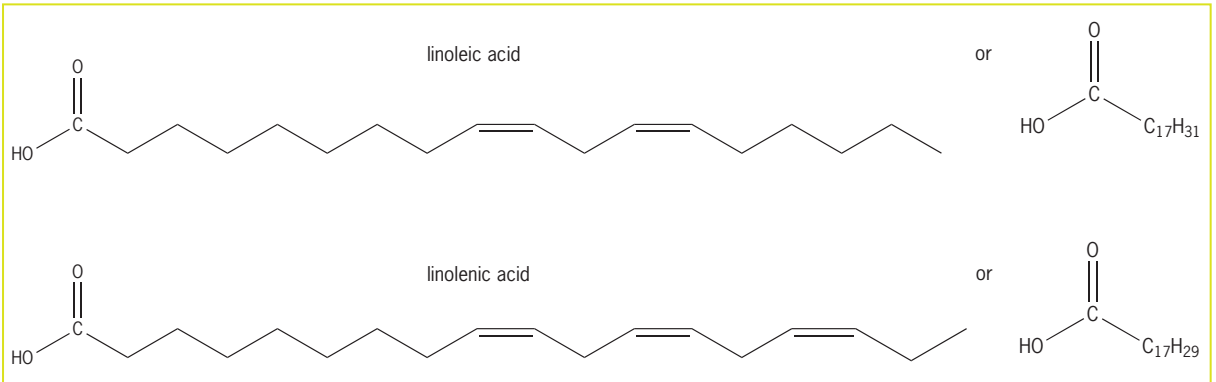


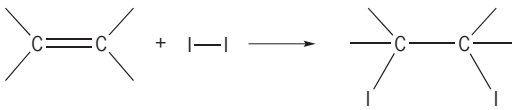
Figure 2.4.8 The structures of linoleic acid and linolenic acid.

Iodine numbers

Halogens undergo addition reactions with unsaturated hydrocarbons and therefore also with unsaturated fatty acids. In particular, reactions between bromine, Br_2, or iodine, I_2, and unsaturated fatty acids are easy to monitor because bromine and iodine are decolourized (become colourless) when they react with unsaturated hydrocarbons. The reaction with iodine is less dangerous than that with bromine, since an aqueous solution of bromine (bromine water) releases fumes of bromine, creating a respiratory hazard. Aqueous solutions of iodine are also more stable than those of bromine and can be standardized using a sodium thiosulfate solution. The reaction between iodine and an unsaturated fatty acid can then be performed as a titration.

AS B.4.5
Define the term *iodine number* and calculate the number of C=C double bonds in an unsaturated fat/oil using addition reactions.

QuickTime CW Video
Testing for unsaturation with bromine



One mole of iodine reacts with one mole of carbon–carbon double bonds. This means that the number of mole of I_2 reacting with one mole of a fat or oil indicates the number of double bonds present in the fat or oil molecule.

The iodine number of a fat or an oil is the mass of iodine that reacts with 100 g of the lipid. The more unsaturated an oil is, the higher its iodine number will be.

Worked example 1

0.010 mol of linoleic acid reacts with 40 cm^3 of a 0.50 mol dm^{-3} of I_2. Determine the number of double bonds present in each linoleic acid molecule.

Solution

$n(I_2) = cV = 0.50 \times 0.040 = 0.020$ mol

In 0.010 mol of linoleic acid there are 0.020 mol of double bonds

$\therefore$ There are $\frac{0.020}{0.010} = 2$ double bonds per molecule.

Worked example 2

The general formula for saturated fatty acids is $C_nH_{2n+1}COOH$. The molecular formula of arachidonic acid is $C_{20}H_{32}O_2$.

a Determine the number of carbon–carbon double bonds in arachidonic acid.

b Determine the iodine number of arachidonic acid.

Solution

a If arachidonic acid was a saturated fatty acid, by matching to the general formula, we can see that its formula would be $C_{19}H_{39}COOH$. However its molecular formula is given as $C_{20}H_{32}O_2$ or $C_{19}H_{31}COOH$.

This formula has 8 less hydrogen atoms in it than a saturated fatty acid would have, indicating that arachidonic acid has 4 carbon–carbon double bonds in each of its molecules.

b The iodine number of a fat or an oil is the mass of iodine that reacts with 100 g of the lipid.

First we must find the number of mole of the arachidonic acid in 100 g.

$$M(C_{20}H_{32}O_2) = 304.52 \text{ g mol}^{-1}$$

$$n = \frac{m}{M} = \frac{100}{304.52} = 0.328 \text{ mol}$$

Since there are four mole of double bonds per mole of arachidonic acid:

then $I_2 : C_{20}H_{32}O_2$
$4 : 1$

$$n(I_2) = 4 \times 0.328$$
$$= 1.31 \text{ mol}$$
$$m(I_2) = n \times M(I_2)$$
$$= 1.31 \times 253.8$$
$$= 333 \text{ g}$$

The iodine number of arachidonic acid is 333.

Lipids in the diet

AS **B.4.7 Describe the enzyme-catalysed hydrolysis of triglycerides during digestion.**

Like proteins and polysaccharides, triglycerides undergo enzyme-catalysed hydrolysis during digestion. Unfortunately, triglycerides do not dissolve in water, so they are not easily broken down by fat-digesting enzymes (lipase) in the watery content of the gastrointestinal tract. Thus fats tend to take longer to digest than carbohydrates or proteins.

Digestion of fats occurs in the small intestine. In the duodenum (the upper part of the small intestine), bile, produced in the liver but stored in the gallbladder, enters via the bile duct. Bile emulsifies fats, dispersing them into small droplets which then become suspended in the alkaline contents of the digestive tract. This process of emulsification allows the enzyme lipase, which enters the duodenum from the pancreas, to gain easier access to the fat molecules and thus accelerates their breakdown and digestion. Lipase catalyses the hydrolysis of the triglycerides into glycerol and fatty acids.

The walls of the small intestine are covered by millions of finger-like projections called villi. Inside each villus is a series of lymph vessels and blood vessels. The glycerol and fatty acids that are the products of hydrolysis are absorbed into the lymph vessels and eventually drain into the bloodstream.

The fatty acids are transported in the bloodstream to adipose cells, where they are stored, or to muscle cells, where they are oxidized for energy. The majority of digested fat is stored as body fat in the adipose cells. The glycerol is transported to the liver where it may be converted into glucose or may be used to help break down glucose to release energy.

In section 2.1 the higher energy value of fats, when compared with carbohydrates, was discussed. The enthalpy of combustion of fat is 37 kJ g^{-1} while that of carbohydrate is 17 kJ g^{-1}. Although this higher energy value is offset somewhat by the difficulty in digesting fats, fats have a higher energy value than carbohydrates.

When the formulas of fatty acids and carbohydrates are compared, it is found that there is a greater proportion of oxygen atoms in the carbohydrates than in fatty acids—their degree of oxidation is greater. Therefore the fatty acids have greater potential for oxidation and the subsequent release of energy. The combustion reactions of a fatty acid and a carbohydrate of similar molar mass can be compared to illustrate this difference.

Let us compare a trisaccharide, made up of three glucose units with a molecular formula of $C_{18}H_{32}O_{16}$ to linoleic acid, $C_{18}H_{32}O_2$.

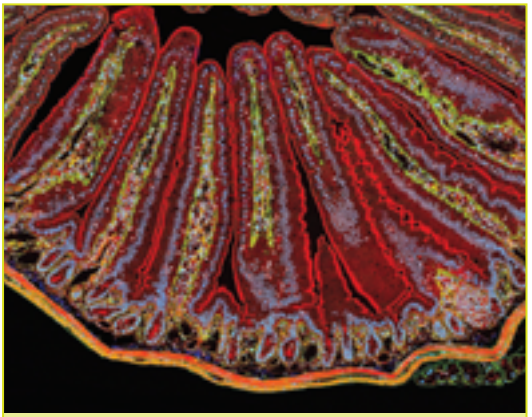


Figure 2.4.9 A fluorescent light micrograph showing the villi, finger-like projections that increase the surface area of the small intestine.

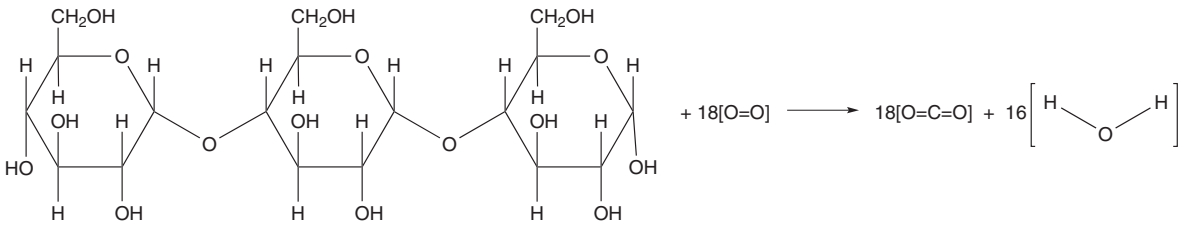
AS B.4.8
Explain the higher energy value of fats as compared to carbohydrates. © IBO 2007

The combustion reactions are:

$$C_{18}H_{32}O_{16} + 18O_2 \rightarrow 18CO_2 + 16H_2O$$
$$C_{18}H_{32}O_2 + 25O_2 \rightarrow 18CO_2 + 16H_2O$$

The standard enthalpy of reaction can be calculated using average bond enthalpy values.

1 Trisaccharide ($C_{18}H_{32}O_{16}$)



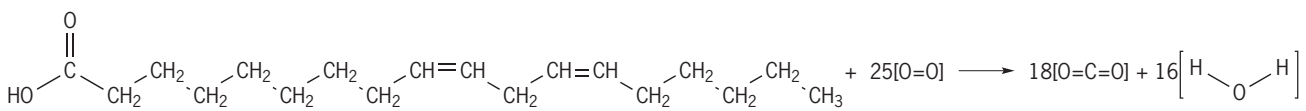
Enthalpy of reaction = ΣD(bonds broken) – ΣD(bonds formed)

where Σ represents the sum of the terms and D represents the bond enthalpy per mole of bonds.

$$\Delta H = [15D_{C–C} + 21D_{C–H} + 21D_{C–O} + 11D_{O–H} + 18D_{O=O}] – [36D_{C=O} + 32D_{H–O}]$$
$$= [(15 \times 348) + (21 \times 412) + (21 \times 360) + (11 \times 463) + (18 \times 496)] – [(36 \times 743) + (32 \times 463)]$$
$$= 35\,453 – 41\,564$$
$$= –6111 \text{ kJ}$$
$$M(C_{18}H_{32}O_{16}) = 504.5$$
$$\text{Energy content per g} = \frac{6111}{504.5} = 12.1 \text{ kJ g}^{-1}$$

The enthalpy of combustion for the trisaccharide $C_{18}H_{32}O_{16}$ is –6111 kJ mol^{-1} and the energy content is 12.1 kJ g^{-1}.

2 Fatty acid ($C_{18}H_{32}O_{16}$)



Enthalpy of reaction = ΣD(bonds broken) – ΣD(bonds formed)

$$\Delta H = [15D_{C-C} + 2D_{C=C} + 31D_{C-H} + D_{C=O} + D_{O-H} + D_{C-O} + 25D_{O=O}] - [36D_{C=O} + 32D_{H-O}]$$
$$= [(15 \times 348) + (2 \times 612) + (31 \times 412) + 743 + 463 + 360 + (25 \times 496)] - [(36 \times 743) + (32 \times 463)]$$
$$= 33\,182 - 41\,564$$
$$= -8382 \text{ kJ}$$

$M(C_{18}H_{32}O_2) = 280.5$

$$\text{Energy content per g} = \frac{8382}{280.5} = 30.0 \text{ kJ g}^{-1}$$

The enthalpy of combustion for the fatty acid $C_{18}H_{32}O_2$ is –8382 kJ mol^{-1} and the energy content is 30.0 kJ g^{-1}.

From the balanced equations (above) for the combustion of linoleic acid (a fatty acid) and the trisaccharide, the ratio of oxygen to fatty acid is 25 : 1 whereas oxygen to trisaccharide is 18 : 1. A great deal more oxygen is required to oxidize the fatty acid than the trisaccharide. As in the comparison between the complete and incomplete oxidation of hydrocarbons, the reaction that requires a greater proportion of oxygen yields a greater amount of energy.

In summary, we can explain the higher energy content of fats than in carbohydrates as being the result of more energy being released in bond making in products than is required for bond breaking in reactants. This can be traced to the smaller number of hydroxyl groups in a fatty acid than in a carbohydrate and to the larger amount of oxygen required to oxidize a fatty acid than a carbohydrate.

Lipids lead a 'double life' with respect to our health. There are many uses in the human body for lipids, but they can easily lead to ill health when they form too great a part of our diet.

WORKSHEET 2.3
Lipids and carbohydrates

B.4.9
Describe the important roles of lipids in the body and the negative effects that they can have on health.

TABLE 2.4.1 IMPORTANT ROLES OF LIPIDS AND NEGATIVE EFFECTS OF LIPIDS

Important roles of lipids in the body	Negative effects of lipids on our health
Lipids are stored as a source of energy that can be used.	Excessive amounts of lipids in the diet can lead to obesity.
They can be oxidized to give large amounts of energy.	Cholesterol present in LDL increases arteriosclerosis.
They are stored in adipose tissue around the body providing insulation and a protective covering for vital organs.	Trans fatty acids increase the risk of coronary heart disease due to increased levels of LDL cholesterol.
Polyunsaturated fats may lower levels of LDL cholesterol.	Saturated fats, in particular lauric (C_{12}), myristic (C_{14}), and palmitic (C_{16}) acids, increase levels of LDL cholesterol.
Cholesterol present in HDL lowers the threat of heart disease.	
Omega-3-polyunsaturated fatty acids reduce the risk of heart disease.	
Phospholipids help to form cell membranes, along with proteins.	
Lipids provide the basis of steroid hormones such as progesterone and testosterone.	

Section 2.4 Exercises

1 Compare the composition of triglycerides, phospholipids and steroids.

2 **a** Outline the difference between HDL and LDL cholesterol.

b Explain how LDL cholesterol is damaging to the health of humans.

3 The reaction between fatty acids and glycerol results in a triglyceride.

a Draw the structure of glycerol.

b Draw a general structure of a fatty acid.

c State the name of the functional group formed during the reaction between a fatty acid and glycerol.

4 **a** Outline one structural difference between unsaturated fats and saturated fats.

b Considering the structural differences you have described above, compare the melting points of saturated and unsaturated fats with similar relative molecular masses.

c Explain the difference in melting point that you stated in your answer to part **b**.

5 **a** Explain what is meant by the term *essential fatty acid*.

b Describe the roles of linoleic acid and linolenic acid in the body.

6 Define the term *iodine number*.

7 Timnodonic acid, or eicosapentaenoic acid, is an unsaturated fatty acid found in fish oil. Its molecular formula is $C_{20}H_{30}O_2$.

a Determine the number of carbon–carbon double bonds in timnodonic acid.

b Determine the iodine number of timnodonic acid.

8 1.693 g of erucic acid (molar mass = 338.64), reacts with 25 cm^3 of a 0.20 mol dm^{-3} of I_2.

a Determine the number of carbon–carbon double bonds present in each erucic acid molecule.

b Determine the iodine number of erucic acid.

9 Palmitoleic acid is an unsaturated fatty acid with the formula $CH_3(CH_2)_5CH{=}CH(CH_2)_7COOH$. Write an equation for the reaction between glycerol and palmitoleic acid to form a triglyceride.

10 **a** Describe the conditions required for the digestion of triglycerides.

b Describe where lipids are stored in the human body.

c List three uses of lipids in the human body.

2.5 MICRONUTRIENTS AND MACRONUTRIENTS

AS B.5.1 Outline the difference between micronutrients and macronutrients. © IBO 2007

Up to this point we have discussed the chemical substances in our diet that we require in relatively large amounts. Proteins, fats and carbohydrates are all classed as **macronutrients**, they make up more than 0.005% of our body weight. Included in this class are minerals such as sodium, potassium (needed for healthy nerve function), magnesium, potassium, phosphorus, sulfur and chlorine.

Many other substances are needed in our diet but in much smaller amounts. These are known as **micronutrients.** Only milligrams or even micrograms of these are needed, but they are still essential as they function as **co-factors** of enzymes, and so are vital to digestion and other important bodily processes. Micronutrients include vitamins and trace minerals such as iron, copper, fluorine, zinc, iodine, selenium, manganese, molybdenum, chromium, cobalt and boron. Micronutrients make up less than 0.005% of our body weight.

Figure 2.5.1 A balanced diet provides sources of a range of macronutrients and micronutrients.

Vitamins

AS B.5.2 Compare the structures of retinol (vitamin A), calciferol (vitamin D) and ascorbic acid (vitamin C). © IBO 2007

AS B.5.3 Deduce whether a vitamin is water- or fat-soluble from its structure. © IBO 2007

Vitamins are chemicals that are vital to the normal functioning of an animal's metabolism and that the animal cannot synthesize itself. Vitamins often function as co-factors of enzymes. The ability of vitamins to be transported and stored in the essentially aqueous environment of the body is important, so vitamins are classified as either fat-soluble or water-soluble. Water-soluble vitamins have more hydroxyl groups than fat-soluble vitamins. These polar hydroxyl groups form hydrogen bonds with water and so enable the vitamins to dissolve in blood and other aqueous environments in the body. Fat-soluble vitamins have few hydroxyl groups and instead can dissolve in non-polar (fatty) environments due to van der Waals' forces between the long non-polar chains in the fat-soluble vitamins and the non-polar fatty acids.

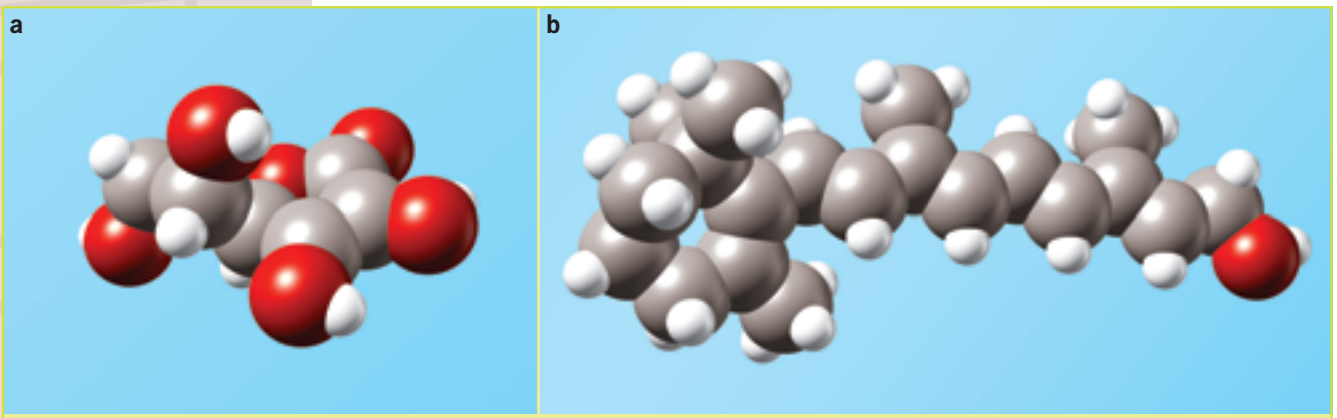


Figure 2.5.2 (a) Vitamin C is water-soluble due to the large number of hydroxyl groups in the molecule. (b) Vitamin A, with a large non-polar chain is insoluble in water.

Pr PRAC 2.3 Determination of vitamin C in fruit juices

Water-soluble vitamins include vitamins B and C (ascorbic acid), while the fat-soluble vitamins are vitamins A (retinol), D (calciferol), E and K. Fat-soluble vitamins can build up in the fatty (adipose) tissues of the body. This has both positive and negative effects. A supply of vitamins A, D, E or K present in the body can compensate for a diet that becomes deficient in these vitamins, while excessive consumption of these vitamins may result in hypervitaminosis.

Water-soluble vitamins are excreted by the body if they are not used, so these must be consumed as a regular part of the diet.

CHEM COMPLEMENT

Hypervitaminosis can kill

Xavier Guillaume Mertz (1883–1913) was a Swiss explorer, who explored the Antarctic with Douglas Mawson on his 1911–1914 expeditions. During one surveying expedition in Antarctica, a tragedy occurred and the sled carrying their supplies and one team member fell through a snow-covered crevasse. Mertz and Mawson found themselves stranded 315 miles from the main base without sufficient food to make the journey. To enable them to reach the safety of the base Mertz and Mawson fed on their sled dogs, but the high levels of vitamin A in the dog livers resulted in Mertz developing stomach pains and finally dying from vitamin A poisoning on 7 January 1913.

If retinol, calciferol and ascorbic acid (see figure 2.5.3) are compared, the major differences in their structures can be seen. The fat-soluble vitamins—retinol and calciferol—have long hydrocarbon chains, and one hydroxyl group on each molecule, making them both alcohols. In contrast, ascorbic acid (vitamin C) has many hydroxyl groups, which contribute to its solubility in water.

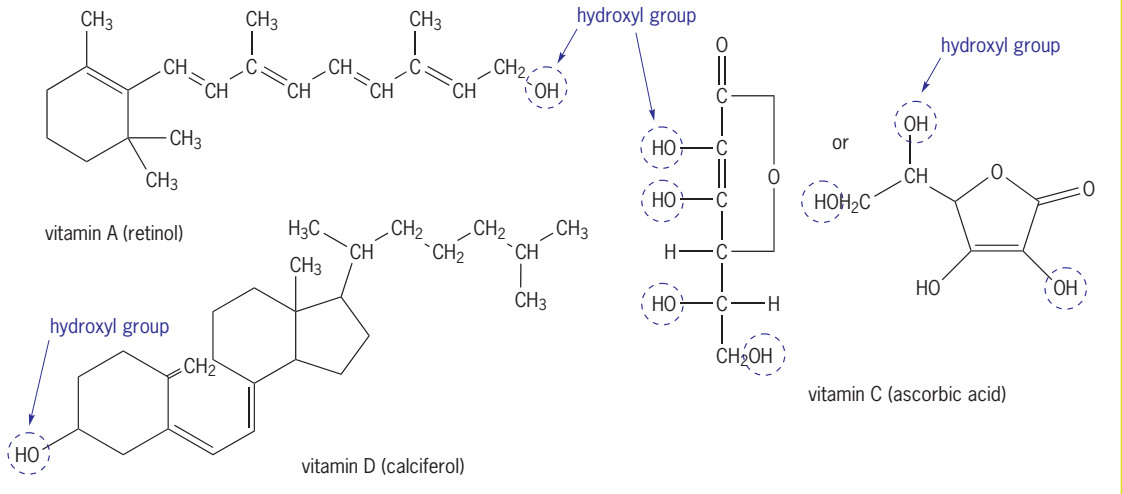


Figure 2.5.3 The structures of vitamins A (retinol), C (ascorbic acid) and D (calciferol).

Vitamin A (retinol) is derived from carotene. It can exist as an aldehyde (retinal), or as an acid (retinoic acid). It is required for the production of rhodopsin, which is the light sensitive material in the rods of the retina. Deficiency symptoms include night blindness, excess skin dryness and a lack of mucous membrane secretions. It is estimated that nearly 3 million preschool children in developing countries are blind because of a deficiency of vitamin A.

Vitamin C (ascorbic acid) is used in the formation and maintenance of collagen. It plays a major role in the formation of bones and teeth, and enhances the absorption of iron from vegetables. Deficiency symptoms include the disease known as scurvy—the bleeding and weakening of gums, tooth decay and the loss of teeth. Scurvy was at one time common among sailors aboard ships at sea for long periods of time. The lack of fruits and vegetables in their diets was found to be the cause by a Scottish surgeon in the British Royal Navy, James Lind. Scurvy often accompanies other forms of malnutrition such as beriberi or pellagra.

AS B.5.4
Discuss the causes and effects of nutrient deficiencies in different countries and suggest solutions. © IBO 2007

Figure 2.5.4 A major symptom of kwashiorkor is a swollen belly in a starving child.

PRAC 2.4
Colorimetric analysis of iron in foods

CHEM COMPLEMENT

The importance of a balanced diet

An early experiment investigating the importance of diet on health was conducted in a Southern US penitentiary (jail) where prisoners, in return for early release, were fed nothing but a corn diet. They were fed corn bread, corn meal and many other corn-based foods. The prisoners soon developed pellagra. Corn is clearly a poor source of niacin (vitamin B_3) and of tryptophan. This was one of the first recorded research experiments into the importance of diet to overall health.

Vitamin D (calciferol) is needed for normal bone formation and the retention of calcium and phosphorus in the body. It protects teeth and bones against the effects of a low calcium intake. Deficiency symptoms include rickets, which is a deformity of the ribcage and skull, as well as causing bowlegs. Excessive vitamin D consumption can cause excessive absorption of calcium and phosphorus and the formation of calcium deposits on major organs.

The primary effect of macronutrient deficiencies will be poor growth and the potential for starvation. Marasmus and kwashiorkor are two diseases that are the result of protein deficiencies. While marasmus affects children under the age of one and is due to a severe deficiency of nearly all nutrients, especially protein and energy-providing foods, kwashiorkor affects slightly older children. Jamaican pediatrician Cicely D. Williams introduced the term 'kwashiorkor' into international scientific circles in an article in the medical journal *The Lancet* in 1935. The name means 'rejected one' and reflects the incidence of the disease in children who have been weaned from breast milk, but have then proceeded on to a diet that is high in carbohydrates and lacks protein. Symptoms of kwashiorkor include a swollen abdomen, as well as weight loss.

Micronutrient malnutrition is practically unknown in developed countries due to inexpensive interventions such as food fortification, supplementation and dietary diversification, but the developing nations are not so lucky. A number of micronutrient deficiencies face the populations of developing countries.

An iron deficiency causes anemia. This results in one out of four maternal deaths in the developing world. The World Bank has estimated that the cost of fortifying flour with iron would be only 20 cents (US) per person per year. While anemia is common in developed countries, it is usually due to bad choices in the diet. An attempt at vegetarianism without consuming the correct balance of non-meat products may result in anemia.

Iodine is necessary for the synthesis of the thyroid hormones triiodothyronine and thyroxine (see section 2.6). In response to low thyroid hormones, the pituitary gland releases thyroid stimulating hormone (TSH), which acts to increase synthesis of the missing thyroid hormones. It also causes the thyroid gland to grow in size by increasing cell division. The result is the condition known as goitre. In addition to causing goitre, a lack of iodine in the diet is also reported to be the world's leading cause of mental retardation. It is estimated that more than 2 billion children suffer from lowered IQ and retardation due to iodine deficiency. A lack of iodine in the diet is easily remedied by adding sodium iodide to table salt (ionized salt). The costs of providing iodized salt are estimated at 10 cents per person per year.

Another unpleasant disease that results from a poor diet is pellagra. Pellagra may occur when the diet is deficient in niacin (vitamin B_3) and proteins containing the amino acid tryptophan. Tryptophan can be converted into niacin, so foods with tryptophan but without niacin, such as milk, can prevent pellagra. The symptoms of pellagra range from diarrhoea, through dermatitis and dementia to death. Pellagra is an endemic disease in Africa, Mexico, Indonesia and China. It does occur in Western countries in patients with very poor diets.

Another vitamin B deficiency is beriberi. In developing countries, a lack of vitamin B_1 may be due to a diet that consists mainly of polished white rice. Such a diet would be very low in thiamine because the thiamine-bearing rice husk has been removed. Thiamine is needed for the conversion of carbohydrates to energy in the body, as well as the correct functioning of the heart, muscles and nervous system. Beriberi is also seen in chronic alcoholics with an inadequate diet. The symptoms of beriberi include weight loss,

emotional disturbances, weakness and pain in the limbs, and periods of irregular heart rate.

How can these diseases be cured? Often the solution is as straightforward as providing dependant people (such as those in refugee camps) with food rations that are composed of fresh and therefore vitamin- and mineral-rich foods. Nutrients can be added to commonly consumed foods, such as iron to flour. Genetic modification of foods may supply an answer by increasing the vitamin and mineral content of crops, although this is often an expensive solution that is not available to the starving people of the world. Nutritional supplements can be useful when a diet does not supply enough of particular vitamins, for example vitamin B_1.

Figure 2.5.5 Despair on the face of a woman in a Somalian refugee camp.

TABLE 2.5.1 SUMMARY OF CAUSES AND EFFECTS OF NUTRIENT DEFICIENCIES

Nutrient	Effect of deficiency	Symptoms	Cause of deficiency
Vitamin A (retinol)	• Reduced rhodopsin production	• Night blindness	Food supplies low in carotene
	• Dryness of mucous membranes	• Excessive skin dryness • Xerophthalmia (dry eyes) • Lack of mucous membrane secretions	
Vitamin C (ascorbic acid)	Reduced formation and maintenance of collagen	Scurvy: bleeding and weakening of gums, tooth decay and loss of teeth	Lack of fruit and vegetables in diet
Vitamin D (calciferol)	Low calcium and phosphorus retention and bone malformation	Rickets: deformity of ribcage, skull and legs	Lack of dairy foods
Proteins	Marasmus	• Dry skin, loose skin folds • Drastic loss of adipose tissue from normal areas of fat deposits	Diet low in all nutrients
	Kwashiorkor	• Swollen abdomen • Weight loss	Weaned from breast milk onto a low protein diet
Iron	Anemia	Low energy	Diet low in meat, dairy or green leafy vegetables
Iodine	Low levels of thyroid hormones	• Goitre (increased growth of thyroid gland) • Mental retardation	Diet low in iodine
• Tryptophan (an amino acid) • Niacin (vitamin B_3)	Tryptophan is converted into niacin, which is needed to help production of energy from food	Pellagra: diarrhoea, dermatitis, dementia and death	Diet low in dairy, meat, eggs, wheat flour, maize flour
Thiamine (vitamin B_1)	• Inefficiencies in converting carbohydrates into energy • Reduced functioning of the heart, muscles, and nervous system	Beriberi: weight loss, emotional disturbances, weakness and pain in limbs, irregular heart beat	• Diet high in polished white rice • Generally inadequate diet

THEORY OF KNOWLEDGE

On 10 December 1948 the General Assembly of the United Nations adopted the Universal Declaration of Human Rights. Article 25 states that: 'Everyone has the right to a standard of living adequate for the health and well-being of himself and of his family, including food, clothing, housing and medical care and necessary social services, and the right to security in the event of unemployment, sickness, disability, widowhood, old age or other lack of livelihood in circumstances beyond his control.' In 2000, a UN resolution adopted the Millennium Development Goals. Article 19 resolves to reduce by half the proportion of people living on less than a dollar a day and reduce by half the proportion of people who suffer from hunger.

- You are a chemist interested in using your expertise to understand world hunger. You have many questions on your mind like: What is hunger? What are the causes of hunger? What are the effects of hunger? What are the solutions? Carefully think about the things you need to know and suggest a possible research question for investigation. Try and narrow your focus. 'Can rice be enriched with vitamin D?' is an example of a good focused question.
- Consider some of the challenges scientists might face from experts in other areas of knowledge in the process of finding scientific solutions to world hunger. What sort of questions might a human scientist (economist, psychologist, sociologist, geographer) working in this field ask? How would these questions differ from those of a chemist?

Section 2.5 Exercises

1 Describe the structural features of vitamin C that make this vitamin water-soluble.

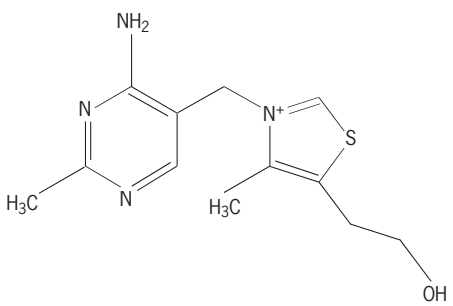
2 Explain why carbohydrates are regarded as a macronutrient.

3 Describe one function of micronutrients.

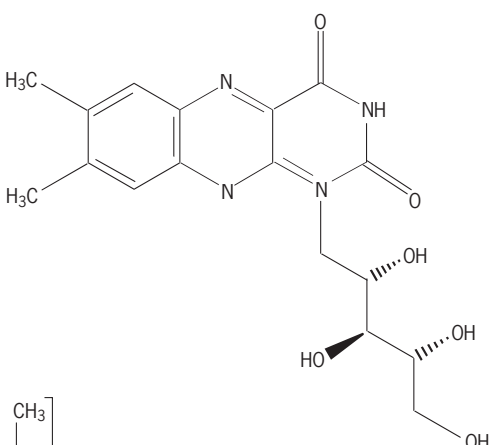
4 Compare the structure of ascorbic acid with that of retinol (vitamin A).

5 Consider the following vitamins and use their structures to deduce whether they are water-soluble or fat-soluble.

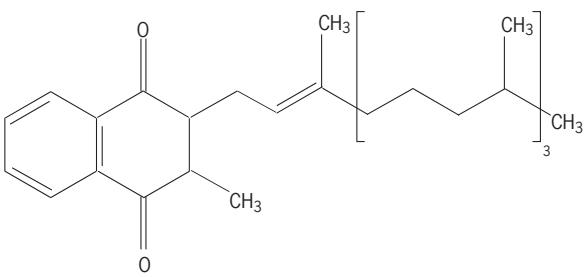
a thiamine



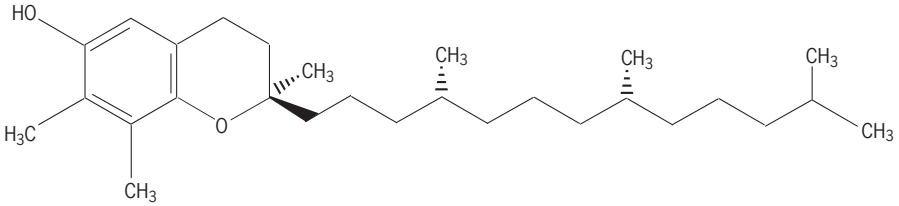
b riboflavin



c phylloquinone



d tocopherol



6 Explain why a child might contract each of the following conditions.

a Scurvy

b Rickets

c Beriberi

d Anemia

e Kwashiorkor

2.6 HORMONES

A **hormone** is a regulatory compound that is released by cells of one tissue and which acts to regulate the metabolism of target cells in another tissue. Hormones are often described as 'chemical messengers'.

AS B.6.1 Outline the production and function of hormones in the body. © IBO 2007

Hormones are secreted directly into the blood by endocrine glands such as the hypothalamus, the pituitary gland, the thyroid gland, the pancreas, the adrenal glands and the ovaries. The range of hormones produced is large, so we will only look in detail at a small number of them.

The **sex hormones**, testosterone, progesterone and estradiol, are the hormones that control sexual function in males and females, including production of sperm, ovulation and support of pregnancy. Secondary sex characteristics such as muscle bulk, hairiness and deepening of the voice in males and the growth of breasts, widening of hips and an increased amount of body fat in hips, thighs, buttocks and breasts in females are also supported by the sex hormones. These hormones are all present in both males and females; however, the relative amounts vary between the sexes. For example, levels of progesterone in adult males are similar to those in women during the follicular phase (the first part) of the **menstrual cycle**.

Estrogens are a group of steroid compounds of which the three naturally occurring members are estradiol, estriol and estrone. They are the primary female sex hormones, promoting the development of female secondary sex characteristics, such as breasts. They are also involved in the thickening of the lining of the uterus and in other regulatory aspects of the menstrual cycle. In males, estrogen plays a role in the maturation of sperm and is thought to be necessary for a healthy libido. Progesterone belongs to a class of hormones called progestogens, but is the major

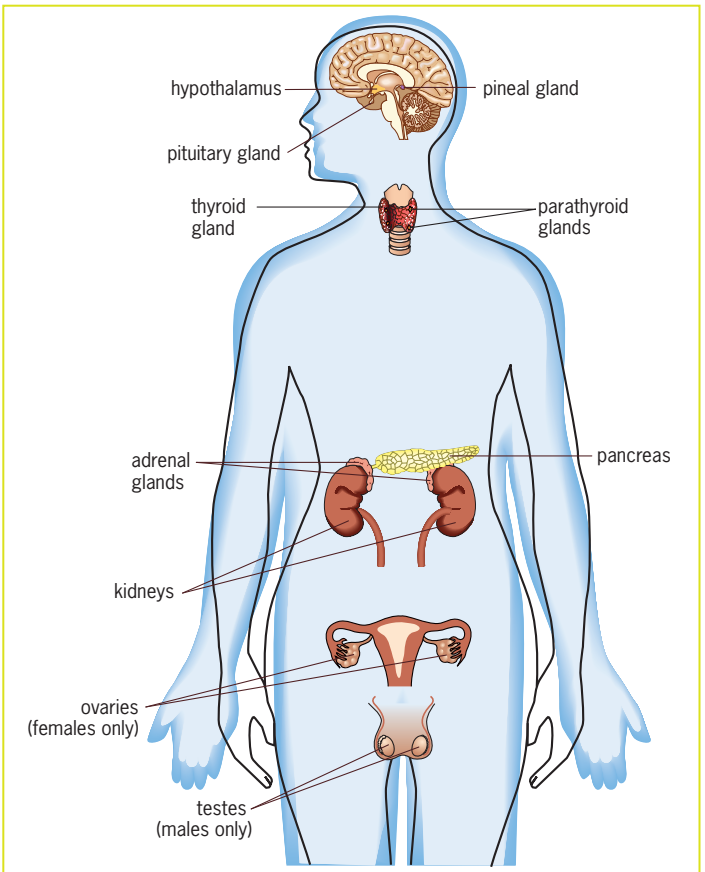


Figure 2.6.1 Hormones are chemical messengers that travel through the bloodstream from endocrine glands to other parts of the body where they act.

naturally occurring member of the class. Progesterone is the hormone that helps pregnancy to continue by preparing the uterus for the fertilized egg. High levels of progesterone are maintained throughout pregnancy.

Testosterone is the primary male sex hormone. It is present in both males and females, but there are much higher levels in males (about 40–60 times more). Females are more sensitive to testosterone than males. Testosterone plays a key role in health and wellbeing as well as in sexual functioning. Testosterone is the precursor of progesterone in the human body.

TABLE 2.6.1 SOME HORMONES AND THEIR FUNCTIONS

Hormone	Produced in	Acts upon	Function of hormone
Luteinizing hormone (LH)	Pituitary gland	Ovaries or testes	• In females stimulates ovulation • In males stimulates secretion of testosterone
Follicle-stimulating hormone (FSH)	Pituitary gland	Ovaries or testes	• In females stimulates maturation of follicles in ovaries • In males stimulates production of sperm
Thyroid-stimulating hormone (TSH)	Pituitary gland	Thyroid gland	Stimulates thyroid gland to secrete thyroxine (T4) and triiodothyronine (T3)
Antidiuretic hormone (ADH)	Pituitary gland	Kidneys	Controls retention of water in the kidneys
Triiodothyronine (T3) Thyroxine (T4)	Thyroid gland	A range of cells throughout the body	Regulates the basal metabolic rate and affects protein synthesis
Insulin	Pancreas	Liver, muscle and fat tissue cells	• Causes glucose to be taken up from the blood and to be stored in liver and muscle • Stops use of fat as energy source
Testosterone	Testes, adrenal glands	Range of organs throughout body	• Growth of muscle mass, increased bone density, growth and strength • Maturation of sex organs, male secondary sex characteristics, e.g. deepening of voice, beard growth
Adrenaline (epinephrine)	Adrenal glands	Heart, liver and many other organs	• Fight or flight response—boosts the supply of oxygen and glucose to the brain and muscles by range of actions • Suppresses immune system and non-emergency bodily processes
Noradrenaline	Adrenal glands	Skeletal muscles	• Fight or flight response • Increases skeletal muscle readiness
Aldosterone	Adrenal glands	Kidneys	Increase reabsorption of ions and water in the kidney
Progesterone	• Ovaries • Placenta (during pregnancy)	Uterus, cervix, and a range of other organs	• Supports pregnancy and prepares uterus for fertilized egg; also performs a range of anti-inflammatory functions • Required for synthesis of aldosterone
Estrogens	Ovaries	Range of organs throughout body	Huge range including stimulating ovulation, promoting formation of female secondary sex characteristics, increasing uterine growth, reducing muscle mass

The structure of cholesterol is similar to that of the sex hormones because it is the precursor of these hormones. The part of the structure that is common to all is the steroid backbone—a series of four fused rings—three of which have six members and one with five members. The difference between these molecules lies in their functional groups. Progesterone and testosterone are very similar in structure; the only difference occurring on the five-membered ring where testosterone has a hydroxyl group while progesterone has a methyl ketone group. Like testosterone, estradiol has a hydroxyl group on the five-membered ring, but its first ring has an aromatic ring structure and a hydroxyl group rather than the carbonyl group found in progesterone and testosterone. Estradiol is also missing a methyl group between the first two rings. Like estradiol, cholesterol has a hydroxyl group on its first ring, but possesses the methyl group that estradiol is missing, and has a hydrocarbon chain in the position on the five-membered ring that is occupied by other functional groups in the sex hormones.

B.6.2
Compare the structures of cholesterol and the sex hormones. © IBO 2007

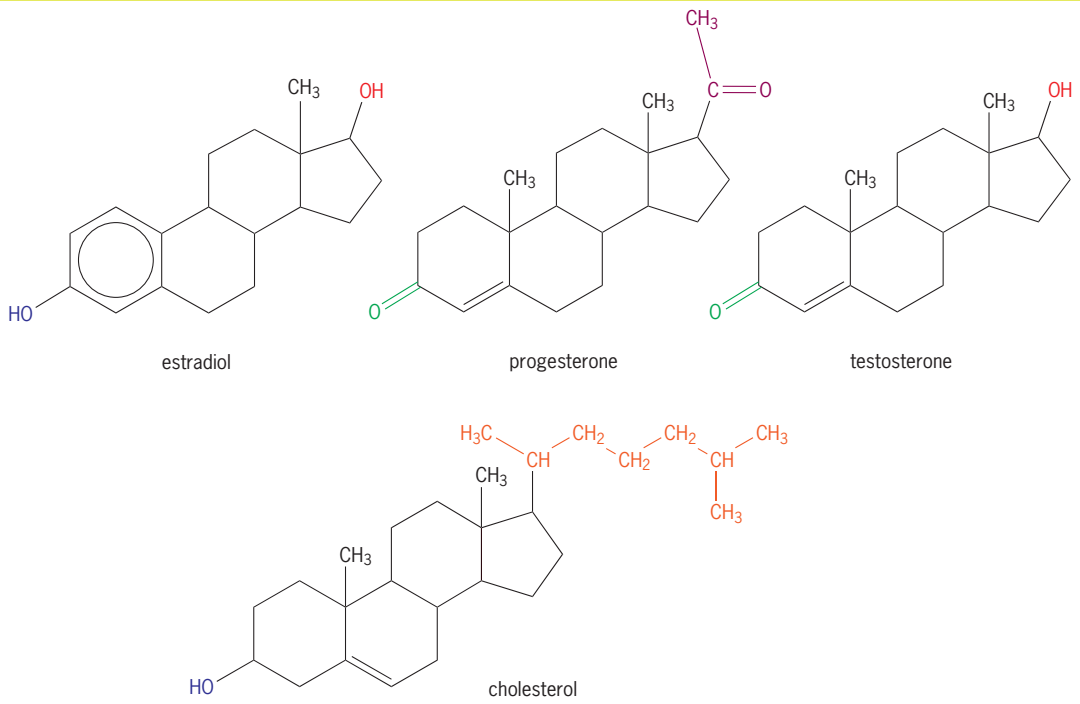


Figure 2.6.2 The structures of estradiol, progesterone, testosterone and cholesterol.

The co-operative actions of progesterone and estrogen regulate the menstrual cycle. At the beginning of the cycle, levels of estrogen and progesterone are low. The release of follicle-stimulating hormone (FSH) by the pituitary gland increases estrogen levels. At the midpoint of the menstrual cycle, the pituitary gland releases luteinizing hormone (LH), which stimulates ovum (egg) release and the increase of progesterone secretion—this is needed for the maintenance of a pregnancy. At the end of the menstrual cycle, hormone production decreases and menstruation occurs.

B.6.3
Describe the mode of action of oral contraceptives. © IBO 2007

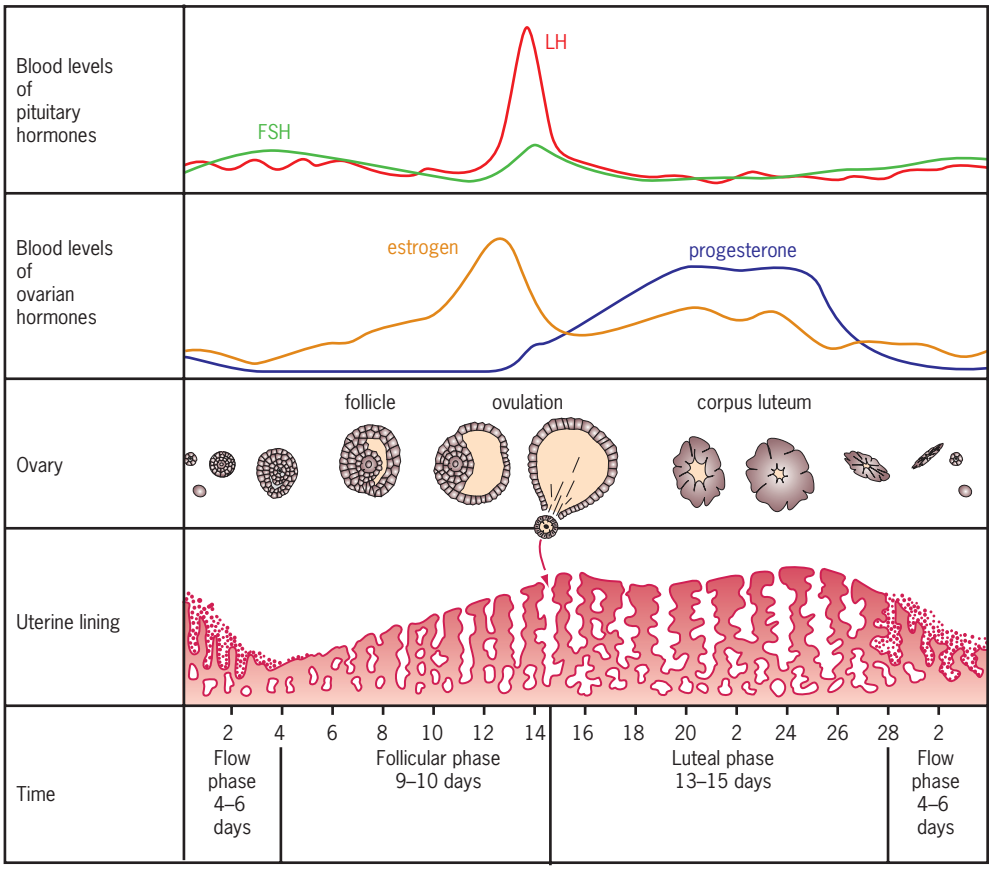


Figure 2.6.3 The relative levels of hormones during the menstrual cycle.

Oral contraceptives work by administering estrogens and progestins (synthetically produced progestogens), which inhibit FSH and LH secretion by the pituitary gland and hence ovulation. Although progestins are the main hormones that inhibit follicle development and ovulation, estrogen has also been found to perform those functions, as well as reducing the incidence of breakthrough bleeding (an unpleasant side effect). A secondary mechanism of action of oral contraceptives is inhibition of sperm penetration through the cervix by changing the composition of the cervical mucus.

The structures of progesterone and synthetically produced progestogens are sufficiently similar that they perform the same function in the human body. In figure 2.6.4 it can be seen that the major difference lies in the additional C_2H group with a carbon–carbon triple bond that is present in the progestin.

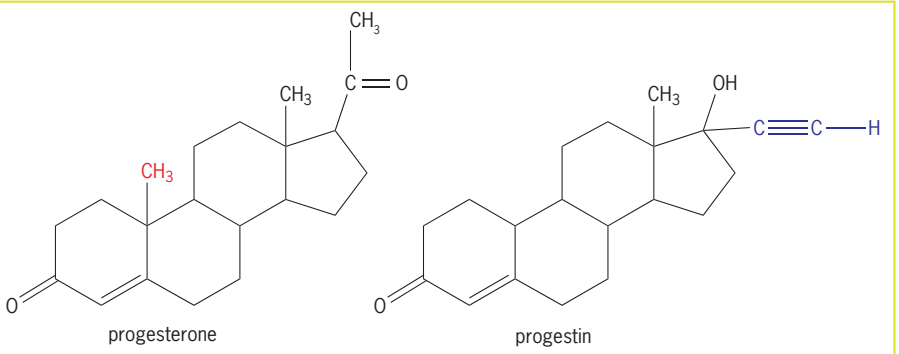


Figure 2.6.4 The structures of naturally occurring progesterone and the synthetic progestin are very similar.

Anabolic steroids are a class of steroid hormones related to testosterone. The structure of these compounds is very similar to that of testosterone (see figure 2.6.5) and they mimic the action of testosterone. They increase protein synthesis within cells, which results in the build up of cellular tissue, especially in muscles. This build up is called anabolism.

AS **B.6.4**
Outline the use and abuse of steroids. © IBO 2007

Anabolic steroids may be used therapeutically to stimulate bone growth and appetite, and to induce puberty in males. It can be used to help patients with chronic wasting diseases such as cancer or AIDS, since it increases the muscle mass and the physical strength of the patient, helping them to resume normal activities. There are health risks involved with the use of steroids. These include an increased level of LDL cholesterol and the accompanying decrease in HDL levels, high blood pressure and liver damage.

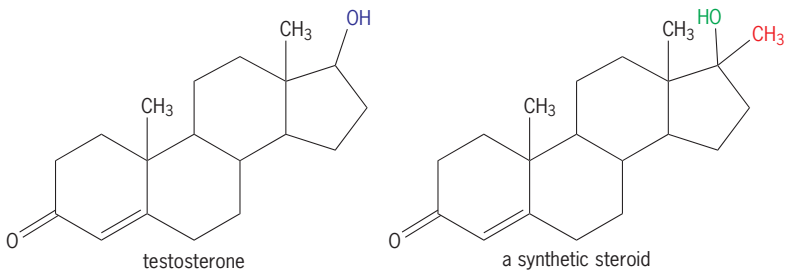


Figure 2.6.5 The structure of anabolic steroids and testosterone is very similar.

Anabolic steroids may be put to controversial use in sports and bodybuilding. They offer an advantage to the user by increasing muscle mass, but have side-effects such as increased bad temper, increased incidence of acne, trembling of the hands, premature balding and a decreased sperm count. Such use is banned by sporting bodies across the world. In a number of countries anabolic steroids are a controlled substance, meaning that their manufacture, possession and use are regulated by the government.

Section 2.6 Exercises

1 **a** State the general role of hormones in the body.
b State the name of the type of gland that controls the production of hormones.

2 Refer to figure 2.6.2 to answer these questions.
a Identify one sex hormone that has a hydroxyl group.
b Identify one sex hormone that has a carboxyl group.

3 For each of the following hormones, describe where it is produced and outline its specific role in the body.
a Testosterone
b Progesterone
c Estradiol
d Adrenaline
e Thyroxine

4 Explain how an oral contraceptive can prevent a pregnancy from occurring.

5 Describe the effect achieved and the circumstances in which steroids may be:
a used effectively
b abused.

2.7 ENZYMES

HL

AS **B.7.1**
Describe the characteristics of biological catalysts (enzymes). © IBO 2007

Enzymes are proteins that act as biological catalysts, increasing reaction rates of biological processes without being used up in the process. Unlike inorganic catalysts, enzymes tend to be very specific in terms of the reactions they can catalyse. Their specificity depends on their tertiary and quaternary structures (see p. 84). The ability of an enzyme to increase the rate of a biological process by as much as 10^7 times means that the concentrations of potentially dangerous substances such as hydrogen peroxide are reduced rapidly. Catalase is an enzyme that increases the rate of the decomposition of hydrogen peroxide, a corrosive by-product in the cells of many organisms. Catalase enables the rapid decomposition of the hydrogen peroxide to water and oxygen. Inorganic catalysts such as manganese dioxide can also catalyse this reaction.

AS **B.7.2**
Compare inorganic catalysts and biological catalysts (enzymes). © IBO 2007

Compared with inorganic catalysts, enzymes:

- have faster reaction rates
- operate under milder conditions, that is low temperatures (37°C) and pressures and within narrow ranges of temperature and pH
- can be rendered inactive at temperatures below their operating temperature and can be denatured at higher than usual temperatures at which the three-dimensional structure of the enzyme is destroyed
- are very selective, generally dealing with only one set of reactants (substrates). For example, lactase breaks down the disaccharide lactose, and a different enzyme, maltase, breaks down the disaccharide maltose.

AS **B.7.5**
Describe the mechanism of enzyme action, including enzyme–substrate complex, active site and induced-fit model. © IBO 2007

One model to explain enzyme action is the lock and key model. An enzyme's shape is seen to match that of the molecule whose reaction it catalyses. Figure 2.7.1 illustrates the reacting substance (the substrate) attaching to the active site (particular part of the enzyme to which the substrate can attach or bind) of the enzyme, forming an enzyme–substrate complex. The significant dipoles and other charged areas in the enzyme interact with the substrate, weakening bonds within the substrate molecules. Reaction therefore occurs more readily than it would without the enzyme. The products detach from the enzyme, leaving it unchanged. This model is a simplification of a very complex process.

Interactive CW Animation
Enzyme catalysis

QuickTime CW Video
Catalysis

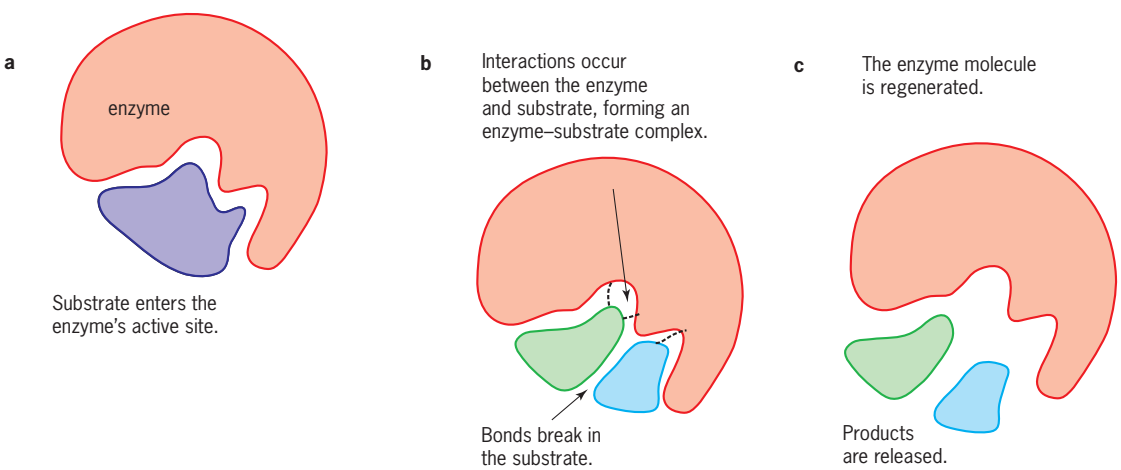


Figure 2.7.1 The sequence of events as an enzyme interacts with a substrate according to the lock and key model.

A more recent model of enzyme action which improves on the lock and key model is the **induced-fit model**. In this model, the enzyme changes shape as the substrate binds to the active site. Since enzymes are rather flexible structures, the active site is continually reshaped by interactions with the substrate. This enables a more precise fit to be achieved between the enzyme and substrate. When the products leave the enzyme, it returns to its original form.

QuickTime CW Video
Energy and enzymes movie

Pr
PRAC 2.5
Investigating sucrase

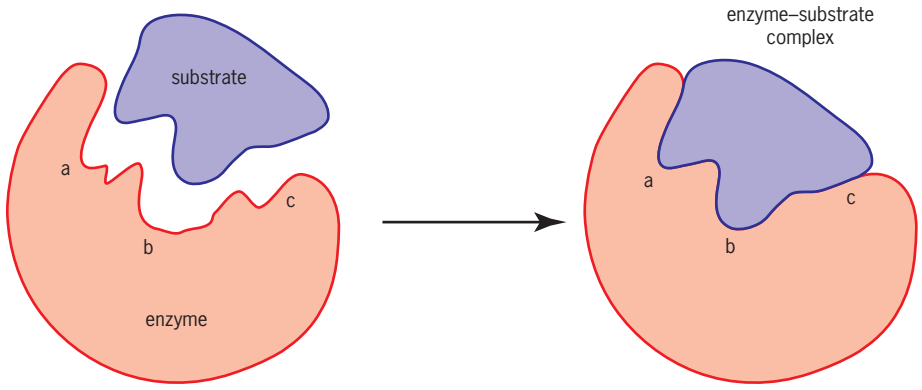


Figure 2.7.2 The induced-fit model of enzyme action.

Enzyme effectiveness is influenced by temperature. An increase in temperature causes an increase in reaction rate (as it does for most chemical reactions), but at elevated temperatures the rate of enzyme-catalysed reactions drops dramatically. At these elevated temperatures the enzyme is denatured, changes shape and so loses its catalytic power. Similarly, enzyme structure, and therefore enzyme action, may be altered by changes in pH. At different pH values the charges on amino acid side groups in the enzyme may alter, affecting the bonds between them, and so denaturing the enzyme. The presence of heavy-metal ions can permanently alter the tertiary structure of an enzyme. Heavy metals such as Ag^+, Hg^{2+}, Pb^{2+} have strong affinities for –SH groups and replace the hydrogen atoms in these groups. As the –SH group is part of the side chain of the amino acid cysteine, it is present in many enzymes, which may then be affected by heavy metals.

AS **B.7.7**
State and explain the effects of heavy-metal ions, temperature changes and pH changes on enzyme activity. © IBO 2007

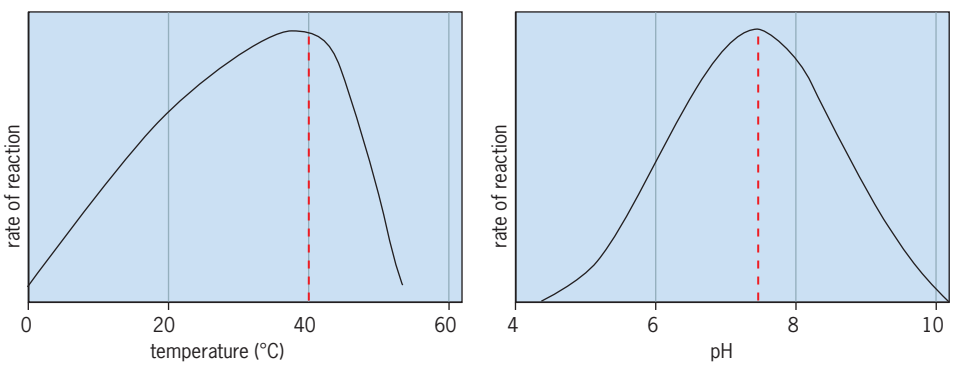


Figure 2.7.3 The activity of enzymes is influenced by changes in temperature and pH.

Each of the enzymes in the body has a function to perform. With so many enzymes, it should come as no surprise that sometimes one or more enzymes are either missing or not performing as they should. When this happens, problems usually result. The significance of the problem varies. One example is Tay–Sachs disease, a fatal inherited disease in which babies are born without the ability to produce hexosaminidase, the enzyme necessary to break down fats in the brain and blood.

Enzymes are also used to catalyse industrial processes. For example, enzymes in yeasts have been used for many hundreds of years in the production of alcohol and bread. Rennet is an enzyme that can be obtained from the stomachs of young cows. Rennet reacts with casein in milk, causing the milk to curdle. This is the first stage of the cheese-making process. Rennet can now be made from yeast, hence some cheeses can be labelled 'vegetarian'. Enzymes are added to detergents to catalyse the breakdown of protein stains, such as blood and grass stains in clothes.

AS **B.7.6**
Compare competitive inhibition and non-competitive inhibition. © IBO 2007

Inhibition of enzymes occurs when a substance prevents the enzyme from doing its job. These inhibitors work in one of two ways.

Competitive inhibitors attach to the active site of the enzyme, blocking the substrate from doing so. The inhibitor has a close structural similarity to substrate, so it binds to the active site of the enzyme. If the inhibition is truly competitive, the binding of substrate and inhibitor are mutually exclusive. (The substrate cannot bind if the inhibitor has already bound to the active site.)

Non-competitive inhibitors also bind to the enzyme, but not at the active site. A non-competitive inhibitor may bind to a free enzyme or even to an enzyme–substrate complex. The substrate and the inhibitor do not mutually exclude each other, but the presence of the inhibitor will reduce the ability of the enzyme to work efficiently. Binding elsewhere causes distortion of the enzyme's three-dimensional shape, reducing or completely removing the enzyme's catalytic activity when the distortion is transmitted to the active site, or the inhibitor overlaps the active site. Many heavy metals such as lead, mercury and chromium will function as non-competitive inhibitors.

These inhibitors are important in the regulation of enzyme activity within cells and may also be exploited in the production of drugs. For example, tamoxifen is a hormone treatment that lowers the risk of breast cancer recurring. It is an inhibitor molecule that fits neatly onto the estrogen receptors of breast cells, preventing estrogen molecules from doing so. Estrogen's interaction with breast cells is thought to be responsible for the development of breast cancer. If tamoxifen can inhibit the reaction of estrogen, breast cancer can be prevented.

AS **B.7.3**
Describe the relationship between substrate concentration and enzyme activity. © IBO 2007

Enzymes cannot act at a distance; the enzyme and substrate must combine to form the enzyme–substrate complex ES:

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

Three rate constants are needed to describe the system. The reaction in which an enzyme and its substrate form an enzyme–substrate complex has a rate constant k_1, and that in which the enzyme–substrate complex releases the products and a free enzyme has the rate constant k_3. The reaction in which the enzyme–substrate complex releases the substrate and leaves the enzyme free has the rate constant k_2.

In most cases the enzyme is present in extremely low concentrations (10^{-8} mol dm^{-3}) while the substrate is in large excess. This leads to a steady-state approximation, i.e. we assume that the rate of formation of the

enzyme–substrate complex, ES, from free enzyme and substrate is exactly balanced by the rate of conversion of ES into free enzyme and product—the concentration of ES remains essentially constant. The high concentration of substrate ensures that the enzyme remains constantly active.

The Michaelis–Menten equation for the initial reaction rate of a single substrate with an enzyme is

$$V = \frac{V_{max}}{1 + K_m / [S]} = \frac{V_{max}[S]}{K_m + [S]}$$

where [S] = substrate concentration

K_m = the Michaelis constant = $\frac{(k_2 + k_3)}{k_1}$

V = rate of conversion of substrate to products

V_{max} = the maximum rate of conversion of substrate to products (enzyme activity)

This equation provides a relationship between the rate of conversion observed at any particular substrate concentration and the maximum rate of conversion that would be achieved at infinite substrate concentrations.

AS B.7.4
Determine V_{max} and the value of the Michaelis constant (K_m) by graphical means and explain its significance. © IBO 2007

V_{max} and K_m are often referred to as the kinetic parameters of an enzyme, and their determination is an important part of the characterization of any enzyme. The Michaelis constant, K_m, is equal to the substrate concentration at which the rate of conversion of substrate to product is half of the maximum rate possible. It indicates the affinity of the enzyme for a substrate, or the **enzyme activity**. When K_m is small, there is a high affinity for the substrate (high enzyme activity) and V_{max} will be approached quickly. As K_m increases, the affinity of the enzyme for the substrate *decreases*.

A plot of the rate of conversion, V, against substrate concentration, [S], gives a value of V_{max}. The value of V approaches V_{max} slowly and asymptotically.

[S] = K_m when $V = \frac{V_{max}}{2}$.

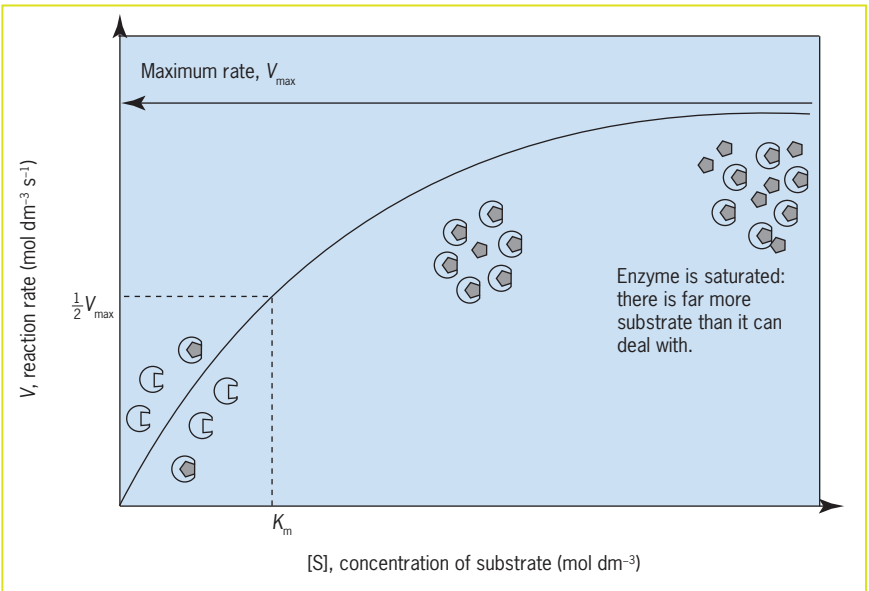


Figure 2.7.4 The relationship between rate of reaction, V, and substrate concentration, [S].

The presence of inhibitors has a significant effect on the maximum rate of conversion of substrate to products and the Michaelis constant (K_m), which indicates enzyme activity. A competitive inhibitor will allow the enzyme to reach the same maximum rate of conversion; however, it will take much longer for this to be achieved. The slope of the graph is much less than for the uninhibited reaction. Since V_{max} is the same for the uninhibited reaction and that inhibited by a competitive inhibitor, $\frac{1}{2} V_{max}$ will also be equal; however, the gentler slope of the inhibited graph means that its K_m will have a greater value. The enzyme affinity will be reduced. This is shown in figure 2.7.5. $K_m(3)$ is the value of K_m for a reaction influenced by a competitive inhibitor.

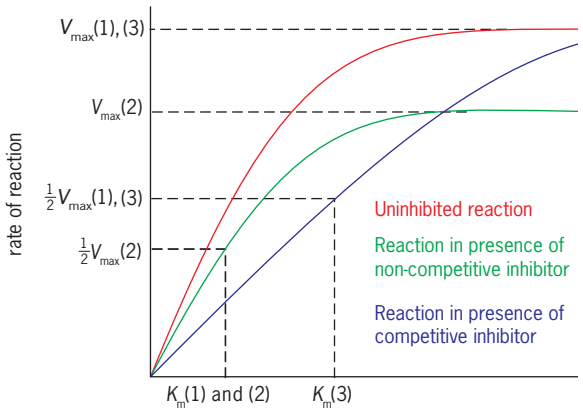


Figure 2.7.5 The effect of substrate concentration on activity of enzyme under the influence of inhibitors.

In comparison, a non-competitive inhibitor will cause a reduction in the maximum rate of conversion, V_{max}. The value of $\frac{1}{2} V_{max}$ will consequently be lower, but the slope of the graph will be such that its K_m will be equal to K_m for the uninhibited reaction. The enzyme activity will be the same for the uninhibited reaction and the reaction inhibited by a non-competitive inhibitor. This is shown in figure 2.7.5 where $K_m(1)$ and $K_m(2)$ are equal, but $V_{max}(2)$ is less than $V_{max}(1)$.

TABLE 2.7.1 EFFECT OF INHIBITORS ON ENZYME ACTIVITY

Inhibitor	V_{max}	K_m
Competitive inhibitor	Equal to V_{max} of uninhibited enzyme	Greater than K_m of uninhibited enzyme
Non-competitive inhibitor	Less than V_{max} of uninhibited enzyme	Equal to K_m of uninhibited enzyme

Section 2.7 Exercises

1 **a** Describe the general function of enzymes in the human body.

b State two differences between an enzyme and an inorganic catalyst.

c Explain what is meant by the primary structure of an enzyme.

2 An important enzyme-catalysed reaction in the body is the decomposition of carbonic acid (H_2CO_3) to carbon dioxide and water by the enzyme carbonic anhydrase.

a Carbonic acid is formed by the reaction of hydrogen carbonate ions with hydrogen ions. Write a balanced chemical equation for this reaction.

b Explain how the tertiary structure of an enzyme such as carbonic anhydrase determines its function.

c If a solution containing the carbonic anhydrase is maintained at 60°C, the enzyme is ineffective as a catalyst. Explain why this is so.

d State whether the primary structure is altered when the enzyme is heated to 60°C and explain your answer.

3 Explain, with reference to the active site, how enzymes are able to catalyse biological reactions.

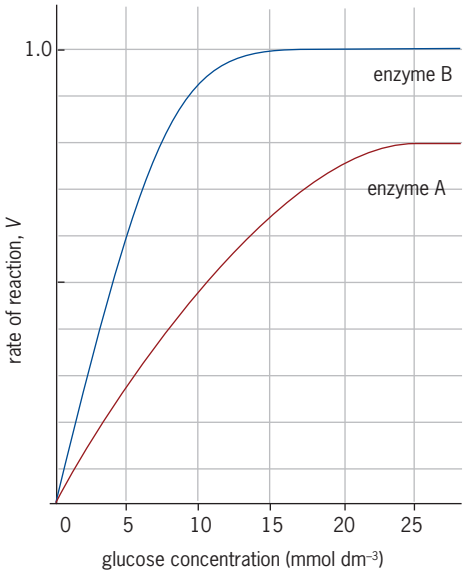
4 Two types of enzymes, enzyme A and enzyme B, catalyse the conversion of glucose to glucose-6-phosphate.

a Explain what is meant by the term K_m and state its relationship with the activity of an enzyme.

b Calculate the value of K_m for enzyme A from the graph.

c Deduce which enzyme has the greater value of K_m.

d Deduce which enzyme has the greater affinity for the substrate.



5 **a** Describe how the value of K_m will change for an enzyme in the presence of a:

i competitive inhibitor

ii non-competitive inhibitor.

b Copy the graph for enzyme A in question **4**, above, and draw a line on the graph to represent the effect of competitive inhibition.

6 Describe three changes to the surroundings of an enzyme that will affect the activity of an enzyme.

2.8 NUCLEIC ACIDS

HL

Chromosomes are found in the nuclei of cells. They carry genetic information and contain macromolecules called nucleic acids. Of all the biopolymers, the nuclei acids are possibly the most important because they are central to the amazing ability of living things to replicate their own kind. **Nucleic acids** are the genetic materials that allow the transmission of hereditary traits from one generation to the next.

Heredity is the set of characteristics you inherited from your parents. The basic heredity unit that controls a particular characteristic is a **gene**. Many thousands of genes are located in each of your body cells. Together, these genes may be thought of as a genetic program, a set of instructions that determine eye colour, body shape and the many features that make you what you are. Each gene is made of the chemical called deoxyribonucleic acid (DNA).

DNA and another nucleic acid called ribonucleic acid (RNA), are biopolymers with relative molar masses of up to 10^9 g mol^{-1} (larger than proteins, which have a maximum molar mass of about 10^6 g mol^{-1}). Both DNA and RNA are condensation polymers formed by monomers called nucleotides. Nucleotides consist of three components covalently bonded together: a pentose sugar (deoxyribose in DNA, ribose in RNA), a phosphate group and one of four nitrogen-containing organic bases (guanine, adenine, thymine and cytosine in DNA; guanine, adenine, uracil and cytosine in RNA). Note that the nitrogen-containing bases are more correctly called purines (adenine and guanine) and pyrimidines (cytosine, thymine and uracil).

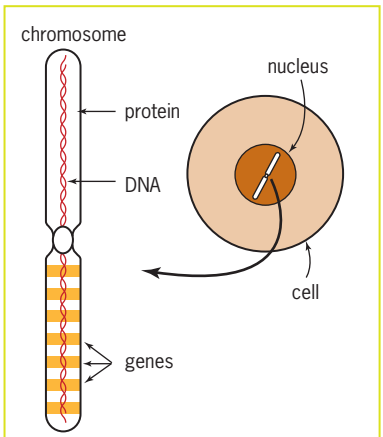
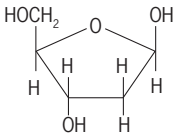
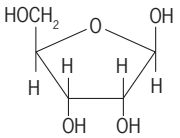
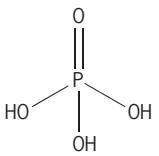
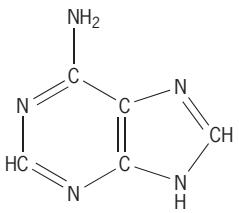
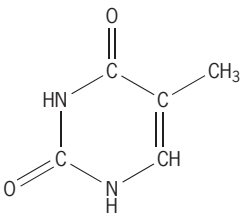
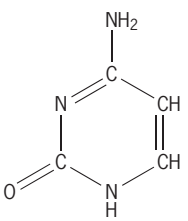
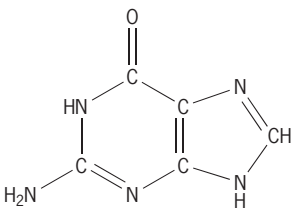
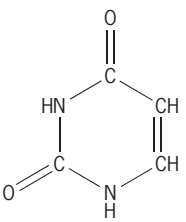


Figure 2.8.1 The relationship between genes, chromosomes and DNA.

AS **B.8.2 Distinguish between the structures of DNA and RNA. © IBO 2007**

AS **B.8.1 Describe the structure of nucleotides and their condensation polymers (nucleic acids or polynucleotides). © IBO 2007**

TABLE 2.8.1 COMPONENTS OF DNA AND RNA

Nucleotide component	Structure	Class of component	Found in
Deoxyribose		Monosaccharide (pentose)	DNA only
Ribose		Monosaccharide (pentose)	RNA only
Phosphate		Phosphate	DNA and RNA
Adenine		Purine	DNA and RNA
Thymine		Pyrimidine	DNA only
Cytosine		Pyrimidine	DNA and RNA
Guanine		Purine	DNA and RNA
Uracil		Pyrimidine	RNA only

Condensation reactions between these components produce a polynucleotide or nucleic acid. A nitrogenous base undergoes a condensation reaction with a sugar to form a nucleoside, which in turn undergoes a condensation reaction with a phosphate group to produce a nucleotide. Further condensation reactions between the phosphate group of one nucleotide and the sugar of another forms a chain of nucleotides—a polynucleotide. In each condensation reaction a molecule of water is formed.

As polymerization continues, a backbone of alternating sugar and phosphate groups is formed, with the nitrogenous bases pointing out from the backbone. Genetically, these nitrogen bases are the critical components, because it is the base sequence along the polymer (i.e. its primary structure) that is the key to the storage of genetic information.

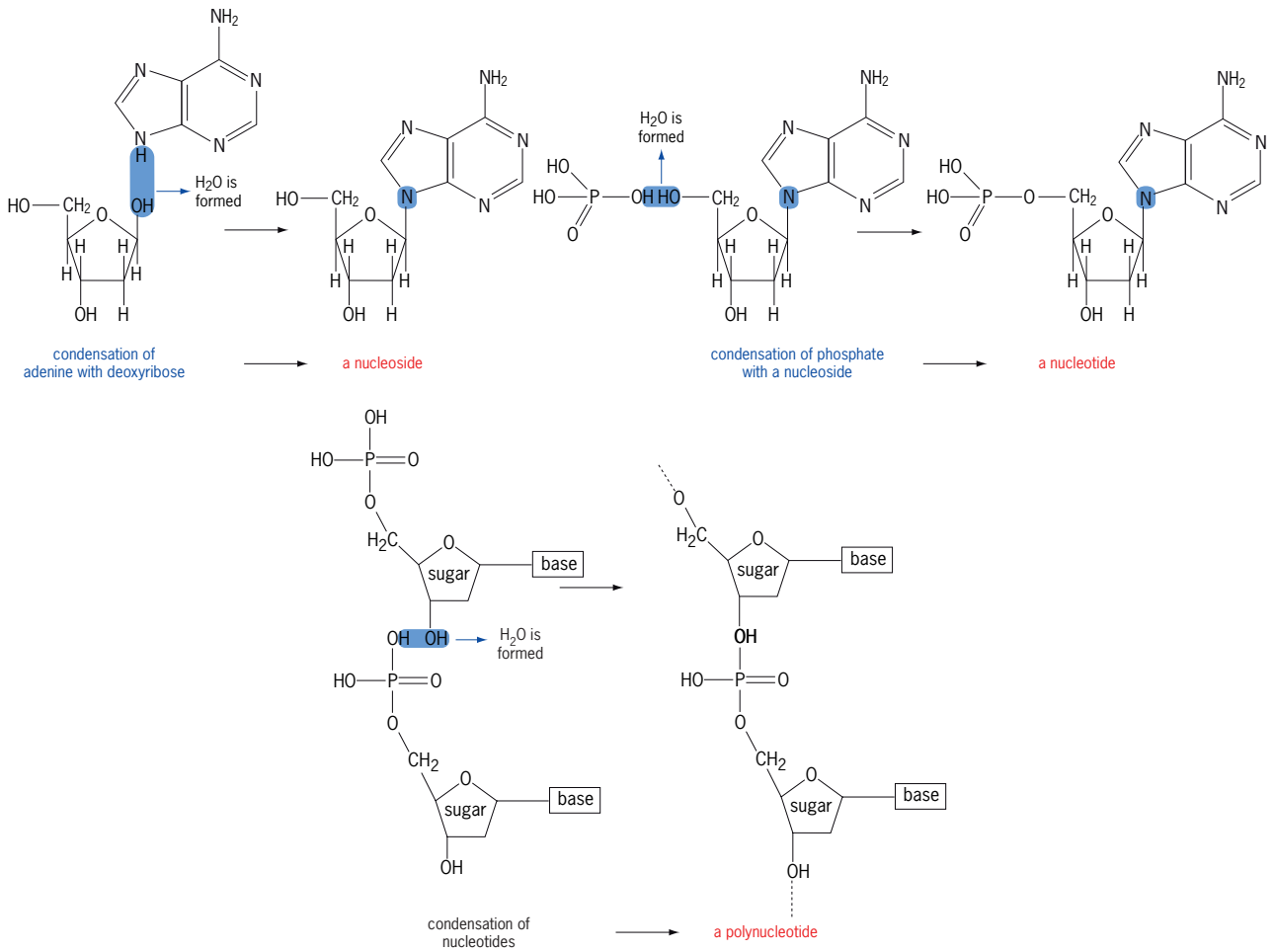


Figure 2.8.2 Condensation reactions result in the formation of a polynucleotide.

AS **B.8.3**
Explain the double helical structure of DNA. © IBO 2007

WORKSHEET 2.4
DNA structure and function

The secondary structure of DNA is that of a **double helix**. The polynucleotide chain winds into a helical shape, with approximately ten nucleotides per complete turn. Two helices are held together by hydrogen bonding to give the characteristic double helix structure. The hydrogen bonding occurs between the nitrogenous bases and is quite specific (see figure 2.8.3). Purines only bond with pyrimidines. This means that cytosine only bonds with guanine, and adenine only bonds with thymine (or uracil in RNA). The specificity of this complementary base pairing is largely due to the sizes of the bases. Only these combinations have the appropriate lengths to bridge the two backbones of the DNA. The strong hydrogen bonding between bases and the twisting sugar–phosphate backbone results in the 'twisted ladder' structure of DNA. The specific base pairings are vital for the genetic role of DNA.

In contrast, most RNA molecules are single stranded; but the long molecule folds upon itself, forming hairpin loops and helical sections, in which base pairs (cytosine with guanine and adenine with uracil) hydrogen bond.

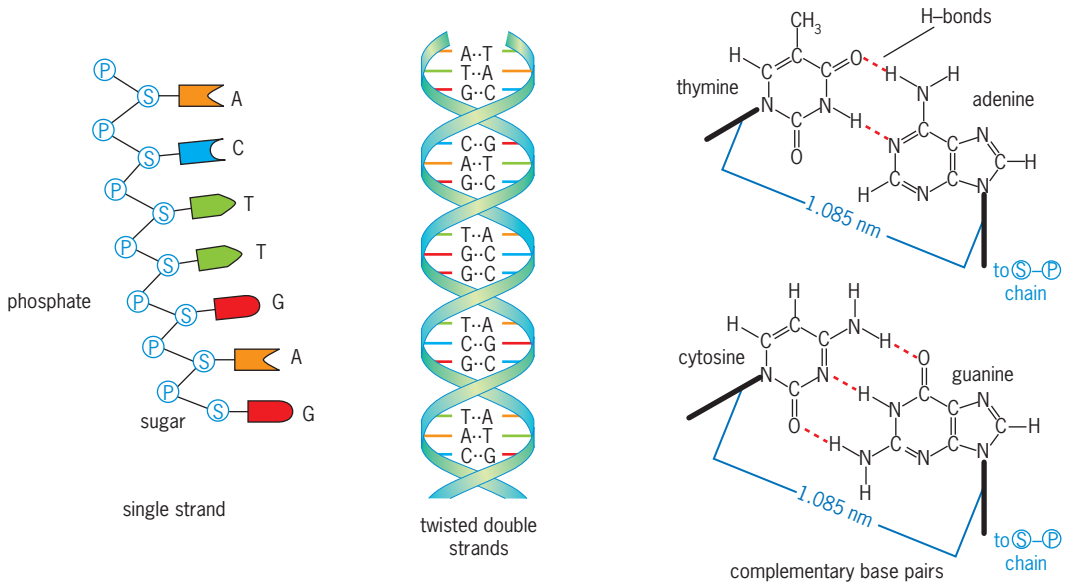


Figure 2.8.3 Hydrogen bonding between base pairs holds the two strands of DNA together.

AS **B.8.4**
Describe the role of DNA as the repository of genetic information and explain its role in protein synthesis. © IBO 2007

DNA and protein synthesis

DNA performs two functions: it reproduces itself prior to mitosis, and it carries the coded information necessary to produce proteins. When cells divide, the genetic information must be copied in a process called **replication**. The DNA strands first 'unzip' as the hydrogen bonds break, new nucleotides come into position, complementing the pattern of nitrogen bases on the original strands of DNA, new sugar–phosphate backbones form and, finally, the pairing of nitrogenous bases results in the formation of two new DNA molecules, each identical to the original.

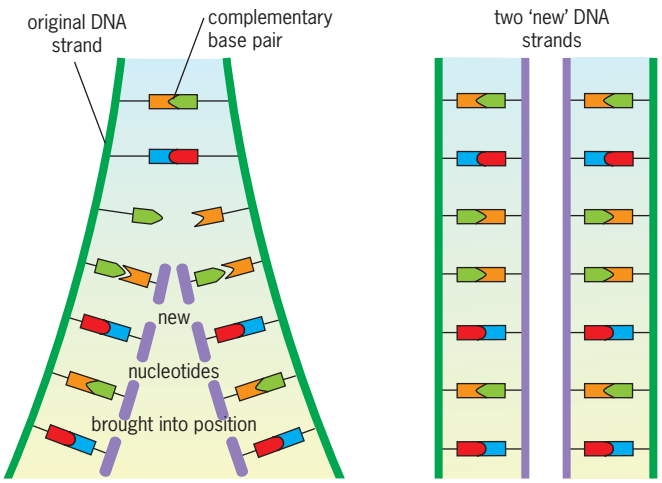


Figure 2.8.4 DNA copies itself in the process known as replication.

The coded genetic information stored in DNA lies in its sequence of bases. Sets of three bases along the length of the DNA strand, called a **triplet code**, provide information that allows the required amino acids to be assembled to make up the proteins of the organism. Sixty-four triplet codes occur (there are 64 possible permutations of the three-base sequences), with each of these triplet codes 'coding for' an amino acid. For example, the triplet code GGC (guanine–guanine–cytosine) codes for glycine, while CGG codes for arginine. Several different triplet codes may represent the same amino acid. The order of triplet codes along a length of DNA therefore codes for the order of amino acids along a length of protein. Different combinations of the 20 amino acids create thousands of different proteins, each with a distinct role to play in the organism.

A look at the cell structure indicates a problem here. DNA is found within the nucleus, while protein synthesis takes place in the ribosomes which are in the cytoplasm—outside the nucleus. DNA cannot pass through the nuclear wall, so how is the coded information passed? A messenger molecule, messenger ribonucleic acid (mRNA), performs that role. In a process called **transcription**, RNA is produced from DNA. RNA has a similar structure to DNA but is single-stranded and can pass into the cytoplasm, where it binds to ribosomes.

The 'final' step in using the DNA code is to produce proteins. In this process, called **translation**, a transfer RNA molecule (tRNA) 'collects' the appropriate amino acid from the surrounding cytoplasm, then base-pairs with the mRNA, adding its amino acid to the growing protein chain. By this rather complex sequence of events, the genetic code carried by the DNA molecule directs the types of proteins synthesized within the cell. All steps—replication, transcription and translation—rely heavily on the specific hydrogen bonding between complementary base pairs, and on condensation reactions between the sugar, phosphate and nitrogenous bases.

CHEM COMPLEMENT

Why a triplet code?

The genetic code must specify 20 amino acids. If the four bases were used in a 'doublet' code, there would only be 16 possibilities—not enough. If the four bases were used in a 'Quartet' code, there would be 256 possibilities—too many. While 64 combinations, with the four bases used in a 'triplet' code, are not exactly 'just right', they do code for the 20 amino acids without a huge level of redundancy. In 2000, the complete sequence of bases in human DNA, known as the human genome, was finally determined.

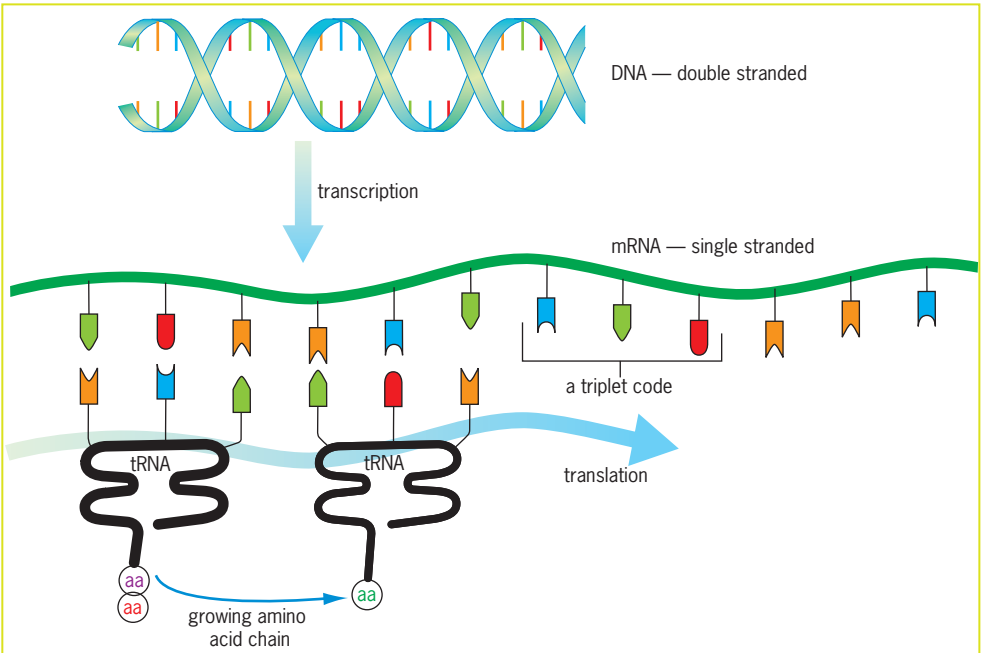


Figure 2.8.5 RNA carries the genetic code from the nucleus to the ribosomes where protein synthesis occurs.

DNA profiling

AS B.8.5 Outline the steps involved in DNA profiling and state its use. © IBO 2007

The base sequence in the nucleus of every one of your billions of cells is identical. However, unless you have an identical twin, your base sequence is different from that of every other person. In 1985, English geneticist Alec Jeffreys invented a technique called **DNA profiling** (or **DNA fingerprinting**), which made use of the uniqueness of each individual's DNA. The use of this technique in criminal investigations is now well known.

DNA profiling is a particularly powerful technique because DNA:

- can be extracted from almost any part of the body, be it blood (the white blood cells, not red blood cells), saliva or tissue
- is resilient; useful examples can be obtained from long-dead organisms
- can be extracted from minute samples
- is inherited from parents, and therefore can be used to determine the paternity of children
- can help to prove someone's innocence or guilt in a crime. By comparing the DNA of a suspect with DNA found at the scene of a crime, that person's innocence or guilt might be established.

DNA profiling does not compare the entire DNA molecule of the 'suspect' with the entire DNA molecule in question. The DNA molecule is too long and complex for this. However, it has particular sections that are more useful than others. The process then is to break the DNA molecules into fragments and to separate these fragments according to the sequence of bases they contain. The process of making a DNA profile is outlined in figure 2.8.6.

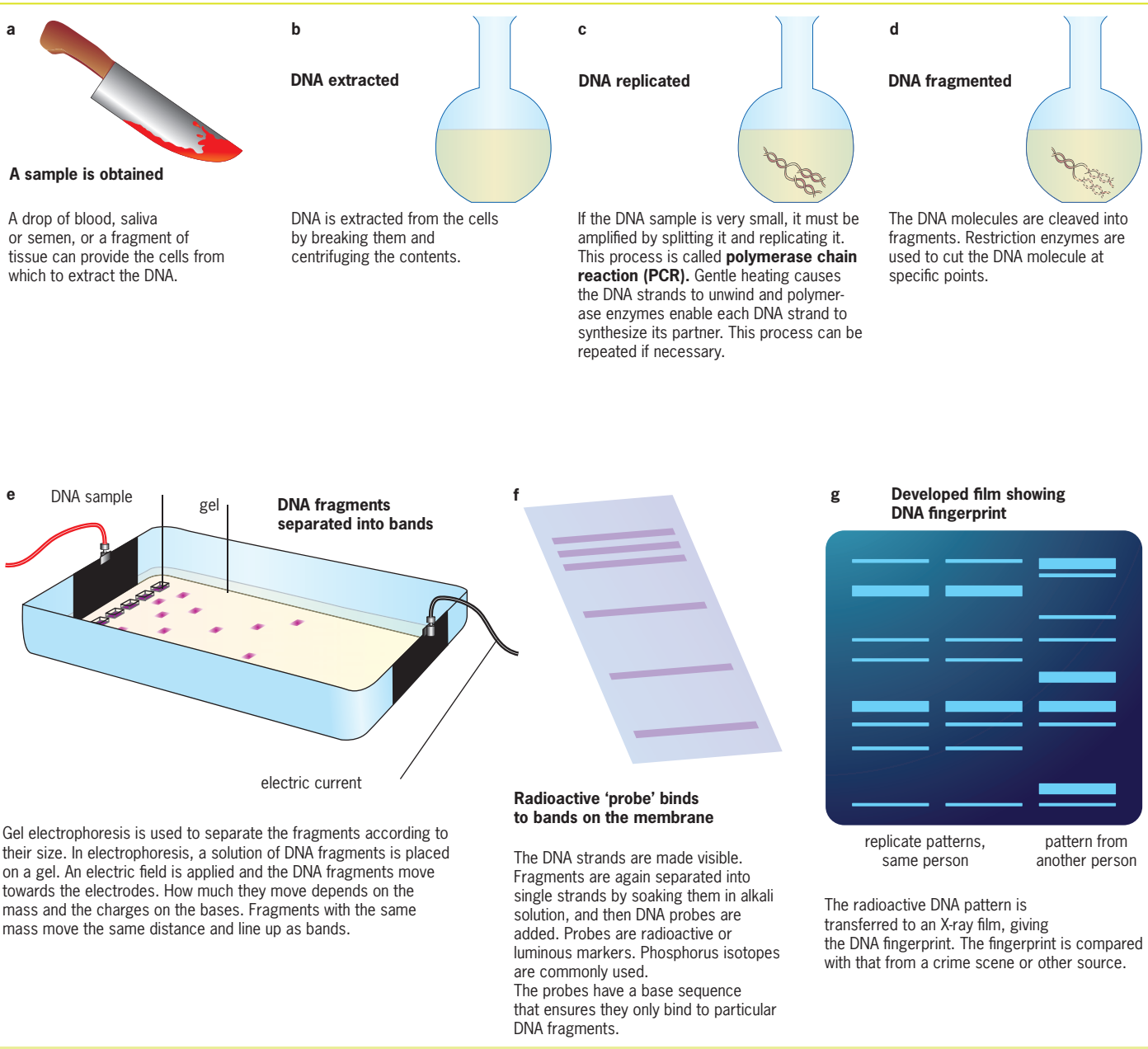


Figure 2.8.6 DNA profiling.

A DNA profile does not deliver absolute certainty, but it does produce a high chance of a correct conclusion. The more probes used (see figure 2.8.6), the more certain the result. The first successful criminal prosecution based on DNA evidence was that of Colin Pitchfork in England in 1987. After two teenagers were murdered in a village in England, police requested the entire male population to submit to DNA testing. Colin Pitchfork tried to get a friend to submit a blood sample for him, but then confessed when this did not work. He was found guilty based on DNA evidence.

Figure 2.8.7 DNA profiles are compared.

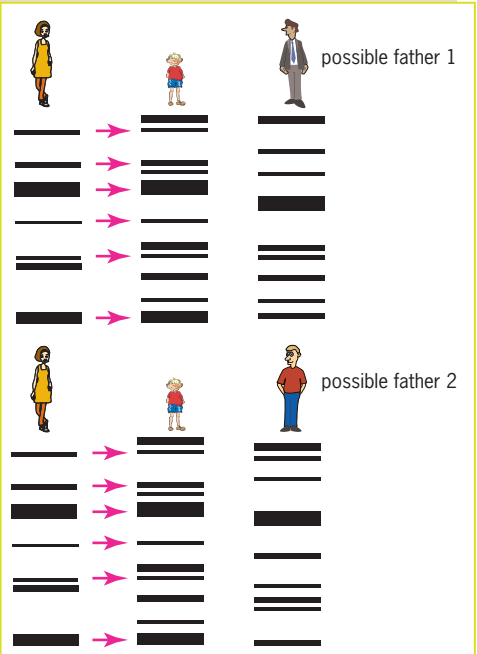


Figure 2.8.8 DNA fingerprinting is used in paternity testing. Who is the father?

Since then, the use of DNA testing has been increasing each year and has extended to such areas as:

- paternity testing or other relationship testing
- authenticating foods such as fish and wines (It is common practice to substitute less expensive fish cuts for more expensive ones.)
- matching organ donors with recipients
- detecting bacteria and other organisms that are present in low proportions
- determining pedigree for livestock breeders
- authenticating souvenirs.

DNA testing can also be used in reverse to prove someone's innocence. David Milgaard was convicted in Manitoba, Canada, in 1969 of rape and murder. When DNA testing became accepted by the legal system, he was exonerated and released after spending 23 years in jail, as his DNA did not match the stored semen samples from the crime scene. DNA testing also led to the arrest of the real culprit, whose DNA was on file from an earlier conviction.

There are several instances of famous royal dynasties in history coming to an end with conjecture over whether there really were hidden heirs to the throne. For example, the son of Louis XVI of France was proven, from preserved remains, to have died in a French prison, contrary to the suggestion that he had been spirited away and replaced with someone else. In 1991, DNA testing allowed pathologists to confirm that remains found in Siberia were indeed those of the Romanov family. Nicholas Romanov was the last tsar of Russia and his family disappeared in 1918 during the Russian Revolution. Several people have had their DNA tested against the remains to see if one of Romanov's daughters really survived the execution.

DNA profiling is also used in paternity cases. Figure 2.8.8 shows the DNA fingerprints of two men in dispute over the paternity of a child. The DNA results for the child and mother are shown and arrows indicate where the child's DNA matches the mother's. Which 'father' has DNA matching the child's?

THEORY OF KNOWLEDGE

A detective arriving at a crime scene finds a small piece of blood-stained cloth on the ground, two metres away from a murder victim. The cloth does not match the clothing the victim is wearing. The detective picks up the cloth with a pair of forceps and places it in a bag labelled 'biological evidence'. Several months later a DNA fingerprint of the blood will be presented in court as evidence in a murder trial.

There is no doubt that DNA fingerprinting has become a reliable source of evidence and since its introduction has been used not only as a way of securing convictions but also to exclude suspects who might otherwise be falsely charged with and convicted of serious crimes. However, the results of DNA fingerprinting are not 100% certain. Before we can consider how trustworthy this type of evidence is, we need to consider some of the issues surrounding its reliability.

The term *DNA fingerprint* implies that, like a fingerprint, the pattern for a given person is unique to them. Actually, a DNA fingerprint just presents a probability that the person in question is indeed the person to whom the pattern belongs. Typically, evidence in court is presented by way of a 'match probability', the probability that the defendant's DNA matches that of the crime sample. Two of the methods of DNA fingerprinting are RFLP (restriction fragment

length polymorphism) and PCR (polymerase chain reaction). The RFLP method gives a much higher probability of a match, but analysis of a sample can take a couple of months, a larger amount of sample (thousands of cells) is needed and the cells have to be fresh. The match with PCR is less probable, but requires a smaller sample of about 50 cells, analysis can be done in a couple of days and the sample does not have to be fresh, it can be decades old. Both PCR and RFLP can only analyse nuclear DNA, DNA found in the nucleus of a cell. However bone, hair and teeth do not have cells with nuclei, so another process involving the analysis of mitochondrial DNA (mtDNA) is used. In the investigation of murder cases that have gone unsolved for many years, where bone, hair and teeth may be all that remain, mtDNA is used.

The actual process of creating a DNA fingerprint is time consuming, and requires meticulous laboratory work. Correct protocols need to be followed so that the sample does not become contaminated. The quality of the lab doing the analysis, and its techniques and hygiene standards can affect the accuracy in the results. This is especially problematic in countries where there are no nationwide standards for the testing of DNA.

Crime scene evidence that has been altered, tampered with or destroyed is not admissible in court, so the method by which DNA evidence is collected and transported is important. When crime scene investigators collect evidence, they have a duty to make sure that the evidence is not compromised. The quality of the sample depends on the competence of these investigators because if the DNA sample is not in perfect condition it can affect the accuracy of the analysis. Extreme heat from fire and other physical conditions at a crime scene can also damage a sample, making it hard to analyse.

So how much can we rely on the results of DNA fingerprinting presented as evidence in court?

Consider the following questions.

- Can DNA evidence be considered strong enough justification to convince a jury beyond reasonable doubt?
- A key factor that distinguishes acceptable from unacceptable justification is reliability. How can the reliability of DNA fingerprinting be established?
- Evidence in criminal law is any item (e.g. weapon, document, clothing, artefact, DNA fingerprint) or testimony (e.g. oral or written statement) that assists in the proof of innocence or guilt. Which types of evidence do you consider to be the most trustworthy in court and why?
- Because DNA fingerprinting is a product of 'science', do you think there might be a tendency for members of a jury to consider the results more reliable than other types of evidence? In what way, if any, do you think the prior beliefs of a jury might be altered in the light of results presented by DNA fingerprinting?
- In court expert witnesses are called to interpret DNA fingerprints. How does the reliability of the evidence presented by a forensic scientist compare with that of a non-expert eye witness? What are the strengths and weaknesses of both forms of evidence?
- DNA fingerprinting provides a probability that the person in question is the person to whom the pattern belongs. Consider that the probability of two individuals having a match is 1 in 1000 with PCR and 1 in 100 000 with RFLP. Would the PCR method be reliable enough to convince a jury beyond reasonable doubt? What about the results from RFLP? Which probabilities would be acceptable to prove guilt?

Section 2.8 Exercises

1 DNA is a polynucleotide.

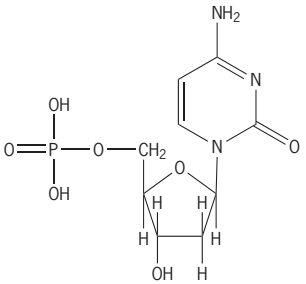
a Refer to figure 2.8.2 for this question. Explain what is meant by a:

i nucleoside **ii** nucleotide.

b Explain how one nucleotide may differ from another.

2 Distinguish between the structures of DNA and RNA.

3 **a** Identify the three components—sugar, base and phosphate—in the nucleotide shown to the right.

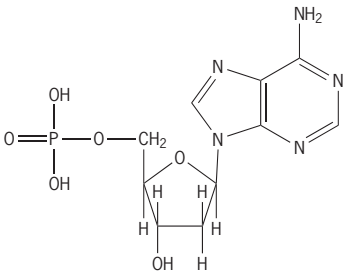


b The base sequence along one strand of DNA is CGTACG. Deduce the sequence of the complementary strand of:

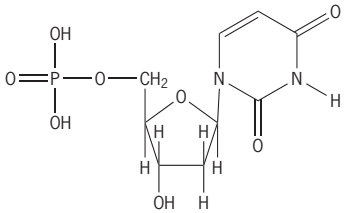
i DNA **ii** RNA

4 For each of the nucleotides shown below, state whether they would be found in DNA or RNA, both DNA and RNA, or in neither.

a



b



5 Explain how the two strands of DNA are held together in a double helix.

6 DNA controls protein synthesis, yet DNA resides within the cell nucleus, and protein synthesis occurs outside the nucleus. Explain how DNA exerts its control over protein synthesis.

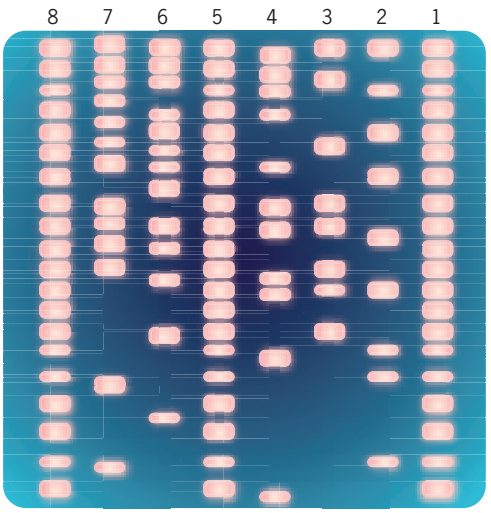
7 State which nucleic acid or acids are involved in the process of:

a transcription

b translation.

8 State three uses for DNA profiling.

9 In the DNA profile shown, three 'lanes' (1, 5 and 8) contain RNA fragments of a virus. The other lanes contain DNA from 'suspects' in criminal and paternity cases. Suggest why lanes 1, 5 and 8 were included in the experiment.



10 Refer to figure 2.8.6 for this question.

a In DNA profiling, state what the letters PCR stand for.

b Suggest why a DNA sample might be subjected to PCR before DNA fingerprinting.

2.9 RESPIRATION

HL

AS **B.9.1**
Compare aerobic and anaerobic respiration of glucose in terms of oxidation/reduction and energy released. © IBO 2007

Glucose is used by cells to obtain energy by the process known as respiration. **Aerobic respiration** requires that the glucose is oxidized by oxygen, whereas anaerobic respiration uses another organic compound as the oxidizing agent. In humans **anaerobic respiration** often occurs in muscles during prolonged exercises. It results in the build up of lactic acid that is accompanied by painful muscular cramping. In yeast, anaerobic respiration produces ethanol and carbon dioxide—a process widely used to produce alcoholic beverages.

The reactions involved in respiration are not simple and involve a large number of steps; however, the overall equations can be used to show the difference in energy being released in the three processes.

The overall equation for aerobic respiration is

$$C_6H_{12}O_6 + 6O_2 \rightarrow 6CO_2 + 6H_2O \qquad \Delta H = -2860 \text{ kJ mol}^{-1}$$

While that for anaerobic respiration in humans is

$$C_6H_{12}O_6 \rightarrow 2C_3H_6O_3 \qquad \Delta H = -120 \text{ kJ mol}^{-1}$$

And in yeast the equation for anaerobic respiration is

$$C_6H_{12}O_6 \rightarrow 2C_2H_5OH + 2CO_2 \qquad \Delta H = -69 \text{ kJ mol}^{-1}$$

Although anaerobic respiration yields less energy per mole of glucose, it is often a quicker process than aerobic respiration.

Aerobic respiration is a redox reaction. The glucose is oxidized by oxygen and oxygen is reduced. This can be seen if we assign oxidation numbers:

ON $\overset{0}{C_6}\overset{+1}{H_{12}}\overset{-2}{O_6} + 6\overset{0}{O_2} \rightarrow 6\overset{+4-2}{CO_2} + 6\overset{+1}{H_2}\overset{-2}{O}$

Similarly anaerobic respiration in yeast is a redox reaction in which the glucose is both oxidized and reduced.

ON $\overset{0}{C_6}\overset{+1}{H_{12}}\overset{-2}{O_6} \rightarrow 2\overset{-2}{C_2}\overset{+1}{H_5}\overset{-2}{O}H + 2\overset{+4-2}{CO_2}$

Anaerobic respiration in humans is not a redox reaction, as there is no change in the oxidation number of any of the elements.

Aerobic respiration consists of three main sequences of chemical reactions.

1 *Glycolysis:* In this series of reactions the six-carbon glucose molecules are converted into two three-carbon pyruvate molecules. Importantly, an additional product is adenosine triphosphate (ATP), which is used as the chemical storage of the energy that is released from glucose during the reactions.

   This may be simplified as

   $$\underset{\text{glucose}}{C_6H_{12}O_6} + 2NAD^+ \rightarrow \underset{\text{pyruvate}}{2C_3H_4O_3} + + 2NADH + 2H^+$$

2 *The TCA (tricarboxylic acid) cycle*. During this cycle two-carbon groups from pyruvate are broken down into carbon dioxide. In addition, ATP and the reduced coenzymes NADH and $FADH_2$ are also produced.

3 *The electron transport (or cytochrome) chain*. NADH molecules from glycolysis and the TCA cycle pass electrons to the chain. Energy released as the electrons are passed down the chain is used to make ATP. Electrons leaving the chain combine with oxygen and hydrogen ions to form water.

Note that detailed equations are not needed for these steps, only the basic ideas need be known.

Glycolysis is also the first step of anaerobic respiration. In humans the pyruvate that is formed during anaerobic glycolysis is converted to lactic acid, whereas in yeast it is converted to ethanal and then to ethanol and carbon dioxide. When a human muscle cell operates anaerobically, it incurs an oxygen debt. This is because the lactic acid still needs to be oxidized back to pyruvate and the electrons passed back to the electron transport chain. To repay this oxygen debt, we continue to breathe hard after we have stopped using our muscles strenuously.

Iron-containing molecules called **cytochromes** are the principal electron-transfer molecules in the electron transport chain (the third stage in respiration). Cytochromes consist of an iron complex (called a heme complex) and a protein. The structure of the heme complex is shown in figure 2.9.1. The planar ligand to which the iron(II) ion is coordinated is called a porphyrin.

AS **B.9.2**
Outline the role of copper ions in electron transport and iron ions in oxygen transport.

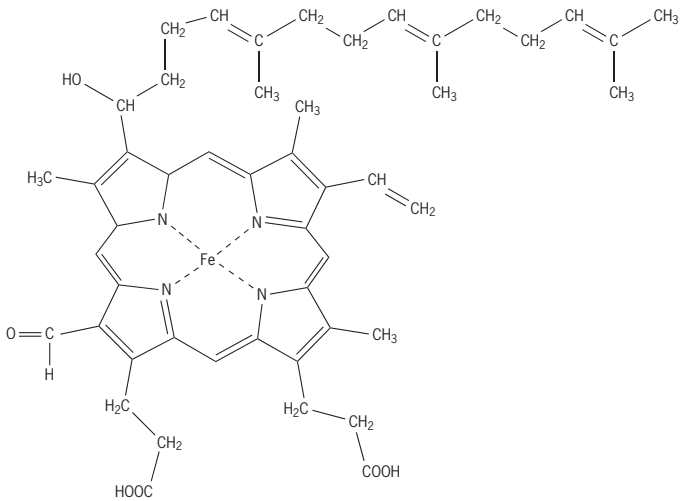


Figure 2.9.1 The heme group of cytochrome oxidase.

Copper(II) ions are also used in aerobic respiration during which glucose is changed into energy (ATP). One step involves the transfer of electrons to cytochromes. As electrons pass from one molecule to the next, copper ions are alternately oxidized and reduced.

$$Cu^{2+} + e^- \rightleftharpoons Cu^+$$

The interconversions of amino acids also require electrons and the oxidation of copper(I) to copper(II) provides these electrons.

Iron also plays a principal role in the transport and storage of oxygen in mammalian blood and tissues. Oxygen is stored in myoglobin, which is made up of a heme complex and a protein in a structure that is very similar to that of cytochromes.

The transport of oxygen in the blood is carried out by hemoglobin. This is made up of four myoglobin-like units and so can bind four oxygen molecules. The hexacoordinate Fe^{2+} ion is coordinated to four nitrogen atoms of the porphyrin ring and to a nitrogen atom of the protein chain. The sixth coordination position is occupied by oxygen. When hemoglobin is carrying oxygen it is a bright red colour due to the splitting effect of the oxygen molecules on the 3d orbitals of Fe^{2+}. When the oxygen molecules are released, water molecules occupy the sixth coordination position and the colour of the complex becomes bluish since the splitting of the 3d orbitals by the water ligand is less than that of oxygen. This results in blood that has already delivered its oxygen to cells having a bluish tint.

Interactive CW Animation
Heme group

Section 2.9 Exercises

1 Compare aerobic and anaerobic respiration of glucose in terms of the energy released by the processes.

2 Describe the aerobic respiration of glucose in terms of the oxidation and reduction of the reactants.

3 Food is oxidized by a series of redox reactions involving the transport of electrons. Identify the ions of two different metals used in these reactions.

4 The structure of the heme group in cytochrome oxidase is shown in figure 2.9.1. Identify the type of bonding between the four nitrogen atoms and the iron.

5 Outline the role of iron ions in oxygen transport in the human body.

Chapter review questions and tests are available on student CD.

Chapter 2 Summary

Terms and definitions

2-Amino acid A carboxylic acid molecule with an amino functional group on carbon number 2 and with one of 20 different side groups also bonded to carbon number 2.

Aerobic respiration The oxidation of glucose to obtain energy, in the presence of oxygen.

Amylopectin A form of starch that occurs as branched chains of α-glucose units.

Amylose A form of starch that occurs as relatively straight, unbranched chains of α-glucose units.

Anabolic steroids A class of steriod hormones related to testosterone, which mimic the action of testosterone.

Anaerobic respiration The oxidation of glucose in the absence of oxygen.

Bomb calorimeter A sealed metal container or 'bomb' in which the food may be combusted.

Calorific value Energy content of a food, usually measured in kJ g^{-1} or kJ mol^{-1}.

Carbohydrate A group of compounds with the empirical formula $C_x(H_2O)_y$, where x and y are whole numbers.

Cellulose A polysaccharide that is the main structural material of plant cell walls. It cannot be digested by humans.

Cholesterol A sterol that is produced by the liver and is found in all body tissues, where it helps to control the permeability of cell membranes.

Chromatography The separation of components of a mixture according to their relative attraction to the stationary and mobile phases.

Chromosomes Structures found in the nuclei of cells which contain the genetic material of an organism.

Co-factor A substance that needs to be present in addition to an enzyme for a certain reaction to take place.

Competitive inhibitor An inhibitor that blocks an enzyme from doing its job because it has a close structural similarity to the substrate and attaches to the active site of the enzyme, preventing the substrate from doing so.

Cytochromes Iron-containing molecules that are the principal electron-transfer molecules in the electron transport chain. They consist of an iron or copper complex and a protein.

Denaturation The destruction of the three-dimensional shape of a protein by physical and chemical agents.

Dietary fibre Predominantly cellulose-based plant material that is not hydrolysed by enzymes secreted by the human digestive tract.

Dipeptide A molecule made by the condensation reaction between two amino acids.

Disaccharide A molecule made by the condensation reaction between two monosaccharides.

DNA profiling (fingerprinting) The process in which DNA molecules are broken into fragments that are then separated according to the sequence of bases they contain. This enables the comparison of one or more DNA sample.

Double helix Two DNA strands held together by hydrogen bonding between the nitrogenous bases to give the characteristic double helix structure.

Electrophoresis The separation of amino acids, proteins or nucleic acids according to their charge and their mass by the application of a potential difference.

Enzyme A biological catalyst that enables biochemical reactions to progress by a lower activation energy pathway.

Enzyme activity A measure of the attraction of the enzyme for a substrate.

Ether linkage An ether functional group (–O–) that joins two monosaccharides in a disaccharide or polysaccharide.

Fats A food group made up of triglycerides.

Fatty acid A carboxylic acid molecule with a long (greater than 8 carbons) hydrocarbon tail.

Gene The basic heredity unit that controls a particular characteristic.

Glycerol ($C_3H_8O_3$) A tri-ol that reacts with three fatty acid molecules to make a triglyceride.

Glycogen The energy storage polysaccharide in animals.

Glycolipids Lipid bound to carbohydrate.

Glycolysis The first step of respiration in which pyruvate is made.

Glycosidic linkage See *ether linkage*.

High density lipoproteins (HDL) Cholesterol-carrying lipoproteins that are small and dense due to a large proportion of protein; also known as 'good' cholesterol.

High performance liquid chromatography (HPLC) A type of column chromatography that uses a solid stationary phase and a liquid mobile phase under pressure.

Hormone A 'chemical messenger' or regulatory compound that is released by cells of one tissue and acts to regulate the metabolism of target cells in another tissue.

Hydrolysis The addition of water to a bond to break the bond and form two molecules (often catalysed by enzymes).

Induced-fit model A model of enzyme action in which the enzyme changes shape as the substrate binds to the active site.

Iodine number The mass of iodine that reacts with 100 g of a lipid.

Isoelectric point (of an amino acid) The pH at which an amino acid occurs as a zwitterion. At this pH the amino acid carries no net electrical charge.

Lipids A group of fat-soluble naturally occurring molecules such as fats, oils, waxes, cholesterol, sterols, fat-soluble vitamins and phospholipids.

Lipophilic A substance that dissolves in non-polar solvents. Literally meaning 'fat-loving'.

Lipoproteins A biochemical assembly that contains lipids bound to proteins. They are used to transport cholesterol in the blood (among other uses).

Low density lipoproteins (LDL) Cholesterol-carrying lipoproteins that are larger and less dense than HDL. LDL molecules contain much more cholesterol than HDL molecules; also known as 'bad' cholesterol.

Macronutrients Substances that make up more than 0.005% of our body weight; includes proteins, fats and carbohydrates.

Menstrual cycle The recurring cycle of physiological changes involving ovulation and associated changes that enable pregnancy to occur.

Micronutrients Substances that make up less than 0.005% of our body weight; includes vitamins and trace minerals.

Monosaccharide The simplest type of carbohydrate, with the empirical formula CH_2O.

Mono-unsaturated fatty acid A fatty acid with one carbon–carbon double bond.

Ninhydrin An organic dye that is sprayed onto colourless parts of a mixture and gives them a purple colour. Commonly used to detect amino acids.

Non-competitive inhibitor A substance that binds to the enzyme, but not at the active site, and prevents an enzyme from doing its job.

Nucleic acid The genetic material that allows the transmission of hereditary traits from one generation to the next. It is a polymer that has a backbone of alternating sugar and phosphate groups, with the nitrogenous bases pointing out from the backbone.

Peptide link An amide functional group that joins two amino acid residues in a dipeptide, tripeptide or polypeptide (protein).

Phospholipid A molecule with a hydrophilic (polar) head consisting of a phosphate group and a hydrophobic tail made up of two long hydrocarbon chains. They are one of the major components of all biological membranes.

Polypeptide A condensation polymer formed by the reaction between many amino acids.

Polysaccharide A condensation polymer formed by the reaction between many monosaccharides.

Polyunsaturated fatty acid A fatty acid with more than one carbon–carbon double bond.

Primary structure (of protein) The number of amino acids, the identity of those amino acids and the arrangement of the amino acids in the protein chain.

Protein A condensation polymer made up of more than 150 amino acids.

Quaternary structure (of protein) The interaction of separate polypeptide chains to give further complexity to the structure and form the 'working shape' of the protein.

Replication The copying of DNA in preparation for mitosis.

Retention (retardation) factor (R_f) A measure of how far a component of a mixture has travelled in comparison to the solvent during paper or thin-layer chromatography:

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

Saturated fatty acid A fatty acid with a single carbon–carbon bond.

Secondary structure (of protein) Bonding (hydrogen, dipole–dipole) of amino acids in the chain to form α-helices or β-pleated sheets.

Sex hormones (testosterone, progesterone and estradiol) Hormones that control sexual function in males and females, such as ovulation, production of sperm, and support of pregnancy.

Starch The plant energy storage polysaccharide.

Steroid A lipid characterized by a carbon skeleton with four fused rings, generally arranged in a 6-6-6-5 fashion.

Sterol An alcohol with a fused system of rings.

Tertiary structure (of protein) Folding to form the particular three-dimensional shape that is responsible for the unique function of that protein. Structure is due to hydrogen bonds, van der Waals' forces between non-polar side groups, ionic attractions between ionized side groups, ion–dipole attractions or covalent bonds (disulfide bridges) that form when sulfur-containing side groups react.

Transcription The production of RNA from DNA.

Translation The process in which a transfer RNA molecule 'collects' the appropriate amino acid from the surrounding cytoplasm, then base-pairs with the mRNA, adding its amino acid to the growing protein chain.

Triglyceride A very large, non-polar molecule that is insoluble in water and is formed by the condensation reaction between three fatty acids and glycerol.

Tripeptide A molecule made by the condensation reaction between three amino acids.

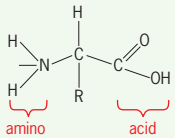
Triplet code The combination of three organic bases on a strand of DNA which codes for a particular amino acid to be incorporated in a protein cell.

Vitamins Chemicals that are vital to the normal functioning of an animal's metabolism and that the animal cannot synthesize itself.

Zwitterion An ion that has both a negative charge and a positive charge on it at the same time.

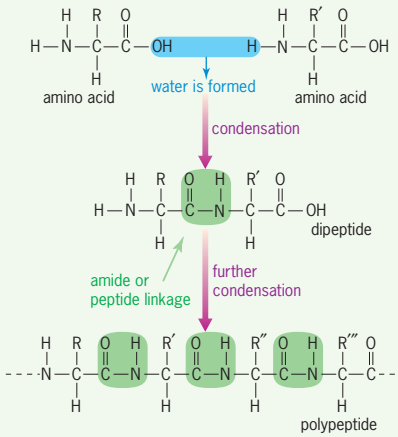
Concepts

- Biochemistry is the study of the chemistry of living things and the reactions occurring within organisms.
- Proteins are an important class of biopolymer. They have various functions in organisms, including as structural material, biological catalysts and energy sources.
- The monomer of proteins is the α-amino acid. There are 20 different amino acids in proteins. Amino acids are both acidic and basic.

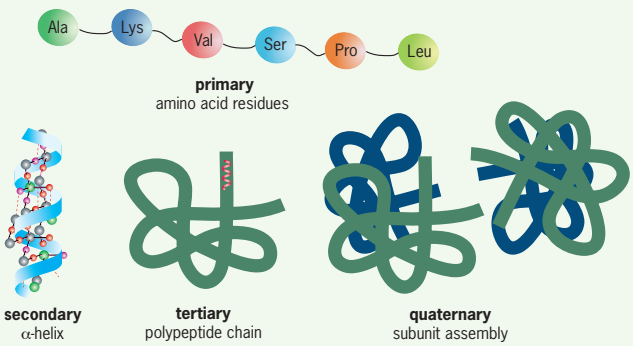


2–amino acid

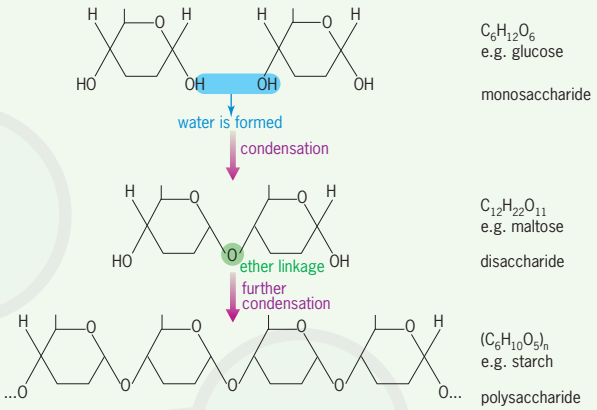
- Amino acids undergo a polymerization reaction to form dipeptides, and further condensation reactions to form polypeptides.



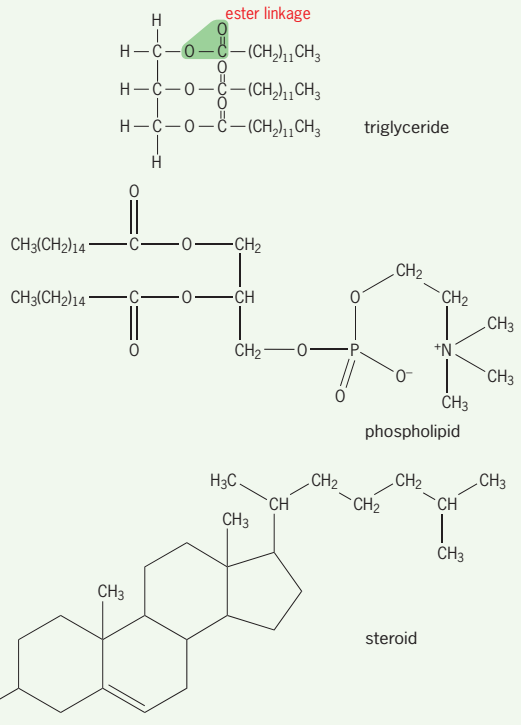
- Proteins have various levels of structure. The complex three-dimensional protein shape is important to its function.



- Denaturation of proteins—the disruption of bonds in the secondary, tertiary and quaternary structures—occurs as a result of heating or of adding chemicals.
- Carbohydrates ($C_x(H_2O)_y$) are an important class of biochemical compounds. They include monosaccharides, disaccharides and polysaccharides. Glucose ($C_6H_{12}O_6$) is the major energy source for most organisms.
- Glucose and other monosaccharides undergo condensation reactions to produce disaccharides. Glucose undergoes condensation polymerization to produce the polymers starch (an energy store), cellulose (a structural material) and glycogen (an energy store in animals).



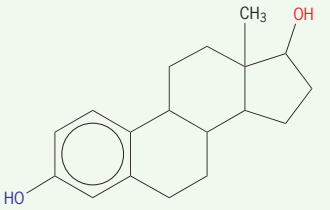
- Higher levels of HDL cholesterol are preferable to LDL cholesterol as LDL cholesterol is linked to cardiovascular disease.
- Calculations of iodine number can be used to determine the degree of saturation of a fatty acid.
- Saturated fatty acids pack together more closely and can bond more effectively, and so have higher melting points and are more likely to be solids than unsaturated fatty acids in which the double bonds create 'kinks' in the hydrocarbon chain.
- Lipids perform a number of important roles in the human body, including energy storage, insulation and protection of organs, as steroid hormones and as structural elements of cell membranes. Examples of lipids include triglycerides, phospholipids and steroids.



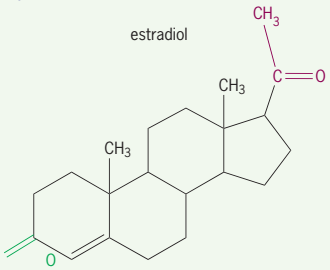
- Vitamins are examples of micronutrients that are vital to digestion and other important bodily processes.
- Vitamins may be fat-soluble (vitamins A, D, E and K) or water-soluble (vitamins B and C), depending on their structure. Water-soluble vitamins have many polar hydroxyl groups; fat-soluble vitamins tend to have long non-polar hydrocarbon chains.
- Deficiencies of vitamins and other micronutrients may cause a range of harmful conditions such as anemia, goitre, night blindness, pellagra, beriberi, scurvy and rickets.
- Deficiencies of macronutrients such as protein in the diet results in malnutrition and associated conditions such as kwashiorkor and marasmus.

- Hormones are chemical messengers that are secreted directly into the blood by endocrine glands.

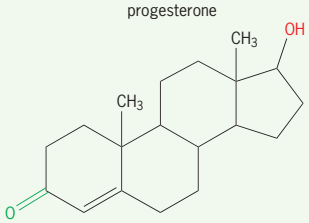
  The structures of cholesterol and the sex hormones are similar due to their common steroid backbone.



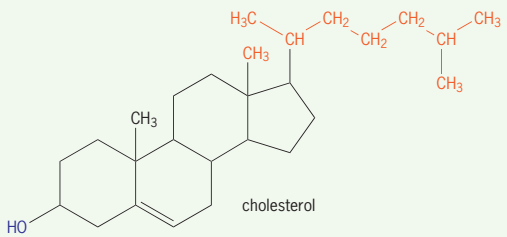
estradiol



progesterone



testosterone

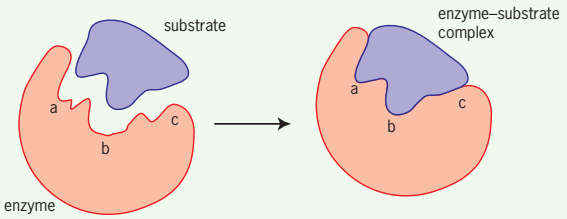


cholesterol

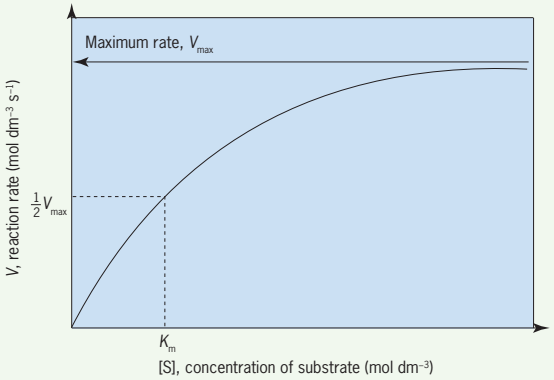
- Enzymes are a specialized group of proteins that act as biological catalysts. They are highly specific, act within narrow pH and temperature ranges and increase reaction rates dramatically.

HL

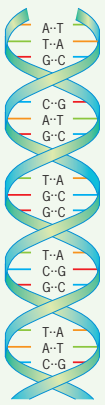
- Enzymes are thought to operate according to the induced-fit model.



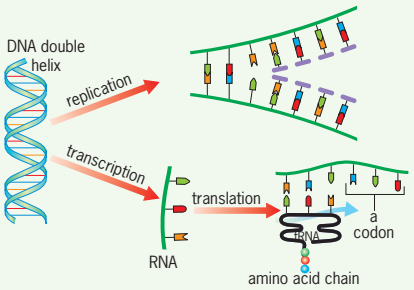
- V_{max} and K_m may be determined by graphical means since the graph approaches V_{max} asymptotically and K_m = [S] at $\frac{1}{2}V_{max}$.



- Nucleic acids are biopolymers that make up the genetic code of organisms. The monomer units of nucleic acids are nucleotides, composed of sugar, phosphate and nitrogenous base units. These components join together by condensation reactions.
- Nucleotides join by condensation reactions to form a long DNA chain. Two DNA chains join by hydrogen bonding between complementary base pairs to form a twisted ladder-like structure known as a double helix.



- DNA replicates (copies itself) in a process that relies on the specific hydrogen bonding between complementary base pairs. The sequence of bases along the DNA forms a genetic code. Sets of three bases (codons) code for the sequence of amino acids. Specific base pairing is critical in the processes of transcription and translation. By these processes, the genetic code carried in the DNA base sequence directs the production of proteins.



- DNA profiling is an identification technique that relies on the uniqueness (except between identical twins) of every individual's DNA base sequence. DNA fingerprinting involves a series of steps, leading to the production of a fingerprint that may be compared with that of a sample taken from a 'suspect' in a forensic or paternity case.
- More energy is released when glucose reacts in aerobic respiration than in anaerobic respiration, although pyruvate is made in both processes.
- Cytochromes are the principal electron-transfer molecules in the electron transport chain (the third stage in respiration). Cytochromes consist of an iron complex called a heme complex and a protein.
- Hemoglobin is responsible for the transport of oxygen in the blood. This is made up of four myoglobin-like units and can bind four oxygen molecules per hemoglobin unit.

3 CHEMISTRY IN INDUSTRY AND TECHNOLOGY

Chapter overview

This chapter covers the IB Chemistry syllabus Option C: Chemistry in Industry and Technology.

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

- identify iron sources
- describe the properties and uses of iron, steel, aluminium and aluminium alloys
- describe how iron ore is converted into steel
- explain how the properties of metals are modified through heat treatment and alloying
- describe how alumina is converted into aluminium
- discuss the environmental impact of iron and aluminium production
- compare alternative uses of oil
- compare three types of cracking
- describe how the properties of polymers depend on structural features and explain how their properties are modified
- discuss advantages and disadvantages of polymer use
- compare homogeneous and heterogeneous catalysts
- discuss factors affecting catalyst selection
- describe how rechargeable batteries and fuel cells work, and discuss their similarities and differences
- describe different types of liquid crystals and outline how liquid-crystal displays work
- define nanotechnology and discuss its implications
- describe the structure and properties of carbon nanotubes

- distinguish between addition and condensation polymers
- describe how condensation polymers are formed
- describe the mechanisms by which LDPE and HDPE are formed
- describe silicon doping and the interaction between sunlight and semiconductors
- identify molecules that show liquid-crystal states
- explain how twisted nematic liquid crystals work
- describe the liquid-crystal properties of Kevlar™
- discuss the production and uses of chlorine and sodium hydroxide, and the environmental impact of the chlor-alkali industry.

HL

IBO Assessment statements C.1.1 to C.12.3

Technology is the practical application of science, while **industry** refers to trading and manufacturing goods. In this chapter, we will explore modern industrial processes, such as metal extraction, oil refining and polymer synthesis. We will also look at the applications of technologies such as **liquid crystals**, which are used for display panels in laptop computers, watches and many other digital devices. No discussion of modern chemical technology would be complete without mentioning **nanotechnology**. The advent of nanotechnology heralded a new and exciting era in science. As yet, no one knows exactly where explorations in the nano-world will take us, but concepts that used to belong strictly to the realm of science fiction are now a reality. In terms of gross earnings, the chemical industry is easily outstripped by many other sectors, including mining, entertainment, banking and insurance. However, advances in chemistry have the power to change the way we live, from the way we make electricity and plastics to the way we treat disease. Can many other industries make the same claim?

Figure 3.0.1 Other industries could not function without the substances provided by the chemical industry.

3.1 IRON, STEEL AND ALUMINIUM

Aluminium and iron are the most abundant metals in the Earth's crust, occurring as compounds rather than pure metals. Iron has been in common use since ancient times, but aluminium has only become available in the last 200 years. The use of metals in history mirrors the relative difficulty involved in extracting them from their ores. Iron can be extracted by heat, but aluminium is a more reactive metal than iron—it can only be obtained through electrolysis, a technique first used to extract metals by Humphrey Davy in 1807.

Sources of iron

AS **C.1.1**
State the main sources of iron.
© IBO 2007

An *ore* is a mineral deposit that can be mined for profit; a *mineral* is a naturally occurring, inorganic, crystalline substance with definite physical and chemical properties. Metallic iron is found in meteorites, but iron mostly occurs in ionic form since it oxidizes easily. The most important iron ores are **haematite** (iron(III) oxide, Fe_2O_3) and **magnetite** (Fe_3O_4). Other iron minerals may also be mined, but greater processing is required as their iron content is relatively low. Limonite is a mixture of hydrated iron oxides, sometimes called brown haematite.

Most iron deposits consist of layers of iron oxides alternating with layers of chert, a rock rich in silica (SiO_2). In addition to silica, iron ore contains impurities such as sulfur, phosphorus, manganese and alumina (Al_2O_3). If iron pyrite (FeS_2) or **siderite** ($FeCO_3$) are used as ore, they must be roasted in air to drive off the sulfur and carbon dioxide, thus obtaining iron oxides for use in the **blast furnace**—the device used to extract iron from its ores. Note that the expense of mining and refining iron ores makes recycled scrap iron or steel a viable and commonly used source of iron for industrial use.

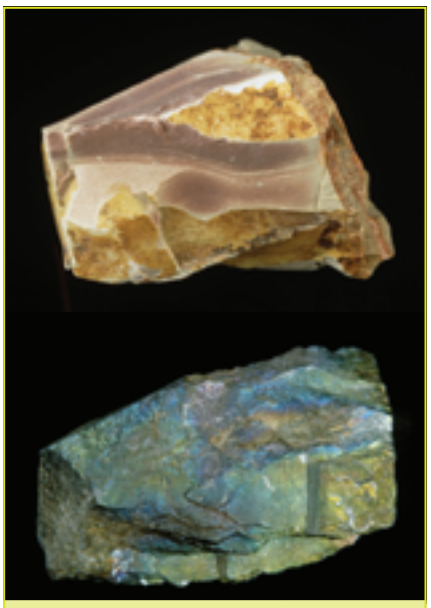


Figure 3.1.1 Haematite (top) and magnetite (bottom) are the most important iron ores.

The extraction of iron

A blast furnace is a huge, tapering steel tower, internally lined with *refractory* bricks. **Refractory materials** are able to withstand extremely high temperatures. The tower sits atop a crucible-style hearth in which the products collect. A blast furnace produces up to 10 000 tonnes of **pig iron** (crude iron) per day. If high-grade ore is used, it is crushed and placed directly into the blast furnace, but low-grade ore requires processing (called beneficiation) to increase its iron content. Gangue is the non-iron-containing part of the ore.

AS **C.1.2**
Describe and explain the reactions that occur in the blast furnace. © IBO 2007

The blast furnace is charged by feeding iron ore, limestone ($CaCO_3$) and **coke** continuously through the top. Coke is coal that has been heated in the absence of air to remove volatile components; its carbon content is greater than 85%. Blasts of hot air at around 700–1200°C are forced into the furnace at the bottom through small holes called **tuyères**. The hot air takes 6–8 seconds to reach the top of the tower and goes through a series of reactions as it rises.

Limestone is used for **slag** formation; slag is the molten, vitreous material that contains the waste from the **smelting** process. When heated, the limestone (calcium carbonate) forms lime (calcium oxide):

$$CaCO_3(s) \rightarrow CaO(s) + CO_2(g)$$

At very high temperatures, molten calcium oxide can react with molten silica, one of the primary impurities in the iron ore:

$$CaO(l) + SiO_2(l) \rightarrow CaSiO_3(l)$$

Smelting is a heating process in which chemical reactions occur to produce a metal from its oxide. The molten products separate into two or more layers. The two layers formed in the blast furnace are molten iron and slag.

In the blast furnace:

1 *The fuel is ignited.*

When the hot blasts of air are sent through the tuyères, the coke ignites in a highly exothermic reaction:

$$C(s) + O_2(g) \rightarrow CO_2(g) \qquad \Delta H = -394 \text{ kJ mol}^{-1}$$

The heat from this reaction increases the temperature at the bottom of the tower.

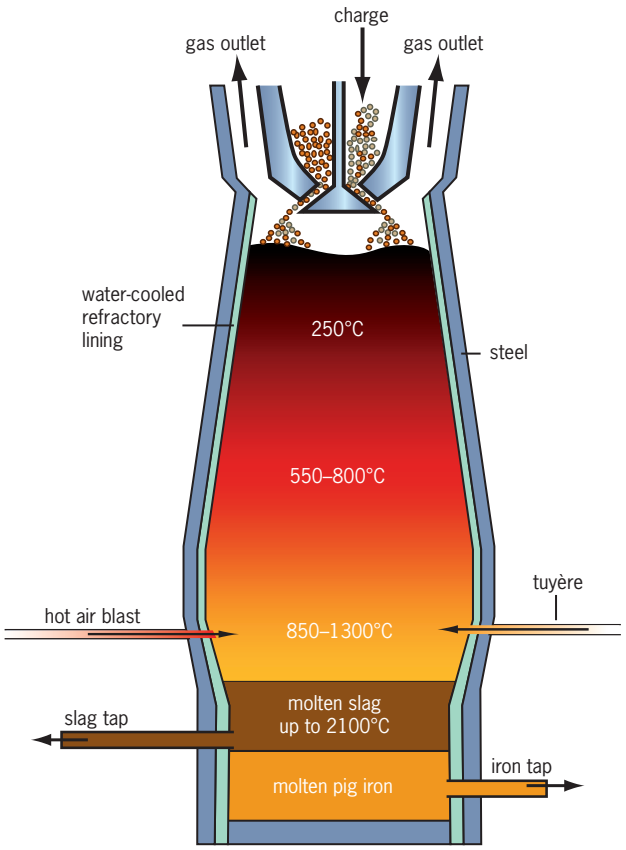


Figure 3.1.2 The blast furnace.

2 *The rising carbon dioxide reacts with more coke.*

In an endothermic reaction, carbon dioxide combines with coke to form carbon monoxide:

$CO_2(g) + C(s) \rightarrow 2CO(g) \qquad \Delta H = +173 \text{ kJ mol}^{-1}$

3 *The carbon monoxide reduces the iron ores.*

A series of reactions initiated at different temperatures reduce the iron ore to metallic iron:

$3Fe_2O_3(s) + CO(g) \rightarrow 2Fe_3O_4(s) + CO_2(g)$

This starts at 450°C.

$Fe_3O_4(s) + CO(g) \rightarrow 3FeO(s) + CO_2(g)$

This starts at 590°C.

$FeO(s) + CO(g) \rightarrow Fe(s) + CO_2(g)$

This starts at 700°C.

Another reaction that can occur is

$FeO(s) + C(s) \rightarrow Fe(s) + CO(g)$

Although the states of the iron and iron compounds have been written as solid, as they descend to hotter temperatures toward the bottom of the furnace, they soften. Iron melts at 1535°C; it trickles down and collects as a liquid in the hearth at the bottom of the tower. Of course, the pig iron formed is not pure; it still contains around 4% carbon from the coke. Iron with this carbon content melts at approximately 1100°C

If water vapour is present in the blast air, alternative reactions are possible. Coke could react with the water:

$C(s) + H_2O(g) \rightarrow CO(g) + H_2(g) \qquad \Delta H = +131 \text{ kJ mol}^{-1}$

The hydrogen could then reduce the iron ore:

$Fe_3O_4(s) + 4H_2(g) \rightarrow 3Fe(s) + 4H_2O(g) \quad \Delta H = +150 \text{ kJ mol}^{-1}$

4 *Slag forms.*

As previously noted, limestone forms lime as it is heated. In addition to reacting with silicates, lime can remove any remaining sulfur from iron compounds:

$FeS(s) + CaO(s) + C(s) \rightarrow CaS(s) + FeO(s) + CO(g)$

Figure 3.1.3 Molten pig iron is tapped off from a blast furnace.

The calcium sulfide becomes part of the slag. The slag is less dense than the molten iron and so floats in a layer on top of it. In doing so, it protects the iron from being oxidized by the incoming blast of air. Other impurities such as Al_2O_3 and MgO also dissolve into the slag, in reactions such as:

$CaO(l) + Al_2O_3(l) \rightarrow CaAl_2O_4(l)$

The iron and slag are periodically tapped off (removed) at different levels. Waste gases escape at the top of the furnace. Once tapped off, the pig iron is transported to on-site steelworks in large ladles, still in the molten state.

Changing iron into steel

About 60% of steel worldwide is produced by the basic oxygen process. 'Basic' refers to the alkaline nature of the slag and refractory lining. The process takes place in a vessel called a converter and can turn tonnes of pig iron into steel in less than an hour. The converter is lined with resin-bonded magnesia–carbon refractory bricks. It can be rotated 360° and operates at around 1600–1700°C.

AS C.1.3
Describe and explain the conversion of iron into steel using the basic oxygen converter. © IBO 2007

Pig iron contains too much carbon to be useful—it is hard and brittle. Its conversion to steel involves extracting some of the impurities from the pig iron, including most of the carbon. In the converter—a large furnace—the impurities are oxidized and extracted as either gases or slag. Pure oxygen, or oxygen mixed with argon, is used as the oxidizing agent. Air can't be used because nitrogen would react with iron, forming iron nitride.

The steps in the process are:

1 *The vessel is tilted 45° and charged.*

Hot, molten pig iron from the blast furnace is added, along with any scrap steel ready for recycling; 20–30% of the charge is made up of scrap metal.

2 *The vessel is returned to an upright position and fluxes are added.*

A lime (CaO) and dolomite (CaO and MgO) flux is dropped onto the charge. These materials form the slag, which gets rid of much of the silicon and phosphorus after they have been oxidized. The presence of MgO reduces the wear of the magnesia (MgO) refractory lining.

3 *The lance is lowered to a metre above the bottom of the vessel.*

Oxygen is blown into the molten metal using a water-cooled oxygen lance, causing the carbon impurities to react in exothermic redox reactions such as:

$$C(s) + O_2(g) \rightarrow CO_2(g)$$
$$2C(s) + O_2(g) \rightarrow 2CO(g)$$

The longer the oxygen is blown in, the lower the carbon content. Monitoring the composition of the exit gases allows control of the carbon content. Other impurities are also oxidized:

$$4P(l) + 5O_2(g) \rightarrow P_4O_{10}(l)$$
$$Si(l) + O_2(g) \rightarrow SiO_2(l)$$
$$2Mn(l) + O_2(g) \rightarrow 2MnO(l)$$

The oxygen is blown in at high pressure, producing a viscous mix of slag, metal and bubbles. The flux materials dissolve non-gaseous oxidized impurities:

$$CaO(l) + SiO_2(l) \rightarrow CaSiO_3(l)$$
$$6CaO(l) + P_4O_{10}(l) \rightarrow 2Ca_3(PO_4)_2(l)$$

Manganese can react with sulfur or iron(II) sulfide:

$$Mn(l) + S(l) \rightarrow MnS(l)$$
$$Mn(l) + FeS(l) \rightarrow MnS(l) + Fe(l)$$

Manganese(II) sulfide tends to dissolve in the slag, which forms a layer on top of the molten steel. The presence of manganese is important. Although iron(II) sulfide dissolves in hot steel, in solid steel it would crystallize out and weaken the alloy.

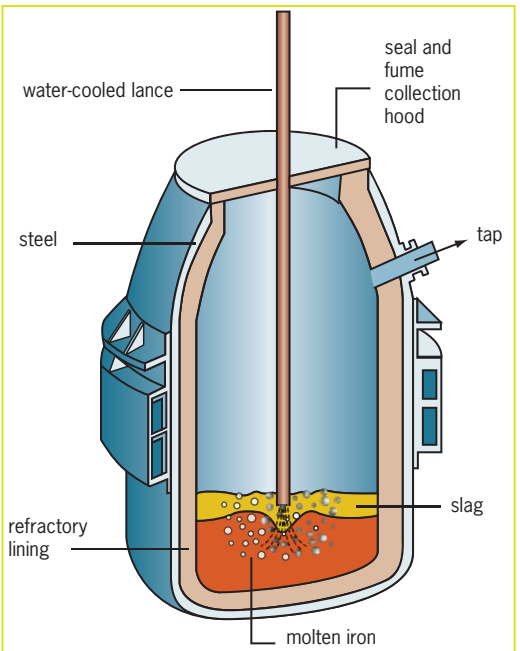


Figure 3.1.4 The basic oxygen converter.

4 *The lance is raised and samples are taken.*

The composition if the steel is analysed to see if it is ready for extraction, or if further oxidation is required.

5 *The steel is removed.*

Tilting the vessel allows the steel to be tapped off into large ladles. At this stage, various types of steel can be made by adding the alloying elements to the molten steel.

6 *The slag is removed.*

Tilting the vessel the other way allows the slag to be tapped off.

Alloys

AS **C.1.4 Describe alloys as a homogeneous mixture of metals or a mixture of a metal and non-metal. © IBO 2007**

In their pure form, most metals have few applications. Alloys are substances with metallic properties that are mixtures of at least two elements, at least one of which must be a metal. Most are homogeneous, meaning that the alloying element is evenly distributed throughout the structure. Table 3.1.1 shows three alloys that are a part of everyday life.

TABLE 3.1.1 COMMON ALLOYS

Alloy	Composition	Properties	Uses
Bronze	90% Cu 10% Sn	Strong, corrosion resistant, easily cast	Sculptures, church bells
Sterling silver	93% Ag 7% Cu	Harder than pure silver, unreactive, lustrous	Jewellery
Duralumin	95% Al 4% Cu 1% Mn and Mg	Low density, very strong	Aircraft bodies, racing bikes

AS **C.1.5 Explain how alloying can modify the properties of metals. © IBO 2007**

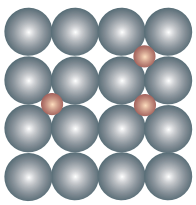


Figure 3.1.5 In steel, carbon resides in the interstitial spaces, blocking slip planes and making the alloy hard.

Generally, alloys are harder than their parent metals and less malleable, with lower melting points and electrical conductivity. Different-sized atoms or ions in the array block slip planes and prevent the metal ions from sliding easily over one another. The 'foreign' atoms and ions also inhibit the easy flow of delocalized electrons through the lattice, reducing electrical conductivity.

Alloying may produce marked changes in chemical properties. For example, the addition of chromium to iron in stainless steel significantly increases its resistance to chemical attack. Chromium forms a thin oxide layer that is impermeable to water and air. If the alloying element and parent metal's atoms are similarly sized, one element takes the place of the other in the structure. If there is a big difference in size, the smaller element fills the spaces between metal ions (interstitial spaces) in the lattice. The filling of interstitial spaces makes the alloy harder, stronger and less malleable.

Alloying iron with carbon to make steel significantly increases iron's hardness. There are more than 3500 different types of steel. Most have been developed in the last 20 years. Engineers like to say that if the Eiffel Tower was rebuilt today, only about a third as much construction material would be needed, thanks to the superior performance of modern steel. Table 3.1.2 shows the effect of adding various elements to steel. One important alloying effect is to increase the metal's **hardenability**—its response to the heat treatments discussed in the next section.

TABLE 3.1.2 EFFECT OF ELEMENT ADDITION ON THE PROPERTIES OF STEEL

Element	Effect on steel properties
Carbon	Harder, stronger, lower ductility and malleability
Manganese	Harder, stronger
Chromium	Corrosion and abrasion resistance
Titanium	High strength at high temperatures
Nickel	Harder (slightly), tougher, stronger, more corrosion and shock resistant
Molybdenum	Harder, stronger
Boron	Increases hardenability
Cobalt	Harder, stronger, improves heat resistance, prevents cracking and distortion

WORKSHEET 3.1
Analysis of iron in iron ore

Heat treatment of steel

C.1.6
Describe the effects of heat treatment of steel. © IBO 2007

If metals or alloys are heated and cooled in a controlled manner, their physical and chemical properties may be altered without changing their shape. This is called **heat treatment**, and steel responds particularly well to it. Metals, like most other materials, do not exist as single crystals; they are composed of many tiny crystals called grains. Close examination of the surface of a polished metal reveals these irregularly shaped grains, separated from each other by distinct boundaries. The properties of metals and alloys depend on the properties of the grain, and on the size, arrangement and interactions between grains.

PRAC 3.1
Modifying metals

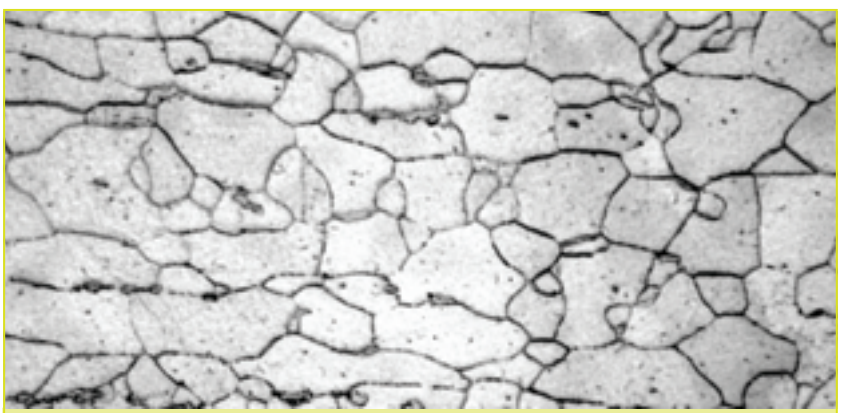


Figure 3.1.6 Metals and alloys are composed of interlocking grains.

A 'perfect' metal crystal with regular rows of cations would be weak because the rows could easily move past each other. Introducing imperfections into the array will decrease this 'slipping', and increase strength. Similarly, movement of ions does not readily occur across grain boundaries. Metals with many small grains have more grain boundaries per unit volume, making the metal stronger.

When steel is to undergo, or has undergone, a great deal of cold working, such as being drawn into wires, it needs to be softened. Softening increases the steel's ductility and reduces its strength, hardness and brittleness. **Annealing** and **tempering** are the two methods by which steel may be softened. **Quenching** is a method used to harden the steel.

Typical steel annealing consists of heating the item at 760–910°C (depending on carbon content) for a period of time that is determined by the item's thickness. The steel is then cooled at a rate of about 20°C per hour. As it cools, relatively large grains form. The result is softer, more ductile steel. Although recrystallization occurs, the alloy does not melt.

Quenching is the fast cooling of a hot, solid alloy in a liquid. This hardens the steel by trapping carbon atoms in certain positions. If quenching only produced hardening at the surface, it would not be effective. By alloying the steel with the right elements, its hardenability—the ease with which the entire structure can be hardened—can be improved. Quenching is normally carried out at room temperature. Quenching agents include water, salt water, oil and polymers.

Tempering is used to reduce the brittleness of quenched steel, and also to remove any internal stresses due to distorted crystal structures. Quenching and tempering can be alternated until the right degree of hardness, strength, ductility, toughness (measured by impact strength) and structural stability are attained. In tempering, the steel is heated to either 150–400°C, or 400–700°C. At the lower temperature, fine grains are formed and the material is still quite hard. At higher temperatures, coarser grains form and the material becomes softer. Thanks to these methods, and other steel treatment methods, it is now possible to fine-tune the properties of a piece of steel to make them closely match the intended use.

Properties of iron and steel

AS C.1.7
Describe the properties and uses of iron and steel.

In pure form, iron has limited use in organic synthesis reactions and metal-working. As ions, iron is essential to living things, particularly as part of hemoglobin, which carries oxygen around the body. As a structural material, steel is much more useful than iron. The major classes of steel are shown in table 3.1.3. Within these classes, there are innumerable variants.

TABLE 3.1.3 TYPES OF STEEL

Steel type	Composition	Properties	Examples of uses
Carbon steels	C 0.10–1.5% Mn 1.65% max Si 0.60% max Cu 0.60% max	• Lower carbon content: soft, malleable, ductile, low tensile strength • Higher carbon content: stronger, less ductile, lower melting point	Car body panels, shafts, axles, springs, wires, nails
Alloy steels	C 0.08–1.1% • Mn, Si and Cu content may exceed the limits above • Generally contain other alloying elements, for example Al, B, Cr (≤3.99%), Co, Ni, Ti, W	Depend on alloying elements, but in general: high hardenability, strength, toughness, ductility, machinability	Rail tracks (Mn-steel), permanent magnets (Si-steel), structural steel (Ni-steel)
Stainless steels	C 0.08–1.1% Cr 14–18% Ni 7–9%	Hard, corrosion and abrasion resistant, tough, good hardenability, shiny	Sinks, kitchen utensils, containers (e.g. in beer and wine-making)

Alloying can increase iron's tensile strength, hardness and resistance to corrosion, while lowering its melting point and ability to conduct heat and electricity. Note that tensile strength is the amount of pulling a fibre can take before it breaks.

CHEM COMPLEMENT

A brief history of iron

The word iron derives from the Scandinavian word iarn; its symbol, Fe, derives from the Latin ferrum. The Ancients used iron from meteorites, which contained 6–8% nickel. The Iron Age (approximately 1000 bc to 50 ad) usurped the Bronze Age (approximately 3000–1000 bc) when hotter furnaces were created. Bronze, an alloy of copper and tin with a melting point of 900–1000°C, was easy to extract and produce using wood fires. In contrast, the extraction of iron needed hotter charcoal fires. In a charcoal fire, the following reactions occur:

$3Fe_2O_3(s) + 11C(s) \rightarrow 2Fe_3C(s) + 9CO(s)$

$Fe_2O_3(s) + 3C(s) \rightarrow 2Fe(s) + 3CO(g)$

The iron products blended together to form cast iron, which was hard and brittle due to its high carbon content. The cast iron was transformed into wrought iron by hammering at 800–900°C. The hammering squeezed out the slag and impurities, but it was a long and tedious process. In days gone by, iron was five times more expensive than gold because it took so much effort to extract it.

Quenching was used at the time, although understanding of the chemistry of the process was centuries away. Quenching agents included water, donkey urine and the urine of red-headed boys. Writings about the famous Damascus steel used by the Crusaders refer to heating the blade of the sword until it glowed like a sunrise, cooling it to royal purple, then quenching it by plunging it into the body of a muscular slave. The idea was that the sword would then carry the strength of the slave.

Although iron revolutionized armaments, brittle cast iron was a poor choice for any use requiring tensile strength. In famous clashes such as the Battle of Trafalgar in 1805, when the British Fleet claimed victory over the combined might of the French and Spanish fleets, many injuries were caused simply by iron guns bursting. For some time, soldiers were at great risk from their own weapons.

In 1855, Henry Bessemer took out a patent on a process for blowing air through molten pig iron to remove impurities. This process, the precursor to today's basic oxygen process, resulted in steel use increasing at a phenomenal rate. In the following years, processes were refined. Modern steels are vastly superior to those used throughout history.

Figure 3.1.7 Sword made from Damascus steel, which came from Damascus, Syria.

The production of aluminium

The annual worldwide consumption of aluminium is over 35 million tonnes. Aluminium is produced by the electrolytic reduction of alumina (Al_2O_3) in a highly energy-intensive process, which is at best 50% energy efficient. **Bauxite** is the principal aluminium ore. It is a hard, reddish rock containing impurities such as water and iron oxides, as well as various aluminium hydroxides. Bauxite undergoes the Bayer process to remove impurities and produce alumina as a fine, white powder.

Aluminium smelting occurs in large, rectangular steel vessels known as **Hall–Héroult cells** (**pots**). They have aluminosilicate refractory lining on the bottom, which acts as a thermal insulator, and are connected in series to form a pot line. At Australia's Point Henry smelter, there are 368 pots connected in three pot lines, with each cell being fed a current of about 150 000 A at a potential difference of 4–5 V. Resistance to this enormous current flow is sufficient to keep the mixture within each cell molten. Some smelters use higher currents of up to 500 000 A.

AS **C.1.8**
Describe and explain the production of aluminium by electrolysis of alumina in molten cryolite. © IBO 2007

The anodes and cathode are made of carbon. The cathode forms a lining on the base of the cell, while 600 kg anode blocks are gradually lowered into the electrolyte. The anodes are consumed in the reaction. They are generally pre-baked, and are made from petroleum coke with coal tar pitch acting as a binder. They are replaced every 20 days on average, and the anode stubs are recycled. Usually, 20 anode bocks are used per cell, with one being replaced each day. This allows for continuous operation. The cathode is made of the same material as the anodes, but the sides of the cell are lined with partially graphitized anthracite, with a coal tar pitch binder.

Alumina (Al_2O_3) normally melts at 2050°C, but electrolysis can be performed at 940–970°C if the alumina is dissolved in molten **cryolite** (Na_3AlF_6). Cryolite melts at 1012°C. Aluminium fluoride (10–12%) and calcium fluoride (4–6%) are also added to help lower the melting point further (to 940–970°C) and increase the conductivity of the electrolyte. Another reason for adding fluoride compounds is to replace the fluorine lost as fluorocarbons and inorganic fluorides. Alumina is periodically fed into the cell.

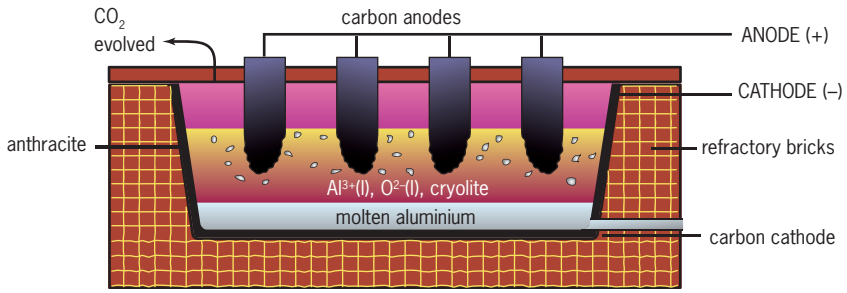


Figure 3.1.8 The Hall–Héroult cell.

The carbon anodes react with oxide ions to form carbon dioxide. At the same time, aluminium ions are reduced at the carbon cathode. The reaction equations are:

Anode(+): $C(s) + 2O^{2-}(l) \rightarrow CO_2(g) + 4e^-$

Cathode(–): $Al^{3+}(l) + 3e^- \rightarrow Al(l)$

The overall reaction is: $3C(s) + 4Al^{3+}(l) + 6O^{2-}(l) \rightarrow 3CO_2(g) + 4Al(l)$

Current leaves the cell through steel collector bars and moves on to the next cell in the pot line.

Aluminium is more dense than the electrolyte; it sinks to the bottom of the cell and is periodically tapped off to a holding furnace prior to casting. Some cryolite hardens on the sides of the cell, forming a ledge that protects the lining. The cryolite also forms a hard crust on top of the cell, which thermally insulates it. Holes have to be punched into the crust to feed fresh alumina in. Exit gases seep through the crust holes and are collected and sent to dry scrubbers to remove fluorides.

The carbon lining of the pots eventually disintegrates; it must be replaced every 3–5 years. Some of the melt also seeps through the carbon lining to the refractory lining. Hence, the refractory lining needs periodic replacement too. The process uses vast quantities of electricity; it takes 5.5×10^{10} J to produce one tonne of aluminium.

Properties and uses of aluminium

Smelters produce aluminium of about 99.7% purity. Super pure aluminium is rarely needed. Aluminium is a pale silver metal easily machined, cast and formed.

AS **C.1.9 Describe the main properties and uses of aluminium and its alloys. © IBO 2007**

Aluminium has many attractive properties. Its low density and corrosion resistance offer great advantages over steel. It has high thermal and electrical conductivity, is non-magnetic, non-toxic, and highly malleable and ductile. Unusually, it maintains its shininess in powder form, making it ideal for use in silver-coloured paints. Aluminium forms a thin (6.35×10^{-7} cm) oxide film on its surface. This film is impermeable to air and moisture, and protects the metal from chemical attack. The film may be thickened using a process called anodizing. Super-pure aluminium is used for electrodes, as a chemical reagent, and in thermally stable, highly reflective mirrors. The latter are used in devices such as X-ray telescopes.

Pr PRAC 3.2
Anodizing aluminium

The only drawbacks to aluminium are its lack of hardness and tensile strength. This can be solved by alloying it with small quantities of other elements. Table 3.1.4 shows three classes of aluminium alloys.

TABLE 3.1.4 THREE CLASSES OF ALUMINIUM ALLOYS

Class	Major alloying elements	Advantages	Example uses
2XXX	Cu	High strength (yield strength up to 520 MPa), reduced ductility and corrosion resistance	Aircraft construction, car bodies
3XXX	Mn	High corrosion resistance, at least 50% stronger than pure aluminium (up to 180 MPa)	Irrigation pipes, cladding, cooking utensils, roofing sheet, drink cans
6XXX	Mg + Si	Good corrosion resistance, excellent extrudability, reasonable strength (yield strength up to 290 MPa), highly weldable	Frames, furniture

Most aluminium is used in the transportation industry (cars, aeroplanes, trains, boats), followed by its use in packaging (foil, cans, cookware), construction, electrical (overhead wires), machinery and consumer durables. The aerospace industry could not exist without aluminium. Space shuttles are 90% aluminium, while a Boeing 747 aircraft contains 75 000 kg of aluminium, representing 80% of its weight. The low density of aluminium makes for faster travel. Japan's Bullet Train has passenger cars made of aluminium alloy. No other metal comes close to offering the great advantages of aluminium, but, at three times the price of steel, aluminium consumption is unlikely to overtake iron consumption any time soon. A large portion of aluminium's price is attributable to the massive amount of electricity required to extract it.

Figure 3.1.9 This Boeing 747 aircraft is 80% aluminium.

Environmental impact of iron and aluminium production

Accessing the benefits of aluminium and iron comes at a price. Both industries consume huge quantities of water and electricity, and result in emissions and wastes. In total, global metal production is responsible for the emission of over 1.5 billion tonnes of greenhouse gases annually. For both iron and aluminium, a four-step process is needed to obtain the end product: mining, mineral

AS **C.1.10 Discuss the environmental impact of iron and aluminium production. © IBO 2007**

processing, metal production, and refining. Each stage has associated environmental issues.

Impacts associated with mining include topsoil removal, changes in topography, decrease in vegetation, noise pollution from drilling and blasting, pollution of surface water, depletion of groundwater, land loss to infrastructure, and vehicle emissions. Processing the raw ore takes energy and produces dangerous wastes. The Bayer process for refining alumina produces red mud, a highly alkaline waste. Some loss to the environment is inevitable at every stage of an industrial process.

Producing coke for the blast furnace and carbon anodes for the Hall–Héroult cell creates significant emissions. Coke produced from wood involves deforestation. Heating coal to make coke drives off volatiles: oils, **particulates**, tars, ammonia, **NO_x**, CO and SO_2. Other pollutants include phenol, toluene (and other volatile organic compounds, **VOCs**) and hydrogen sulfide. Although the waste gases are **scrubbed**, some losses occur. Baking the coke again to produce anodes causes hydrocarbons to be released, which can be collected and burned.

Carbon monoxide and particulates leak from the blast furnace, which is also a generator of sulfur and nitrogen oxides. Hoods have been installed to reduce loss during tapping. Slag from the blast furnace is crystallized and used in construction, for example in cement and road beds. Slag from the basic oxygen steelmaking process is not suitable for construction use and so mostly ends up as **landfill**. Wastewater from all processes is contaminated, but is recycled as much as possible.

Alloying and surface treatments produce emissions that result in **acid rain** and **photochemical smog**. The very fine particulate matter (<2.5 μm) released represents a significant health hazard. Once lodged in the lungs, the particles are not easily removed. The aluminium industry is a major contributor of greenhouse gases emissions, which exacerbate the global warming problem. Carbon dioxide, **perfluorocarbons** (potent greenhouse gases), polycyclic aromatic hydrocarbons, sodium fluoride particulates and sulfur dioxide are released. Recycling aluminium also contributes to greenhouse gas emissions, but not nearly as much as primary aluminium production. Recycling requires only 5% of the energy needed to produce aluminium from alumina. About 70% of aluminium cans and 41% of steel cans are recycled; the rest are sent to landfill.

Because the anodes contain residual hydrocarbons, reactions occur in the Hall–Héroult cell producing hydrogen fluoride. The gases emitted from the cells are collected and sent through dry scrubbers, where alumina removes the fluoride compounds. This alumina is then fed into the cell and the fluoride is recycled. Some is still lost, but emissions of fluoride compounds have decreased dramatically since dry scrubbing was introduced. Another problem is what to do with the spent cell lining. The lining contains aluminium carbide, cyanides, fluorides, and is highly alkaline. Some is powdered and used as a fluoride source in cement kilns, but most ends up in landfill.

Despite the problems, metal industries are slowly reducing their environmental impacts. In the past 25 years, the release of air and water pollutants from metal smelting operations has decreased by 90%. In the same time, energy efficiency has increased dramatically. Current industry concerns are related to the fact that developed countries have been forced to clean up by rigorous emissions control, but developing countries, which are rapidly increasing their metal outputs, have no such emissions guidelines.

Section 3.1 Exercises

1 Explain why iron has been used since ancient times, but aluminium has only been in common use for about 150 years.

2 **a** Identify the two main iron ores.

b Calculate the maximum mass of iron that could be extracted from two tonnes of siderite.

3 Define the following terms.

a Ore

b Mineral

c Refractory

d Alloy

e Slag

f Smelting

4 Outline, using equations, how these reactions occur in a blast furnace.

a Carbon monoxide is formed.

b Iron ore is reduced by carbon monoxide.

c Alumina impurities dissolve in slag.

5 Explain why pig iron needs further processing before it can be used.

6 Explain, using equations, how pig iron is purified in a basic oxygen converter.

7 Describe the effect of the following alloying elements on the properties of steel.

a Carbon

b Nickel

c Chromium

d Cobalt

8 List the heat treatments that:

a soften alloys

b harden alloys.

9 Compare the usefulness of pure iron and steel.

10 Construct a table showing the main parts of a Hall–Héroult cell and their functions.

11 **a** Write the equations for the electrode reactions in a Hall–Héroult cell.

b Describe the functions of cryolite, aluminium fluoride and calcium fluoride in the Hall–Héroult cell.

12 Explain why most aluminium is used in alloy form, not pure.

13 Discuss the relative environmental impacts of the steel and aluminium industries.

3.2 THE OIL INDUSTRY

Fossil fuels include crude oil, natural gas and coal. They are non-renewable resources formed over millions of years from the remains of plants and animals. Historically, the term **petroleum** has been used to refer to any naturally occurring mixture of hydrocarbons in the solid, liquid or gaseous state. It is now common practice to use the terms crude oil and petroleum interchangeably. The oil industry encompasses exploration, mining, transport, purification, sales and marketing, and the production of **petrochemicals**—chemicals derived from petroleum.

Figure 3.2.1 Fossil fuels are needed to run all kinds of transport. Fuel could be conserved if more people chose a hybrid vehicle over a 'petrol-guzzling' 4WD.

The fate of oil

AS C.2.1
Compare the use of oil as an energy source and as a chemical feedstock. © IBO 2007

The advent of motor vehicles and aeroplanes has created a huge global market in the fuels needed to run them. Presently, over 90% of crude oil derivatives are used for fuel. Some of this fuel goes into keeping humans on the move; the rest goes towards heating. Crude oil is a complex mixture of hydrocarbons, together with small quantities of other compounds containing sulfur, nitrogen, oxygen and metals. The hydrocarbons found in crude oil include straight chain alkanes, branched alkanes, **aromatics** and asphaltenes—large, polycyclic molecules.

De DEMO 3.1
Fractional distillation of an alcoholic drink

Oil is not used in its raw state. The crude oil is transported to oil refineries, where it undergoes **fractional distillation** in a fractionating tower. The crude oil is heated to 600°C in the absence of air. The components with boiling points higher than this temperature remain in viscous liquid form; they sink to the bottom of the tower. The vaporized components rise, cooling as they ascend. At regular intervals going up the tower, there are increasingly cooler horizontal trays containing structures called bubble caps, which help the cooling and condensation process. When components in the vapour reach a tray that is close to their boiling point, they condense and are collected as a mixture of alkanes with similar boiling points called a fraction. Figure 3.2.2 shows some typical fractions.

The smaller molecules have weaker van der Waals' forces between them; they have the lowest boiling points and are collected at the top of the tower.

The world is currently facing an energy crisis. We have burned through most of our fossil fuel reserves at an astonishing rate. Oil reserves are dwindling and the price per barrel is rapidly increasing. As well as being a fuel source, crude oil is used to make many important chemicals such as polymers, solvents and resins.

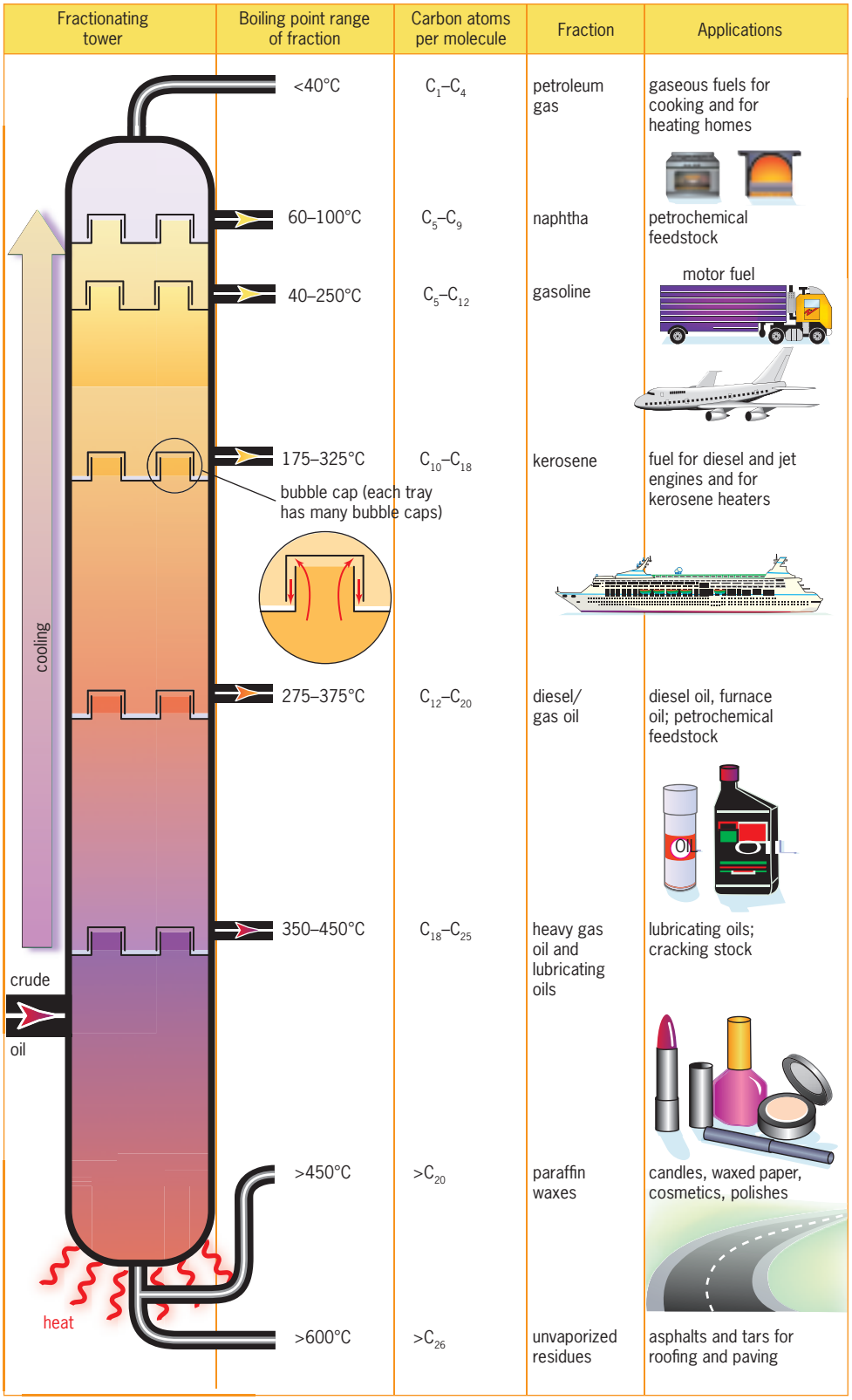


Figure 3.2.2 Fractions found in crude oil.

Currently, less than 5% of oil is used for **petrochemical feedstocks**—defined as liquid or gaseous hydrocarbons used to make intermediate or primary chemicals. Figure 3.2.3 shows some common chemicals derived from petroleum, and the feedstocks used to create them. If you look around your house, or even your school, petrochemicals will be evident in every room. At the moment, there is enough oil for both the fuel and petrochemical industries, but that won't always be the case. If economically viable renewable alternatives can't be found, people will be forced to make drastic lifestyle changes.

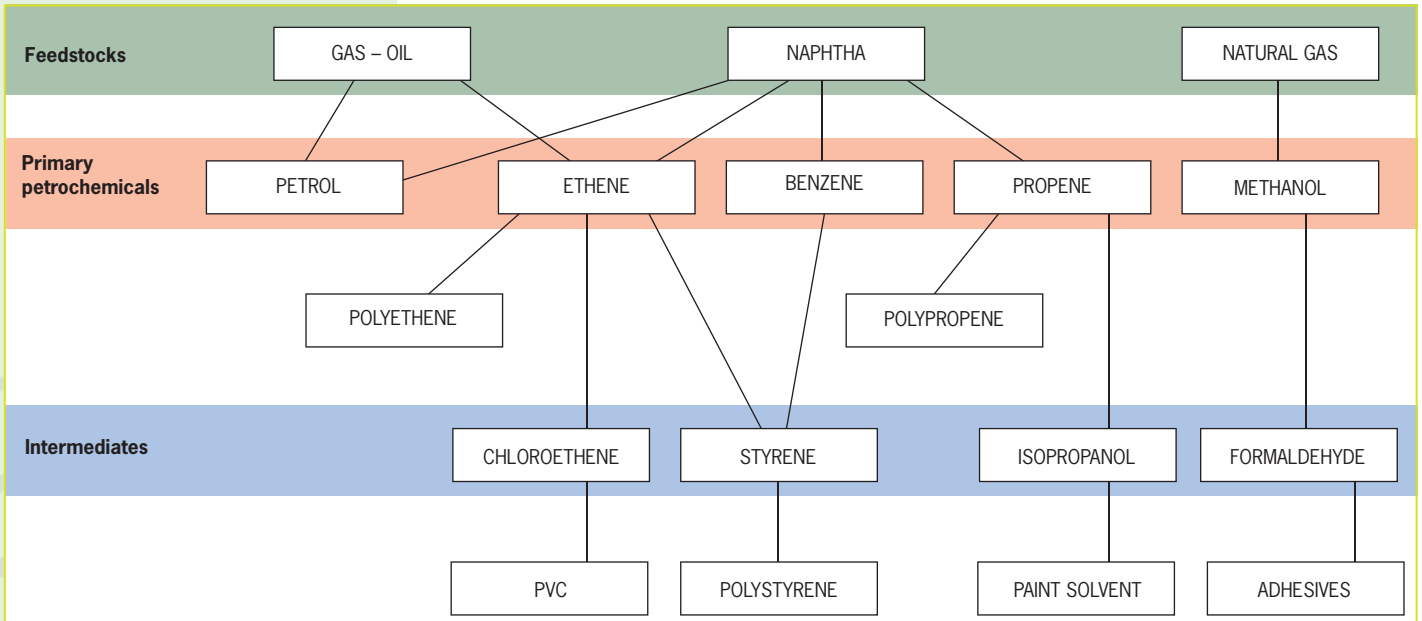


Figure 3.2.3 Petrochemicals are a prominent part of everyday life.

TABLE 3.2.1 ADVANTAGES AND DISADVANTAGES OF USING CRUDE OIL DERIVATIVES FOR ENERGY

Why crude oil should be used for energy	Why it shouldn't
• Infrastructure already in place • Relatively inexpensive • Easy to transport	• Contributes to air pollution and global warming • There are alternatives available, such as nuclear and wind energy

TABLE 3.2.2 ADVANTAGES AND DISADVANTAGES OF USING CRUDE OIL DERIVATIVES FOR FEEDSTOCKS

Why crude oil should be used for feedstocks	Why it shouldn't
• Infrastructure already in place • Easier to find energy alternatives than petrochemical alternatives	• 80% goes into producing polymers, many of which are not recyclable or biodegradable • there are alternatives available in the form of biopolymers

Cracking

After the crude oil has been fractionally distilled, further refinement processes take place. A vital reaction related to the petrochemical industry is *cracking*. In **cracking reactions**, larger, saturated hydrocarbons are broken down into smaller molecules, including some unsaturated carbon compounds. This reaction allows large quantities of ethene, the most important building block of petrochemicals, to be produced. Various fractions obtained in the distillation process can be used as feedstock for cracking, the naphtha and gas–oil fractions in particular. Sometimes, smaller molecules such as ethane are also used.

AS **C.2.2 Compare catalytic cracking, thermal cracking and steam cracking. © IBO 2007**

The cracking reactions are endothermic and reversible. Typical reactions include:

$$CH_3(CH_2)_{11}CH_3(g) \rightleftharpoons C_2H_4(g) + CH_3(CH_2)_9CH_3(g) \quad \Delta H = +93.5 \text{ kJ mol}^{-1}$$
$$C_2H_6(g) \rightleftharpoons C_2H_4(g) + H_2(g) \quad \Delta H = +137.0 \text{ kJ mol}^{-1}$$
$$C_3H_8(g) \rightleftharpoons C_2H_4(g) + CH_4(g) \quad \Delta H = +81.3 \text{ kJ mol}^{-1}$$

In all types of cracking, after the feedstock has been quickly passed through the cracker, it is rapidly cooled in heat-recovery boilers to limit further reactions. The gases are then further cooled and compressed. The products are separated by fractional distillation. Some fractions may be passed back for another turn in the cracker.

Cracking of petroleum fractions was first achieved by heating the fraction to very high temperatures in the absence of air. This process, called **thermal cracking**, is expensive due to the energy required to maintain high temperatures. It is also difficult to control which products form, as the feedstock molecules can theoretically break in any position. In this process, an aqueous solution of an alkaline salt is introduced downstream of the feedstock to suppress the formation of solid carbon (coke), which would reduce efficiency and clog up pipes. The wastewater from the overhead accumulator in the fractionator can contain oil, sulfides, ammonia and phenol.

Pr PRAC 3.3
Cracking paraffin oil

TABLE 3.2.3 THERMAL CRACKING SUMMARY

Feedstock fraction	Naphtha (most commonly)
Temperatures used	450–750°C
Pressures used	70 atm
Catalyst	None
Mechanism	The molecules split producing highly reactive free radicals. Further reactions between the free radicals and other molecules results in the products.
Environmental impact	The major impact comes from the massive amount of energy needed—6300 MJ per tonne of naphtha. The production of this electricity creates pollution, especially if it comes from fossil fuels. Also, waste alkali liquor and oil-saturated solid sludge are produced. Emissions of VOCs occur from the stack, storage vessels and during waste-water treatment.

A variation on thermal cracking is **steam cracking**. In this type of cracking, a mixture of steam and feedstock is passed through coiled metal tubes in the cracker at very high temperatures. The residence time in the cracker is in the order of milliseconds.

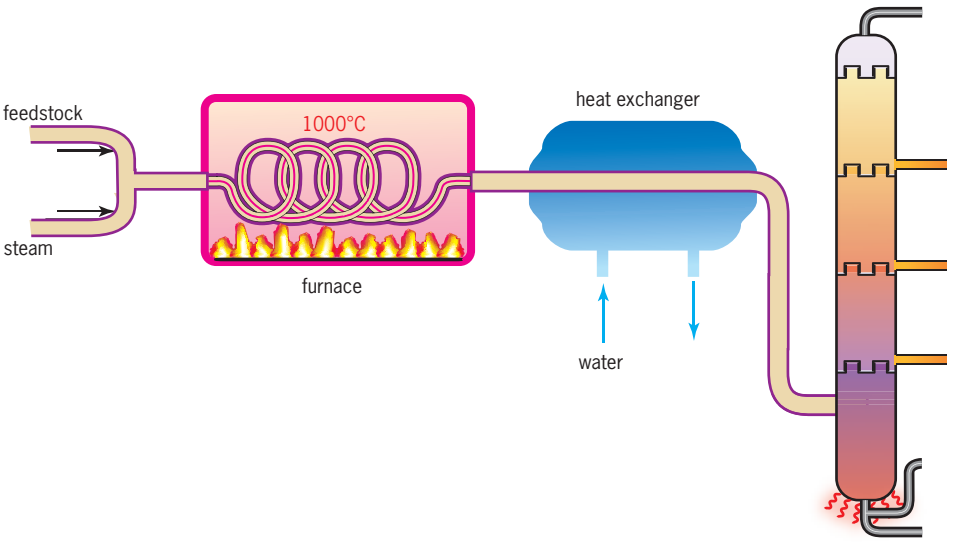


Figure 3.2.4 The main steps in the steam cracking process.

TABLE 3.2.4 STEAM CRACKING SUMMARY	
Feedstock fraction	Naphtha, gas–oil, ethane or propane
Temperatures used	750–900°C
Pressures used	Slightly above atmospheric
Catalyst	None
Mechanism	Helped by collision with water molecules, the hydrocarbon molecules split, producing highly reactive free radicals. Further reactions between the free radicals and other molecules results in the products.
Environmental impact	Again, a lot of energy will be needed to generate high temperatures. The feedstocks are frequently contaminated, especially if they were stored in metal containers during transport. These contaminants promote coke formation. Decoking cycles can releases particulates into the air. Emissions of VOCs occur.

Since these are reversible reactions with more molecules on the product side, low pressures would favour the formation of products, according to Le Chatelier's principle. However, working with gases at pressures below atmospheric is hazardous, since a leak would cause air to be drawn in, possibly forming an explosive mixture. The steam allows relatively low pressures to be used without danger. It also helps to prevent carbon deposits forming in the tubes:

$$C(s) + 2H_2O(g) \rightleftharpoons CO_2(g) + 2H_2(g)$$

An important development in cracking was the introduction of catalysts, which allowed the process to be carried out at much lower temperatures. This process is called **catalytic cracking**. The three main types of catalysts used are acid-treated natural aluminosilicates; amorphous, synthetic alumina-silica combinations; and crystalline, synthetic alumina–silica compounds called **zeolites**.

Zeolites have a honeycomb-like structure of tiny pores. Reactant molecules are adsorbed in these pores, where their reactions are catalysed. Zeolites are the most used catalysts because they offer higher activity, higher yields and less overcracking. Better control of products is achieved in catalytic cracking compared to the other methods, and shorter residence times in the cracker are possible.

Only 15% of the total catalyst mass needs to be zeolite for a good result. The zeolites give high product percentages of compounds 5–10 carbons long, which is good for petrol production. High proportions of branched alkanes and benzene also result.

TABLE 3.2.5 CATALYTIC CRACKING SUMMARY

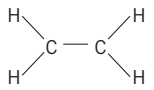
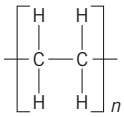
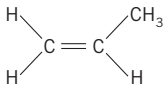
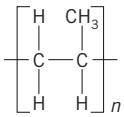
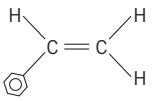
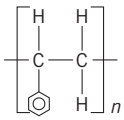
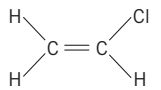
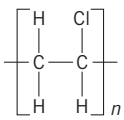
Feedstock fraction	Naphtha, kerosene, gas–oil
Temperatures used	450–500°C
Pressures used	Moderately low
Catalyst	Zeolites and similar
Mechanism	Zeolite catalyst removes a hydrogen and pair of electrons, creating a positively charged carbocation. These react to form the various products.
Environmental impact	Energy needed to produce high temperature. Wastewaters are alkaline and contain sulfides and phenol. In the catalyst regeneration process, carbon monoxide and tiny catalyst particulates can be released to the atmosphere. If leaks occur, VOCs will be released to the atmosphere.

Structure and properties of addition polymers

Cracking reactions produce a great deal of ethene (ethylene), a highly prized chemical. Ethene is the precursor to a wide variety of **addition polymers** such as polyethene and polyvinylchloride (see *Chemistry: For use with the IB Diploma Programme Standard Level*, chapter 11). Cracking reactions also produce propene (propylene), the **monomer** used to make polypropene. In polymerization, thousands of monomers are linked together to form a long chain. Addition polymers are formed from monomers containing a double carbon–carbon bond. Table 3.2.6 shows some common addition polymers.

WORKSHEET 3.2
Ethene

TABLE 3.2.6 SOME ADDITION POLYMERS

Monomer	Polymer	Polymer name	Some uses
Ethene 		Polyethene	Plastic bags, bottles, toys
Propene 		Polypropene	Indoor-outdoor carpeting, bottles, luggage
Styrene 		Polystyrene	Styrofoam insulation, cups, packing materials
Chloroethene 		Polyvinyl chloride (PVC) (polychlorethene)	Plastic wrap, plumbing, garden hoses

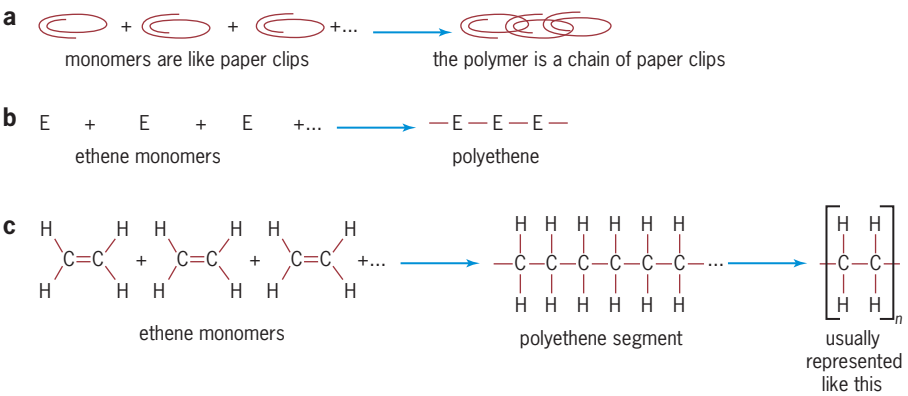


Figure 3.2.5 Ethene (ethylene) undergoes polymerization to produce polyethene (polyethylene).

AS C.3.1 Describe and explain how the properties of polymers depend on their structural features.

Various structural features will alter the physical properties of a polymer. Table 3.2.7 outlines some of these.

Polyethene is the most-used synthetic polymer. More than 50 million tonnes of polyethene are produced per year worldwide. When ethene is polymerized at 200°C and 2000 atmospheres pressure in the presence of organic peroxide catalysts using a benzene solvent, low density polyethene (LDPE) is formed. The chains form rapidly and are highly branched. Because they cannot pack closely, there are only weak van der Waals' forces present between chains and low crystallinity. The polymer melts at about 115°C. It is transparent, flexible, unreactive and waterproof.

TABLE 3.2.7 STRUCTURAL FEATURES AFFECTING POLYMER PROPERTIES	
Feature	**How it affects physical properties**
Branching	Highly branched chains can't pack closely together. This means that there is low crystallinity and thus there are more amorphous (non-crystalline) regions. These highly branched polymers will have relatively low melting points and low tensile strength. Conversely, mostly unbranched chains will pack closely and exhibit high crystallinity and high melting points.
Cross-linking	Cross-linking between polymer chains increases the hardness and rigidity. With many cross-links, the polymer becomes quite brittle. These polymers will char rather than melt.
Chain length	The longer the chain, the stronger the intermolecular forces. This results in higher melting points and greater strength as the chain length increases.
Intermolecular forces	The strength of the intermolecular forces between polymer chains will depend on the length of the chain, the degree of branching and the nature of the side groups. If the side groups are small and polar, intermolecular bonding strength (and hence melting point and strength) will increase.

If ethene is polymerized at 60°C at close to atmospheric pressure in the presence of **Ziegler–Natta catalysts** using a hydrocarbon solvent, high density polyethene (HDPE) is formed. This polymer has few branches. The chains pack closely together and exhibit high crystallinity. It melts at about 135°C. While still waterproof and unreactive, it is less transparent and more rigid than LDPE.

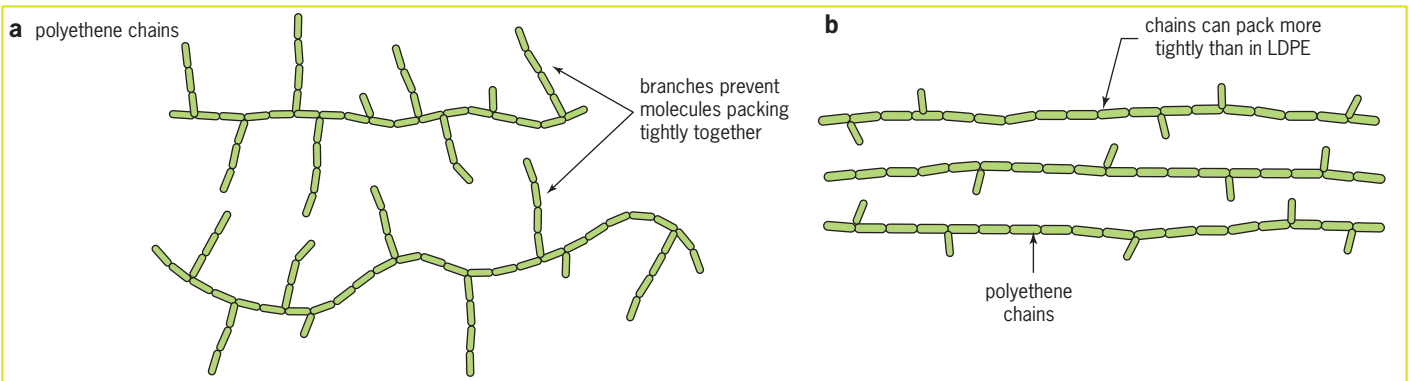


Figure 3.2.6 (a) LDPE has chains with significant branching. (b) HDPE is mostly unbranched.

Ziegler–Natta catalysts can also be used to create polypropene from propene monomers.

The structure and properties of polypropene change, depending on the positions of the methyl groups. The three differently oriented structures are called *isotactic*, *atactic* and *syndiotactic*. **Isotactic polypropene** is the most common type. In this type, the methyl groups are arranged very regularly. The chains pack closely and the dispersion forces between chains are quite strong. This makes isotactic polypropene quite strong. Hence it is used to make objects such as ropes and milk crates.

In **atactic polypropene**, the methyl groups are randomly arranged. Because of this, the chains cannot pack as closely together. It is softer than isotactic polypropene and has a lower melting point. It is used in sealants and adhesives. **Syndiotactic polypropene** is a relatively new form of polypropene. It has a regular, alternating arrangement of methyl groups. The chains can pack more closely than in the atactic form, but not as closely as in the isotactic form. Hence, it has intermediate properties. It is used in packaging, for example shrink wrap.

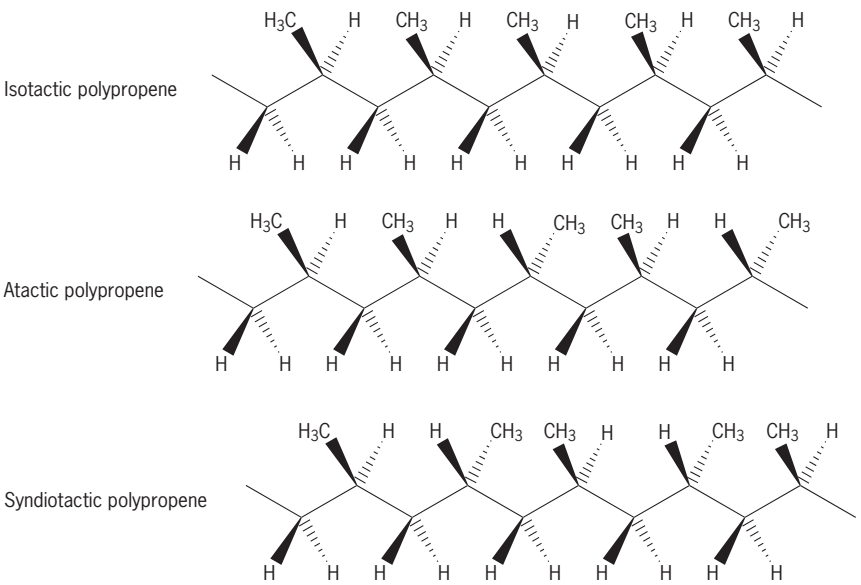


Figure 3.2.7 The three forms of polypropene.

Modifying addition polymers

AS C.3.2 Describe the ways of modifying the properties of addition polymers. © IBO 2007

Polymer scientists have a remarkable range of tools at their disposal to modify or enhance the properties of a polymer. Some of these are listed:

- *Molecular weight*—Extra-long molecules lead to higher melting points and greater toughness. Ultra-high molecular weight polyethene (UHMWPE) is an example of this. The long molecules used in this polymer add such toughness that it has replaced **Kevlar™** in bulletproof vests, and Teflon in artificial hip joints. Another reason for the use of UHMWPE in hip joints is that it has high biocompatibility. Biocompatibility refers to the degree of acceptance of a material by the human body.
- *Glass transition temperature* (T*g*)—Polymers are the only materials that have a glass transition temperature point (*T*g). When a polymer is below this temperature it is hard and brittle, like glass. At temperatures above *T*g, the polymer is softer and more flexible. When a polymer needs to be hard, such as in aeroplane windows, a polymer like Perspex is used that is below its softening temperature. Rubber in car tyres is an example of a polymer being used above its softening temperature.

- *Additives*—Commercial polymers often have many additives, for example, colours, antioxidants, flame-retardant chemicals and **plasticizers**. Plasticizers are oils that make a polymer more flexible and easier to handle. The percentage of plasticizer in some products can be more than 10 per cent. Phthalates like diisononyl phthalate are used to plasticize PVC.
- *Functional group alignment*—In the manufacture of polypropene, catalysts are used to control the position of the methyl group and hence the polymer's properties. This concept can be applied to many other polymers, not just polypropene.
- *Foam*—If a gas is released during polymerization, the gas may be trapped in the polymer, creating a foam. In expanded polystyrene, this is done by including pentane in the beads of polymer. When the beads are heated, the pentane—a volatile hydrocarbon—turns to a gas. In polyurethane, the condensation reaction releases carbon dioxide gas and this is trapped in the foam. Foam is very popular as it means the manufacturer does not need to use as much polymer.
- *Copolymers*—If one polymer's properties are not quite right, an option is to blend in a percentage of a second polymer. This is called *copolymerization*. For example, polystyrene is a hard but brittle plastic. A small percentage of butadiene rubber added to it greatly enhances the toughness of the polymer. The product is called high-impact polystyrene (HIPS).
- *Fibres*—The addition of glass or carbon fibres enhances the toughness of a polymer significantly. Fibreglass is an example of a reinforced polymer. The liquid monomer is poured over a mat of woven glass and then allowed to polymerize. The strength of the product makes it popular in kayaks, car bodies and tennis racquets. Fibre-reinforced polymers are referred to as composites.

Polymers are highly durable and chemically inert. In addition, many have antioxidants added to them during processing to further limit their susceptibility to deterioration. Polymers offer many advantages over natural materials such as wood and metals. They are strong, good insulators, lightweight, easily moulded, versatile and affordable. But there is a downside to the durability of plastics. When a plastic item is no longer useful, it is difficult to dispose of because it does not rust or decompose. A few biodegradable polymers are used, but they are more costly.

AS **C.3.3 Discuss the advantages and disadvantages of polymer use. © IBO 2007**

Given the large volume of plastic used by society today, the environmental issues of plastics dominating landfill are contentious. Burning polymers is not an environmentally acceptable option, as the fumes released are harmful. Some polymers are still burnt. The gases produced can be passed through scrubbers to remove soluble substances such as HCl and NH_3, and filtered to remove soot. Current research is focused on developing green solutions to the polymer disposal problem.

One solution is for society to recycle as many polymer items as possible. Polymers are numbered to help consumers and recyclers identify the polymer they are dealing with. An internationally accepted numbering system is used. See figure 5.8.2 and pages 296–297 for recycling of polymers. To be recycled, a polymer must be thermoplastic. Thermosetting materials will not melt to form a new product.

Figure 3.2.8 Various solutions to the problem of disposing of polymers: shopping bags (a) and biscuit trays (b) that decompose, (c) refunds on plastic containers and (d) reduced use of plastics.

Recycling is limited by a number of factors:

- Sorting of recycled material is costly.
- Polymers have a low density, so a high volume of material must be recycled to obtain a useful mass.
- The recycled material must be cleaned.
- Recycling facilities and personnel are required.
- The public has to support these schemes.
- The recycled plastic is limited in its appearance and uses, for example, non-food-grade containers.

Section 3.2 Exercises

1 List the three major types of fossil fuels.

2 Define the following terms.

a Petrochemical

b Feedstock

3 Discuss the use of petroleum as a fuel source and as a feedstock.

4 Write three different cracking equations, each producing ethene as a product.

5 Compare the three different types of cracking reactions in terms of feedstock, reaction conditions and environmental impact.

6 Explain the effect of the following on the physical properties of a polymer.

a Degree of branching

b Amount of cross-linking

c Chain length

7 Sketch the three different types of polypropene, clearly showing the positions of the methyl groups.

8 Describe two ways in which the properties of a polymer may be modified.

9 List three advantages and three disadvantages of polymer use.

3.3 CATALYSTS, FUEL CELLS AND RECHARGEABLE BATTERIES

Recall from *Chemistry: For use with the IB Diploma Programme Standard Level*, chapter 7, that catalysts are substances that increase the rate of a chemical reaction, but do not undergo permanent chemical change. They work by providing an alternate reaction pathway with a lower activation energy, E_a. Catalysts are extremely important in both nature and industry; they are at work in the soil, air and oceans, our bodies, and in over 80% of all industrial reactions.

Catalysts can be classified according to the state (solid, liquid or gas) they are in, relative to the reactants they are combined with. *Homogeneous catalysts* are in the same state as the reactants; for example, sulfuric acid catalyses esterification reactions, such as the reaction between ethanol and ethanoic

acid. All three substances—ethanol, ethanoic acid and sulfuric acid—are miscible liquids:

$$CH_3COOH(l) + CH_3CH_2OH(l) \xrightarrow{H_2SO_4(l)} CH_3COOCH_2CH_3(l) + H_2O(l)$$

Heterogeneous catalysts are in a different state to the reactants. Often, the catalyst is solid and the reactants are gaseous or in solution. One example in which the catalyst is solid and the reactant is in aqueous solution is the decomposition of hydrogen peroxide using manganese(IV) oxide:

$$2H_2O_2(aq) \xrightarrow{MnO_2(s)} 2H_2O(l) + O_2(g)$$

Mechanisms

AS C.4.1 Compare the modes of action of homogeneous and heterogeneous catalysts. © IBO 2007

Homogeneous catalysts will typically react to form an intermediate compound, which then goes on to participate in further reactions. The catalyst is reformed in the latter stages of the reaction cycle. The hydrogen peroxide decomposition reaction can be catalysed by several different substances. One homogeneous catalyst for this reaction is the iodide ion, I^-. In the initial step, a hydrogen peroxide molecule reacts with an iodide ion to form water and IO^-:

$$H_2O_2(aq) + I^-(aq) \rightarrow H_2O(l) + IO^-(aq)$$

The IO^- ion (an intermediate substance) then reacts with another H_2O_2 molecule:

$$H_2O_2(aq) + IO^-(aq) \rightarrow H_2O(l) + O_2(g) + I^-(aq)$$

Thus, although the I^- catalyst undergoes chemical change during this catalytic cycle, it is not *permanently* altered, since it is regenerated in the second step. Both steps in this reaction will have lower activation energies than the E_a of the uncatalysed reaction. The first step is the rate-determining step, so its E_a will be higher than that of the second step.

Heterogeneous catalysts are usually solid metals or metal oxides. They are most effective when the surface area is high, so it is common to find them in pellet form spread over large beds. The first step in this type of catalysis is the adsorption of one or more reactants onto the catalyst's surface. The specific places where reactants stick are called **active sites**. The greater the number of active sites, the more effective the catalyst. The number of active sites will depend on the type of catalyst, its particle size, whether or not it is clean of dust and other particles, and how it was prepared prior to use.

The adsorbed reactant interacts with the catalyst, which weakens its bonds. If the reaction involves another reactant, it may stick to the catalytic surface itself, or collide with the reactant already on the surface and react. The products then **desorb** (break away) from the surface, and the same active site can be used again.

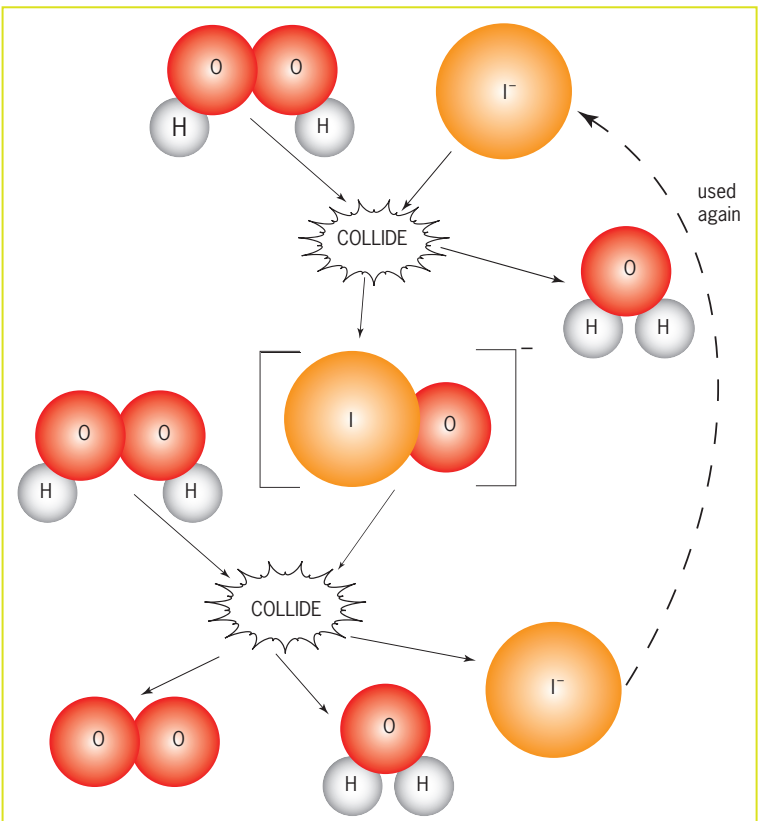


Figure 3.3.1 The I^- catalyst can be continually reused until all the hydrogen peroxide has reacted.

QuickTime CW Video
Catalysis
Action at a catalyst surface

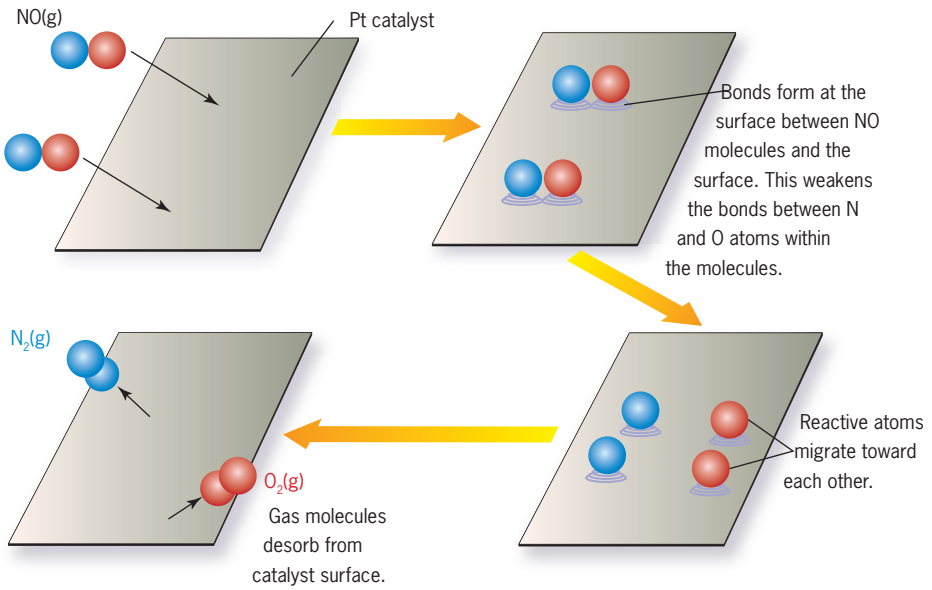


Figure 3.3.2 The decomposition of NO(g) catalysed by platinum.

Many transition metals and their ions are effective catalysts. For example, an iron catalyst is used in the Haber process, and platinum is used in hydrogenation reactions. Not all transition metals are appropriate; tungsten adsorbs substances too strongly to be effective, while silver adsorbs them too weakly. Most transition metals have multiple oxidation states, and this is important in their catalytic roles. As substances react at the surface of the catalyst, it temporarily changes oxidation state and, in doing so, weakens or breaks the bonds in the reacting molecule.

An example of this type of catalysis is the converter reaction in the production of sulfuric acid via the contact process. In this reaction, sulfur dioxide is oxidized to sulfur trioxide in the presence of air and a vanadium(V) oxide catalyst, which is spread in pellet form over three or four separate beds. First, the sulfur dioxide adsorbs to the surface, then it reacts with the catalyst:

$$SO_2(g) + V_2O_5(s) \rightarrow SO_3(g) + V_2O_4(s)$$

In this step, vanadium(V) is reduced to vanadium(IV).

In the next step, oxygen in air oxidizes the vanadium back to its original state:

$$V_2O_4(s) + \tfrac{1}{2}O_2(g) \rightarrow V_2O_5(s)$$

Advantages and disadvantages

AS C.4.2 Outline the advantages and disadvantages of homogeneous and heterogeneous catalysts.

Both types of catalysts have advantages and disadvantages that need to be considered when selecting a catalyst for a particular process. One factor to consider is the ease of **catalyst poisoning**. This term refers to the catalyst being deactivated via reaction with a substance not involved in the reaction. The ions used as homogeneous catalysts tend to be quite reactive, so can easily become deactivated if other reactive species are present. Solid catalysts are

harder to deactivate fully, but active sites on beds of catalysts can become clogged with dust and other particles if due care is not taken to purify the reactant mixture. Cleaning a catalyst bed is a costly process.

TABLE 3.3.1 ADVANTAGES AND DISADVANTAGES OF HOMOGENEOUS CATALYSTS

Advantages	Disadvantages
• Optimal contact between catalyst and reactants • Easier to work with (rate proportional to amount added) and modify • More selective • Easy to add to the reaction system • Since contact is so good, generally mild conditions (low temperature and pressure) can be used • Mechanisms are better understood because they can be studied spectroscopically	• Difficult to separate products and catalyst for reuse—could require distillation, which might destroy the catalyst • Highly susceptible to poisoning

In industry, the hardier heterogeneous catalysts are more commonly used. These too have advantages and disadvantages.

TABLE 3.3.2 ADVANTAGES AND DISADVANTAGES OF HETEROGENEOUS CATALYSTS

Advantages	Disadvantages
• Harder to irreversibly poison • Easy to separate products from catalyst by filtration and thus recycle catalyst • Recovery of catalyst easier • Can be used at very high temperatures	• Reaction only occurs at the surface, so contact is reduced compared to homogeneous catalysts • Mechanisms not as well understood as homogeneous catalysts • Manufacture can be costly • Can be hard to regenerate • Can be difficult to maintain constant temperature

Choice of catalyst

AS C.4.3 Discuss the factors in choosing a catalyst for a process.

There are many things to consider when choosing a catalyst. The following points summarize the major considerations:

- *Selectivity*—Selectivity refers to how specific the catalyst is for the desired reaction, its activity towards particular reactants. In industry, there is always the possibility of side reactions that will reduce the overall yield. By choosing a catalyst that is highly specific, the yield can reach close to 100%.
- *Efficiency*—Various means exist for measuring catalyst efficiency. One measure is the number of reactant (substrate) molecules converted to products per molecule/particle of catalyst before the catalyst loses its activity. The higher this number, the better the catalyst.
- *Homogeneous versus heterogeneous*—The advantages and disadvantages outlined in the previous section will influence the decision here. If the reaction is in the liquid phase, it is possible that a homogeneous catalyst will offer the greatest selectivity. In practice, most industrial reactions use heterogeneous catalysts since they are much easier to reuse and can be used at high temperatures.

- *Durability*—Some industrial reactions need to be performed at extremely high temperatures and/or pressures. The chosen catalyst needs to be able to withstand the particular conditions required. It should be tested prior to use to ensure that it does not degrade. Also, in any reacting system there will be a range of conditions. Even within one catalyst bed there tends to be considerable temperature fluctuation. Not only does a catalyst need to physically withstand the reaction conditions, it also needs to retain high levels of selectivity and efficiency over the full range of conditions present.
- *Availability*—If a constant supply of catalyst is required, one must be chosen that can be reliably obtained from a number of sources.
- *Cost*—Although cost seems to govern just about everything else in industry, it is not a major consideration when choosing a catalyst. As long as the catalyst can be reused many times, its expense can be justified.
- *Environmental impact*—These days, factories must adhere to strict emissions guidelines. Some very effective catalysts, such as the neurotoxin mercury, which was used extensively in batteries and the chlor-alkali industry, can no longer be used because of the environmental consequences. The catalyst should not result in unacceptable environmental impacts in its manufacture, use or regeneration. It also should not compromise the health of workers.
- *Susceptibility to deactivation*—Some highly effective catalysts are not able to be used because they are too easily deactivated. Catalysts can be deactivated by poisoning (by substances such as sulfur, hydrogen sulfide, lead and mercury), solid particles depositing on the surface of the catalyst and blocking active sites and pores (such as carbon), degradation of the surface leading to fewer active sites, or evaporation of catalytic particles from the surface as various compounds. The conditions and substances likely to be present in a reaction system must be considered together with the particular catalyst. If any of the substances will easily deactivate the catalyst, another catalyst must be chosen, or the reaction conditions must be changed. One example is the catalytic converter in a car. These converters are only able to be used with unleaded petrol because lead would poison the platinum catalyst.
- *Recovery for reuse*—Most homogeneous and some heterogeneous catalysts will not be recoverable. However, if they are very cheap, or if they increase the rate drastically, this will not be important. Sulfuric acid, the most produced chemical in the world, is cheap. After it catalyses esterification reactions, it cannot easily be recovered from the product mix. The process to recover it would cost more than the catalyst, so it is easier to discard it. The new-generation Ziegler–Natta catalysts used for producing polymers such as polypropylene are very expensive and amazingly effective; half a kilogram of catalyst can produce 35 tonnes of product. Their expense is more than justified and their recovery not worthwhile. At times, the expense of recovering and recycling a catalyst is justified, particularly if there is not an unlimited supply of the substance. This is the case for metals such as platinum and palladium. Thus, in choosing a catalyst, engineers must consider whether it is appropriate to discard or recover the catalyst, weighing up their decision against the cost.

Fuel cells

One application of catalysts is in fuel cell **electrodes**. **Fuel cells** are electrochemical devices that convert chemical energy directly to electrical energy. Since the conversion is direct, fuel cells offer much greater efficiency than the common electricity-generating methods involving steam turbines. Efficiencies of up to 70% are common for fuel cells, compared to about 30–40% if the same reaction is used to make electricity using turbines and a generator.

AS **C.5.1 Describe how a hydrogen–oxygen fuel cell works.**

Pr PRAC 3.4 Modelling a hydrogen–oxygen fuel cell

Fuel cells consist of an **electrolyte** (conductive) layer in contact with two porous electrodes. Gaseous fuels such as hydrogen or methane are fed continuously into the anode compartment, while an oxidizing agent (usually oxygen in air) is fed continuously through the cathode compartment. At a certain point within the electrode, the fuel comes in contact with the electrolyte layer—this is where the reaction takes place. Thus, unlike a standard battery (chemical storage device), a fuel cell should be able to keep supplying energy indefinitely, with no loss in voltage, as long as reactants are supplied. In reality, the electrodes and other components degrade, so the fuel cells will have a finite life span.

The functions of the electrolyte are to:

1 provide a physical barrier to stop the reactants mixing

2 complete the circuit by conducting ions into and away from electrodes

3 enable the reaction to take place by participating in the reactant–catalyst–electrolyte interface.

The functions of the electrodes are to:

1 catalyse the reactions

2 provide a surface on which the reaction can occur

3 provide a physical barrier between the reactants and the electrolyte

4 conduct electrons and ions.

Theoretically, any exothermic redox reaction in which the reactants are gaseous at the fuel cell's operating temperature could be used in a fuel cell, as long as appropriate electrodes and electrolytes are available. The most common and highly developed cell is the hydrogen–oxygen fuel cell. New technologies are currently in development, but the first generation hydrogen–oxygen fuel cells use either a phosphoric acid or potassium hydroxide electrolyte.

In the acidic version, the liquid phosphoric acid is contained within a Teflon-bonded silicon carbide matrix. The cell uses porous carbon catalysts impregnated with platinum.

At the anode (negative polarity), hydrogen is oxidized:

$$H_2(g) \rightarrow 2H^+(aq) + 2e^-$$

The hydrogen ions are transported away from the anode, towards the cathode. The electrons travel through the external circuit to the cathode (positive polarity), where oxygen is reduced:

$$O_2(g) + 4H^+(aq) + 4e^- \rightarrow 2H_2O(l)$$

Excess water and heat are extracted from the cathode compartment.

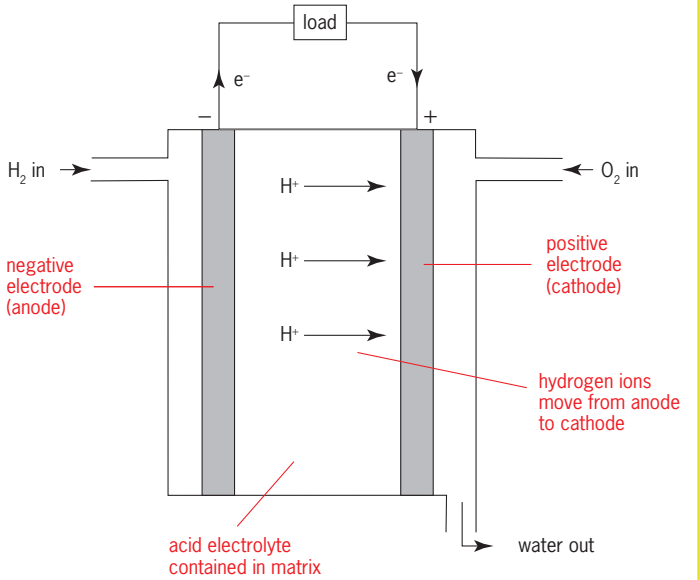


Figure 3.3.3 The acid hydrogen–oxygen fuel cell.

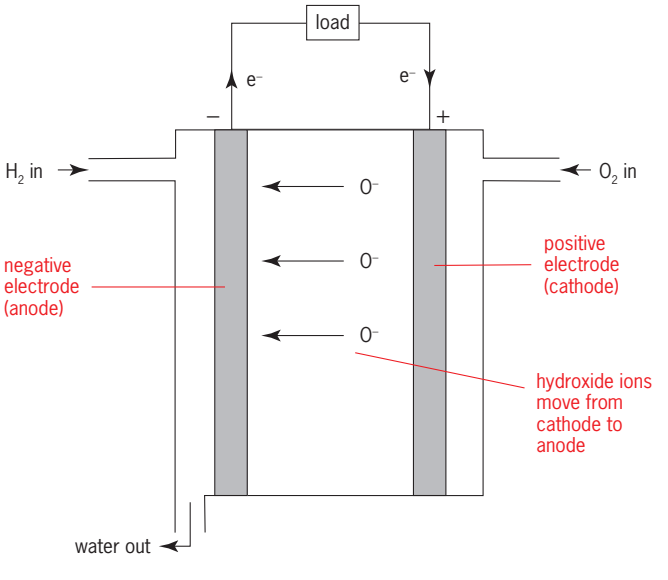


Figure 3.3.4 The alkaline hydrogen–oxygen fuel cell.

In the alkaline fuel cell, the electrolyte consists of concentrated, aqueous potassium hydroxide in a stabilizing matrix. Many different electrode materials are used, including carbon, non-precious metals (such as nickel), or precious-metal impregnated polymers.

Interactive CW Animation
Primary cells

At the anode:

$$H_2(g) + 2OH^-(aq) \rightarrow 2H_2O(l) + 2e^-$$

At the cathode:

$$O_2(g) + 2H_2O(l) + 4e^- \rightarrow 4OH^-(aq)$$

Hydroxide ions travel from the cathode to the anode.

In each case, the overall reaction is:

$$2H_2(g) + O_2(g) \rightarrow 2H_2O(l)$$

and the cell potential = +1.23 V.

Rechargeable batteries

AS C.5.2
Describe the workings of rechargeable batteries.

A battery is an electrochemical device that converts chemical energy to electrical energy by means of a voltaic cell. A battery consists of one or more voltaic cells. **Primary cells** are not rechargeable; **secondary cells** are rechargeable. A requirement of these cells is that the products of the discharge reaction must stay in contact with the electrodes. Rechargeable cells are usually designed to have a lifetime of between 100 and 1000 recharge cycles. They are most appropriate for devices that see regular periods of use and non-use, for example mobile phones and cars.

The major similarities and differences between the discharging and recharging processes are summarized in table 3.3.3.

TABLE 3.3.3 COMPARISON OF DISCHARGING AND RECHARGING CYCLES	
During discharge	**During recharge**
ANODE: site of oxidation CATHODE: site of reduction ANODE POLARITY: negative CATHODE POLARITY: positive Electrons travel anode → cathode Chemical energy → electrical energy Cell potential = positive	ANODE: site of oxidation CATHODE: site of reduction ANODE POLARITY: positive CATHODE POLARITY: negative Electrons travel anode → cathode Electrical energy → chemical energy Cell potential = negative

Three commonly used rechargeable cells are described below.

The **lead–acid battery** (lead–acid accumulator) is a secondary cell used in cars. It is a 12 V battery consisting of six 2 V voltaic cells in series. The cathode is lead(IV) oxide packed on a metal grid. The anode is lead. The electrolyte is sulfuric acid (around 4 mol dm^{-3}). During discharge, the lead at the anode is oxidized and the lead(IV) oxide at the cathode is reduced. In both cases, lead(II) sulfate is the product. This adheres to the electrodes as it is formed.

Discharge reactions:

Anode(–):	$Pb(s) + SO_4^{2-}(aq)$	$\rightarrow PbSO_4(s) + 2e^-$
Cathode(+):	$PbO_2(s) + SO_4^{2-}(aq) + 4H^+(aq) + 2e^-$	$\rightarrow PbSO_4(s) + 2H_2O(l)$
Overall reaction:	$PbO_2(s) + Pb(s) + 4H^+(aq) + 2SO_4^{2-}(aq)$	$\rightarrow 2PbSO_4(s) + 2H_2O(l)$
Cell potential:	+2.04 V (per cell)	

Because the reactants are solid, there is no need to separate the cell into anode and cathode compartments. Wood or glass-fibre spacers are put in between the electrodes so they can't touch. As both reactants and products are solids, the cell emf is fairly constant during discharge.

In the recharging process, the reactions are reversed. What was the cathode during discharge now becomes the anode, although its polarity remains the same. The anode during discharge becomes the cathode.

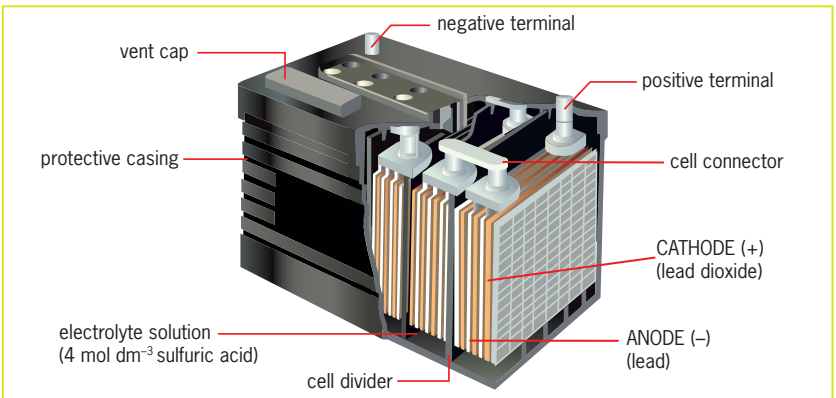


Figure 3.3.5 The lead–acid cell.

Recharge reactions:

Anode(+):	$PbSO_4(s) + 2H_2O(l)$	$\rightarrow PbO_2(s) + SO_4^{2-}(aq) + 4H^+(aq) + 2e^-$
Cathode(–):	$PbSO_4(s) + 2e^-$	$\rightarrow Pb(s) + SO_4^{2-}(aq)$
Overall reaction:	$2PbSO_4(s) + 2H_2O(l)$	$\rightarrow PbO_2(s) + Pb(s) + 4H^+(aq) + 2SO_4^{2-}(aq)$

Degradation of the electrodes occurs with each discharge–recharge cycle, so these batteries have a limited lifespan.

The **nickel–cadmium cell** can be used in virtually any situation requiring a portable energy source, including cordless drills and phones. They are more expensive than dry cells, and tend to lose charge easily. They also suffer from 'memory effect', caused by charging the battery before it is fully discharged. This can reduce its lifespan. Disposal of these batteries can pose a problem due to the presence of the heavy metal cadmium.

In this cell, the nickel-plated case is the negative terminal, collecting current from the cadmium anode. The cell cover is the positive terminal, collecting current from the nickelic hydroxide cathode. During discharge, cadmium(II) hydroxide forms at the anode and nickel(II) hydroxide forms at the cathode. The electrodes are kept apart by a porous fabric separator and the electrolyte is potassium hydroxide. The electrodes and separator are rolled up in a cylindrical arrangement. The cell supplies 1.25 V.

During discharge, the reactions are:

Anode(–):	$Cd(s) + 2OH^-(aq)$	$\rightarrow Cd(OH)_2(s) + 2e^-$
Cathode(+):	$NiOOH(s) + H_2O(l) + e^-$	$\rightarrow Ni(OH)_2(s) + OH^-(aq)$
Overall reaction:	$Cd(s) + 2NiOOH(s) + 2H_2O(l)$	$\rightarrow Cd(OH)_2(s) + 2Ni(OH)_2(s)$

There is an excess of cadmium present so that hydrogen is never formed in a side reaction. Oxygen is formed from the oxidation of water when the nickel plate is fully charged. This reacts with metals within the cell. The major recharge reactions are:

Anode(+):	$Ni(OH)_2(s) + OH^-(aq)$	$\rightarrow NiOOH(s) + H_2O(l) + e^-$
Cathode(–):	$Cd(OH)_2(s) + 2e^-$	$\rightarrow Cd(s) + 2OH^-(aq)$
Overall reaction:	$Cd(OH)_2(s) + 2Ni(OH)_2(s)$	$\rightarrow Cd(s) + 2NiOOH(s) + 2H_2O(l)$

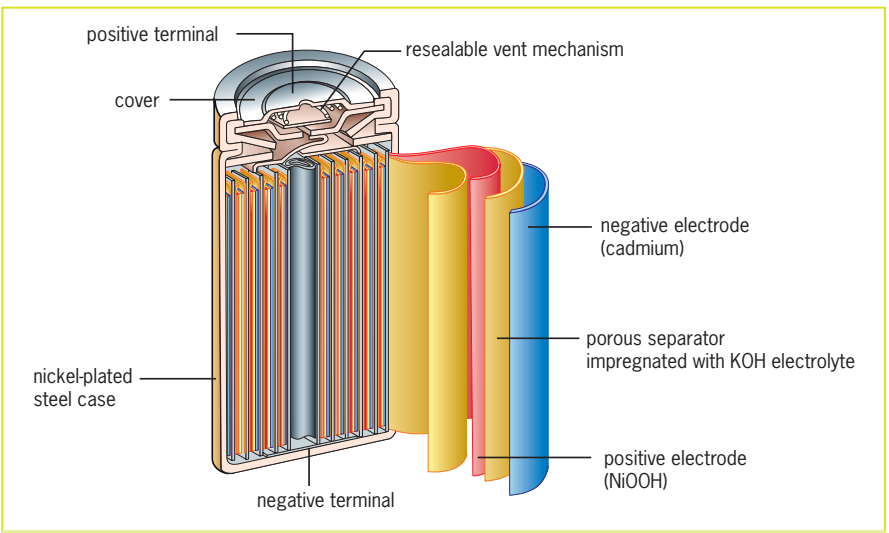


Figure 3.3.6 The nickel–cadmium cell.

Lithium ion cells are the most advanced, but also the most expensive, of the recently developed commercial secondary cells. These cells produce a potential difference of about 3.7 V and have a greater charge density than nickel–cadmium cells. They also lose less charge over time, do not contain toxic substances and don't have the memory effect problems of other cells. These cells cannot contain moisture because it would react violently if any solid lithium was

formed during overcharging. In these batteries, lithium ions move from the negative electrode to the positive electrode while discharging, and move the opposite way when recharging. The negative electrode consists of graphite. There are a variety of electrode materials to choose from for the positive electrode, including layered transition metal oxides such as cobalt(III) oxide. The electrolyte is a lithium salt contained in an organic solvent such as ether.

Using a cobalt(III) oxide electrode, the discharge reactions may be represented as:

Anode(–):	$CLi_x(s)$	$\rightarrow C(s) + xLi^+ + xe^-$
Cathode(+):	$Li_{1-x}CoO_2(s) + xLi^+ + xe^-$	$\rightarrow LiCoO_2(s)$
Overall reaction:	$CLi_x(s) + Li_{1-x}CoO_2(s)$	$\rightarrow C(s) + LiCoO_2(s)$

The opposite process occurs in recharging.

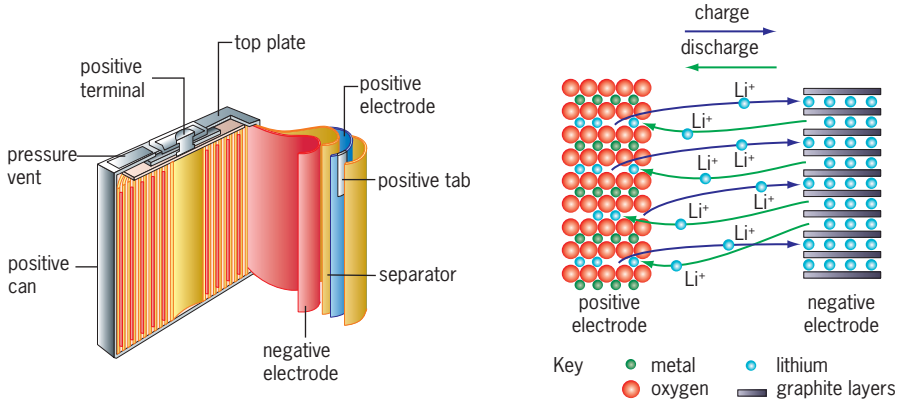


Figure 3.3.7 The lithium ion battery.

WORKSHEET 3.3
A rechargeable battery

Interactive Animation
Cross-section of dry cell

Comparing and contrasting fuel cells and secondary cells

Both fuel cells and secondary cells offer advantages for particular applications.

Similarities

1 They both convert chemical energy to electrical energy.
2 Electrons move from anode to cathode.
3 A large range of sizes are available.
4 Both have high efficiencies because the conversion is direct.
5 The efficiency of both will decrease rapidly once outside normal operating temperatures.

Differences

1 Fuel cells do not store reactants and so could theoretically keep running indefinitely (no maximum to the running time).
2 Fuel cells are generally more expensive.
3 Fuel cells usually run at higher temperatures.
4 Fuel cells are less portable because they often require pumps and cooling systems.
5 In secondary cells, the products stay adsorbed to the electrodes. In fuel cells, the products are removed.

AS **C.5.3 Discuss the similarities and differences between fuel cells and rechargeable batteries.**

6 Secondary cells can be recharged, but it is not relevant to talk about 'recharging' a fuel cell since reactants are constantly supplied.

7 Secondary cells will experience voltage drop during discharge, but fuel cells won't.

Section 3.3 Exercises

1 For each of the following reactions, state whether the catalyst being used is homogeneous or heterogeneous.

a $N_2(g) + 3H_2(g) \rightleftharpoons 2NH_3(g)$
Fe catalyst

b $2MnO_4^-(aq) + 5(COOH)_2(aq) + 6H^+(aq) \rightleftharpoons 2Mn^{2+}(aq) + 10CO_2(g) + 8H_2O(l)$
$Mn^{2+}(aq)$ catalyst

c $2C_2H_4(g) + 4HCl(g) + O_2(g) \rightleftharpoons 2C_2H_4Cl_2(g) + 2H_2O(g)$
$CuCl_2(s)$ catalyst

2 In one stage of the industrial production of sulfuric acid, sulfur dioxide gas is oxidized to sulfur trioxide gas in the presence of a vanadium(V) oxide catalyst. The catalyst is in pellet form and the reacting gases are passed over several catalyst beds. The equation for the reaction is:

$2SO_2(g) + O_2(g) \rightarrow 2SO_3(g)$

a Describe the advantages of the catalyst being in pellet form.

b Is the catalyst homogenous or heterogeneous?

c Suggest why the gases are passed over several catalyst beds.

3 Explain why iodide ions are considered a catalyst for the decomposition of hydrogen peroxide, despite the fact that they undergo chemical change.

4 Discuss the advantages of heterogeneous catalysts compared to homogeneous catalysts.

5 List four factors that must be considered when selecting a catalyst for an industrial process, and explain why these factors are important.

6 Compare the reactions occurring in acid and alkaline hydrogen–oxygen fuel cells.

7 Explain why car batteries are called 'secondary cells'.

8 Evaluate the usefulness of fuel cells as opposed to secondary cells.

3.4 LIQUID CRYSTALS

AS C.6.1
Describe the meaning of the term liquid crystals.

Not all substances move smoothly from their solid to liquid states when heated. A small number of molecular materials pass through a phase with properties intermediate between those of a solid and a liquid. These liquid-crystal phases may take numerous forms. In all forms, the substance is able to flow like a viscous liquid, but retains some crystalline order; within the substance, the molecules are oriented in a particular direction. The orientation of the molecules determines the physical properties of the material such as their electrical conductivity, optical properties and elasticity. The liquid-crystal phase is sometimes referred to as the 'fourth state of matter'.

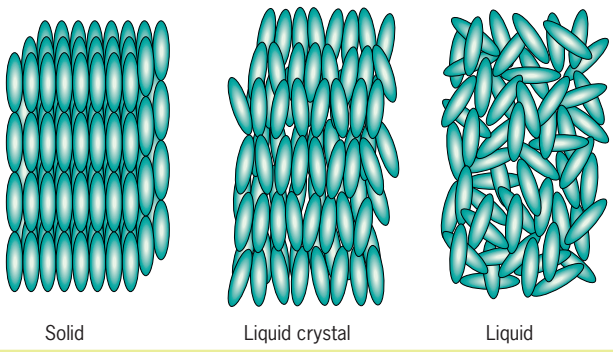


Figure 3.4.1 The particle arrangement in a liquid crystal has elements of both a solid and a liquid.

Liquid crystals were first discovered by Austrian botanical physiologist Friedrich Reinitzer in 1888. When studying cholesterol in carrots, he discovered that cholesterol benzoate does not have a simple melting curve. Instead, it has two separate melting points. At 146°C, it becomes a milky liquid, which turns clear at 179°C. He found that the milky liquid had crystalline properties. Little happened in the liquid-crystal field for decades. It is only relatively recently that Reinitzer's scientific curiosity has been turned into a lucrative commercial field.

Liquid-crystal display devices are now ubiquitous. Liquid crystals are also used in very accurate thermometers. Due to their relatively weak intermolecular forces, certain kinds of liquid crystals are highly sensitive to small changes in temperature. The liquid-crystal polymer Kevlar™ is used extensively in bulletproof clothing. Both inorganic and organic substances have been identified which exhibit liquid-crystal properties under certain conditions. Table 3.4.1 outlines a few of these. Approximately 0.5% of organic compounds have liquid-crystal phases.

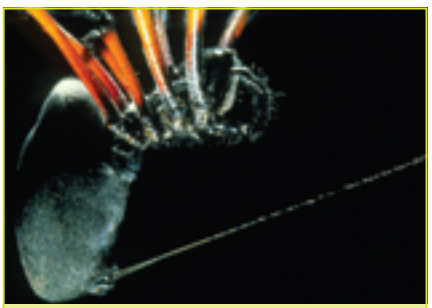


Figure 3.4.2 Spider silk has liquid-crystal properties.

TABLE 3.4.1 EXAMPLES OF LIQUID CRYSTALS

Substance	Description
Graphite	All types of carbon appear to pass through a liquid-crystal phase in their formation. In addition, carbon fibres may be spun out of mesophase (liquid crystal) carbon formed by heating pitches.
Cellulose	Many pure and substituted celluloses form liquid crystals in solution. They can be woven into fibres and are being investigated as possible replacements for viscose rayon. Cellulose liquid crystals may also be used in thermometers.
Kevlar™	In solution, the polymer chains naturally line up next to each other in a liquid-crystal formation. This makes the material crystalline and strong. It is woven into fibres while still in this liquid-crystal state, which reinforces the strength.
Spider web solution	While spinning their webs, the spiders force the solution of silk proteins through a liquid-crystal phase in which the molecules line up to form nanocrystals. Spider web silk is stronger than Kevlar™.
DNA	Both long and short strands of DNA in solution line up to form a liquid-crystal state. The liquid-crystal state of DNA appears to have some role in its functioning.
Soap	The opalescent, gooey liquid at the bottom of a soap dish is a liquid crystal. The soap molecules are arranged in layers in this slimy substance.

AS C.6.2
Distinguish between thermotropic and lyotropic liquid crystals. © IBO 2007

Types of liquid crystals

All liquid crystals fall into one of two categories—*thermotropic* or *lyotropic*. **Thermotropic liquid crystals** are pure substances that pass into a liquid-crystal phase within a certain temperature range, in a reversible process. **Lyotropic liquid crystals** exist in solution, with only a particular concentration range exhibiting liquid-crystal behaviour.

The molecules that form thermotropic liquid crystals are either rod-shaped or have flat, disc-like structures. The latter have linked aromatic rings at the centre. The shapes of these molecules allow them to align closely next to each other. Biphenyl nitriles (shown in figure 3.4.3) are examples of molecules exhibiting this flat structure. Thermotropic liquid crystals are used in liquid-crystal displays.

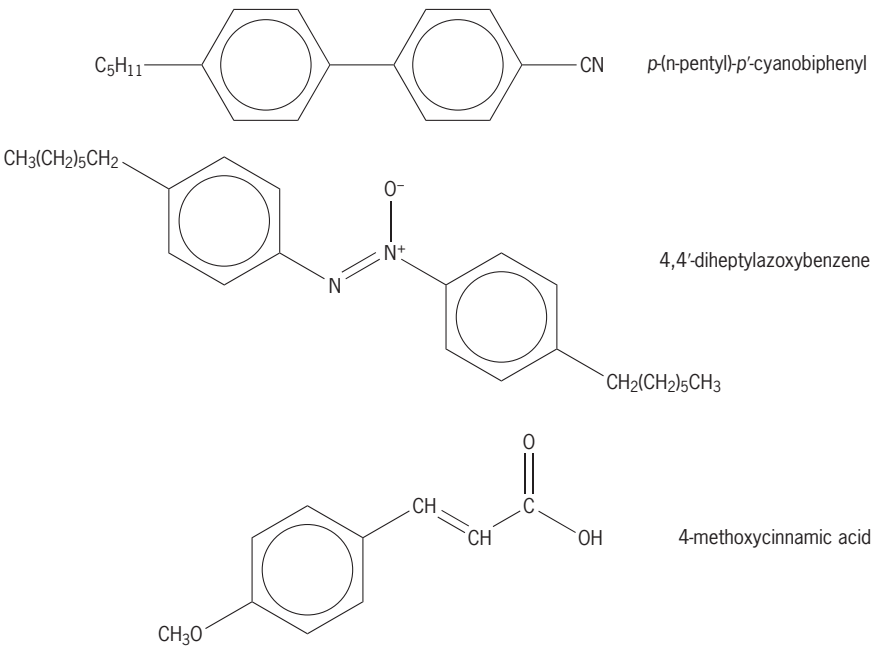


Figure 3.4.3 Examples of substances that exhibit thermotropic liquid-crystal behaviour.

All molecules that exhibit lyotropic liquid-crystal behaviour are surfactants; they have a polar (hydrophilic) head and a non-polar (hydrophobic) tail. Soap is one example. Another example is the lipid bilayer that makes up cell membranes. In water, soap molecules at sufficient concentration will arrange themselves into hollow spheres and rods called **micelles**. At the surface of the micelles, the polar head is oriented outwards to interact with the polar water.

At concentrations of at least 40% w/v (40 g per 100 cm^3), the micelles start to arrange themselves into a loose crystalline structure. At very high concentrations, the liquid crystal may take the form of a bilayer. Applications of lyotropic liquid crystals in medicine include use as a coating for oral pharmaceuticals to prevent them being destroyed in the stomach, and to dissolve hydrocortisone in ointments.

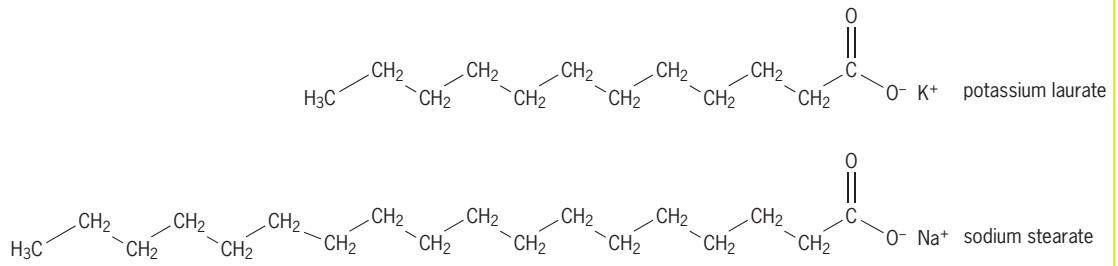


Figure 3.4.4 These soap molecules are capable of forming lyotropic liquid crystals.

Possible arrangements of molecules in liquid crystals

AS C.6.3 Describe the liquid-crystal state in terms of the arrangement of the molecules and explain thermotropic behaviour. © IBO 2007

In a solid, particles are held rigidly in fixed positions. In a liquid, there is a high degree of disorder. In a liquid crystal, the molecules have enough kinetic energy for some intermolecular bonds to be broken. The weakest bonds break first, leaving the others mostly intact. In this state, there is disorder in some directions and order in others. This is the basis of thermotropic behaviour. Once enough thermal energy has been added, there will be disorder in all directions and the liquid crystal will no longer exist.

The possible molecule arrangements within a liquid crystal are described as *nematic*, *smectic* or *cholesteric*. In **nematic liquid crystals**, the molecules are aligned along one axis, but have no order in any other direction. 'Smectic' comes from the Greek word for soap. The slime at the bottom of a soap dish, mentioned earlier, is an example of a **smectic liquid crystal**. In this arrangement, the molecules are more ordered than in the nematic arrangement. Not only are they aligned along a major axis, they are also arranged in layers that can slide past each other. There are various types of smectic arrangements, depending on the orientation of the molecules within the layers.

In **cholesteric liquid crystals**, the molecules are aligned along a main axis in layers. In each successive layer, the molecule orientation is different. Often the twists in orientation are only very slight. The twists give rise to colour, and this type of liquid crystal is used in thermometers.

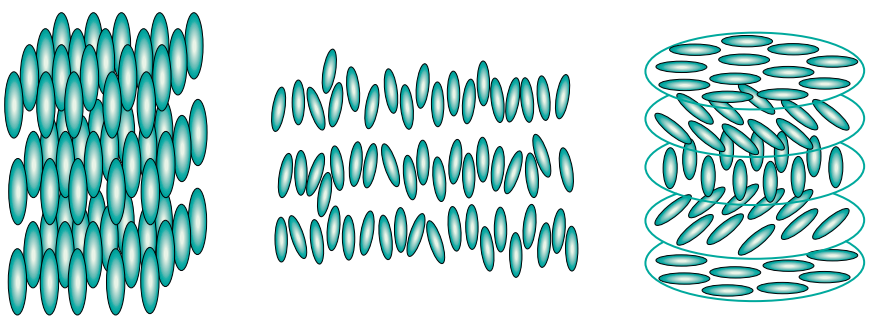


Figure 3.4.5 Possible molecule arrangements in liquid crystals.

Liquid-crystal display devices

C.6.4 Outline the principles of the liquid-crystal display device. © IBO 2007

First developed in 1972, flat screen liquid-crystal displays (LCDs) have gradually crowded out the older, bulkier, less energy-efficient cathode ray tube (CRT) displays. LCDs are seen in many devices, including calculators, computers and digital watches. They are particularly popular for battery-powered devices, since they use only a small amount of power.

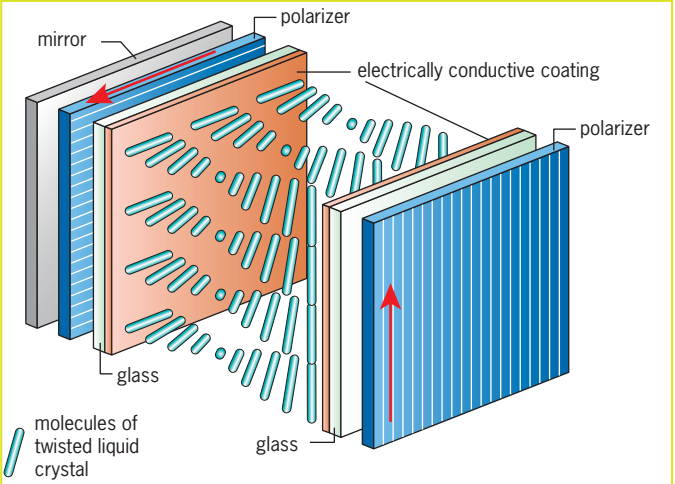


Figure 3.4.6 Device used in a liquid-crystal display.

LCDs take advantage of the fact that a small electric current can produce changes in the orientation of the molecules within a liquid crystal, which change its optical properties. In the usual design, a very thin layer (no more than 20×10^{-6} m) of nematic liquid-crystal material is placed between two transparent electrodes. On the other side of each electrode is a polarizer. At one end, the polarizer is horizontally orientated; at the other it is vertically orientated. Without the liquid-crystal layer, the polarizers would block out all the light.

Within the liquid crystal, the molecules at the end with the horizontal polarizer are oriented horizontally. At the other end, they are oriented vertically. In between, a special process is used to cause the molecules to 'twist' in a regular way, so that the orientation of the molecules gradually varies from one end to the other. This arrangement is known as **twisted nematic**.

Figure 3.4.7 This display appears bright due to the light being reflected by the mirror at the back of each cell.

As light passes through the device, the plane of polarization of the light turns by 90°, thanks to the action of the liquid-crystal substance. This means that the light's polarization now matches the polarizer at the other end, so it can pass through. In devices such as watches and calculators, a mirror reflects the light back, so the display looks bright. If an electric potential is applied across the electrodes, it changes the orientation of the molecules. This stops incoming light from being rotated 90° and thus it cannot pass through the rear polarizer. This makes the display look black. Hence, selectively charging certain display areas can create patterns. Computer screens use backlighting instead of reflected light. They also use red, blue and green filters and sub-pixels to create full-colour displays.

Properties of substances used in LCDs

C.6.5 Discuss the properties needed for a substance to be used in liquid-crystal displays. © IBO 2007

The thermotropic materials chosen for LCDs have to possess certain properties:

- *Wide range of operating temperatures*—It is important that the material stays in the correct liquid-crystal state over a range of temperatures. Most substances used in LCDs work from –40°C to 85°C. Some newer materials can be used at temperatures of up to 125°C, but this is less common.
- *Appropriate voltage threshold*—The material has to respond correctly to the applied voltage; this changes with temperature, so manufacturers have to carefully consider how the material will respond over a range of voltages. If the temperature increases, and the voltage threshold drops significantly, the image may ghost.
- *Chemical stability*—The substance should show high stability so that the device will have a long life. If the material breaks down easily, the device will not last long and the consumer will be most displeased!

- *Fast response to voltage changes*—The display often needs to change very quickly—think about how often your computer display changes each minute. The liquid-crystal material must be able to react quickly. The large, high-resolution screens now available employ materials that move twice as quickly as any in previous models.
- *Precise response*—Not only must the material respond quickly, it must also respond precisely. If the produced molecule orientation isn't just right, the quality of the display will be compromised.
- *Good optical transmissivity*—Obviously, light needs to be able to pass through the material, so optical transmissivity is a required property, although the light is changed on its way through.
- *Polarity*—Charges have to be present within the substance, so that it can respond to electrical fields and thus have its orientation changed.

Generally, mixtures of organic molecules are employed in LCDs, rather than pure substances. This can offer advantages in that each mixture component can be picked because of its superior performance in one area. The disadvantage is that the molecules might not all move in unison—think of a crowded bus braking suddenly. All the people standing will not necessarily all fall at the same speed in the same orientation. The challenge for engineers has been to find the best blends. Only a small quantity is needed. The average 12″ (30 cm) LCD screen contains only 230 mg of liquid-crystal material.

Section 3.4 Exercises

1 Describe how the liquid-crystal state differs from that of a solid or a liquid.

2 List five devices that incorporate liquid crystals.

3 **a** Distinguish between lyotropic and thermotropic liquid crystals.

b Identify one example of a substance that can act as a:

i thermotropic liquid crystal

ii lyotropic liquid crystal.

4 Explain how soap molecules orient themselves in a liquid-crystal solution.

5 Outline the major differences and similarities between nematic, smectic and cholesteric liquid-crystal arrangements.

6 Explain how an LCD on a calculator shows a number.

7 Describe three necessary properties that a liquid-crystal material used in an LCD must possess.

3.5 NANOTECHNOLOGY

A relatively new and emerging field is nanotechnology, which promises many exciting innovations such as stronger, lighter, stain-resistant materials, better skin-protecting and healing cosmetics, more efficient solar cells, and safer and more effective drug-delivery methods.

AS **C.7.1 Define the term nanotechnology. © IBO 2007**

How small is small?

An excellent definition of nanotechnology comes from the 2005 book *Nanotechnology for Dummies*. In it, the authors, Booker and Boysen, describe nanotechnology in the following manner:

> Nanotechnology involves research and technology development at the 1 nm to 100 nm range. Nanotechnology creates and uses structures that have novel properties because of their small size. Nanotechnology builds on the ability to control or manipulate at the atomic scale.

A nanometre (nm) is equal to a billionth (10^{-9}) of a metre. This is incredibly small, but not quite on the atomic scale. The radii of atoms is generally measured in picometres. 1 pm = 10^{-12} m. A large atom such as uranium has a radius of 139 pm (0.139 nm), a red blood cell is 2500 nm wide and hair is at least 60 000 nm wide.

Particles that exhibit certain properties on a macro scale can exhibit quite different properties when broken down to nano-size. These curious effects are called **quantum effects**. Strange things can happen; an insulator on the macro or micro level can become a conductor, and something that was soft can become very strong. Researching, developing and using these properties is at the core of nanotechnology.

It is only recently that scientists have been able to conduct experiments at the nano-level. **Scanning tunnelling microscopes** and a range of other tools and techniques have allowed for a much greater understanding of the nano-world. Another important development was the discovery of **fullerenes** and **nanotubes**—new forms of carbon. Complex nano-materials may be produced from the bottom up—literally one atom at a time, or from the top down, where substances are manufactured on a macro scale then whittled down to nano-size.

CHEM COMPLEMENT

The right tools for the job

A range of new and sophisticated microscopes have been developed, capable of viewing and manipulating tiny particles. The atomic force microscope (AFM) uses a tiny tip connected to a cantilever to travel over the surface of a material. When the tip is attracted to or repelled from the surface, the deflection of the cantilever is measured by a laser. The AFM then produces a profile of the features of the material's surface. The AFM tip can also be used to move small objects such as nanotubes.

The scanning electron microscope (SEM) creates images by bombarding a material with a stream of electrons, which dislodges more electrons from the material's surface. Signals created by the dislodged electrons are used to create an image of the material's surface. In the transmission electron microscope (TEM), a stream of electrons is passed through the material and the pattern of transmitted electrons is used to produce an image of the material.

The instrument with the highest resolution is the scanning tunnelling electron microscope (STEM or STM). This uses an electric current (a tunnelling current) that flows when a very sharp tip moves near a surface and maintains a fixed distance of about 1 nm from it. The tip sits on a piezoelectric tube (it expands or contracts depending on the applied voltage). Applying a small voltage allows tiny adjustments to keep the tunnelling current constant, and the tip at a constant distance from the surface. Analysis of the data produces an image of the surface, and even allows individual atoms to be seen. Using another voltage, applied between the tip and the surface, atoms in the surface layers can be manipulated and construction of materials one atom at a time becomes a possibility. The STM was created at IBM's Zurich Laboratory. In 1989, IBM scientists were the first to successfully manipulate atoms using the STM—they wrote 'IBM' using 35 individual xenon atoms.

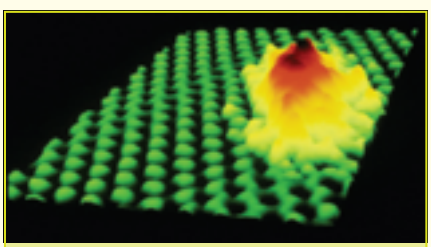


Figure 3.5.1 This STM image shows gold atoms on a graphite surface.

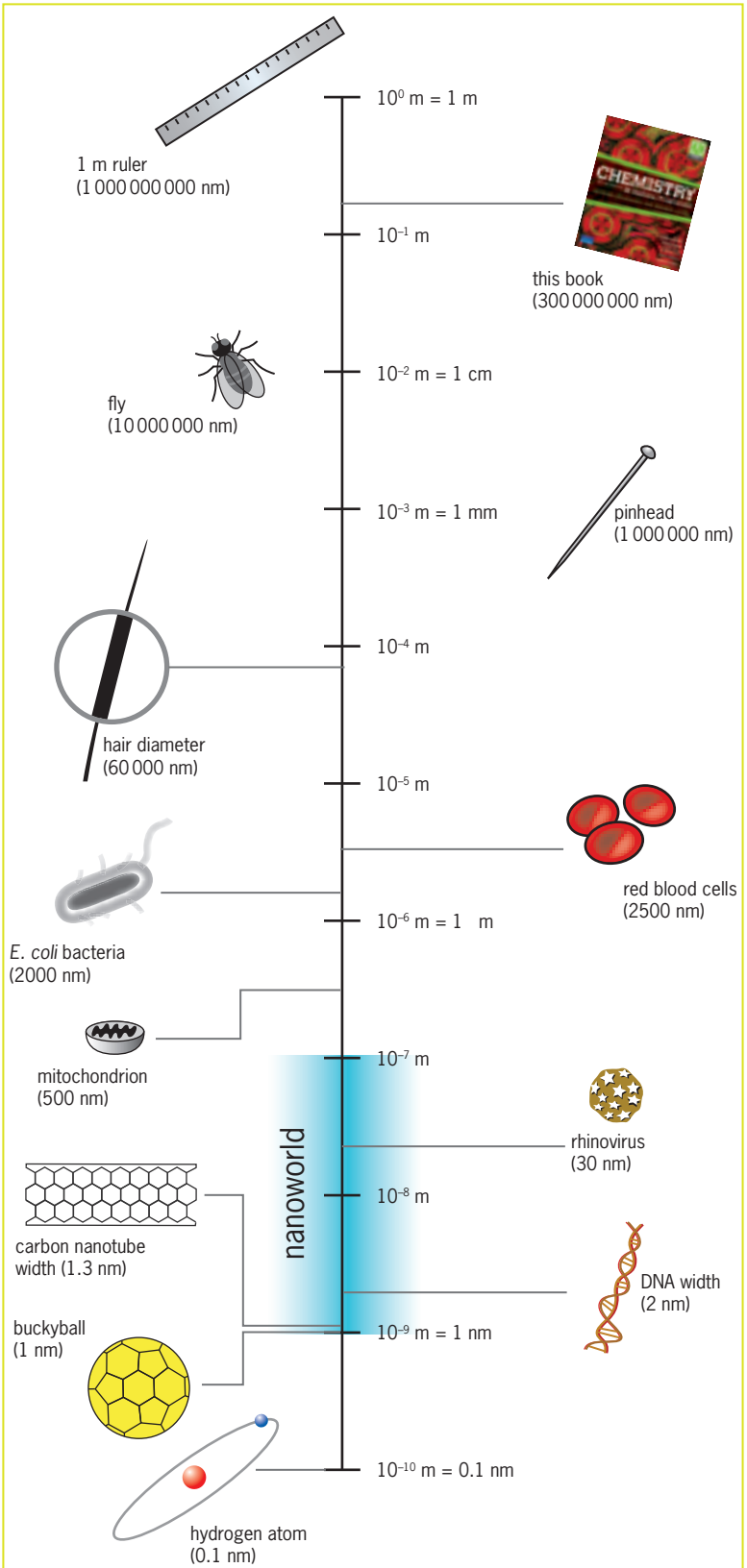


Figure 3.5.2 How small is a nanometre? At each step down the chart, the objects are ten times smaller.

THEORY OF KNOWLEDGE

Our fives senses give us valuable information about the world, but each sense has a limited range of sensitivity, and captures only certain kinds of data. Nevertheless we have no doubts about the existence of the sky, the floor under our feet or the chair we are sitting on. So how do we know then that atoms exist independently of our experience of being able to see them?

Three main theories explain the relationship between perception and reality. The theory of commonsense realism states that 'what you see is what is there'. In scientific realism, the world exists as an independent reality that is different from the way we perceive it. The table we perceive is different from the table that is actually made up of atoms. The last theory, phenomenalism, says that things do not exist outside our experience of them.

The discovery of technologies such as the scanning tunnelling microscope has changed our view of reality and given us access to knowing about the world beyond that which our sense of sight allows. We now know that carbon nanotubes exist independently of our everyday perceptions and experience. However, much of this technology represents nanoparticles, using computer-model simulations. So does this technology blur the distinction between what is a simulation and what is reality?

Manipulating atoms

AS C.7.2 Distinguish between physical and chemical techniques in manipulating atoms to form molecules. © IBO 2007

The STM is the main tool used for manipulating atoms on the nano-scale. The related tools that have been developed are known collectively as scanning probe microscopes (SPM). For example, the tip of an STM could be dipped into a silver crystal, thus coating the tip. By gently touching the tip to another part of the surface, a nano-size crystal can be deposited. This would represent a physical technique, since no chemical reaction has occurred. Physical techniques used to create nano-particles include grinding solids down to the appropriate size and using inert gas condensation.

TABLE 3.5.1 EXAMPLES OF PHYSICAL TECHNIQUES FOR PRODUCING NANO-PARTICLES

Method	How it works	What sort of particles has it been used to produce?
Grinding	Ultrafine grinders are used to produce tiny particles.	Various metal oxides
Evaporation	Self-assembly of molecules into nanostructures can be induced by slowly evaporating the solvent.	Polymerosomes
Inert gas condensation	Particles form from the vapour phase when they collide with inert gas at low pressure.	Metals and metal oxides

Chemical methods for producing nano-particles include precipitation, chemical vapour deposition and spray pyrolysis. All of these techniques involve chemical changes—the breaking of existing chemical bonds and the formation of new ones.

TABLE 3.5.2 EXAMPLES OF CHEMICAL TECHNIQUES FOR PRODUCING NANO-PARTICLES

Method	How it works	What sort of particles has it been used to produce?
Precipitation	Solutions are mixed together in controlled conditions to produce tiny particles.	Ionic salts
Chemical vapour deposition	Precursor gaseous substances react with laser-heated nanoparticles on an inert surface.	Various metal oxides
Spray pyrolysis	Reactants are mixed in a solvent and then sprayed through a flame.	Various metal oxides
Microbial synthesis	Fungi and bacteria produce nano-particles in extracellular reactions.	Gold and silver nano-particles produced by fungi

Nanotubes

AS C.7.3 Describe the structure and properties of carbon nanotubes. © IBO 2007

Carbon nanotubes are an allotrope of carbon and belong to the fullerene family, which also includes buckyballs. In 1991, Japanese scientist Sumio Ijima found a new form of carbon while attempting to synthesize fullerenes. The carbon he found had a structure like a graphite layer that had been rolled into a cylinder. The cap at each end of the cylinder was a half fullerene molecule. The cylinder had a very small diameter—about a nanometre, hence the name carbon nanotube.

Nanotubes can have a width about one ten-thousandth that of a human hair, yet its length can be millions of times greater than its width. Both single-walled (SWNT) and multi-walled nanotubes (MWNT) have been produced. Some nanotubes are closed at the end; others are open. The hexagonal lattice can be oriented differently, producing three kinds of nanotubes. Within the nanotube, the main cylinder is purely hexagons of carbon atoms, as in graphite, while at the curved ends there are alternating hexagons and pentagons, as found in buckyballs. There are many potential applications of nanotubes, including in electronics and textiles, and as catalyst

supports. Nanotubes have a number of unique properties. Bundles of nanotubes:

- are strong due to their strong covalent bonding—their tensile strength is around 100 times that of steel of the same diameter
- are elastic and lightweight—their density is about a quarter that of steel
- can be manufactured as either an insulator or highly conductive fibre depending on the length of the tube and the number of layers—the delocalized electrons in the 2p orbitals of the carbon atoms allow them to conduct
- have high thermal conductivity and can be used in computer circuits to help dissipate the heat generated in the silicon chips
- can act as test tubes and be filled with gases or other molecules—the storage of hydrogen in a nanotube might offer possibilities for hydrogen-powered cars, while the storage of biological molecules is also possible
- can be used as light bulb filaments in place of tungsten
- can be used in place of graphite and glass as a fibre for reinforcing plastics—the toughness of the reinforced plastic is incredibly high. They can even be used in a similar way to strengthen concrete. Carbon nanotubes are the first synthetic material to have a greater strength than spider silk
- attract each other strongly, making it possible to spin threads from them.

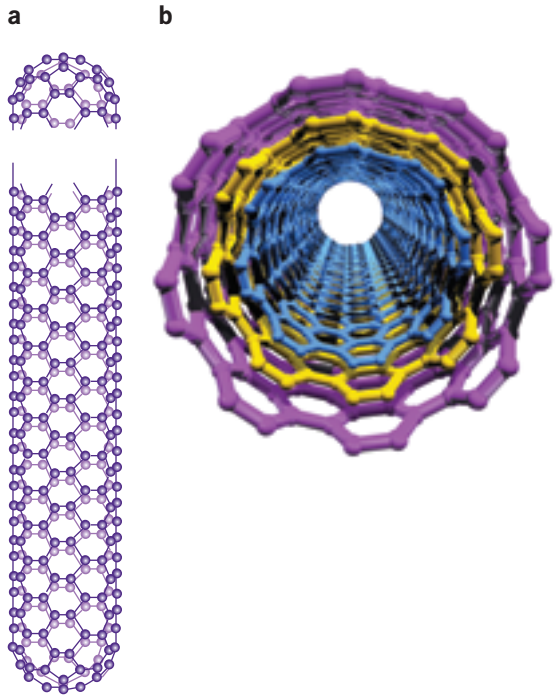


Figure 3.5.3 (a) Single-walled and (b) multi-walled nanotubes.

THEORY OF KNOWLEDGE

Sumio Ijima, while conducting electron microscope research at Arizona State University in the 1970s, observed a spherical carbon structure that he could not identify. It was not until the 1980s when he read a paper published by Smalley and Kroto, the discoverers of C_{60} fullerene molecules that he realized there might be a connection. Ijima wrote to Smalley and Kroto, and Smalley encouraged him to present a paper on his findings. In 1991, after nearly 30 years of research, Ijima finally observed the tubular carbon nanotubes he was looking for. His findings were published in the November 1991 publication of *Nature*, and since then have generated a lot of intense interest and research.

Ijima's work highlights some important aspects of the nature of science.

- Access to up-to-date technologies, such as a good electron microscope, and the freedom to explore and innovate allows scientists to expand their knowledge and understanding.
- The type of research undertaken is influenced by many factors. For Ijima, his job in the research and development department of NEC provided the funding. His own personal interest was in part directed by wanting to know more about the unidentified structures he saw under the electron microscope.
- The commitment to peer review is a crucial part of the process by which new scientific knowledge becomes accepted. By publishing papers and attending conferences scientists stay up to date with recent developments that help inform their research.
- Expectations and their experiences both past and present influence what scientists know.
- From time to time, a major advancement stimulates rapid development into new fields of knowledge.
- Finding strong evidence is important for all scientists. When the verification needed can't be found, they collaborate with others to strengthen their evidence.
- Curiosity is a trait valued in all scientists. It motivates them to ask questions, seek answers and test not only their own ideas but those of other scientists as well.

Implications of nanotechnology

AS C.7.4
Discuss some of the implications of nanotechnology.

Is there a down side to this exciting new nano-world? To quote the director of the lobby group GeneEthics, 'Each type of nanoparticle may be as deadly as asbestos, so the worker and public health challenge is huge.' There is no doubt that negative impacts must be considered, but the possible benefits of nanotechnology are seemingly limitless.

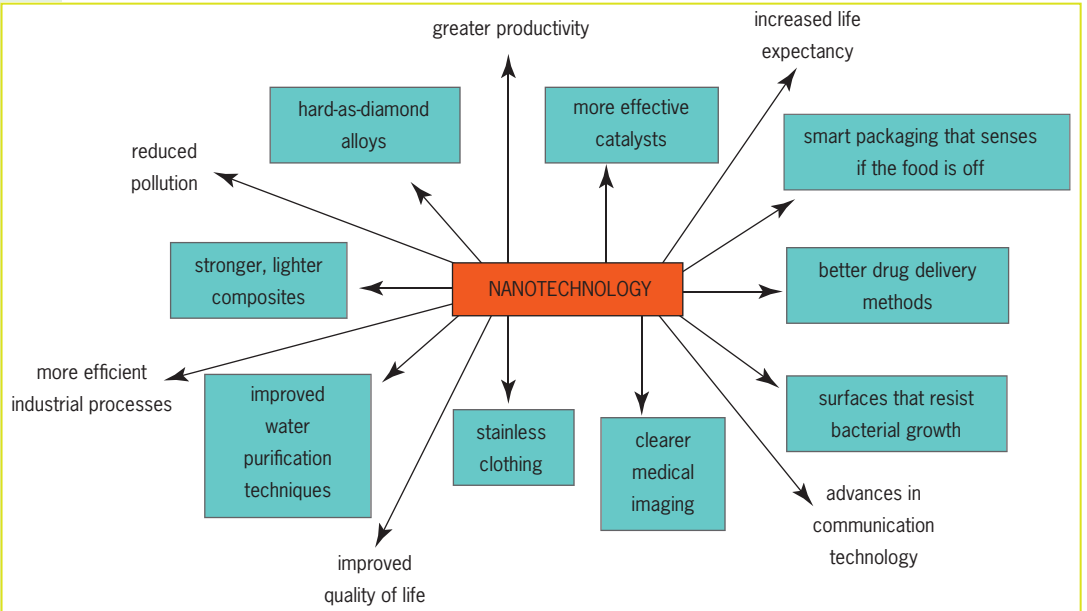


Figure 3.5.4 What nanotechnology could mean for society.

It is highly likely that you already own something which uses nanotechnology—possibly a computer display, stain-resistant pants or a beauty cream. Other applications include:

- increased digital storage capacity, creating smaller and faster computers and other electronic devices
- new drug delivery systems
- improved implants and prosthetics with surfaces that could biologically interact with the body
- bottom-up manufacturing that would use less raw materials and create less pollution
- stronger, lighter construction materials
- more efficient solar cells
- improved waste-treatment facilities.

Because this is a relatively new field, the effects of nano-particles on human health are unknown. Exposure occurs during manufacture, transport and use. There have long been calls for regulation of this burgeoning industry, but how do you produce regulations when so many questions are still unanswered? It is estimated that the nanotechnology industry could be worth a trillion dollars by 2011. The industry has powered ahead in a very short time—it is really only in the last 15 years that nanotechnology has come into its own. Companies have been accused of putting dollars before safety, continuing with product development and distribution before the risks of nanotechnology are fully understood.

Figure 3.5.5 Nanotechnology is already widely used, even though the health effects are largely unknown.

At the moment, material safety laws really only relate to the properties of substances on a macro or micro level. Applying these regulations to the same material on a nano-level is not really valid due to quantum effects. At the moment, many governments have strategies in place for programs that will eventually lead to laws and regulations for the nano-world. An example of this is the Australian Government's National Nanotechnology Strategy.

Research has shown that nano-sized particles can have extremely toxic effects, and could be especially damaging if lodged in the lungs. Some people have suggested that as much care should be taken when dealing with nano-particles as is taken in handling highly radioactive or biohazard substances. Again, regulations are not in place due to a lack of conclusive data. It is possible that nano-particles could pose a more significant threat than bacterial or viral invaders. The normal means by which our immune systems deal with pathogens may not be effective against nano-particles.

Although few studies into the hazards of nano-particles exist, there is enough reason for concern to accept the validity of the arguments to halt the mass production of consumer goods incorporating nanotechnology until the risks have been studied and we know how to handle them. There is a chance, for example, that nano-particles could cross the brain barrier, or be passed from mother to foetus. Considering how different the properties of a substance such as the widely used titanium dioxide (used in sunscreens) might become on the nano-scale, we cannot make the assumption that substances normally considered benign will still be harmless in the nano-world.

There is great concern that companies will continue to ignore the potential hazards of nano-particles, and by the time we discover what they can do, they will already have spread into every home. Although governments have strategies in place, what they really need is firm data. Since this is a new area of science, at the beginning there were no techniques available for determining toxicological effects of such particles. These are being developed, but it will still be a long time before the true safety concerns have been determined.

THEORY OF KNOWLEDGE

Advances in what we know about food, agriculture, manufacturing, communications, sanitation, medicine and transport have been responsible for dramatic changes in how people live and work. Because of this, what is valued in society influences what research and development is valued and funded in science. For example, when society values the need to find alternatives for fossil fuels, wants ways to recycle or extend the life of batteries and know for sure if nano-particles are bad for your health, research in these areas will take place. In addition, economic and political forces influence the direction of science, what research will be carried out and what new technologies will be developed and used.

You are on the selection committee of an organization that offers grants for scientific research into new technologies. What sort of questions would you ask someone who is applying for funding for medical research into a drug-delivery system that uses nano-particles to control the release of antibiotics to the site of an infection?

Section 3.5 Exercises

1 Define *nanotechnology* and explain, in terms of metres, how small the nano-scale is.

2 Describe the role of scanning tunnelling microscopes in the development of nanotechnology.

3 Explain how atoms can be manipulated by both physical and chemical means.

4 **a** Describe the structure of a carbon nanotube.

b List the special properties of carbon nanotubes.

5 Discuss the safety implications of using nanotechnology, and the validity of the statement, 'Each type of nanoparticle may be as deadly as asbestos.'

3.6 MORE ABOUT POLYMERS

HL

Earlier in this chapter, we looked at addition polymers. Now we will examine the other important class of polymers—condensation polymers. We will also look more closely at the mechanisms by which low-density polyethene (LDPE) and high-density poylethene (HDPE) are manufactured.

Condensation polymers

AS C.8.1 Distinguish between addition and condensation polymers in terms of their structures. © IBO 2007

AS C.8.2 Describe how condensation polymers are formed from their monomers. © IBO 2007

In contrast to addition polymerization, monomers used to form condensation polymers such as polyester and nylon do not need to contain a carbon–carbon double bond; instead, they must contain reactive functional groups. A **condensation reaction** is one in which larger molecules join covalently. Each new bond causes the formation of a small molecule, normally water. An example is ethanoic acid and ethanol undergoing an esterification reaction to form an ester, ethyl ethanoate, and water:

$$CH_3COOH(l) + CH_3CH_2OH(l) \xrightarrow{H_2SO_4(l)} CH_3COOCH_2CH_3(l) + H_2O(l)$$

If monomers were available that each contained two functional groups, aligned at opposite ends of the molecules, the process could be used to form long chains.

If a monomer with two hydroxyl groups reacted with a monomer with two carboxyl groups, the esterification reaction continues and *polyester* is formed. In situations such as this, where two different monomers are used to form the polymer chain, the product is called a **copolymer**.

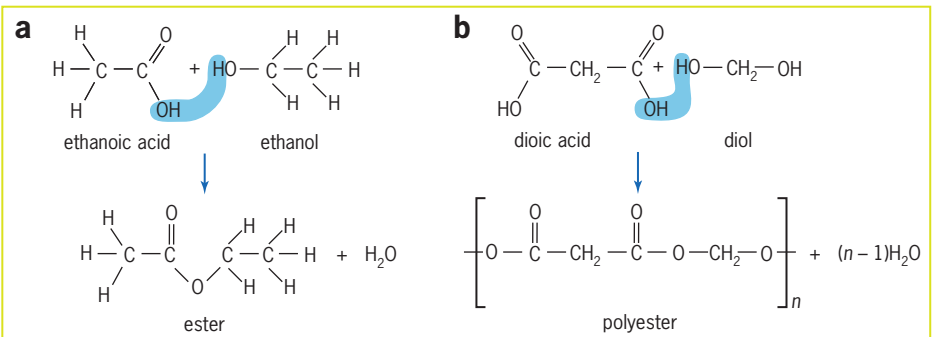


Figure 3.6.1 (a) Ester and (b) polyester formation.

Natural examples of condensation polymers include proteins and cellulose. In addition polymers, the backbone of the long chain is composed of carbons linked by single carbon–carbon bonds. In contrast, the condensation polymer has a backbone of mostly carbon, broken up at regular intervals by the link formed in the condensation reaction that formed it. Commonly, this is an ester, ether or peptide linkage.

Polyurethane foam is formed from a two-part mixture. The first part contains a diol or triol (most commonly glycerol), a catalyst, a blowing agent (low boiling point liquid) and a silicone surfactant. The other part contains a diisocyanate. The polymerization reaction is rapid, and the water causes some of the isocyanate to decompose, forming carbon dioxide that produces the foam. Unfoamed polyurethane is spun into Spandex fibres and sold as Lycra. An example synthesis pathway is shown in figure 3.6.2.

QuickTime CW Video
Synthesis of nylon 6–10

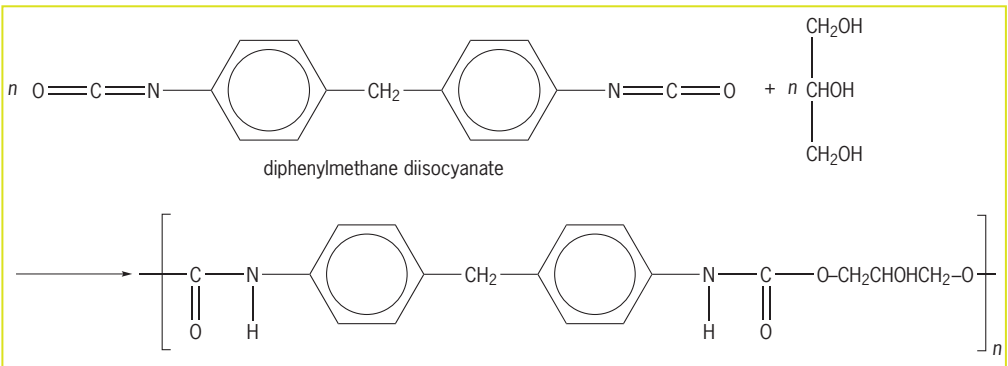


Figure 3.6.2 The formation of polyurethane.

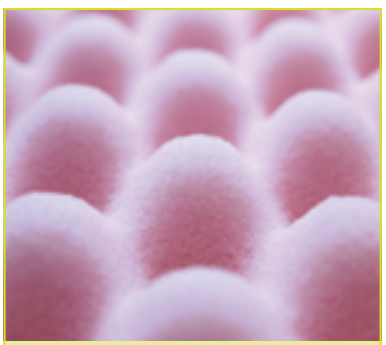


Figure 3.6.3 Polyurethane is used in footwear, automobile parts and insulation.

Polyethylene terephthalate (PET or PETE, polyethylene terephthalic ester) is a type of polyester used in clothing, fibres and drink bottles. It is one of the five most-produced polymers in the world. PETE fibres are called Dacron and Fortrel; PETE film is called Mylar. To make PETE, the diol ethylene glycol (1,2-ethanediol) is reacted with the diester dimethyl terephthalate ($CH_3OOCC_6H_4COOCH_3$) or terephthalic acid (TPA). The latter is more commonly used. In the TPA process, the intermediate bis-hydroxyethyl terephthalate is first formed in aqueous solution then polymerized using reduced pressure, heat and a catalyst.

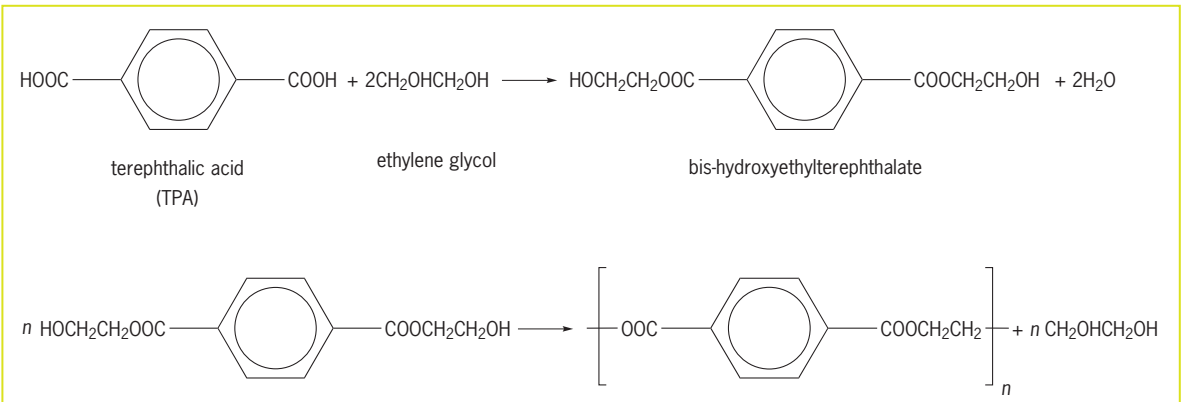


Figure 3.6.4 The formation of polyethylene terephthalate.

Phenol–methanal plastics, also called Bakelite, are the oldest synthetic polymers. They are unusual in that the monomers each contain only one functional group. Phenol (C_6H_5OH) joins with methanal (CH_2O) to form the copolymer. The reaction relies on the fact that the hydrogens in the *ortho* position (i.e. in the position next to the hydroxyl group) in phenol may react with the methanal, with the formation of a water molecule. The polymer formed is strong, black and resistant to heat, so is often used in electrical applications.

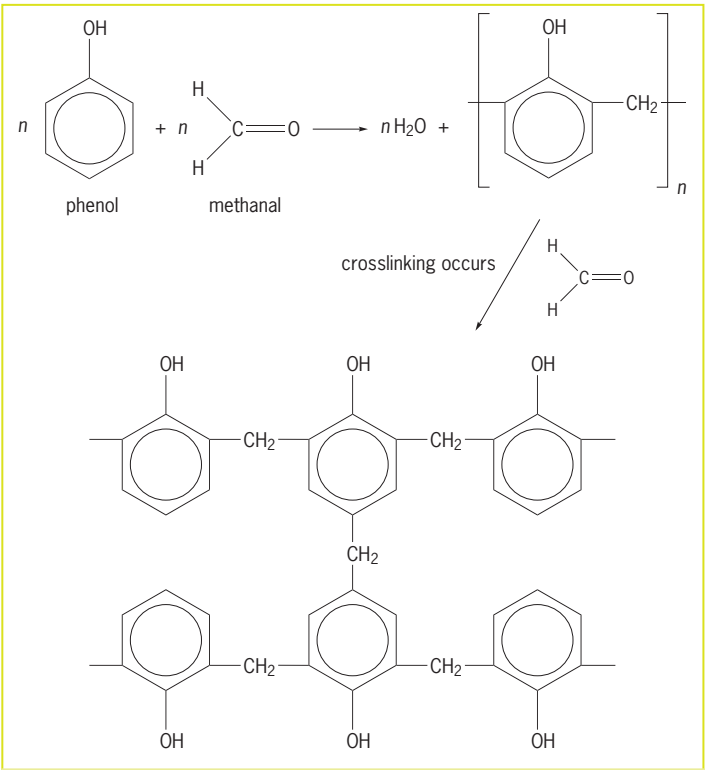


Figure 3.6.5 The formation of phenol–methanal polymers.

Structure and properties

AS C.8.3
Describe and explain how the properties of polymers depend on their structural features.
© IBO 2007

As described in section 3.2, the properties of polymers will depend on the strength of their intermolecular forces, which rely on chain length, degree of branching and polarity, and the presence or absence of cross-links. Bulky side groups will inhibit the molecules from approaching too closely and will decrease the amount of crystallinity in the structure.

Due to its amide linkages, polyurethane chains are able to hydrogen bond to each other. This increases the strength of their intermolecular forces, compared to a polymer such as PETE. Similarly, Kevlar™ chains are able to hydrogen bond to each other. Like polyurethane, they also contain amide linkages between monomers in the chain. The hydrogen bonding between Kevlar™ chains causes them to line up in a regular fashion. The strong bonding within the polymer molecules, the strong bonding between the chains, and the regular arrangement of the chains in the material combine to make Kevlar™ incredibly strong. It is used for ropes, fireproof material, bulletproof vests and sports gear. Cross-linking greatly increases the strength of a polymer, and changes a thermoplastic polymer into a thermosetting one. In the initial formation of phenol–methanal plastic, a linear chain is produced. Cross-links may then be formed using the *para* hydrogen (i.e. in the position directly opposite the hydroxyl group) in phenol. This results in a strong, black plastic. Hydrogen bonds are also possible between chains due to the hydroxyl group on phenol.

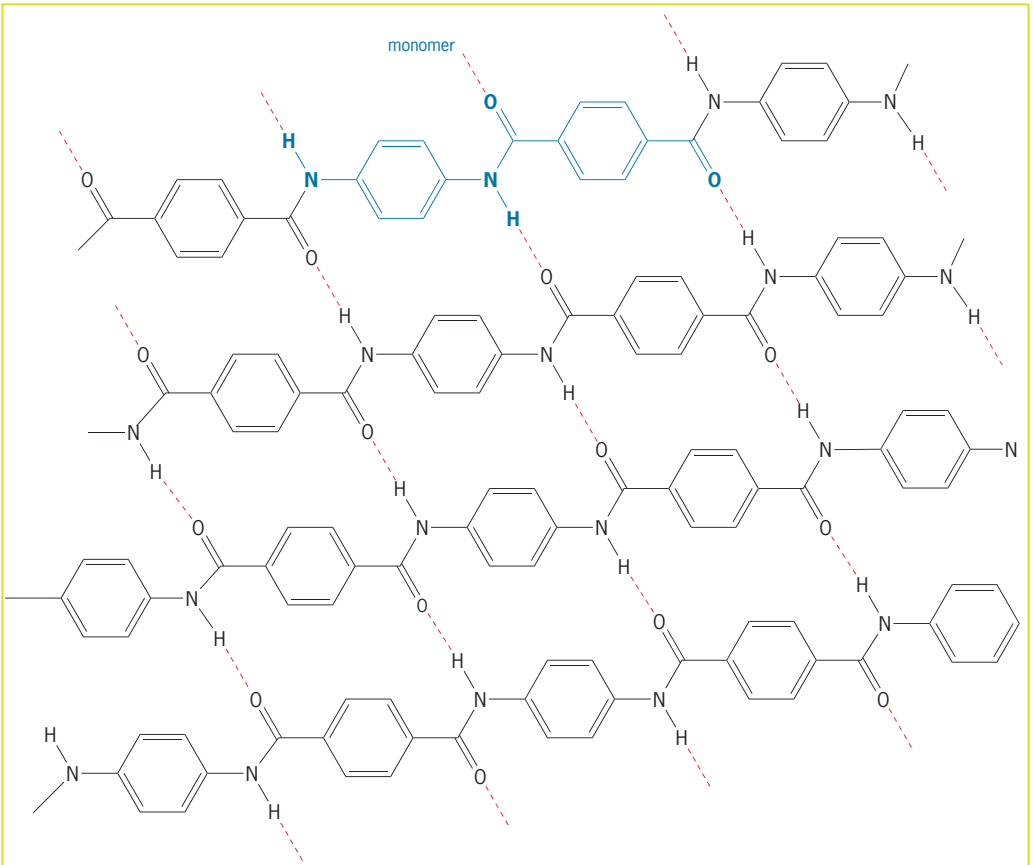


Figure 3.6.6 The structure of Kevlar™.

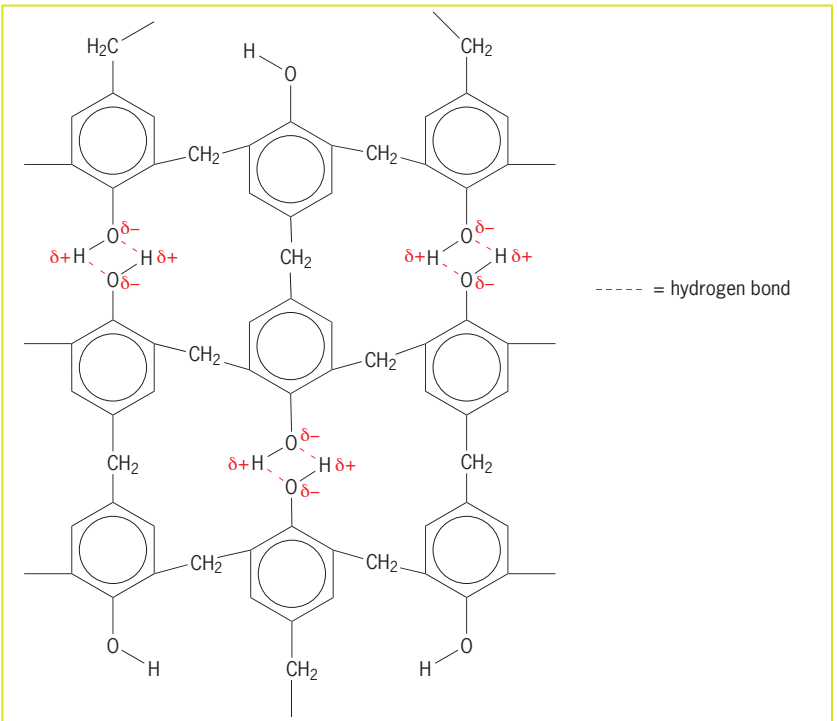


Figure 3.6.7 Cross-links in a phenol–methanal polymer.

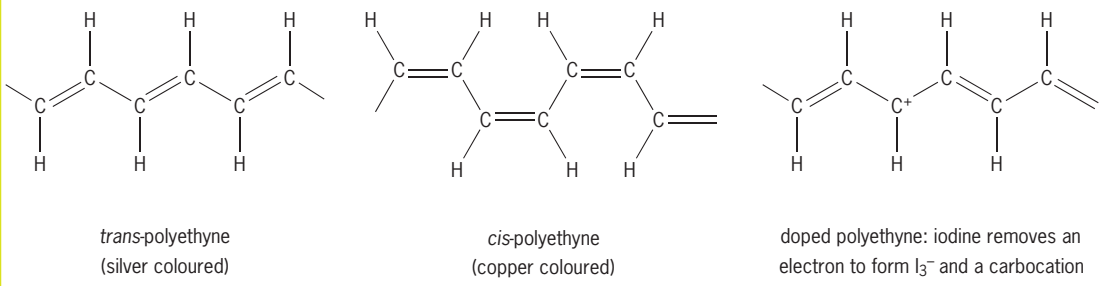
Modifying polymers

AS C.8.4 Describe ways of modifying the properties of polymers. © IBO 2007

The modification of addition polymers has previously been described. This section describes a few more ways in which polymers may be modified. Many additives are currently available, including UV absorbers, dyes, antimicrobials, flame retardants, plasticizers, antistatic agents, discolourant inhibitors, antioxidants and cross-linking agents.

Polyurethane has the toughness of steel and better elasticity than rubber. Polyurethane foam is widely used, but can be susceptible to ageing effects. To combat this, additives such as the UV stabilizer hydroxybenzotriazole and antioxidants such as polymeric hindered phenols are employed. The bubbles of gas in the foam improve the insulation capacity. Normal blowing agents are water (which reacts to form carbon dioxide) and methylene chloride. If air seeps into the foam, it reduces its insulation capacity.

Polyethyne belongs to a class of polymers called conducting polymers. Due to the conjugated double bonds in the structure, it is able to conduct electricity. It was discovered accidentally when a student added 1000 times too much catalyst in a reaction. The red cis form is unstable, but the silver-blue trans form is stable.



trans-polyethyne (silver coloured) | *cis*-polyethyne (copper coloured) | doped polyethyne: iodine removes an electron to form I_3^- and a carbocation

Figure 3.6.8 The forms of polyethyne.

Doping involves adding an agent that either adds electrons to (reduces) the polymer or takes electrons from (oxidizes) the polymer. It has been discovered that doping polyethyne with iodine oxidizes the polymer, increasing its conductivity from 10^{-3} S m^{-1} to 3000 S m^{-1}. The iodine attracts pi-bond electrons, creating 'holes' in the structure. This makes electrons jump from one end to the other to fill the 'holes'. This discovery earned Alan MacDiarmid, Alan Heeger and Hideki Shirakawa the 2000 Nobel Prize in Chemistry.

The popular material polyester is strong, but has some drawbacks. It is difficult to dye, and feels warm and heavy against the body. High temperatures and pressures are needed to print and dye polyester. To address these problems, polyester is usually blended with other fibres, such as cotton at a 60% cotton, 40% polyester ratio. It is also sometimes blended with wool. Techniques have been developed to produce more comfortable polyester fibres, so 100% polyester clothing is not uncommon, despite its lack of breathability.

Advantages and disadvantages of polymers

AS C.8.5 Discuss the advantages and disadvantages of polymer use. © IBO 2007

Polymers overtook natural materials for many uses over 50 years ago. They are readily produced and their properties can be designed to match a particular need. For example, polyurethane comes in a wide range of hardnesses. Plastics are lightweight, are excellent thermal and electrical insulators, and are resistant to chemical attack. Polymer materials are tougher than those made

of natural fibres and last longer. Most are derived from petroleum, a non-renewable resources. Many can be recycled, but much recyclable plastic still ends up in landfill.

Disadvantages of polymers include the toxic substances used in their manufacture. The diisocyanates used to make polyurethane are known carcinogens. Chlorofluorocarbons, which were formerly used as flowing agents are known to deplete ozone. Another disadvantage of polymer use is their lack of stability to heat. Thermoplastic polymers melt easily, while thermosetting polymers char, often giving off toxic fumes. Polyurethane foams are also not biodegradable. However, they are incredibly tough, with excellent wear properties.

PETE is widely used in beverage bottles and fibres. It is strong, unreactive, and has excellent transparency. Much PETE is recycled, and the uses for recycled PETE include making polarfleece fabric. It is not biodegradable and ends up in landfill if not recycled. The strong, heat resistant phenol–methanal plastics are distinctively black. While their strength and excellent resistance to heat and electricity make them ideal for electrical insulation, they are not recycled and also generally end up in landfill. In landfill, the great stability of plastics makes them unlikely to produce groundwater, soil or air pollutants. Nonetheless, they represent wasted resources.

Producing polyethene—reaction mechanisms

AS **C.9.1 Describe the free-radical mechanism involved in the manufacture of low-density polyethene. © IBO 2007**

Low-density polyethene, as described earlier, is produced at high temperatures and pressures in a reaction initiated by organo-peroxides such as benzoyl peroxide. The reaction occurs in three steps: initiation, propagation and termination.

1 *Initiation*—In this mechanism, benzoyl peroxide is used as the catalyst. Initiation involves the catalyst molecule being cleaved in two. The two parts are identical, so this is called **homolytic cleavage**. When the covalent bond holding the two parts together is broken, one electron from the bond goes to each part. This means that each part is now a **free radical**—a highly reactive species with an unpaired electron. It is the desire of the radical for another electron that drives the next part of the reaction.

2 *Propagation*—In the next step, the electron-hungry free radicals attack the double bond of ethene. The radical attracts an electron from the double bond to form a single bond with the carbon at one end of the molecule. The other carbon now has an unpaired electron, so the molecule is still a radical. It then attacks another ethene molecule. This process occurs again and again. Note that the chains grow from only one end.

3 *Termination*—Eventually, an event occurs that stops the chain from growing any larger. The normal mechanism of termination is for the radical ends of two growing chains to react together in a coupling reaction. When this happens is a matter of chance, so many different polymer chain lengths are found in any polyethene sample.

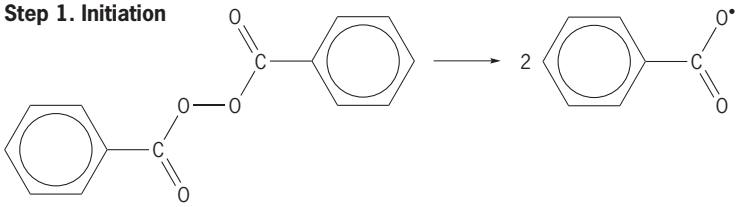
There is actually another way in which termination can occur, called *disproportionation*. In this process, two growing chains meet up and something curious happens; one of the free radical ends takes a hydrogen and an electron from the carbon atom that is second last in the chain. The chain that has lost a hydrogen ends up with a double bond. The unpaired electron at the end of a growing chain sometimes collides with the middle of another chain and pair itself with an electron from a carbon–hydrogen bond, stealing the hydrogen in the process. This is called 'chain transfer to polymer' and results in a high degree

AS **C.9.2**
Outline the use of Ziegler–Natta catalysts in the manufacture of high-density polyethene.

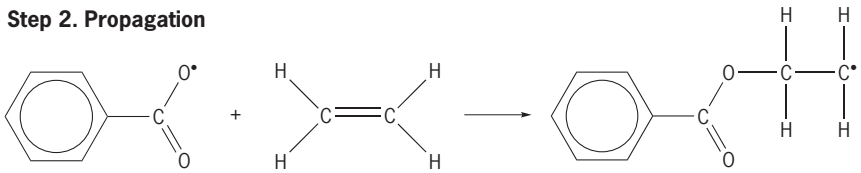
of branching, since it leaves an unpaired electron in the middle of a chain. This occurs often in LDPE, which is why the chains have so much branching.

Until the 1950s, LDPE was the only type of polyethene available. That situation changed thanks to the German scientist Karl Ziegler. In his attempts to produce polyethene at atmospheric pressures, Ziegler initially used aluminium triethyl, but the chains produced were too small. He tried reacting the metal alkyl with an organometallic catalyst, titanium tetrachloride. At 100°C and atmospheric pressure, this produced polyethene with little branching. Ziegler's colleague Giulio Natta used these catalysts to produce polymers, and they were jointly awarded the Nobel Prize in Chemistry in 1963. The catalysts have become known as Ziegler–Natta catalysts.

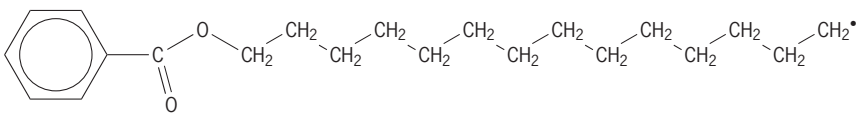
Step 1. Initiation



Step 2. Propagation



The new radical formed hits another ethene monomer and the process continues, building larger and larger chains.



Step 3. Termination

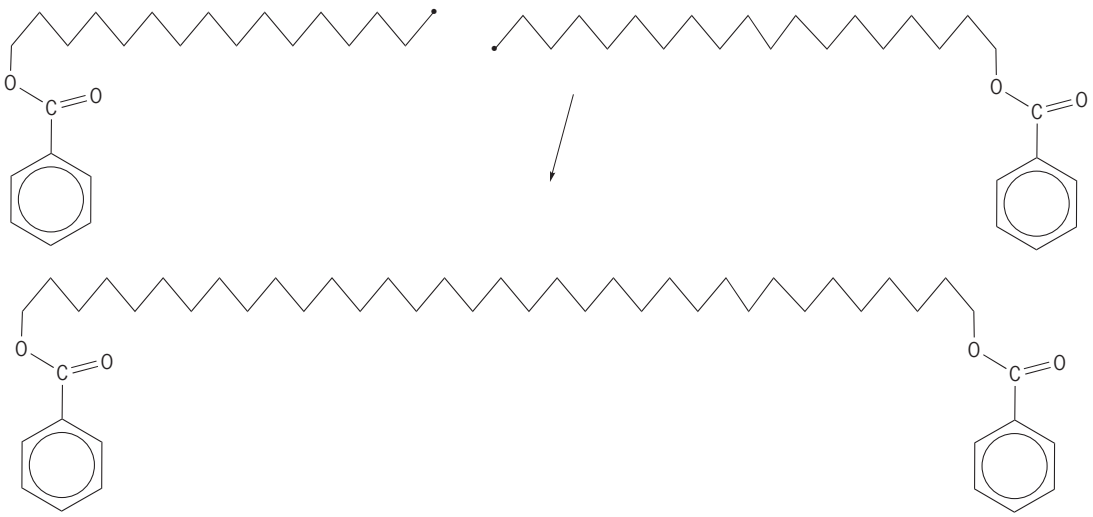


Figure 3.6.9 LDPE production reaction mechanism.

At the start of the HDPE manufacturing process, a metal alkyl and an organometallic compound are reacted together to form the catalyst. The reaction forms an organometallic complex, which contains an active site. Ethene gas reacts with the active site to form the catalyst. This mechanism is not fully understood. If, as suggested by Natta, the active site is the Ti–C bond, the mechanism might look something like that shown in figure 3.6.10.

$Ti–R + CH_2CH_2 \rightarrow Ti–CH_2CH_2R$

$Ti–CH_2CH_2R + CH_2CH_2 \rightarrow Ti–CH_2CH_2CH_2CH_2R$

and so on.

One way to terminate chain growth is to add H_2:

$Ti–CH_2CH_2CH_2CH_2CH_2CH_2R + H_2 \rightarrow Ti–H + CH_3CH_2CH_2CH_2CH_2CH_2R$

Figure 3.6.10 Proposed polymerization mechanism for HDPE.

In this process, some branching still occurs, but much less than is found in LDPE.

Section 3.6 Exercises

1. Identify the characteristics that the monomers making up a condensation polymer must have that differ from those that make up an addition polymer.
2. List the functional groups needed to form links between monomers in:
   - **a** polyester
   - **b** phenol–methanal plastic
   - **c** polyurethane.
3. Describe how polyurethane foam is formed.
4. Describe the structural features that make Kevlar™ one of the hardest known substances.
5. Explain how iodine doping changes the properties of polyethyne.
6. List three advantages and three disadvantages of condensation polymer use.
7. Outline the mechanism by which LDPE is produced using organo-peroxide catalysts.
8. Explain why less branching forms in HDPE than in LDPE.

3.7 SILICON AND VOLTAIC CELLS, AND LIQUID CRYSTALS

HL

Semiconductors were first studied almost two centuries ago. With charge-carrying properties between those of pure conductors and pure insulators, semiconductors underpin the field of electronics. Silicon, a semiconductor, was first isolated by Jons Jakob Berzelius in 1924.

Doping silicon

AS **C.10.1**
Describe the doping of silicon to produce p-type and n-type semiconductors. © IBO 2007

As a semiconductor, silicon's capacity to conduct can be manipulated by adding impurities. This process is called doping. The difference between conductors, insulators and semiconductors is best explained using **band theory**. Within a sample, there is some splitting between allowable energy levels of electrons such that, rather than discrete levels being present, the split levels blend to form bands. The valence electrons can be visualized as occupying a valence band. If energy is added to the substance, an electron may move from the valence band to the next band up—the conduction band.

Electrons can only move up to the conduction band if the gap in energy between this band and the valence band is relatively small. In conductors, the gap is small and easily bridged. In insulators, the gap is large and few electrons can move up to the conduction band. Semiconductors sit somewhere in between. They are poor conductors at low temperatures, but at higher temperatures the thermal energy helps electrons to make the leap to the conduction band.

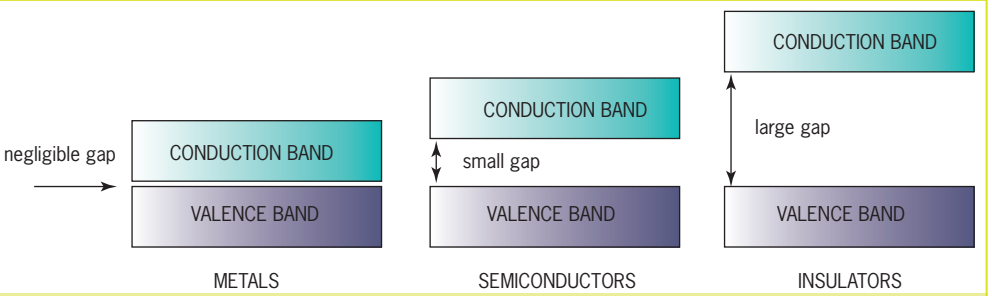


Figure 3.7.1 Gaps in energy between the valence and conduction bands in the three types of materials.

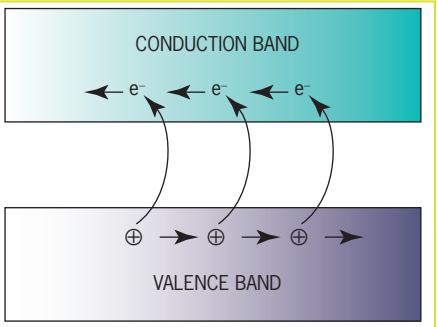


Figure 3.7.2 Holes and electrons move in opposite directions when an electric field is applied.

Both positive and negative charge carriers are present in semiconductors. When an electron moves up to the conduction band, it leaves behind a 'hole', which is simply a vacant site in the crystal. This hole can migrate; a nearby electron can move to fill the hole, thus leaving a hole where that electron used to be, and so on. If an electric field is applied, the holes move in the direction of the field and the electrons move in the opposite direction. Pure crystals in which the number of electrons in the conduction band equals the number of holes are known as an **intrinsic semiconductor**.

If a small quantity of an element that has at least one more valence electron than silicon is added, an **n-type** (negative type) **semiconductor** is created. Examples of such elements are arsenic, antimony and phosphorus. Four of the valence electrons of the doping agent bond with silicon, but the fifth is 'left over'. The energy level of this 'free' electron lies somewhere in the gap between the valence band and the conduction band. The electron can easily move up to the conduction band when energy is supplied. The added element is called a **donor atom** because it contributes free electrons. This type of doped semiconductor is called a negative type (n-type) because the charge carriers are electrons.

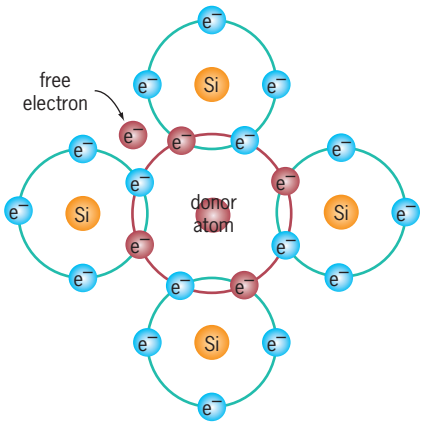


Figure 3.7.3 The pentavalent donor atom (a group 5 element) contributes free electrons.

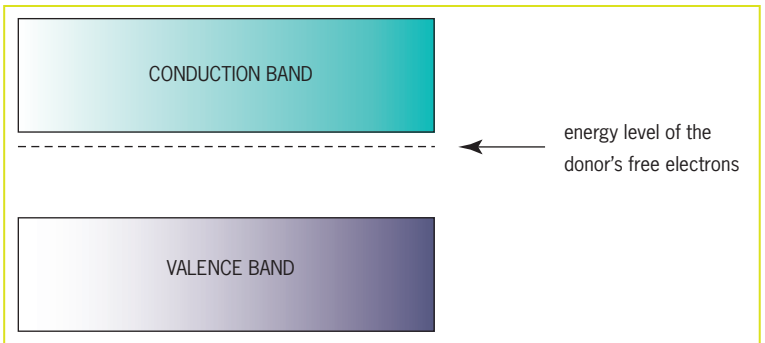


Figure 3.7.4 An n-type semiconductor.

When silicon is doped with an element such as aluminium or indium that has at least one less valence electron, a **p-type** (positive type) **semiconductor** is formed. The added element bonds to silicon, but doesn't have enough valence electrons to pair up with all of silicon's valence electrons. This creates 'holes', or electron deficits. The energy level of the holes sits above the valence band. Electrons can move from the valence band into the holes. This type of semiconductor is called a positive type (p-type) because the charge carriers are essentially positive holes. If the conduction characteristics are largely determined by the doping agent, the semiconductor is known as an **extrinsic semiconductor**.

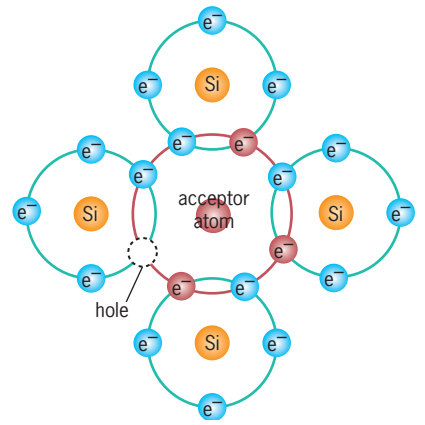


Figure 3.7.5 The trivalent acceptor (a group 3 element) creates positive holes.

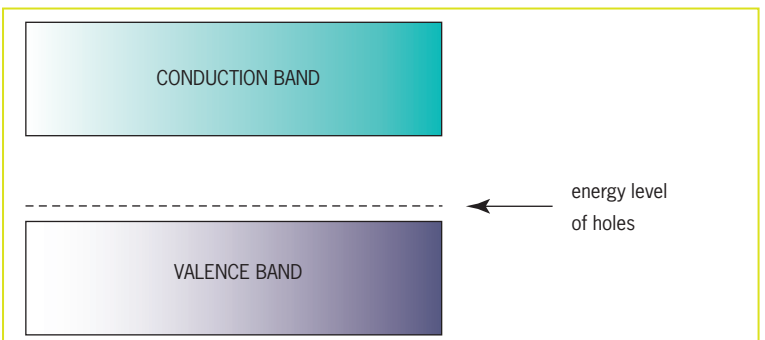


Figure 3.7.6 A p-type semiconductor.

Sunlight and semiconductors

Sunlight is composed of photons of electromagnetic radiation. Light striking a semiconductor may be transmitted, reflected or absorbed. Solar cells consist of a layer of a p-type semiconductor covered by a layer of n-type semiconductor. Where the two layers meet, a depletion layer exists with excess positive charge on the n side and excess negative charge on the p side.

AS **C.10.2 Describe how sunlight interacts with semiconductors.**

When sunlight falls on this junction, light with wavelength less than 1.13×10^{-6} m causes electrons to move to the n-type side, while the holes move to the p-type side. If a load is connected across the junction, electricity flows. Advantages of solar cells over conventional power stations include higher efficiency, no noise and no pollution. A drawback is their expense.

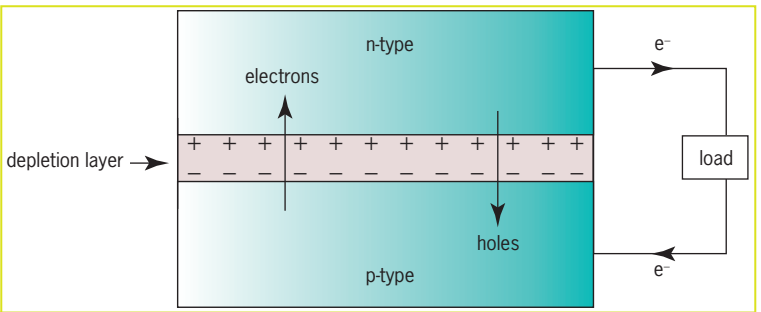


Figure 3.7.7 The movement of charge in a photovoltaic cell.

Liquid crystals

AS **C.11.1 Identify molecules that are likely to show liquid-crystal properties, and explain their liquid-crystal behaviour on a molecular level. © IBO 2007**

Many different substances that exhibit liquid-crystal properties have now been identified. Because liquid-crystal behaviour relies on the molecules aligning themselves along a particular axis, liquid-crystal molecules are generally rod-shaped. In addition, they have some degree of polarity. A more or less rod-shaped structure allows the molecules to align themselves and pack reasonably closely together. Polarity increases the strength of the intermolecular forces and allows the molecules to respond to electric fields.

The structure of the first-known liquid crystal, cholesterol benzoate, seems complicated at first glance, but it is essentially a molecule with a polar head and non-polar tail. The twisted tail produces a particular type of liquid crystal. The successive layers in the liquid crystal are orientated differently from each other. This is the *cholesteric arrangement*. This is an example of a discoid liquid-crystal molecule. The other major class of liquid-crystal molecules is the rod-shaped molecule.

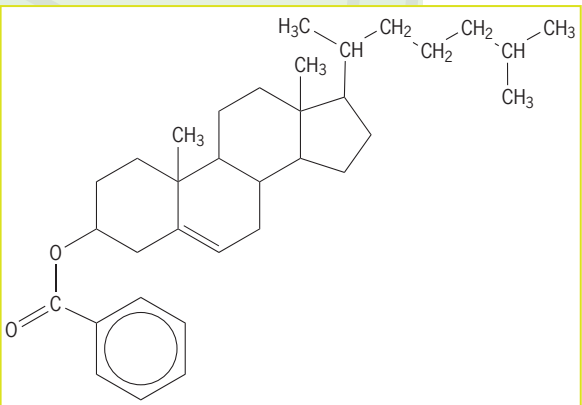


Figure 3.7.8 Cholesterol benzoate forms a cholesteric liquid crystal.

A class of liquid-crystal molecules much-used in liquid-crystal displays are the biphenyl nitriles. If the phenyl groups could rotate, the intermolecular bonds would not be as strong. Having the phenyl groups directly joined to each other keeps the structure flat. The nitrile group makes one end of the molecule polar. An alkyl group is generally attached at the other end, although sometimes a polar group may be located at that end too. Figure 3.7.9 shows examples of biphenyl nitrile compounds. Note the rod-like structure that is common to all of them. This means that the molecules will clearly prefer to line up along their long axis. It is important that most of the molecules are non-polar, so that the intermolecular forces are not so strong that the substance solidifies too easily. Note that sometimes a linking group is placed between the phenyl groups. If this is the case, it must be a group that allows the planar, non-rotatable nature of the phenyl groups to be maintained.

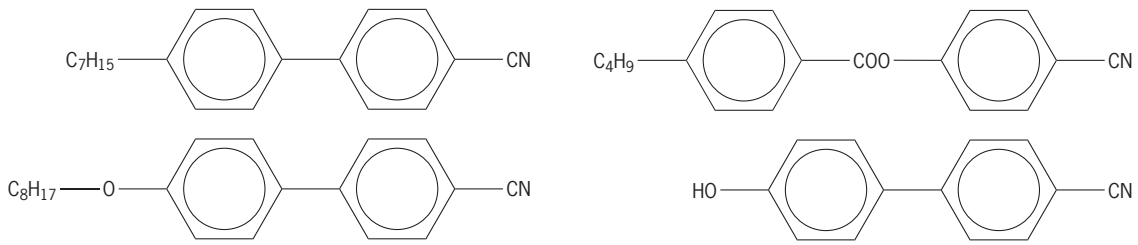


Figure 3.7.9 Examples of biphenyl nitrile compounds.

Twisted nematic liquid crystals

Twisted nematic liquid crystals have previously been described. These are the sorts of arrangements found in liquid-crystal displays. The liquid crystal used is generally a mixture of molecules. When a monochrome display is made, a thin glass panel is coated in metal oxide, then covered by a layer of polymer. The polymer is rubbed with a cloth in one direction, to produce microscopic grooves. A second glass panel prepared in the same way as the first is adhered to the first one, with spacers in between. The grooves in the two glass panels run at 90° to each other. The space between the panels is evacuated and the liquid-crystal mixture is injected. Polarizers are placed at either end, in alignment with the grooves. As previously described, the molecules are arranged in layers, each slightly twisted with respect to the layers above and below. This arrangement is where the term 'twisted nematic' comes from. Recall that 'nematic' means that the molecules are all aligned along a major axis.

AS C.11.2 Describe and explain in molecular terms the workings of a twisted nematic liquid crystal. © IBO 2007

As seen in figure 3.7.10, from one side of the very thin layer of liquid crystal to the other, the molecules twist by 90°. This allows the light to pass through the polarizers, since it is also twisted at 90°. The efficiency of the light rotation is nearly 100%. Light travels in the direction of the optical axis of the molecules, which is the long axis.

When an electric potential is passed across the two glass panels, the polar molecules respond. Figure 3.7.11 shows how the molecules change.

As can be seen, the twist is destroyed, so light passes through the liquid crystal, but is blocked by the polarized filter at the end. These devices are desirable for displays because they are light and non-bulky, unlike the cathode ray tube design.

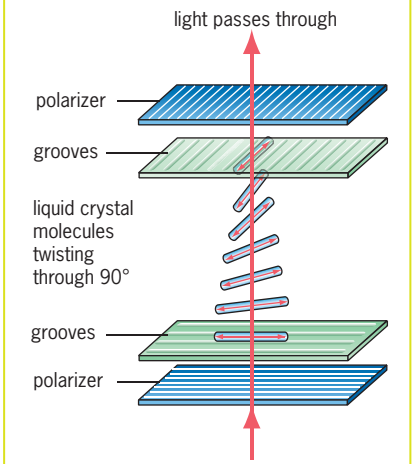


Figure 3.7.10 The arrangement of molecules in a twisted nematic liquid crystal.

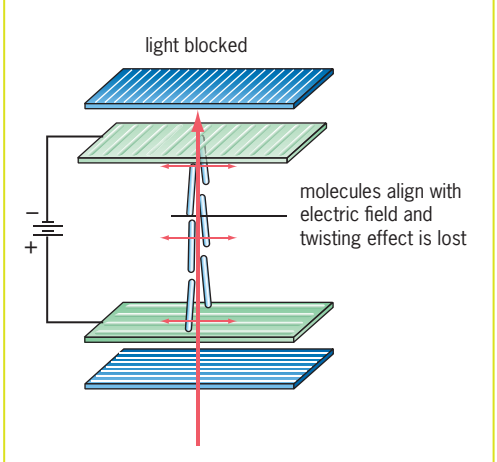


Figure 3.7.11 Response of nematic liquid-crystal molecules to current.

AS C.11.3
Describe the liquid-crystal properties of Kevlar™, and explain its strength and its solubility in concentrated sulfuric acid. © IBO 2007

Kevlar™

Kevlar™ was discovered by a young scientist at DuPont laboratories, Stephanie Kwolek, who was studying the polymer polyparaphenylene terephthalamide. The long, straight molecules packed closely together, resulting in strong intermolecular bonding. Later, this polymer was used to create the fibre Kevlar™. Until the discovery of carbon nanotubes, Kevlar™ was the toughest synthetic material known.

The repeating units in a Kevlar™ chain are phenyl groups linked by amide bonds. The strange thing about this polymer was that when it was dissolved in tetramethylurea and calcium chloride, the molecules lined up along a major axis. This was strange because molecules in solution are normally disordered. The order of the Kevlar™ molecules make it a liquid crystal. Kevlar™ was an exciting discovery because it was the first known polymer liquid crystal.

In Kevlar™, rotation is possible, but the molecules are basically linear. Amide bonds on adjacent chains can hydrogen bond to each other, so the intermolecular forces are quite strong. This is why the material is so tough. The structure retains some flexibility. In making consumer items such as bulletproof vests, the fibres are twisted when spun, then interwoven to create a layered net. If a person is shot, the energy is spread over all the layers. The victim feels the impact and possibly some pain, but the bullet won't pierce the tough Kevlar™ vest. The spun fibre has extraordinary tensile strength and low density.

During the spinning process, Kevlar™ must be kept in solution. Concentrated sulfuric acid achieves this aim. It prevents the molecules from hydrogen bonding to each other by protonating the amide links.

Sulfuric acid is a nasty reagent to use, but without it the Kevlar™ molecules would bond so strongly via hydrogen bonds that they would be impossible to spin.

Section 3.7 Exercises

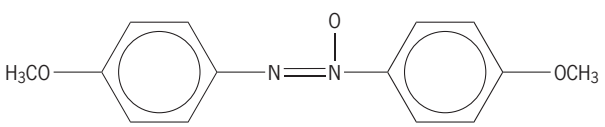
1 Define the term *doping*.

2 Describe how doping is used to create two different kinds of silicon semiconductors.

3 Distinguish between intrinsic and extrinsic semiconductors.

4 Explain the sequence of events that occurs when sunlight falls on a solar cell.

5 Consider the following molecule:



Would you expect this molecule to exhibit liquid-crystal properties? Explain your answer.

6 Describe the arrangement of the molecules in a twisted nematic liquid crystal.

7 Kevlar™ is five times stronger than steel. Describe the elements of its structure that give it this property.

3.8 THE CHLOR-ALKALI INDUSTRY

HL

The **chlor-alkali industry** produces two of the world's most important chemicals—chlorine and **caustic soda** (sodium hydroxide)—via the electrolysis of **brine**, a concentrated sodium chloride solution. There are 650 chlor-alkali plants worldwide, with a chlorine production capacity of more than 55 million tonnes per year; 1.1 tonnes of 50% caustic soda solution is produced for every tonne of chlorine. There are three types of electrolytic cells used to produce chlorine and caustic soda: the **diaphragm cell**, the **mercury cell** and the **membrane cell**.

AS **C.12.1 Discuss the production of chlorine and sodium hydroxide by the electrolysis of sodium chloride.** © IBO 2007

The diaphragm cell is the most-used cell in Russia, China and the US. Japan uses virtually 100% membrane cells. Europe has been slowly phasing in membrane cells over more than a decade, and 2007 was the first year in which more membrane cells than mercury cells were used in the region. The European chlor-alkali industry has made a voluntary commitment to phase out all mercury cells by 2020. Interestingly, while the harmful effects of mercury are beyond dispute, the chlor-alkali industry actually contributes less than 1% of all mercury emissions. Coal-fired power stations and the incineration of wastes release considerably more mercury to the environment.

Pr PRAC 3.5 Electrolysis of sodium chloride

When an electric current is passed through an aqueous sodium chloride solution, several products are possible.

At the anode(+), both water and chloride ions are capable of being oxidized:

$$2H_2O(l) \rightarrow O_2(g) + 4H^+(aq) + 4e^- \qquad E^\ominus = +1.23\ V$$
$$2Cl^-(aq) \rightarrow Cl_2(g) + 2e^- \qquad E^\ominus = +1.40\ V$$

Under standard conditions, these reduction potentials would imply that oxygen will be preferentially produced at the anode. In reality, these two processes will compete with each other. By using brine (around 30 per cent by mass, ≈ 5 mol dm^{-3}), and therefore non-standard conditions, the chloride reaction will be the dominant anode reaction.

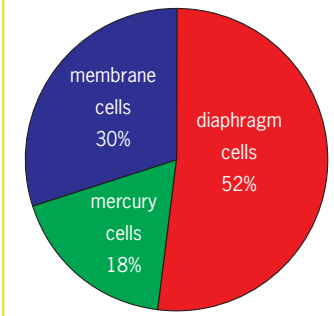


Figure 3.8.1 World use of the different chlor-alkali cells.

At the cathode(–), only water is capable of being reduced, because sodium ions cannot be reduced in the presence of water:

$$2H_2O(l) + 2e^- \rightarrow 2OH^-(aq) + H_2(g) \qquad E^\ominus = -0.83\ V$$

The hydrogen is collected and used as fuel for the co-generation of electricity.

The overall reaction is thus:

$$2H_2O(l) + 2Cl^-(aq) \rightarrow Cl_2(g) + 2OH^-(aq) + H_2(g)$$

CHEM COMPLEMENT

Minamata

Japan has good reason for choosing membrane cells over mercury cells. Between 1932 and 1968, the Chisso Corporation, which manufactured organic compounds, released around 27 tonnes of mercury compounds into Minamata Bay. Minamata is a small town a few hours southwest of Tokyo. The people of the town regularly ate fish from the bay. Many got sick and this outbreak of mercury poisoning became known as 'Minamata disease'. It first became obvious in the early 1950s, with cats acting strangely, and townsfolk becoming unable to walk, talk or think clearly. Decades later, some victims are still seeking compensation. Others are long dead. Because of the publicity this case received, the Japanese are naturally wary of industrial processes involving mercury.

The diaphragm cell

Chlor-alkali plants consist of many cells connected in series. Since the 1970s, activated titanium anodes have been used in diaphragm and mercury cells, replacing the old graphite electrodes. The cathode is made from steel. To separate the anode and cathode compartments, a porous diaphragm is used. This prevents the product gases mixing.

When forming the diaphragm, a fibrous polymer, polytetrafluoroethylene (PTFE or Teflon) is fused with asbestos. Asbestos is cheap and abundant. It is also unreactive. Unfortunately, it is a significant health hazard. Asbestos fibres can lodge in the lungs causing asbestosis, a degenerative lung disease. To increase safety, white asbestos is used, which is somewhat less toxic than blue asbestos.

In this cell, brine is added to the anode compartment. It then flows through the diaphragm to the cathode compartment, from which the sodium hydroxide product is extracted. This hydraulic gradient prevents backflow of hydroxide ions into the anode compartment.

The reactions that occur in this cell are:

At anode:

$$2Cl^-(aq) \rightarrow Cl_2(g) + 2e^-$$

At cathode:

$$2H_2O(l) + 2e^- \rightarrow 2OH^-(aq) + H_2(g)$$

If hydroxide ions contaminated the anode area, the following reactions could occur:

$$4OH^-(aq) \rightarrow O_2(g) + 2H_2O(l) + 4e^-$$

or

$$Cl_2(g) + OH^-(aq) \rightarrow Cl^-(aq) + HOCl(aq)$$

followed by

$$HOCl(aq) \rightarrow H^+(aq) + OCl^-(aq)$$
$$2HOCl(aq) + OCl^-(aq) \rightarrow ClO_3^-(aq) + 2Cl^-(aq) + 2H^+(aq)$$

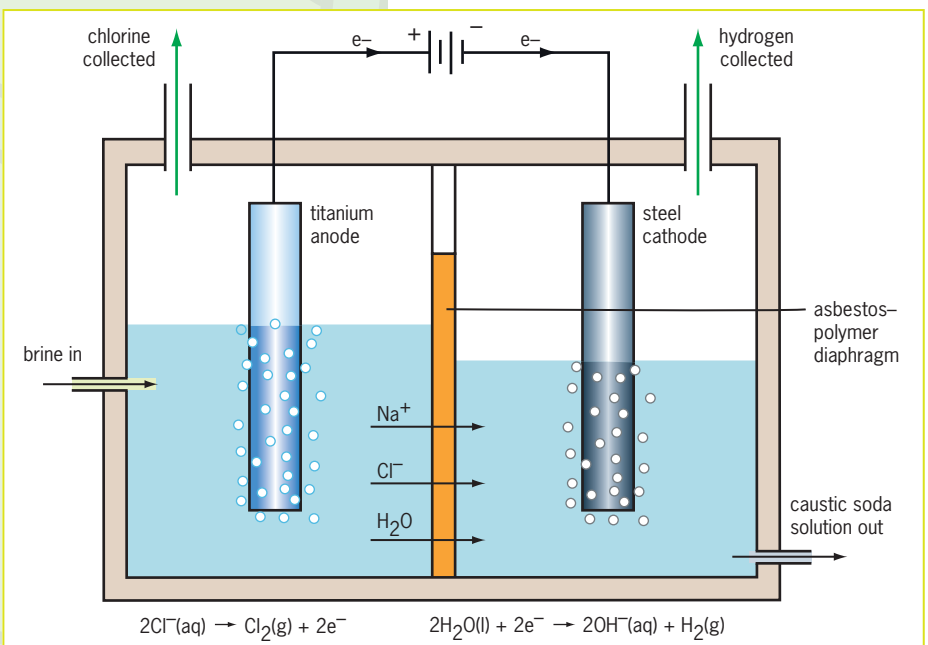


Figure 3.8.2 The diaphragm cell.

The above reactions show how the product can become contaminated with oxygen, or hypochlorite and chlorate ions. In the diaphragm cell, much of the brine remains unreacted, so the product extracted from the cathode compartment contains a high percentage of sodium chloride. The sodium hydroxide concentration is only about 12%, which falls far short of the 50% level needed to meet sales specifications.

A number of evaporators are used to purify the product. Sodium chloride is less soluble than sodium hydroxide (35.9 g 100 g^{-1} versus 114 g 100 g^{-1}). As the amount of water decreases, the sodium chloride crystallizes and is removed by filtration. Unfortunately, the final product still contains up to 1% NaCl, so it can't be used for manufacturing processes in which the presence of chloride ions will have a negative effect. The diaphragm cell consumes about 2900 kW h of electricity per tonne of chlorine produced, but another 700 kW h per tonne is needed to purify the caustic soda product. The chlorine gas is water-saturated when first formed. It is cooled and some moisture is removed. The remaining moisture is removed using concentrated sulfuric acid, a powerful dehydrating agent. The chlorine is then further purified, compressed and cooled. It is liquefied prior to sale.

The mercury cell

The mercury cell uses a titanium anode and a liquid mercury cathode, which flows along the bottom of the cell. The one big advantage of the mercury cell is that it produces a very pure caustic soda solution that is already at the required 50% concentration, so no further purification is needed.

The reactions occurring in the mercury cell are a little different from those of the other cells. Brine is added to the anode compartment, where chlorine is produced.

$$2Cl^-(aq) \rightarrow Cl_2(g) + 2e^-$$

Mercury flows over the bottom of the steel base of the anode compartment, acting as the cathode. At the surface of the mercury, sodium is produced in preference to hydrogen. The sodium immediately dissolves in the mercury, forming an amalgam:

$$Na^+(aq) + e^- \rightarrow Na(l)$$
$$Na(l) + Hg(l) \rightarrow Na–Hg(l)$$

The melting point of sodium is only 98°C, and the melting point of the amalgam is even lower. The sodium amalgam then flows to a secondary vessel containing catalytic carbon balls. In this secondary compartment, sodium in the amalgam reacts with water, forming hydrogen and a sodium hydroxide solution:

$$2Na(l) + 2H_2O(l) \rightarrow 2NaOH(aq) + H_2(g)$$

Because the sodium reacts with pure water in the secondary vessel, the final product has very high purity. The mercury can be recycled and returned to the electrolytic vessel. The spent brine has more salt added to it and is recycled. In this process, mercury, a potent neurotoxin, can be lost in wastewater, gas ventilation, and in the products. The electrolytic cell needs to be well-ventilated to prevent the build-up of mercury vapour. It is also possible that some of the sodium reacts with water in the first vessel, creating an explosive hazard.

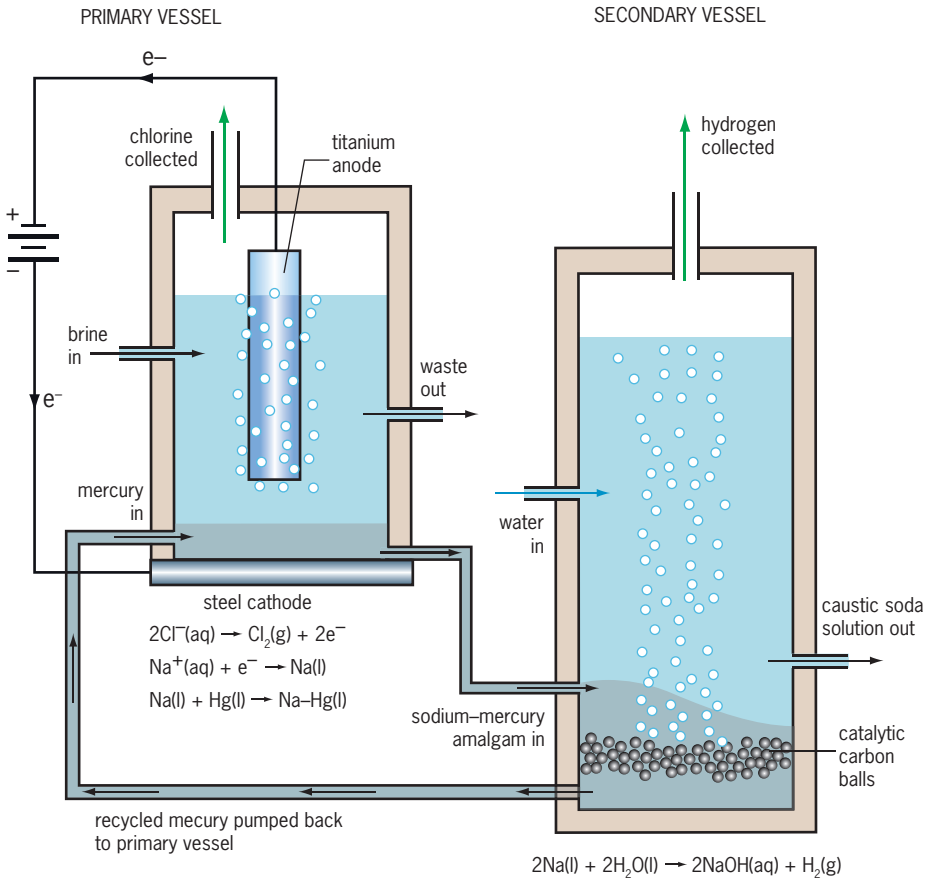


Figure 3.8.3 The mercury cell.

The membrane cell

Unlike the previously discussed cells, the membrane cell does not contain any extremely hazardous substances. Another advantage is the fact that it consumes less energy per tonne of caustic soda. The raw caustic soda product is of high purity, but it does require evaporation to get it to the right concentration for sale. The anode (titanium) and cathode (nickel) reactions are the same as for the diaphragm cell. Brine is added to the anode compartment continuously. The spent brine is recycled to have more salt added before it goes back into the anode compartment. Water flows into the cathode compartment and caustic soda solution flows out.

The reactions occuring in the membrane cell are:

At anode:

$$2Cl^-(aq) \rightarrow Cl_2(g) + 2e^-$$

At cathode:

$$2H_2O(l) + 2e^- \rightarrow 2OH^-(aq) + H_2(g)$$

The membrane prevents product gases from mixing. It is cation-selective, which means that water and positively charged ions can pass through, but not negatively charged ions: sodium ions can pass from the anode compartment to the cathode compartment, but hydroxide ions cannot travel the other way. This results in a highly pure product. Because the caustic soda product requires evaporation, another 250 kW h is needed per tonne of caustic soda.

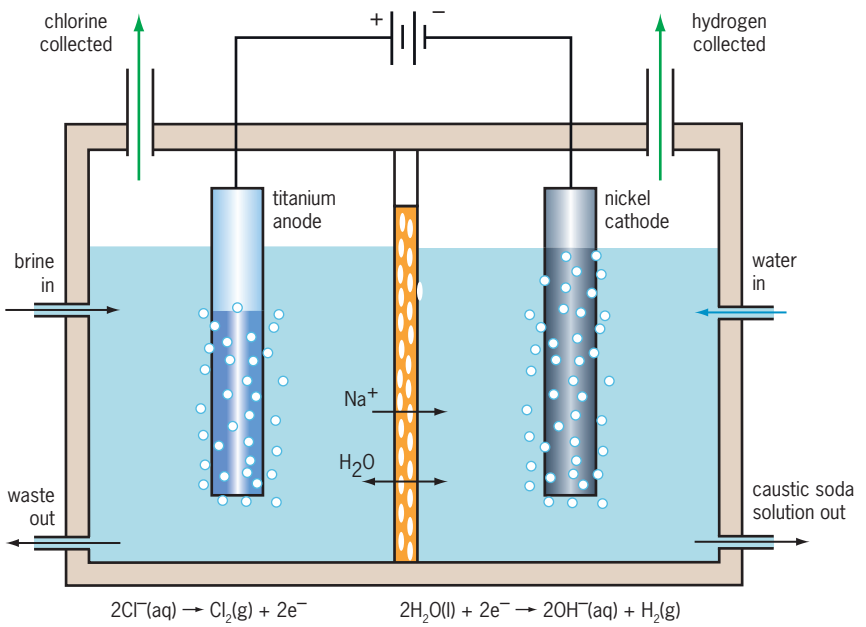


Figure 3.8.4 The membrane cell.

Interactive CW Animation
Semipermeable membrane

TABLE 3.8.1 PROCESS SUMMARY FOR THE THREE CHLOR-ALKALI CELLS

	Membrane cell	Mercury cell	Diaphragm cell
Cell operating temperature	88–90°C	Up to 95°C	Maximum = 85°C
Cathode	Nickel	Mercury	Steel
Anode	Titanium	Titanium	Titanium
Separator	Film of perfluorosulfonate polymer, Teflon reinforcing fabric and a perfluorocarboxylate polymer	Two different vessels used	Asbestos mixed with Teflon
Separator consumption	Negligible	No separator present	800 g per tonne of NaOH
Anode consumption	Negligible	Negligible	Negligible
Cathode consumption	Negligible	0.2 kg per tonne of caustic soda	Negligible
Average energy consumption per tonne of NaOH (kW h)	2500	3600	2700
Concentration of NaOH in raw product (% w/w)	Up to 30	50	10–12
Concentration of NaCl in raw product (% w/w)	0.001–0.002	0.002–0.003	Up to 15

AS C.12.2
Outline some important uses of the products of this process. © IBO 2007

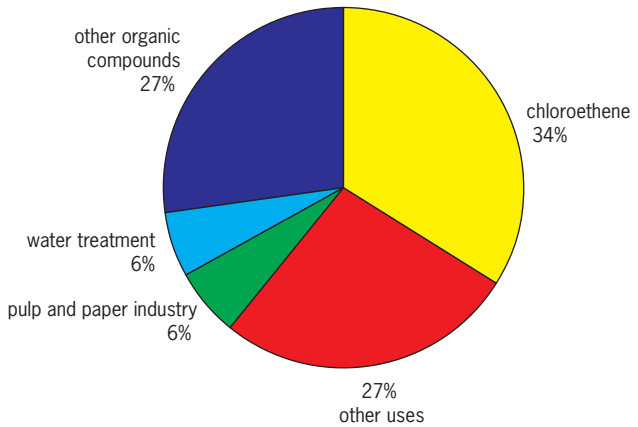


Figure 3.8.5 Chlorine uses.

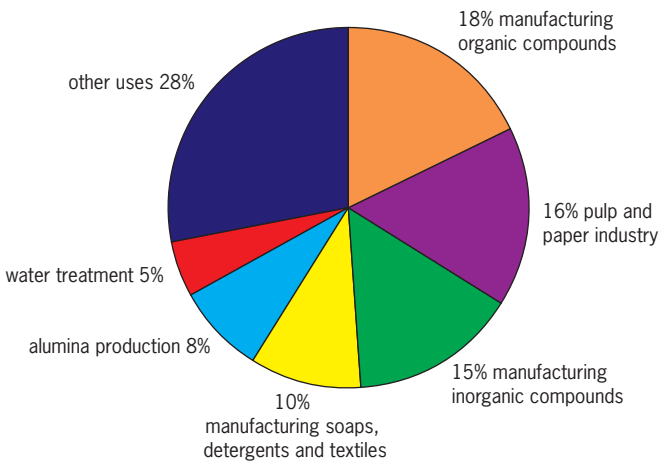


Figure 3.8.6 Caustic soda uses.

Ws
WORKSHEET 3.4
The chlor-alkali industry

AS C.12.3
Discuss the environmental impact of the processes used for the electrolysis of sodium chloride. © IBO 2007

Uses of chlorine

It is difficult to overstate chlorine's importance to industry. Chlorine or chlorine compounds are used in more than 50% of all industrial processes; they are contained in, or used in the manufacture of, 85% of pharmaceuticals and 96% of pesticides. By far the biggest use of chlorine is in creating PVM—polyvinyl monomer, or chloroethene, the monomer used to make polyvinylchloride (PVC). The demand for phosgene ($COCl_2$) is growing—it is needed to make polycarbonates, used for bulletproof glass and aircraft windows. In contrast, the pulp and paper industry is using less and less chlorine for paper whitening. It and many other industries are looking for ways to reduce their chlorine dependence due to the highly toxic nature of chlorine and many of its compounds. As well as their direct effect on humans, chlorine compounds such as chlorofluorocarbons (CFCs) and hydrochlorofluorocarbons (HCFCs) are potent greenhouse gases. In the stratosphere, chlorine radicals break off and catalyse the destruction of ozone.

Uses of caustic soda

Like chlorine, sodium hydroxide has many industrial uses. It is used in caustic scrubbers to remove acidic gases, in the manufacture of soap, in the textile industry, and as an intermediate in the production of pharmaceuticals such as aspirin. These are just a few of its many uses. As well as in chemical production, both organic and inorganic, it is used in the preparation of alumina prior to electrolysis. It is also used in petroleum and natural gas refining. Figure 3.8.6 outlines caustic soda's major applications.

Environmental impact of the chlor-alkali industry

The environmental impact of all the end products made from caustic soda and chlorine are considerable, but we will just look at some of the impacts of the electrolytic processes. In these processes, brine is the feedstock. While that is not harmful, various agents must be used to cool and liquefy the products, and remove impurities. Hydrofluorocarbons and hydrochlorofluorocarbons are two classes of compounds that might be used. These are powerful greenhouse gases, but have limited ozone-depleting potential because they tend to break down in the troposphere.

Major chlor-alkali pollutants are the loss of Cl_2 to the air, cooling agents, impurities removed from products, and spent acids. The loss of neurotoxic mercury from mercury cells is a cause of concern. Pressure from environmental lobbies is leading to a total phase out of their use in the developed world. Exposure of diaphragm cell workers to asbestos is also an issue, as is the loss of asbestos to the environment. These concerns have led to membrane cell technology becoming the preferred process method.

The fact that the membrane process is also more energy efficient makes it an attractive option. The massive electricity input needed to produce chlorine and caustic soda is another environmental issue, especially if the electricity is being made from fossil fuels, as this contributes to global warming. To negate this effect and reduce the electricity costs, plants use the hydrogen formed to produce electricity to run the plant. Waste heat can also be used to generate electricity.

One problem of mercury cell plants being shut down has been what to do with the mercury. No waste plants can deal with this amount of pure mercury. For now, most has been placed into hermetically sealed vessels and dumped in deep underground salt mines.

Section 3.8 Exercises

1 List at least three uses of:
- **a** caustic soda
- **b** chlorine.

2 Explain the advantages of:
- **a** the membrane cell over the diaphragm cell
- **b** the mercury cell over the membrane cell
- **c** the diaphragm cell over the mercury cell.

3 Write the anode and cathode reactions occurring in the:
- **a** mercury cell
- **b** diaphragm cell
- **c** membrane cell.

4 Compare the three chlor-alkali processes in terms of environmental impact.

Chapter review questions and tests are available on student CD.

Chapter 3 Summary

Terms and definitions

Acid rain Rain with a pH less than 5.

Active sites Places on a catalyst's surface where reactant molecules adsorb.

Addition polymer Polymer made from alkenes and alkene derivatives.

Annealing Softening heat treatment method.

Aromatics Molecules containing at least one benzene ring.

Atactic polypropene Isomer of polypropene with randomly orientated methyl groups.

Band theory Theory of semiconductors, which states that electrons move between a conduction band and a valence band.

Bauxite Principal aluminium ore.

Blast furnace Furnace used to extract iron from its ores.

Brine Concentrated sodium chloride solution.

Catalyst poisoning Deactivation of a catalyst by reaction with contaminants.

Catalytic cracking Cracking reaction using catalysts such as zeolites.

Caustic soda Common name for sodium hydroxide.

Chlor-alkali industry Industry in which caustic soda and chlorine are made.

Cholesteric liquid crystal Type used in thermometers. Molecules are aligned along a main axis in layers, and molecule orientation of each successive layer is slightly different. The twists give rise to colour.

Coke Coal that has been heated to drive off volatile components.

Condensation reaction A reaction in which molecules bond together by the reaction of functional groups, with the elimination of one small molecule (usually water) for each new bond formed.

Copolymer Polymer formed from two different monomers.

Cracking reaction Reaction in which larger hydrocarbons are broken down into smaller, less saturated hydrocarbons.

Cryolite (Na_3AlF_6) Substance added to alumina to lower the melting point of the electrolyte during aluminium production.

Desorb To detach from a surface.

Diaphragm cell Chlor-alkali cell in which an asbestos divider is used between the compartments.

Donor atom Doping element that provides conduction electrons.

Doping The addition of small quantities of another substance to a semiconductor.

Electrode Conductive material; site of oxidation or reduction.

Electrolyte Substance that conducts electricity if molten or dissolved.

Extrinsic semiconductor Semiconductor with properties largely determined by the doping agent.

Fossil fuels Coal, oil and natural gas.

Fractional distillation The separation of a mixture by heating and vaporization and subsequent cooling and condensation.

Free radical Reactive substance with an unpaired electron.

Fuel cell Voltaic cell with at least one gaseous reactant supplied continuously.

Fullerenes Allotrope of carbon.

Haematite Iron ore; Fe_2O_3.

Hall–Héroult cell (pot) Cell used to extract aluminium from alumina.

Hardenability Measure of how well a material responds to heat treatment.

Heat treatment Treatments using heat to soften or harden a piece of metal without melting it.

Homolytic cleavage The splitting of a molecule into two identical parts.

Industry The trade and manufacture of goods.

Intrinsic semiconductors Pure crystals.

Isotactic polypropene Isomer of polypropene with methyl groups arranged on the same side of the asymmetrical carbon.

Kevlar™ Liquid-crystal substance that can be woven into fibres; very tough.

Landfill Disposal site for solid waste.

Lead–acid battery Rechargeable cell used in cars.

Liquid crystal Substance that flows like a liquid, but has some crystalline order like a solid.

Lithium ion cell Rechargeable cell containing lithium ions.

Lyotropic liquid crystal Liquid crystal that only exists in solution.

Magnetite One of the major iron ores; Fe_3O_4.

Membrane cell Chlor-alkali cell in which a cation-selective membrane is used.

Mercury cell Chlor-alkali cell in which mercury is used as an electrode.

Micelle Cluster of surfactant molecules.

Monomer Molecule from which polymers are made.

Nanotechnology Field of technology dealing with devices ranging in size from 1 nm to 100 nm in size.

Nanotubes Allotrope of carbon with a closed cylindrical structure.

Nematic liquid crystal Liquid crystal that has order in only one direction.

Nickel–cadmium cell Rechargeable cell used in laptops and mobile phones.

NO_x Pollutant oxides of nitrogen.

n-type semiconductor Semiconductor with group 5 doping agent (i.e. donor atoms); n refers to the (negative) charge of the extra electrons.

Particulates Tiny, solid pollutant particles suspended in air.

Perfluorocarbons Organic compounds in which all hydrogens have been replaced by fluorine; potent greenhouse gases.

Petrochemical Chemical derived from petroleum.

Petrochemical feedstock Liquid or gaseous hydrocarbons used to make intermediate or primary chemicals.

Petroleum Another name for crude oil.

Phenol–methanal plastics These are the oldest synthetic polymers. They are unusual in that the monomers each contain only one functional group.

Photochemical smog Brownish mixture of chemicals in which sunlight-initiated reactions occur.

Pig iron Impure iron produced in the blast furnace; contains about 4% carbon.

Plasticizers Chemicals added to make polymers softer and more flexible.

Pot Another name for a Hall–Héroult cell.

Primary cell A non-rechargeable battery.

p-type semiconductor Semiconductor with group 3 doping agent (i.e. doped with acceptor atoms).

Quantum effects Effects that are only noticeable on the atomic scale, not in the macro world.

Quenching Fast cooling of a heated metal by plunging it into a liquid.

Refractory material Substance able to withstand extremely high temperatures.

Scanning tunnelling microscope Specialized microscope used to visualize and manipulate atoms.

Scrubbing Process of passing flue gases over chemicals to purify them.

Secondary cell A rechargeable battery.

Siderite Iron ore; $FeCO_3$.

Slag The silica-rich waste from the smelting process.

Smectic liquid crystal State in which the molecules are arranged along a major axis and in layers.

Smelting Extraction of metals using heat, resulting in two layers.

Steam cracking Type of cracking which uses steam to split the molecules.

Syndiotactic polypropene Isomer of polypropene with methyl groups alternating on the either side of the asymmetrical carbon.

Technology The practical application of science.

Tempering Softening type of metal heat treatment.

Thermal cracking Type of cracking reaction in which heat is used to split molecules.

Thermotropic liquid crystal Substance that shows liquid-crystal properties in a pure state.

Tuyère Small holes at the bottom of the blast furnace through which blasts of hot air are injected.

Twisted nematic liquid crystal Type of liquid crystal with layers twisted by regular increments; used in liquid-crystal displays.

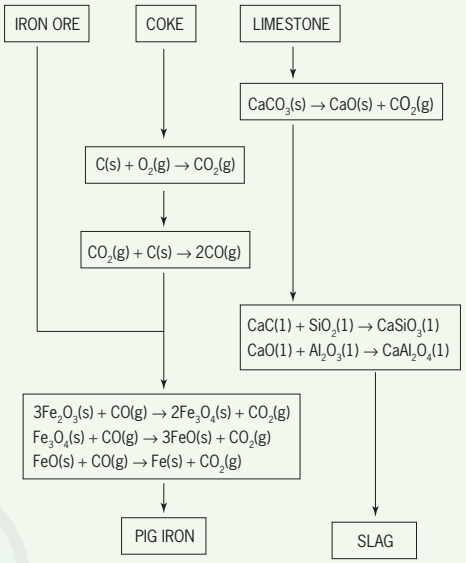
VOCs Volatile organic compounds.

Zeolites Porous type of catalyst used in cracking reactions.

Ziegler–Natta catalysts Organometallic complexes used to make HDPE.

Concepts

- Haematite (Fe_2O_3) and magnetite (Fe_3O_4) are the major iron ores.
- In a blast furnace, a series of reactions occur, summarized by the flow chart below.



- Crude iron is converted to steel via the basic oxygen process.
- Alloys are mixtures with metallic properties; at least one component must be a metal.
- Most alloys are homogeneous.
- Alloying can make a metal harder, stronger and more corrosion resistant.
- Heat treatments such as annealing, tempering and quenching can be used to soften or harden alloys such as steel.
- Steel is stronger, harder and more versatile than iron, which is soft and corrodes easily.
- Aluminium is extracted from alumina using a Hall–Héroult cell. Cryolite is used in this process to lower the melting point of the electrolyte. The overall reaction is:
  $3C(s) + 2Al^{3+}(l) + 6O^{2-}(l) \rightarrow 3CO_2(g) + 2Al(l)$
- Aluminium alloys are light and strong and are used extensively in the transport industry.
- Both the aluminium and iron industries have significant environmental impact, particularly in the consumption of energy.
- 90% of crude oil is used to make fuels.
- Chemicals derived from petroleum are called petrochemicals.
- Large hydrocarbon molecules undergo steam, thermal or catalytic cracking reactions to produce smaller, more useful hydrocarbon molecules in reactions such as $C_3H_8(g) \rightleftharpoons C_2H_4(g) + CH_4(g)$
- Addition polymers are formed from monomers containing a C=C double bond.
- Factors such as molecule size, degree of cross-linking and extent of branching affect the physical properties of polymers.
- There are many ways to modify the properties of polymers, such as using additives like plasticizers, which soften the material.
- Polymers offer considerable advantages over natural materials such as metal and wood.
- Catalysts may be homogeneous (in the same state as the reactants) or heterogeneous (in a different state to the reactants).
- Heterogeneous catalysts have active sites that allow reactant molecules to adsorb to the surface.
- When choosing a catalyst, factors such as selectivity, durability, availability and environmental impact must be considered.
- Fuel cells convert chemical energy to electrical energy while at least one gaseous reactant is supplied continuously.
- Primary cells are not rechargeable; secondary cells are.
- Secondary cells include the lead–acid, nickel–cadmium and lithium ion cells.
- Liquid crystals flow like viscous liquids, but retain some crystalline order.
- Thermotropic liquid crystals are pure substances, while lyotropic liquid crystals only exist in solution.
- Nematic liquid crystals have order along one axis only.
- Liquid crystals are used in LCDs.
- Nanotechnology refers to technological developments in the 1–100 nm range.
- Carbon nanotubes are a synthetic allotrope of carbon. Bundles of nanotubes are even stronger than Kevlar™. They have many potential applications.
- The health effects of exposure to nano-particles are not yet known.
- Condensation polymers form due to the reactions between functional groups on the monomers.
- LDPE is formed using organo-peroxide compounds as catalysts.
- HDPE is formed using Ziegler–Natta catalysts.
- Silicon is 'doped' with group 3 (to produce p-type) or group 5 (to produce n-type) elements to increase its conductivity.
- Twisted nematic liquid crystals are used in LCDs. Their orientation changes when an electric field is applied.
- Kevlar™ has liquid-crystal properties. The fibres are spun from solution in concentrated sulfuric acid.
- The chlor-alkali industry produces chlorine and caustic soda from brine by mercury, membrane and diaphragm electrolytic cells.
- The membrane cell is the cell of choice because it uses the least energy and has no toxic substances in it.

4 MEDICINES AND DRUGS

Chapter overview

This chapter covers the IB Chemistry syllabus Option D: Medicines and Drugs.

By the end of this chapter, you should be able to:

- describe general aspects of medicines and drugs such as their effects, methods of administration, terminology and the stages involved in research and development of new drugs
- discuss the use of antacids for treating excess acidity in the stomach
- compare the use, effects and structures of strong and mild analgesics
- compare the use, effects and structures of depressants, including ethanol
- describe the techniques used for the detection of ethanol in breath, blood and urine
- compare the use, effects and structures of stimulants, including nicotine and caffeine
- discuss the history, structure, method of function and effect of overprescription of penicillins as an example of antibacterial medicines
- state how viruses differ from bacteria, and describe the different ways in which antiviral drugs work
- discuss the difficulties related to the spread of AIDS and the solution to this worldwide problem

- discuss and explain the importance of structural features such as chirality and geometric isomerism in the action of drugs
- describe aspects of drug design such as compound libraries, combinatorial chemistry, computer modelling, structural modification and chiral auxiliaries
- describe the effects and structural similarities and differences of the mind-altering drugs LSD, mescaline, psilocybin and THC
- discuss the legalization of cannabis.

HL

IBO Assessment statements D.1.1 to D.10.3

By definition, a drug is a substance that affects how the body works. This may be for better or for worse. A medicine has the function of improving health. Drugs and medicines can be seen as responsible for renewed health, but also for debilitating addiction and even death in some cases. We have come to rely on medicines to help us overcome diseases ranging from the slightest ailments to deadly diseases such as cancer. Vaccines prevent us from contracting diseases that were deadly to earlier generations, and improvements in our overall health and wellbeing are constantly being made by progress in this area.

Figure 4.0.1 Paintings in the Lascaux caves were done by Cro-Magnon man, an early European culture of modern humans.

The use of plants as healing agents has been traced back through many thousands of years. Cave paintings found in Lascaux, France, depict the use of herbal cures and have been dated by radiocarbon dating to as early as 19000 BC. In the Middle Ages, the tradition of the apothecary arose in England. An apothecary was like a modern-day pharmacist and doctor all rolled into one. The apothecary prepared and sold drugs and other compounds for medicinal purposes and diagnosed illness, as well as prescribing remedies. While it is estimated that 80 per cent of the world's population still relies upon the use of herbal medicine for health care, the manufacture of synthetic drugs is a huge industry, and computers are used to help determine the structure of possible medicines and to speed us in our search for the cure for elusive viruses such as HIV.

4.1 PHARMACEUTICAL PRODUCTS

AS **D1.1**
List the effects of medicines and drugs on the functioning of the body. © IBO 2007

A pharmaceutical can be defined as a substance that is used in the treatment of a disease. The word comes from the Greek *pharmakeutes* meaning a preparer of drugs, or a witch!

Today we understand pharmacy to refer to the preparing and dispensing of medicines according to the prescription supplied by a doctor. The pharmacist is a scientist with a detailed knowledge of medicines, their effects and their side-effects, as well as a keen eye for accuracy— most important when medicines are being dispensed.

Generally a drug or medicine is any chemical that does one or more of the following:

- alters incoming sensory sensations experienced by the user
- alters the mood or emotions of the user
- alters the physiological state, including consciousness, activity levels or physical coordination of the user.

In the testing of drugs to determine their effectiveness, a **placebo** is often used. Tests are carried out in which some patients receive the actual drug in the required dosage, while others receive a placebo. This is made of a pharmacologically inert material, and has no chemical reason for helping the ailment; yet it often produces some benefits in the patient. It appears that a psychosomatic response occurs to the substance given. This '**placebo effect**' is explained in terms of the brain's ability to control the body's natural healing processes—if the patient believes he or she is going to get well, then they sometimes will! In this situation the body's natural healing processes appear to take over; however, not all conditions appear to respond to a placebo, and the placebo effect often does not give lasting relief, but rather a temporary improvement in symptoms brought about by positive thinking.

Developing new pharmaceutical products

AS **D.1.2**
Outline the stages involved in the research, development and testing of new pharmaceutical products. © IBO 2007

The research, development and testing of new pharmaceutical products is a costly and time-consuming process. It is rigidly controlled by governments. There are a number of time-consuming stages that must be passed through before a new pharmaceutical can be released to the public. All of these are important, since the maintenance of public health is an important responsibility of any government.

Research into observed phenomena such as the medicinal effects of a particular plant will initially lead to the isolation of a pharmaceutical. This pharmaceutical is tested in the laboratory on cell cultures first and then on animals. There is much debate on animal testing, and legislation involving the use of animals in this way varies from country to country. Such testing is initially quite broad and is gradually refined to determine effective doses and to investigate side-effects. At this stage the occurrence of detrimental side-effects may cause the investigation to cease.

If animal testing proves safe then clinical trials on volunteer humans can begin. A clinical trial is an organized test of medicines and of new treatment options involving patient and non-patient human volunteers. Clinical trials confirm whether medicines are safe and effective to be used as new treatments for a particular disease or condition. It is important that the volunteers understand this process. They may not be actually given the pharmaceutical that is being tested. Placebos are used as part of the testing process, so people who are desperate for a cure may not make any progress, while others in the group are benefiting from the new pharmaceutical. At this stage, previously unforseen side-effects may be detected and will need to be assessed to determine whether they are sufficiently harmful to prevent the pharmaceutical from being tested further. It is interesting to note that clinical testing of new pharmaceuticals on children is a difficult issue. Often the pharmaceutical is going to be used for children; however, the testing stage poses an ethical dilemma. Is it right to place a child at risk by giving the child a new pharmaceutical, yet how can correct doses be determined if children are not tested?

Government organizations, such as the Food and Drugs Administration (FDA) in the USA and the Therapeutic Goods Administration in Australia, have the final say as to whether a pharmaceutical may finally be approved for marketing. Very few pharmaceuticals make it to the market place. Global pharmaceutical research and development (R&D) expenditure reached US$45 billion in 2002, reflecting the intense pressure within the industry for new medicines.

The average cost of discovering and developing a new medicine is more than US$1 billion. The average development time for new medicines is 12–15 years. On average, only three out of ten medicines will recoup the costs of research and development, clinical trials, registration and marketing. However, the time in which costs are recouped for 'blockbuster' drugs can be much shorter. A drug is classified as a 'blockbuster' if it generates more than $1 billion of revenue for its owner each year. One such drug is Lipitor, a cholesterol-lowering medication marketed by Pfizer with sales of US$12.2 billion per year. With such high costs involved, patent and intellectual property protection has become a major issue for the pharmaceutical industry.

Pharmaceutical companies are constantly working to bring new effective medicines to patients. Globally there are more than 1000 new medicines in clinical trials for Alzheimer's disease, stroke, cystic fibrosis, arthritis and many other diseases.

AS **D.1.3 Describe the different methods of administering drugs. © IBO 2007**

Drugs may be administered by a variety of different methods in order to achieve different effects on the body, or to meet the needs of patients.

1 *Oral*—The drug is taken by mouth. It may be as a tablet, a capsule, a liquid, or a powder that is dissolved in water and then taken. The advantage of this method is that it is very convenient. Liquids may be given to small children who cannot yet swallow a pill, while pills and capsules have the advantage of being portable. The disadvantage of taking a drug by mouth is that the rate of absorption in the stomach is low, with most absorption occurring in the small intestine (like food). This means that the rate of absorption will be fairly low in comparison to other methods.

2 *Rectal*—Some drugs, particularly anticonvulsive, or antipyretic (fever reducing) medication can be effectively given to a patient by insertion of a suppository into the rectum. This method is very effective in case of a patient who is experiencing nausea or vomiting or convulsions. It may also be used after surgery in which the patient's ability to swallow has been affected. A fever in a small child may result in febrile convulsions, so the reduction of the fever needs to be achieved quickly. Many small children will not swallow liquids or tablets, so antipyretic medication, such as paracetamol, can be administered rectally. Absorption of the drug is through the large intestine (rectum) directly into the blood stream, so the effect of the drug is seen much more quickly than through oral administration.

3 *Inhalation*—Asthma is a condition that is becoming more common across the world. Often medication for respiratory conditions such as asthma is administered by inhalation. A 'puffer' is placed in the mouth of the patient and a measured dose of the medicine is inhaled. This method allows rapid absorption since the medicine passes directly from the lungs into the network of blood vessels that meets the lungs. This method of medication may also be used when the effect of the medicine is required on the brain as well as the rest of the body, e.g. in general anaesthesia.

4 *By injection*—An injection may be used to deliver medication in three possible ways. An **intravenous** injection occurs when the needle is inserted directly into a vein. The advantage of this method is that the medicine takes

effect very quickly and precise amounts can be administered. Intravenous injections are difficult to administer and training is needed to be able to give them accurately. **Intramuscular** injections are given directly into a muscle. Many vaccinations are delivered in this way. Anaphylactic responses to allergies are treated with a dose of epinephrine (adrenaline) using an EpiPen and delivered as an intramuscular injection. This method is relatively safe, allowing a large volume of drug to be administered. **Subcutaneous** injections are delivered to the region directly under the skin. Dental injections are typically given subcutaneously. The effects of a subcutaneous injection are not experienced as quickly as intravenous injections, but more quickly than intramuscular.

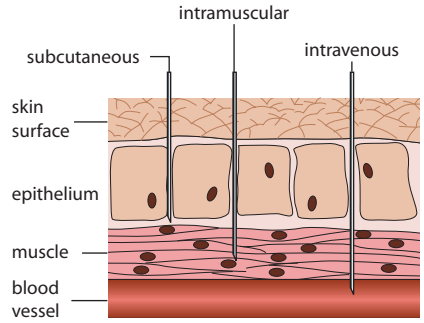


Figure 4.1.1 The depth to which a needle penetrates in the delivery of subcutaneous, intramuscular and intravenous injections.

The dose of a medicine is often crucial. Many medicines are effective in achieving the purpose for which they are intended at their prescribed dosage, but in much greater quantities could kill the user. The term **effective dosage** (**ED_{50}**) describes the amount of a pharmaceutical that would produce a therapeutic response in 50% of the population (meaning that in the other 50%, it would either be too much or not enough), while **lethal dosage** (**LD_{50}**) is used to denote the size of the dose of a pharmaceutical that would be enough to kill 50% of the population if that amount were taken. The ratio of LD_{50} to ED_{50} is known as the **therapeutic window**. This is a measure of the safety range of the drug. When LD_{50}:ED_{50} is large, there is a wide therapeutic window, and toxicity occurs well above the dosage required to achieve the desired efffect. When LD_{50}:ED_{50} is small, the therapeutic window is narrow and dosages must be measured carefully to avoid overdosing and nearing toxic levels.

AS **D.1.4**
Discuss the terms *therapeutic window, tolerance* and *side-effects*. © IBO 2007

In the case of some drugs, the body becomes increasingly able to absorb the drug without any effect being experienced. This is known as **tolerance**. The user requires increasing amounts of the drug in order to experience the effect originally obtained by a smaller dose. This is thought to occur because receptors affected by that drug register that a higher level than usual is present and through regular use of the drug, the brain is able to adjust by turning off some of the receptors for that drug. The system may also change the level of chemicals that it produces in order to compensate for the increased levels of whatever drug is being taken. For example, if a chemical enzyme produced in the body breaks down the drug, the system may increase production of that enzyme, resulting in more of the drug being broken down more quickly. Consequently, the effect of the drug is not experienced (either through lack of receptors or increased enzymes) and more is needed to create the same therapeutic effect. Tolerance results in a potentially dangerous situation as the amount of drug consumed increases and approaches the lethal dosage (LD_{50}).

Unfortunately, many drugs produce more than one effect when they are absorbed into the body. The main effect is the one for which the drug is intended, however, often other effects, **side-effects**, are also experienced. These occur in a lower percentage of cases than the main effect, but may be distressing to the user. For example, morphine is often used as a pain-killer, with intestinal constipation being a side-effect. However, for a person with diarrhoea, the constipation induced becomes the main effect, and the pain relief becomes a side-effect. The relative importance of the two effects becomes the main issue here. The risk to benefit ratio should be also be considered. Is the risk of the side-effects greater than the benefit obtained by taking the particular medicine?

Section 4.1 Exercises

1 Compare the terms *drug* and *medicine*.

2 Explain what is meant by the term *placebo effect*.

3 Describe a disadvantage of oral administration of some medicines.

4 Describe a method of dosage that might be used to provide medicine to a person who was experiencing nausea and vomiting.

5 Compare the three different ways in which drugs can be injected into the body, giving particular attention to the advantages of each method.

6 Compare the following terms: *effective dosage*, ED_{50}, and *lethal dosage*, LD_{50}.

7 Explain what is meant by the term *therapeutic window*.

8 Explain what is meant by the term *side-effect*.

9 In the case of some drugs, the body becomes increasingly able to absorb the drug without any effect being experienced.

a State the term used to describe this phenomenon.

b Explain how the phenomenon you named in part **a** could increase the risk of the drug to the user.

4.2 ANTACIDS

AS **D.2.1**
State and explain how excess acidity in the stomach can be reduced by the use of different bases. © IBO 2007

Heartburn is a painful or burning sensation in the oesophagus, just below the breastbone caused by regurgitation of gastric acid (hydrochloric acid with a pH of 1–2). The pain often rises in the chest and may radiate to the neck, throat, or angle of the jaw. This is due to excess acidity in the stomach, which may be caused by overeating or eating spicy foods. The way in which these factors affect the stomach varies from person to person. Peptic ulcers, which may occur in the duodenum (small intestine) or the stomach, had long been thought to owe their painfulness to the action of stomach acid on nerves exposed by the ulcer. They have now been shown to have other causes. (See Theory of Knowledge, p. 214.)

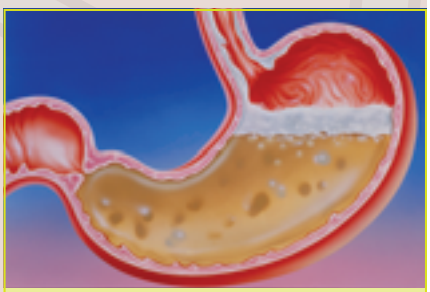


Figure 4.2.1 A foaming antacid forms a raft on top of the contents of a stomach.

These conditions are relieved by **antacids**. These are bases that neutralize the excess acidity and relieve the pain associated with heartburn and peptic ulcers. Antacids are taken orally as tablets or dissolved in water. They are usually an oxide or hydroxide of aluminium, magnesium or calcium, although sodium hydrogencarbonate can also act as an antacid. There are some possible side-effects from taking antacids, such as constipation, increased thirst and decreased appetite, but these are only likely when large quantities of antacids are consumed.

Antacids are often combined with **alginates**, polymers that act as low-temperature gelling agents and are derived from brown seaweeds. The purpose of alginates in an antacid mixture is to produce a neutralizing layer that will prevent acid in the stomach from rising into the oesophagus and causing heartburn. Since one of the products of this neutralization reaction is carbon dioxide, they are also combined with **anti-foaming agents** such as dimethicone. These anti-foaming agents reduce the surface tension of gas

bubbles, causing them to coalesce (come together). As a result the patient produces a 'burp' rather than vomiting, which could occur when so many gas bubbles are present. Dimethicone is often given to babies who gulp in air during feeding to produce a burp rather than a vomit!

The bases present in antacids neutralize the stomach acid (HCl) and in most cases produce a soluble salt and water. In some cases carbon dioxide is also produced.

PRAC 4.1
Investigating antacids

TABLE 4.2.1 CHEMICAL EQUATIONS REPRESENTING THE ACTION OF ANTACIDS

Antacid ingredient	Equation for reaction between antacid and HCl
Magnesium oxide	$MgO(s) + 2HCl(aq) \rightarrow MgCl_2(aq) + H_2O(l)$
Magnesium hydroxide	$Mg(OH)_2(aq) + 2HCl(aq) \rightarrow MgCl_2(aq) + 2H_2O(l)$
Aluminium hydroxide	$Al(OH)_3(s) + 3HCl(aq) \rightarrow AlCl_3(aq) + 3H_2O(l)$
Calcium carbonate	$CaCO_3(s) + 2HCl(aq) \rightarrow CaCl_2(aq) + H_2O(l) + CO_2(g)$
Sodium hydrogencarbonate	$NaHCO_3(aq) + HCl(aq) \rightarrow NaCl(aq) + H_2O(l) + CO_2(g)$

Notice in table 4.2.1 that the molar ratios of antacid to acid vary from 1 : 1 in the case of sodium hydrogencarbonate to 1 : 3 for aluminium hydroxide. We can see that an antacid containing aluminium hydroxide would be more efficient (per mol of $Al(OH)_3$) in neutralizing the acid than one containing sodium hydrogencarbonate; however, a calculation is required to confirm this efficiency per gram of antacid.

Worked example 1

Write the equations for the reaction of aluminium hydroxide and magnesium hydroxide with hydrochloric acid and hence identify which antacid neutralizes the greater amount of hydrochloric acid if 10.0 g of each antacid is used.

Solution

$Al(OH)_3(s) + 3HCl(aq) \rightarrow AlCl_3(aq) + 3H_2O(l)$

$Mg(OH)_2(aq) + 2HCl(aq) \rightarrow MgCl_2(aq) + 2H_2O(l)$

$$n(Al(OH)_3) = \frac{10.0}{78.01} = 0.128 \text{ mol}$$

$$n(Mg(OH)_2) = \frac{10.0}{58.33} = 0.171 \text{ mol}$$

Molar ratios:	$Al(OH)_3$	:	HCl		$Mg(OH)_2$	:	HCl
	1	:	3		1	:	2
	0.128	:	0.384		0.171	:	0.342

10.0 g of aluminium hydroxide will neutralize 0.384 mol of hydrochloric acid, whereas 10.0 g of magnesium hydroxide will only neutralize 0.342 mol of hydrochloric acid. Aluminium hydroxide can be considered to be more effective in neutralizing stomach acid.

THEORY OF KNOWLEDGE

The two tenacious recipients of the 2005 Nobel Prize for Physiology or Medicine, Australians Robin Warren and Barry Marshall, challenged a long-accepted belief that the cause of gastric (stomach) and duodenum (small intestine) ulcers and inflammation was excess acid created by stress, anxiety or eating too much spicy food.

In the late 1970s Warren, in the course of his routine work as a pathologist, observed something interesting: small unknown bacteria colonizing the stomach in patients suffering from ulcers. Curious to find out more, Warren and his colleague Marshall set out to cultivate these *Helicobacter pylori* to see if they were the cause. It took 10 years of painstaking work before they were able to make an irrefutable case that these bacteria could cause excess acidity, and antibiotics not antacids cure most cases. The medical community, however, was slow to accept their findings despite the weight of evidence. Their colleagues were sceptical and tried to convince them to not pursue what seemed to be pointless research. After all, it was widely accepted that bacteria could not grow in the low pH environment of the stomach. Furthermore, Marshall and Warren found it difficult to get their research published in a reputable medical journal until scientists in the UK were able to duplicate their results. In 2000, more than 20 years after their initial observations, their findings were finally accepted by the medical community and pharmaceutical industry. It is now firmly established that *H. pylori* causes more than 90% of duodenal ulcers and up to 80% of gastric ulcers. Marshall and Warren not only created new knowledge, but in the process increased our understanding of the connection between infection, inflammation, cancer and other diseases.

- Why do you think it took so long for Marshall and Warren's findings to be accepted?
- Why do you think Marshall and Warren persevered with their research despite the lack of support from the medical community?
- Rather than waiting for approval from a medical ethics committee for human testing, Marshall used himself as a subject to test the effectiveness of antibiotic treatment. Today ethics committees would not give approval for scientists to carry out research on themselves. Consider some of the reasons why.
- Why was it so important for Warren and Marshall to have their research findings published?

Section 4.2 Exercises

1 State the names of three different metal compounds that are commonly used as antacids.

2 Give an equation for the neutralization of hydrochloric acid in the stomach by:

a sodium hydrogencarbonate

b magnesium hydroxide

c calcium carbonate.

3 **a** Explain how heartburn is caused.

b Describe how antacids can reduce the unpleasant effects of heartburn.

4 State the name of the anti-foaming agent that is added to some antacids.

5 Determine which would be more effective as an antacid, 500 mg of magnesium hydroxide or 500 mg of sodium hydrogencarbonate. Support your answer with balanced equations.

6 Explain why sodium hydroxide is not used as an antacid.

4.3 ANALGESICS

AS D.3.1
Describe and explain the different ways that analgesics prevent pain. © IBO 2007

Pain is part of the body's defence system. It usually triggers a reflex reaction to retract the affected part of the body from a painful stimulus. It can also help to adjust our behaviour to increase avoidance of that particular harmful situation in the future. Pain receptors are specialized nerve endings located throughout the body. They transmit pain signals from injury, disease, movement or environmental stress to the spinal cord and the brain. When tissues are damaged, white blood cells flood to the site to try to minimize tissue destruction. Prostaglandins are produced as a result. In turn the prostaglandins activate the inflammatory response—production of pain and fever. The pain receptors of spinal neurons are sensitized by the prostaglandins and they send a message indicating pain through the nervous system to the brain. Prostaglandins can also constrict blood vessels, increasing the body temperature and have a direct effect on the body's heat regulating centre, the hypothalamus, producing fever.

An **analgesic** is a pharmaceutical that relieves or prevents pain. Mild analgesics intercept the pain stimulus at the source, often by interfering with the enzyme-controlled production of substances such as prostaglandins that cause pain, swelling or fever. If the release of prostaglandins is restricted, the pain receptors are not sensitized, blood vessels are not constricted and the body temperature does not increase. The source of the injury or infection is not cured, but the symptoms that are produced by the injury or infection are hidden because the brain does not get the message that there is pain, and no fever occurs. Mild analgesics include aspirin, paracetamol (acetaminophen) and ibuprofen, also known as actiprofen®. Aspirin and ibuprofen are known as non-steroidal anti-inflammatory drugs (NSAIDs).

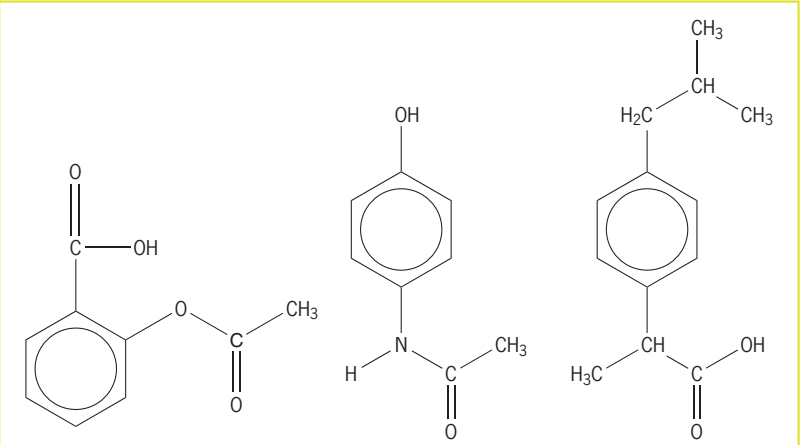


Figure 4.3.1 A comparison of the structures of aspirin, paracetamol and ibuprofen.

Strong analgesics act in a very different manner to mild analgesics. They temporarily bond to receptor sites in the brain, preventing the transmission of pain impulses without depressing the central nervous system. This method of pain relief copies that of the body's own chemical painkillers—endorphins and enkephalins—which are called endogenous opioid peptides. They are able to produce analgesia and a sense of well-being. Enkephalins bind to neuroreceptors in the brain and produce relief from pain. They give a temporary effect immediately after an injury. A strong analgesic such as morphine gives the same effect as the enkephalins, but is many times more effective.

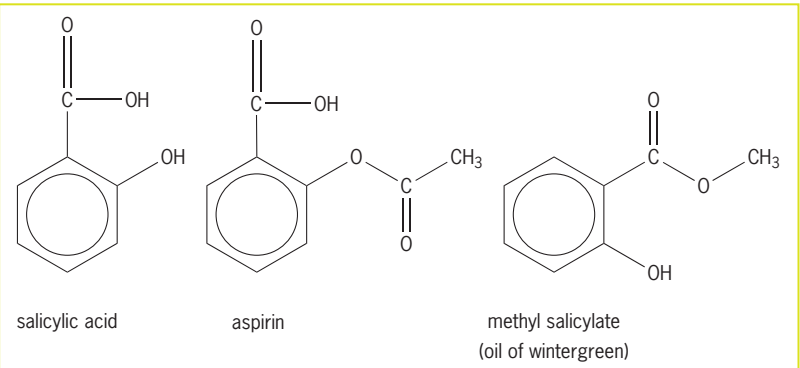


Figure 4.3.2 Aspirin (acetylsalicylic acid) and methyl salicylate are derivatives of salicylic acid.

AS D.3.2
Describe the use of derivatives of salicylic acid as mild analgesics, and compare the advantages and disadvantages of using aspirin and paracetamol (acetaminophen).

Pr PRAC 4.2
Preparation and analysis of aspirin

The analgesic effects of the esters of salicylic acid have been known for centuries. Oil of wintergreen is used in heat rubs to relieve muscle soreness. Salicylic acid acts as an analgesic, but it irritates the stomach, while acetylsalicylic acid (aspirin) is less acidic and less irritating. Aspirin passes through the acidic stomach contents unchanged to be hydrolysed in the alkaline conditions of the intestines to sodium salicylate, which is absorbed into the blood stream. Aspirin is an ester formed by the reaction of salicylic acid with ethanoic anhydride. It may also be prepared by the reaction of ethanoic acid with salicylic acid. In 1899 the Bayer Company in Germany marketed acetylsalicylic acid for the first time.

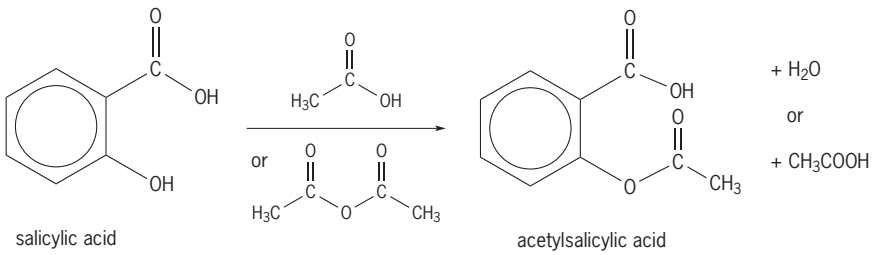


Figure 4.3.3 The formation of aspirin.

Ws WORKSHEET 4.1
Aspirin

Salicylic acid and aspirin have similar properties, due to the similarity of their structures. Another mild analgesic is paracetamol, also known as acetaminophen, with quite a different structure to that of aspirin (see figure 4.3.1): paracetamol is an amide; aspirin is an ester. Paracetamol is more likely to be prescribed to pregnant women for mild pain relief since it lacks many of the side-effects of aspirin.

TABLE 4.3.1 THE ADVANTAGES AND DISADVANTAGES OF ASPIRIN AND PARACETAMOL (ACETAMINOPHEN)

	Advantages	Disadvantages
Aspirin	• Shows antipyretic action as well as anti-inflammatory effects. • Useful in preventing the recurrence of heart attacks because it thins the blood.	Can cause: • ulceration and stomach bleeding • allergic reactions • Reye's syndrome (a potentially fatal liver and brain disorder in children).
Paracetamol (acetaminophen)	• Shows antipyretic action. • Very safe in correct dose for children, adults and pregnant women.	• Does not have much anti-inflammatory action. • Can (rarely) cause blood disorders and kidney damage. • An overdose may cause serious liver damage, brain damage and even death.

THEORY OF KNOWLEDGE

Pain has two elements. The nerve endings of sensory neurons receive the stimuli and the brain translates this information into a feeling of pain. There are, however, limitations associated with relying too heavily on our perception of pain as a way of knowing. For example, our brain can either ignore or recognize the pain based on past associations or expectations without us consciously being aware this is happening. If we expect to feel pain we will and if we don't expect it we won't. Studies show that when patients are given a placebo pill made of sugar and flour and told that it will decrease their sensitivity to pain they actually experience less pain. Hypnosis can be used to control not only the brain's capacity to control pain, but also to alleviate anxiety, nausea, fatigue and emotional distress. Hypnosis is a physiological state caused by a combination of a highly focused attention with deep physical relaxation. You can fall into a hypnotic state when you are watching a movie and get so absorbed in it that you think what you are seeing is real. In theatre arts this is referred to as the 'suspension of disbelief'.

- Can you describe a situation in which your expectations or past associations affected your perception of pain?
- Dr David Speigel of Stanford University said 'People who are hypnotized see what they believe. They don't just tell you what it is—it actually looks that way to them.' If this is the case, to what extent can we rely on our senses as a way of knowing about the world as it really is?

Opium is a **narcotic** that comes from the sap of immature seed pods of opium poppies. It is a mixture of compounds including morphine and codeine, which are therefore known as opiates. Cultivation of opium poppies for food, anaesthesia, and ritual purposes dates back to at least the Neolithic Age. Many ancient civilizations made widespread use of opium, which was the most potent form of pain relief then available, allowing their surgeons to perform prolonged surgical procedures. Opium is mentioned in important medical texts of the ancient world.

Narcotics are analgesics that produce euphoria (a feeling of peace and tranquillity) as well as causing loss of feeling or paralysis. Both morphine and heroin, which is synthesized from morphine, are narcotics.

The structures of morphine, codeine and heroin all contain the ether, alkene (carbon–carbon double bond) and tertiary amine functional groups in exactly the same locations. They differ only in the functional groups found on the aromatic ring and on the ring containing the carbon–carbon double bond. These are circled in figure 4.3.5. The two functional groups in morphine are hydroxyl groups. The reaction in which heroin is produced involves a

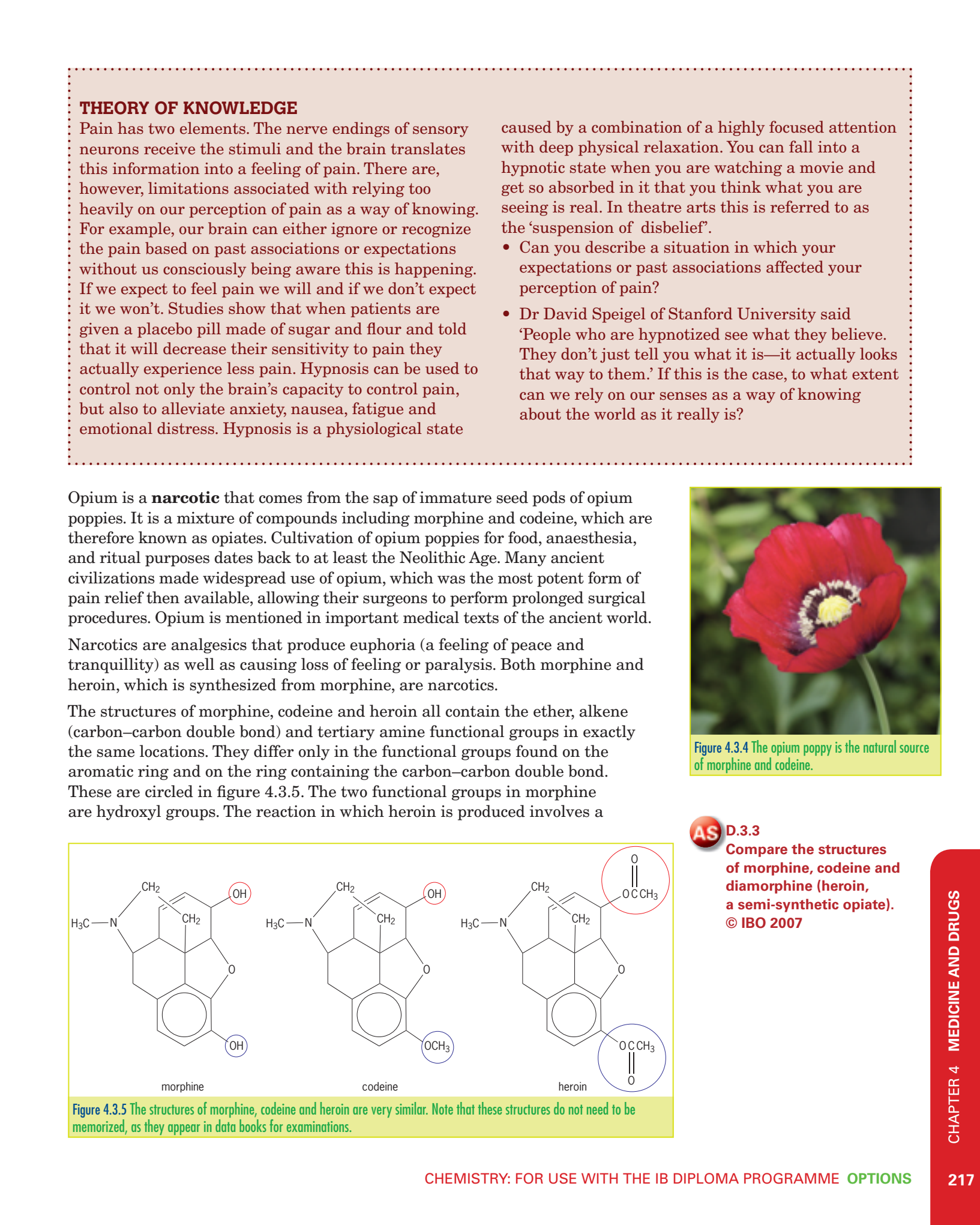


Figure 4.3.4 The opium poppy is the natural source of morphine and codeine.

AS **D.3.3 Compare the structures of morphine, codeine and diamorphine (heroin, a semi-synthetic opiate).**

Figure 4.3.5 The structures of morphine, codeine and heroin are very similar. Note that these structures do not need to be memorized, as they appear in data books for examinations.

condensation reaction between morphine and ethanoic acid. In this reaction these hydroxyl groups will react with the carboxyl group of ethanoic acid and an ester linkage will be formed. The hydroxyl groups are replaced by ethanoate groups. Codeine has one hydroxyl group in common with morphine, but the second hydroxyl group (on the aromatic ring) has been replaced by a methoxy group ($-OCH_3$).

TABLE 4.3.2 THE ADVANTAGES AND DISADVANTAGES OF MORPHINE AND ITS DERIVATIVES AS STRONG ANALGESICS

Advantages	Disadvantages
• Fifty times as potent as aspirin • Pain-relieving strength is superior to all others	• Very addictive • Addicts suffer mental fogginess and mood changes • Must be injected

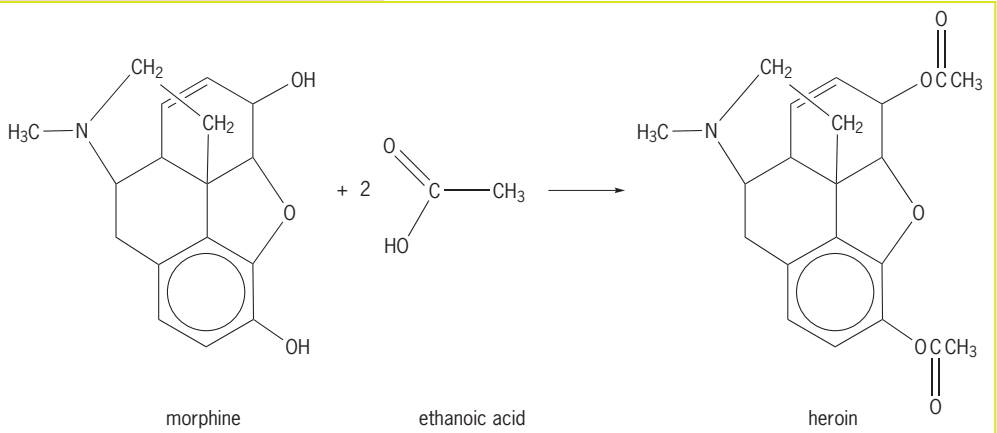


Figure 4.3.6 The reaction between morphine and ethanoic acid produces heroin.

Studies of the efficacy of various opioids have indicated that no other narcotic analgesic is more effective or superior to morphine in the management of severe pain.

Morphine was first isolated in pure form in 1806 by a German pharmacist's assistant called Frederich Sertürner. He gave it the name morphine from the name Morpheus, the Greek god of dreams, after frequently testing it on himself and nearly dying of an overdose at one point. Morphine use developed slowly over the next 50 years, but when hypodermic needles became available in the 1850s it became more accessible. It was widely used in the American Civil war as a pain-killer during surgery and, more commonly, it was used indiscriminately in large doses to relieve the pain and discomfort of dysentery. Morphine, however, is highly addictive, with tolerance and physical and psychological dependences developing very rapidly. It is estimated that as many as 400 000 soldiers became addicted to morphine during this time.

AS **D.3.4 Discuss the advantages and disadvantages of using morphine and its derivatives as strong analgesics.**

Codeine is about 10% as effective as morphine as a pain-killer, but it does not cause addiction. Although codeine can be extracted from opium, most codeine is synthesized from morphine. Codeine can be taken by mouth and is included in combination with other ingredients in tablets that are marketed as strong pain-killers. It is available over the counter in some countries and only by prescription in others. It is the least potent of the opiates.

Heroin is much more potent than morphine, and much more addictive. It depresses the nervous system, causing drowsiness, respiratory depression and decreased gastrointestinal movement, which leads to constipation, nausea and vomiting. Frequent administration has a high potential for causing addiction, and tolerance occurs over a very short period. Heroin is rejected for use by the medical profession.

Section 4.3 Exercises

1 Aspirin and acetaminophen (paracetamol) are commonly used mild analgesics.

a State one advantage of using aspirin rather than acetaminophen for reducing pain.

b Other than reducing pain, state one common effect of using either acetaminophen or aspirin.

c State one unwanted side-effect of using aspirin.

2 **a** State the names of three strong analgesics that you have studied in this option.

b Name two functional groups that are common to these three analgesics.

3 If a doctor were to prescribe morphine for a patient after surgery:

a describe the main effect that would be sought from this drug

b describe a major side-effect that might be experienced by the patient.

4 Consider the following table.

Drug	LD_{50} (mg kg^{-1} body mass)
Morphine	20
Heroin	4

a In a population of 100 000 people, for how many people would a dose of 20 mg of morphine per kg body mass be lethal?

b Identify which of the two substances (morphine or heroin) is less toxic.

5 Mild and strong analgesics have very different methods of action. Describe the difference in the method of action of the two types of analgesic.

6 **a** The nitrogen-containing functional group in acetaminophen is an amide group. Draw the structure of this functional group.

b Name and draw the structure of the nitrogen-containing functional group in codeine.

7 The conversion of morphine into heroin by reaction with ethanoic acid is a condensation reaction. Refer to the structural formulas in figures 4.3.5 and 4.3.6.

a Name and state the formula of the functional group in morphine that reacts with the ethanoic acid in this reaction.

b State the formula of the functional group in heroin that is produced in this reaction.

c Morphine can also be converted into codeine.

i Calculate the difference in relative formula mass between codeine and morphine.

ii Name the functional group that is not changed in the reaction in which morphine is converted to codeine, but is changed in the reaction in which morphine is converted to heroin.

8 Aspirin is a derivative of salicylic acid.

a Draw the structure of salicylic acid.

b State the general names of the two functional groups attached to the benzene ring in a molecule of salicylic acid.

c One of the functional groups in salicylic acid is changed in the reaction with acetic anhydride (or ethanoic acid) to produce aspirin. State the general name of this new functional group in aspirin.

4.4 DEPRESSANTS

Depressants and antidepressants are commonly confused in common language. While depressants are chemical agents that diminish the function or activity of a specific part of the body, antidepressants are psychiatric medications or other substances (nutrients or herbs) used for alleviating depression. Nonetheless depressants are often described as antidepressants because they relieve depression. Depressants are known by range of other names. Among these names are tranquillizers, sedatives, anxiolytics, soporifics and sleeping pills. All of these drugs create a similar effect.

AS **D.4.1 Describe the effects of depressants. © IBO 2007**

Depressants calm and relax the central nervous system (CNS). Many depressants acting on the CNS do so by increasing the activity of a particular neurotransmitter known as gamma-aminobutyric acid (GABA). They slow down the activity of the brain, reduce the rate of breathing and generally dull emotional responses. Depressants may take a while to get used to and side-effects are similar to the effect of too much alcohol—slurred speech, dizziness and loss of coordination.

At low doses, depressants may have little or no effect, while at moderate doses they induce sedation, giving a soothing effect and a reduction of anxiety.

At higher doses depressants may induce sleep. It is in this capacity (to relieve anxiety or insomnia) that depressants are frequently prescribed to patients.

AS **D.4.2 Discuss the social and physiological effects of the use and abuse of ethanol. © IBO 2007**

The use of depressants at very high doses is extremely dangerous, as the very slow breathing and slow heart rates that result may cause death. The most common medically used depressants generally fall into two classes: the barbiturates and the benzodiazepines. Other depressants include alcohol and narcotics.

CHEM COMPLEMENT

The great Austrian wine scandal

Austria is currently considered together with France, Spain and Italy, one of the great wine-making countries of Europe. Despite vineyards having been part of the Austrian landscape since 700 bc, the reputation of Austrian wine took a steep downward plunge in 1985. At that time it was discovered that some Austrian winemakers had been adding diethylene glycol, a close relative of ethylene glycol, which is used in antifreezing agents, to their wine. The effect of the diethylene glycol was to create a sweeter and more full-bodied taste; however, diethylene glycol is a toxic substance. The LD_{50} of diethylene glycol in small mammals is between 2 and 25 g kg^{-1}, however the danger from the wine was small since to reach such levels an adult would have to drink more than 28 bottles per day for two weeks! Nonetheless, the Austrian wine export industry was devastated by the revelation of this fraud. As a result, stricter regulations were imposed on Austrian winemakers, and the industry shifted its emphasis from the production of bulk, sweet wines to lower yields of higher quality, drier wines. These new laws helped Austria to regain its position of respect among the winemakers of Europe.

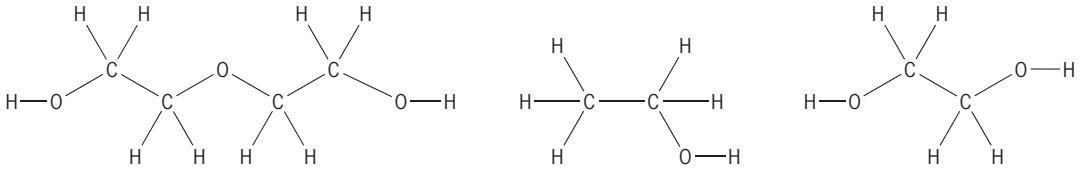


Figure 4.4.1 Diethylene glycol and ethylene glycol are diols. Each molecule has two hydroxyl groups.

Ethanol, C_2H_5OH, is the alcohol that is dissolved in alcoholic drinks. No other member of the alcohol homologous series is used. Alcohol (ethanol) is the most widely used drug in the world. The brewing of alcoholic beverages can be traced in history as far back as ancient Egypt. Throughout Egypt there are many tomb paintings illustrating the gathering and pressing of grapes and making them into wine. Descriptions of various beer recipes can be found in Sumerian writings, some of the oldest known writing of any sort.

PRAC 4.3
Determination of the alcohol content in wine

Figure 4.4.2 A wall painting in the tomb of Queen Nefertari of Egypt, who lived around 1300-1255 BC, shows her offering jars of wine to the gods.

Alcohol is a depressant of the central nervous system, so it makes a person feel relaxed. In large amounts it has a serious effect on mental and physical performance, slowing response times and making activities such as driving a car or operating heavy machinery very dangerous.

Ethanol is an addictive drug. The condition of being addicted to ethanol is called alcoholism. This can cause hardship to the families of addicts. Their financial situation can become perilous as the family income is spent on large amounts of alcoholic drinks instead of food and other essential items. The physical wellbeing of the family is often at risk since many people become bad tempered and violent under the influence of alcohol. A lack of productivity and the subsequent loss of employment by the alcoholic also puts their families under pressure.

Alcohol has many short-term health effects:

- It has an anaesthestic effect on the brain, which prevents quickness of thought and reactions. Reaction time is slowed and coordination is impaired.
- Heart rate speeds up slightly and small blood vessels close to the surface of the skin expand, allowing more blood to flow closer to the surface and lowering blood pressure at the same time.
- Vision can become blurry and speech slurred.
- The liver becomes overloaded if too much ethanol is consumed at once, resulting in headaches and the general feeling of illness the next day that is known as a 'hangover'.
- The diuretic effect of ethanol can lead to dehydration as the body uses large amounts of water to try to rid itself of the 'poison'.
- It can cause stomach upsets, heartburn, sickness and diarrhoea.
- Unconsciousness and possible heart failure and death can result. Alcohol that is already in the stomach will continue to be absorbed even after the person is unconsciousness and so can reach lethal levels. An unconscious person could also be sick and suffocate on their vomit.

Because the liver is primarily involved in breaking down ethanol, it is liver damage that is on the top of the list of long-term health effects from ethanol use and abuse. Although a small percentage of alcohol is digested in the stomach, most is absorbed as it is and is transported through the body in the blood stream. The liver is the only organ that produces enough of the enzyme alcohol dehydrogenase to oxidize alcohol at an appreciable rate. Normally the liver is responsible for transforming fatty acids into triglycerides; however, when alcohol is present, the liver cells are forced to first metabolize the alcohol, letting the fatty acids accumulate, sometimes in large amounts. The structure of the liver cells is permanently changed to reflect their new purpose, so long-term heavy drinkers become less able to metabolize fats. The liver is able to metabolize about one standard alcoholic drink (about 15 g of ethanol) per hour. If more alcohol arrives in the liver than the enzymes can handle, the excess alcohol travels to all parts of the body, circulating until the liver enzymes are finally able to process it.

Long-term damage to health includes:

- cirrhosis of the liver
- cancer of the liver, as well as of many other organs
- increased risk of coronary heart disease and high blood pressure
- increased risk of stroke
- anxiety and depression
- poor eating habits that could even lead to malnutrition. Although the energy content of ethanol is quite high, there are no nutrients in ethanol, so the more energy an individual consumes as ethanol, the less likely they are to eat enough food to obtain adequate nutrients
- changes in physical appearance due to poor nutrition, as well as ruddiness in the face due to expanded blood vessels close to the surface and even a purple, bulbous 'drinker's nose'
- alcoholic hepatitis and fatty liver.

In pregnant women, excessive drinking of alcohol can cause miscarriage, low birthweight babies and foetal abnormalities (foetal alcohol syndrome).

When alcohol is taken at the same time as other drugs, the performance of the other drug is often enhanced significantly. This enhanced effect is known as a **synergistic effect**. For example, alcohol taken with sleeping pills will increase the sedative effects of the pills and can produce coma and death, and alcohol taken with aspirin increases the risk of stomach bleeding. Many prescription medications come with a warning that they are not to be taken with alcohol.

AS **D.4.4 Describe the synergistic effects of ethanol with other drugs. © IBO 2007**

Detection of ethanol in the breath, the blood and urine

The consumption of alcoholic drinks results in the rapid absorption of ethanol into the bloodstream. This travels through the body and makes its way rapidly to the brain, impairing the reflexes and the ability of an individual to drive safely. For this reason, methods have been developed to quickly assess the blood alcohol levels of drivers, so that those with levels that could prove dangerous are discouraged from driving.

Figure 4.4.3 A roadside breath test uses a hand-held breathalyser to measure the motorist's blood alcohol level.

In some countries, such as Australia, all drivers and passengers over the age of 16 involved in motor vehicle accidents attended by the police are required to have a blood sample taken. This is routinely screened for the presence of ethanol at the Forensic Science laboratories using gas–liquid chromatography. If the sample has a blood alcohol concentration (BAC) that is over the maximum allowable level, the police look into the matter further and charges may be pressed.

Ethanol is quite volatile, so when an alcoholic drink has been consumed, some ethanol passes into the lungs from the blood. It is this ethanol that is detected in the 'breathalyser' test performed on motorists by police (on a random basis in Australia). The amount of ethanol in the breath is directly proportional to the amount of ethanol in the bloodstream. The BAC is measured as a percentage mass per volume. For example, a BAC of 0.05% indicates a concentration of 0.05 g ethanol per 100 cm^3 (or 50 mg ethanol per 100 cm^3) of blood.

A tolerance for alcohol can be built up and although the outward signs of alcohol consumption may be reduced by a tolerance, the long-term health effects are not. In people who have not built up a tolerance, a BAC rating of 0.20% represents very serious intoxication (most first-time drinkers would be unconscious by about 0.15%) and 0.35% represents potentially fatal alcohol poisoning. The accepted LD_{50} for alcohol is 0.40%.

Another factor affecting the accuracy of blood alcohol testing is the assumption that the individual being tested is average in various ways. An assumption is made in terms of the ratio between the alcohol present in the blood to alcohol present in the breath. This amount is 2100 to 1; however, the actual ratio in any given individual can vary from 1300 : 1 to 3100 : 1, or even more widely. This ratio varies not only from person to person, but within one person from moment to moment. This can result in a reading being too high, or too low, depending on the person. In addition, it is assumed that all the alcohol has been absorbed, when it is quite possible that the motorist still has some alcohol in their stomach. The blood alcohol level reaches its highest point one hour after the alcoholic beverage has been drunk.

AS **D.4.3 Describe and explain the techniques used for the detection of ethanol in the breath, the blood and urine. © IBO 2007**

Figure 4.4.4 In the breathalyser orange potassium dichromate is reduced to green chromium(III) ions.

The breathalyser contains potassium dichromate, $K_2Cr_2O_7$. This orange compound is reduced to green chromium(III) ions, Cr^{3+}, when it reacts with ethanol. The two half equations for this redox reaction are:

$$Cr_2O_7^{2-}(aq) + 14H^+(aq) + 6e^- \rightarrow 2Cr^{3+}(aq) + 7H_2O(l)$$
orange green

In this reduction reaction, the oxidation number of chromium decreases from +6 in $Cr_2O_7^{2-}$ to +3 in Cr^{3+}.

$$C_2H_5OH(aq) + H_2O(l) \rightarrow CH_3COOH(aq) + 4H^+(aq) + 4e^-$$
ethanol

The breathalyser uses fuel cell technology in which the current generated by the flow of electrons from reducing agent to oxidizing agent gives an indication of the alcohol concentration. The greater the electron current, the greater the concentration of alcohol in the breath. This device is calibrated to allow for the ratio of blood alcohol to breath alcohol.

The accuracy of this method is not high, so if a reading close to 0.05 is recorded, more accurate tests are performed using an intoximeter that employs the technique of infrared spectroscopy. This supplies evidence that will stand up in court. In an intoximeter, infrared radiation is passed through the breath sample and the reference sample. The hydroxyl group in ethanol causes absorption of a specific wavelength of infrared radiation, with the amount of absorption depending upon the amount of ethanol in the breath.

If in a roadside test the driver disagrees with the BAC determined by IR spectroscopy, a blood sample may be taken and tested using gas–liquid chromatography, which not only detects that ethanol is present, but it also accurately measures the amount of ethanol present, thus giving an accurate BAC.

DEMO 4.1
The breathalyser

For more information on chromatography see section 1.6.

Gas–liquid chromatography (GLC) is a technique in which the components of a mixture are separated according to their relative attractions to a stationary phase and a mobile phase. The following stages occur in the analysis of a blood sample for ethanol using gas–liquid chromatography.

1 The blood is heated to make its volatile components (such as ethanol) become gases, and these are swept along the column by the carrier gas —this is the mobile phase.

2 The stationary phase is a liquid, coated onto the surface of solid inert particles that pack the column. The column is long and narrow, allowing very accurate separation of the components of the mixture (blood). The relative solubility of the components in the stationary phase separates them.

3 A sample of ethanol (the reference sample) is run through the GLC to establish its retention time, then the blood sample is run through. A peak with the same retention time as the reference sample will correspond to the ethanol in the sample of blood. The area under the peak gives the amount of ethanol present.

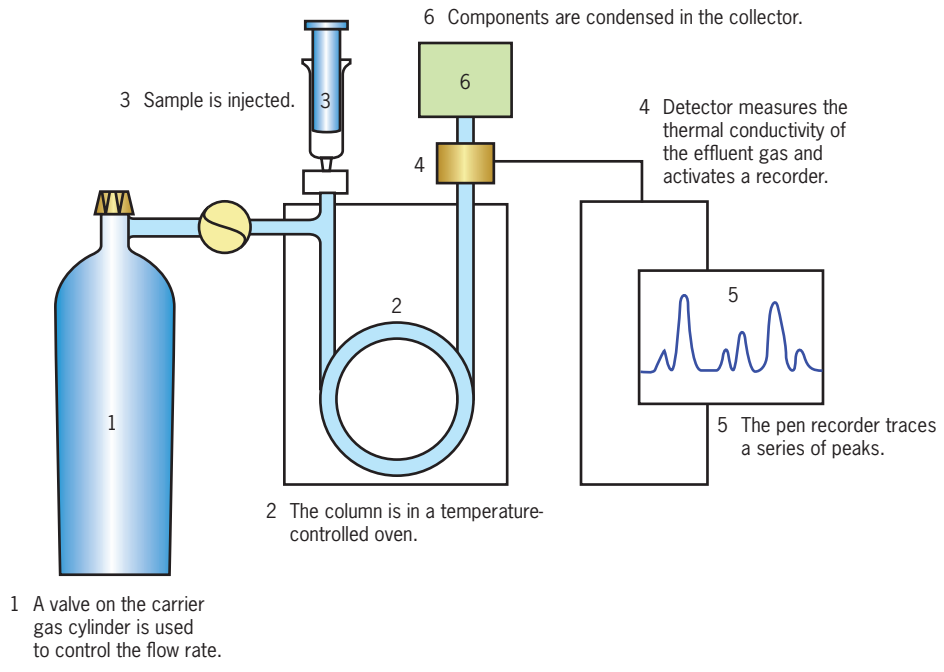


Figure 4.4.5 The presence of ethanol and its exact amount can be determined by gas–liquid chromatography.

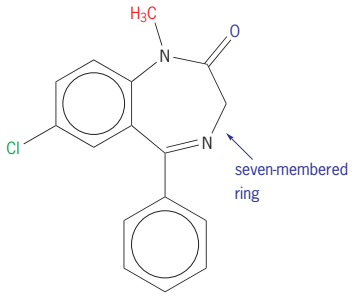
The level of ethanol in urine can be measured using test strips like those used for the testing of other drugs. If the person has not had a drink for 45–60 minutes, their body will have started to eliminate the ethanol in urine, and the urine alcohol level will be about 1.3 times the blood alcohol level. The method of testing urine for alcohol content is one that is used in the workplace; however, the urine alcohol level may not correlate with the degree of intoxication observed at the time of the test. At other times (i.e. when the person is still drinking, or has just stopped) it is impossible to correlate the urine alcohol concentration with the blood concentration. For this reason analysis of a urine sample is not regarded as reliable and cannot be used as evidence in a court case.

AS D.4.5 Identify other commonly used depressants and describe their structures. © IBO 2007

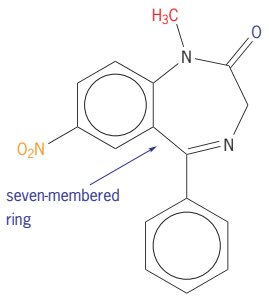
The majority of depressants can only be obtained with a prescription from the doctor. They are known by pharmacists as **anxiolytics** because they treat states of anxiety such as panic attacks, phobias and more generalized anxiety disorders.

- Diazepam (Valium®) is a depressant of the central nervous system. It is prescribed as a tranquillizer for many complaints such as muscular disorders and spasms and to promote sleep.
- Nitrazepam (Mogadon®) is used in sleeping pills.
- Fluoxetine hydrochloride (Prozac®) is an *antidepressant* that is used to treat clinical depression, obsessive compulsive disorder, bulimia nervosa and panic disorders. It is currently the most prescribed antidepressant.

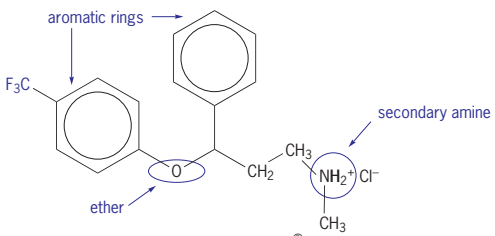
Diazepam and nitrazepam are members of a group of compounds called benzodiazepines. These compounds all have a seven-membered ring that was synthesized by accident in 1961. This ring was found in testing to have pharmacological activity and so the family of benzodiazepines began.



diazepam (Valium®)



nitrazepam (Mogadon®)



fluoxetine hydrochloride (Prozac®)

Figure 4.4.6 The structures of diazepam, nitrazepam and fluoxetine hydrochloride.

The basic structure of benzodiazepines such as diazepam and nitrazepam is that of a seven-membered ring fused to an aromatic ring, with four main substituent groups. In diazepam, these groups are the carbonyl and methyl groups and an aromatic ring on the seven-membered ring, plus a chloro group on the fused aromatic ring. In nitrazepam, the carbonyl and methyl groups and the aromatic ring are still present, however the chloro group on the fused aromatic ring has been replaced by a nitro, NO_2, group.

The structure of the antidepressant fluoxetine hydrochloride is significantly different from that of the benzodiazepines. It has two aromatic rings bonded to an ether group. A CF_3 group is present on one aromatic ring, while there is a secondary amine at the opposite end of the molecule present as NH_2^+ forming a salt with a Cl^- ion.

Section 4.4 Exercises

1 Describe the effects of depressants when given in:

a low doses

b high doses.

2 A simple test for the presence of ethanol in breath is carried out in the breathalyser.

a State the name of the substance with which the ethanol reacts in the breathalyser.

b Describe the colour change that occurs in this reaction.

c State the type of reaction that is occurring in the breathalyser.

3 A more accurate estimation of the concentration of ethanol in the blood can be found using an intoximeter. Explain how an intoximeter works.

4 Describe how the concentration of ethanol in a sample of blood is determined using gas chromatography.

5 Using the specific example of aspirin, describe what is meant by the 'synergistic effect' of drinking alcohol (ethanol) while taking other drugs.

6 In some cases, depressants are described as antidepressants. Explain why this is an accurate description.

7 **a** Describe three specific long-term health problems that may arise from regularly drinking large amounts of ethanol.

b Describe two specific social or economic problems that may arise from regularly drinking large amounts of ethanol.

8 Nitrazepam and diazepam are very similar in their structures. Referring to figure 4.4.6 describe two similarities and one difference between the two structures.

4.5 STIMULANTS

Stimulants generally make the user feel less tired and more alert. In particular this is noticed when the user is already tired. They stimulate the brain and central nervous system.

Other effects include:

- an increase in heart rate and blood pressure
- an increase in breathing rate
- dilation of the pupils of the eyes
- constriction of arteries
- sweating
- a reduction in appetite.

AS D.5.1
List the physiological effects of stimulants. © IBO 2007

Amphetamines stimulate the central nervous system in a similar manner to adrenaline, a hormone that is released by the adrenal gland.

AS D.5.2
Compare amphetamines and epinephrine (adrenaline). © IBO 2007

The action of adrenaline is known as the 'fight or flight' response. This reaction allows an animal to address a threat by either standing up for itself (fight) or escaping the threat (flight). The supply of oxygen and glucose to the brain and muscles is boosted by a range of actions. The immune system is suppressed, as are non-emergency bodily procedures. The ability of amphetamines to work in this way causes them to be known as **sympathomimetic** drugs. They mimic the action of adrenaline on the body by releasing noradrenaline from the nerve terminals in the brain.

They are addictive. Their main effects are to:

- increase locomotor stimulation (they make the user move around more quickly)
- induce a sense of euphoria and excitement
- temporarily reduce appetite.

In addition, amphetamines and adrenaline increase the heart rate, blood pressure, rate of respiration and wakefulness, causing restlessness and insomnia. They allow the body to use reserve energy, but such use of reserve energy may be followed by sudden exhaustion leading to blackout or collapse.

The structure of amphetamines is similar to that of epinephrine (adrenaline). They both have an aromatic ring substituted with a hydrocarbon chain containing an amine group. In the case of amphetamine, this amine is a primary amine. In epinephrine the amine is a secondary amine and there are three hydroxyl groups present bonded to the aromatic ring and to the hydrocarbon chain. Both structures are derived from the phenylethylamine structure.

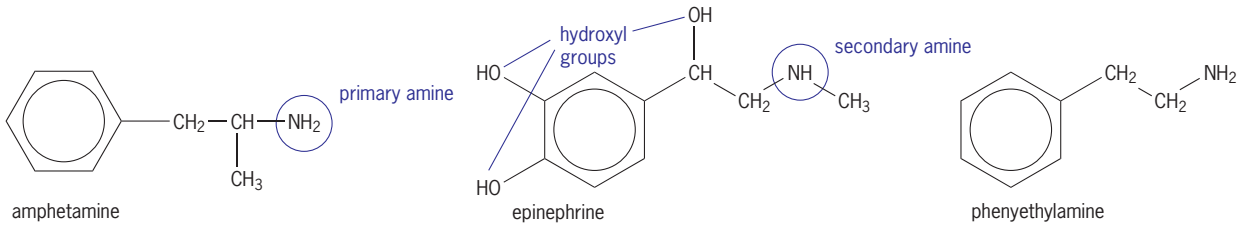


Figure 4.5.1 A comparison of the structures of amphetamine , epinephrine (adrenaline) and phenylethylamine.

Medical use of amphetamines includes treatment of attention deficit or hyperactivity disorder (ADHD) in children and the treatment of narcolepsy, a condition in which the patient suddenly and unpredictably falls asleep at frequent intervals during the day.

Nicotine

AS D.5.3
Discuss the short- and long-term effects of nicotine consumption. © IBO 2007

Nicotine is found in the leaves of the tobacco plant and is therefore present in cigarettes and cigars. It is the most commonly used stimulant across the world. According to the World Health Organization, in 2002 around one-third of the world's adult male population smoked and smoking-related diseases kill one in ten adults globally, or cause four million deaths per year. The statistics involving youth were also startling (see Chem Complement). In its pure form, nicotine is an extremely toxic drug that acts as quickly as cyanide. The LD_{50} of nicotine can be as low as 0.5–1.0 mg/kg for adult humans. In the tobacco plant, nicotine serves to drive away and even poison insects that would eat the leaves.

Figure 4.5.2 Tobacco leaves are harvested, dried and used to make cigarettes and cigars.

CHEM COMPLEMENT

How common is smoking in young people?

Data from the World Health Organization in 2002 revealed the following statistics:

- Among young teens (aged 13–15) worldwide, about one in five smokes.
- Between 80 000 and 100 000 children worldwide start smoking every day—roughly half of whom live in Asia.
- Evidence shows that around 50% of those who start smoking in adolescent years go on to smoke for 15–20 years.
- Peer-reviewed studies show teenagers are heavily influenced by tobacco advertising.
- About a quarter of youth alive in the Western Pacific (covering East Asia and the Pacific) Region will die from smoking.

When people smoke for the first time, the nicotine stimulates the sensory receptors in the stomach, inducing nausea and vomiting. Diarrhoea may also occur as the nicotine first stimulates, then inhibits glandular secretions. While initially only a small amount of nicotine can be absorbed before nausea and vomiting begins, a rapid tolerance occurs in persistent smokers. The smoke from one cigarette contains about 6 mg of nicotine, of which about 0.8–1.5 mg is absorbed into the body and remains in the blood plasma for up to 100 minutes after smoking.

Nicotine is a stimulant, speeding up the heart by about 20 beats per minute with every cigarette; it raises blood pressure, and is a **vasoconstrictor**, making it harder for the heart to pump through the constricted arteries.

TABLE 4.5.1 SHORT-TERM AND LONG-TERM EFFECTS OF NICOTINE CONSUMPTION	
Short-term effects	**Long-term effects**
• Increased alertness	• Increased risk of heart disease and coronary thrombosis
• Reduction of anxiety and tension	• Physical and psychological dependence (addiction)
• Increased heart rate	• Other toxic chemicals are present in cigarette smoke that increase risk of mouth, throat and lung cancer.
• Increased blood pressure	• CO in cigarette smoke reduces the ability of the blood to carry O_2
• Constricts blood vessels, putting stress on the heart	• Increased risk of bronchitis and emphysema
• Reduces urine output	• Cigarettes are costly, so a nicotine addiction increases financial strain on families

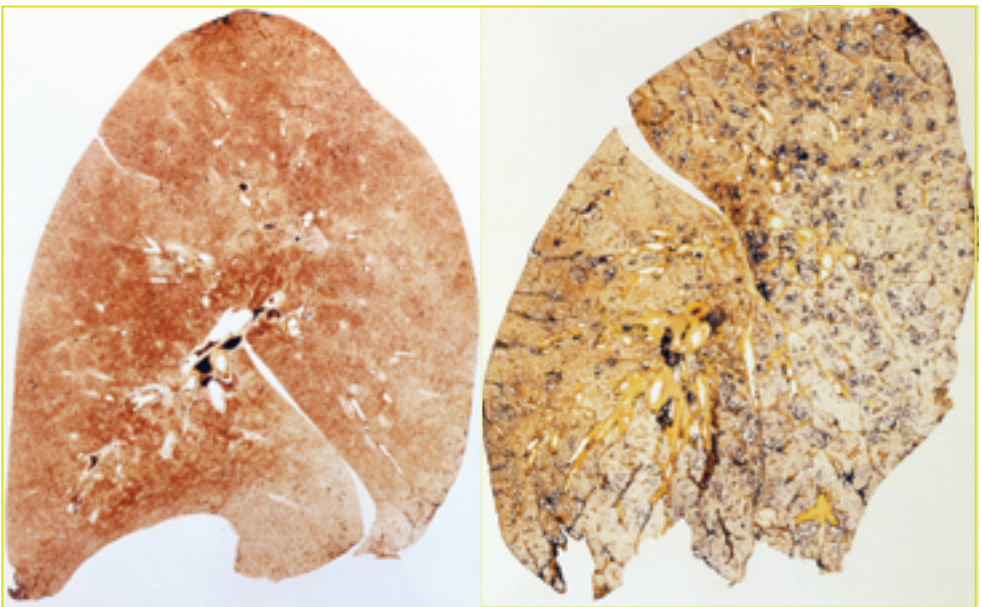


Figure 4.5.3 A comparison of a healthy lung and the lung of a smoker. The black areas in the smoker's lung are carcinogenic tar deposits.

Cigarette smoke condenses to form tar. The lungs of a smoker receive about 100 g of tar a year. This tar has been found to be carcinogenic in all kinds of tissue. The other toxic chemicals in cigarette smoke are responsible for the large range of long-term effects of cigarette smoking such as cancers of the lung, mouth and larynx; bronchitis and emphysema.

Nicotine is an extremely addictive drug. It is easier to become addicted to nicotine than to alcohol or barbiturates. As a result, nicotine produces psychological and physical dependence in users. This creates the need for the addict to keep buying cigarettes or tobacco in other forms and carries with it a large financial burden. One packet of 25 cigarettes costs as much as five loaves of bread and is often smoked by a person in one day.

WORKSHEET 4.2
Depressants and stimulants

AS D.5.4
Describe the effects of caffeine and compare its structure with that of nicotine. © IBO 2007

Caffeine

Caffeine is found in many drinks that we consider relatively harmless. A cup of coffee contains between 40 and 176 mg of caffeine, while a cup of tea contains much less caffeine: from 8 to 91 mg. Chocolate and cola drinks also contain caffeine, with a can of cola drink containing around 40 mg of caffeine. Energy drinks containing 4–5 times the recommended daily intake of caffeine in just one drink are becoming more popular, especially with children. The legal limit for caffeine in soft drink varies across the world and is about 145 to 200 mg kg^{-1}.

Caffeine stimulates the central nervous system by increasing the metabolic rate of nerve cells. In moderate doses caffeine enhances alertness, wellbeing, energy, motivation and concentration. It is found to sustain intellectual effort when tiredness sets in.

In higher doses, it can affect the physical coordination of the user, and it can also cause sleeplessness, anxiety and irritability.

Contrary to the effects of nicotine which reduces urine output, caffeine is a weak diuretic; it increases urine output. As a result it has a dehydrating effect on the user. While caffeine is not thought to be physically addictive, it can lead to mild psychological addiction. Many people have routines in their day that centre around their cups of coffee or tea. Caffeine constricts the cerebral (in the brain) blood vessels. Quitting a several cups of coffee a day 'habit' may produce a powerful headache as the constricted blood vessels dilate. This is one of the best known withdrawal symptoms of caffeine.

The two stimulants, caffeine and nicotine have some similarities in their chemical structures. Both molecules contain tertiary amine groups and both contain six- and five-membered rings. In the case of caffeine, these rings are fused and the six membered ring also contains two carbonyl (C=O) groups.

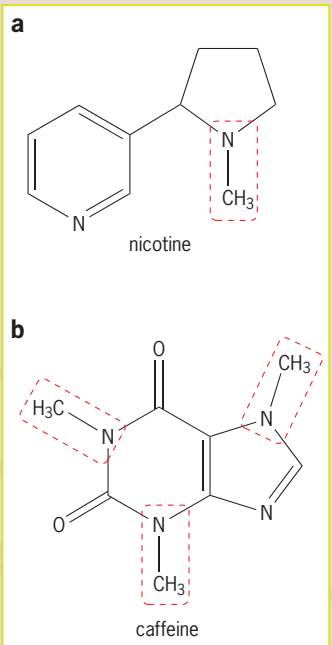


Figure 4.5.4 The structures of (a) nicotine and (b) caffeine. Both molecules contain tertiary amine groups (indicated).

Section 4.5 Exercises

1 Although stimulants make the user feel less tired and more alert, there are many other effects that are observed in a user of stimulants. List four other physiological effects on the body of using stimulants.

2 Amphetamine is a sympathomimetic drug. It has a structure that is very similar to that of adrenaline (see figure 4.5.1).
   - **a** Considering the functional groups in the two molecules, describe two structural similarities in amphetamine and adrenaline.
   - **b** Describe a structural difference between amphetamine and adrenaline.

3 Describe three long-term physical effects of smoking tobacco.

4 Identify the chemical present in tobacco that is a stimulant and is responsible for addiction to smoking.

5 Describe three psychological, social or economic effects of smoking tobacco.

6 Caffeine is a common stimulant with a fused ring structure. Draw the structure of caffeine and circle:
   - **a** a tertiary amine functional group
   - **b** a secondary amine functional group
   - **c** a carbonyl functional group.

7 Describe two effects of consuming caffeine in large amounts.

8 Considering the functional groups present in both compounds, compare the structures of nicotine and caffeine.

4.6 ANTIBACTERIALS

Bacteria are single-celled micro-organisms. They are typically a few micrometres in length and have a wide range of shapes, ranging from spheres to rods to spirals. Bacteria can be found in every habitat. Before the early 20th century, the action of bacteria was responsible for many deaths. **Pathogenic** bacteria caused diseases such as cholera, tuberculosis, syphilis, anthrax, leprosy and bubonic plague. Surgery was not a safe procedure, with the post-surgical possibility of infection almost being more of a problem than the need for the surgery. The first antibacterial (antibiotic) medicines were derived from living organisms.

Penicillin is an antibiotic that was developed in the early 20th century as the result of a serendipitous discovery by Scottish scientist Alexander Fleming in 1928. Initially its development looked to fail; however, further research by Australian scientist Howard Florey and German-born, British biochemist Ernest Chain working at Oxford University, England, set this drug firmly as one of the most important in pharmaceutical history.

AS **D.6.1 Outline the historical development of penicillins.**
© IBO 2007

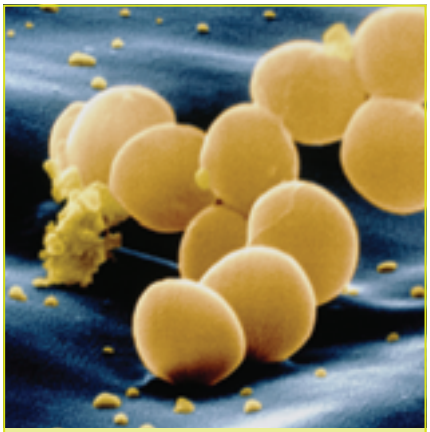


Figure 4.6.1 *Staphylococcus aureus* bacteria, (golden staph) are Gram-positive bacteria that cause skin infections.

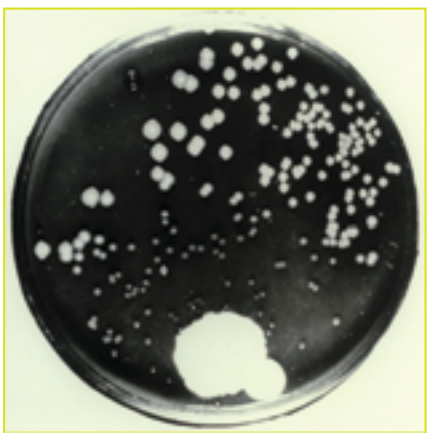


Figure 4.6.2 The original culture plate of the fungus *Penicillium notatum*, made by the Scottish scientist Sir Alexander Fleming in 1928.

TABLE 4.6.1 THE HISTORICAL DEVELOPMENT OF PENICILLINS

Date	Event
1890s	Scientists found that certain fungi killed bacteria.
1928	Alexander Fleming found that mould growing on a Petri dish stopped the bacteria he was investigating from growing. He concluded that the mould had produced a compound, which he called penicillin, that inhibited bacterial growth. He gave up the project when he found penicillin too difficult to extract.
1939	Howard Florey and Ernest Chain renewed the research and went on to discover penicillin's therapeutic action. (They treated some mice with penicillin after injecting them with deadly bacteria and the mice survived.) They also discovered its chemical composition. Chain worked out how to isolate and concentrate penicillin. He also theorized the structure of penicillin, which was confirmed by X-ray crystallography done by Dorothy Hodgkin.
1941	A patient at the Radcliff infirmary, Oxford, England, who had become seriously ill after just a scratch from a rose thorn, was given an intravenous infusion of 200 mg of penicillin. Within 24 hours, his temperature had dropped, his appetite had returned and the infection had begun to heal. However, only a small quantity of penicillin was available, and by the fifth day Florey and his colleagues had run out. The patient died. Florey and his team decided only to work on sick children who did not need such large amounts of penicillin, until their methods of production improved.
1942	John Bumstead and Orvan Hess, working in Oxford, England, became the first in the world to successfully treat a patient using penicillin.
1943	Massive development program started in US growing strains of penicillin.
1944	A large enough supply became available for everybody who needed it — saving thousands of lives in World War II.
1945	Fleming, Florey and Chain received the Noble Prize for Medicine for their work on penicillin.

THEORY OF KNOWLEDGE

In the process of scientific inquiry we often think that controlling the conditions in an experiment and focusing on a single variable is the only way to get reliable results. We seldom think about chance or serendipity as a contributing factor. But are these chance discoveries totally accidental or lucky accidents that come from thorough planning?

In 1945 three men Alexander Fleming, Ernst Chain and Howard Florey shared the Nobel Prize in Physiology or Medicine for the discovery, isolation and mass production of the antibiotic penicillin, a drug that would revolutionize medicine and lead to the development of many lifesaving antibiotics.

Fleming's accidental discovery was built on a body of knowledge of micro-organisms going as far back as 1500 BC where written records first described the use of moulds and fermented materials as therapeutic agents. It was not until the late 19th century that scientists tried to identify and isolate these substances. Williams Roberts, Louis Pasteur, Jules Francois Joubert and Joseph Lister all tried growing cultures of micro-organisms that had the potential to cure human disease, but growing pure cultures proved extremely problematic because they could not control contamination from airborne bacteria and moulds. What did become clear though was that certain moulds could inhibit bacterial growth.

The stage was set for the 'chance' discovery.
In 1928 Fleming, a bacteriologist, was researching the properties of staphylococci bacteria and, like the other scientists of the time, was frustrated with not being able to control contamination. One day while observing a culture plate he noted that the staphylococci bacteria cells immediately surrounding an invading mould had swollen and burst. He assumed that something in the mould must be inhibiting the growth of the bacteria and set about extracting a sample of the active ingredient, which he named penicillin. He published his discovery in 1929 in the *British Journal of Experimental Pathology*. However, Fleming was unable to produce a sufficient amount of pure penicillin needed for further research into its effectiveness in treating of human infections. This challenge was picked up by German Ernst Chain and Australian Howard Florey at Oxford University in England.

In 1939 Chain and Florey developed a method of isolating and purifying penicillin. The first human trial took place in 1941. The patient, a man with septicaemia (blood poisoning), responded well but later died, when the limited supply of penicillin ran out. Despite the sad ending to this initial penicillin treatment, interest in its potential lifesaving benefits soared with the onset of the Second World War. These events gave great urgency to the need to find a way to mass produce penicillin to treat ever increasing numbers of people wounded in the war. To solve this problem Chain travelled to the US and found that a process very similar to that used to brew beer could be used. This technological advancement enabled large amounts of usable penicillin to be produced, saving tens of thousands of wounded soldiers who would otherwise have died.

Regardless of whether you believe chance discoveries are totally accidental, medical science has many examples of unexpected discoveries that have changed the world. All these events however have one thing in common: each was observed and then acted upon by scientists with open, creative and curious minds. Examples include the drug Viagra to treat impotence, the sedative potassium bromide, smallpox vaccination, botulinum toxin or Botox, Librium and Valium for the treatment of mental illnesses, quinine for the treatment of malaria and insulin for the treatment of diabetes to name a few.

- Fleming was modest about his part in the development of penicillin, and he praised Florey and Chain for transforming it into a practical drug. Why do you think that Fleming is credited as the discoverer of penicillin?
- *Time* magazine named Fleming as one of the 100 most influential people of the 20th century. Is he a worthy recipient? Justify your answer.
- What did Louis Pasteur mean when he said 'Chance favours only the prepared mind.'
- In 1897, 32 years before Fleming presented his paper on the treatment of staphylococcus with penicillin, Ernest Duchesne, a 23-year-old French medical student, wrote in his doctoral thesis of the successful treatment, with the penicillin mould, of animals with bacterial infections. Duchesne died in 1912. Why do you think that Duchesne is rarely given credit for his discovery?

Bacteria may be classified as Gram-positive, with a relatively simple cell wall made up of various polymers including a small proportion of proteins and polysaccharides, or Gram-negative, with a much more complex many-layered cell wall. This classification is linked to whether the bacteria can be stained by a particular stain called Gram's stain, but their difference is better reflected in the difference in the structure of their cell wall. In particular, the outer layer of the cell wall of Gram-negative bacteria is complex and made up of polysaccharides of varying natures, depending on the bacteria, as well as of proteins. Difficulty in penetrating this complex outer layer is thought to be why some antibiotics are less active against Gram-negative bacteria than against Gram-positive bacteria.

Penicillin kills bacteria by interfering with their ability to synthesize cell walls. Crosslinks within the cell walls are not formed. Water enters the cell by osmosis, the cells burst and the bacteria die. Some forms of penicillin, such as benzylpenicillin, are only able to do this with the Gram-positive bacteria, while others, such as amoxicillin, can work on Gram-negative bacteria.

Penicillins are a member of a group of antibiotics known as beta-lactams (β-lactams). The general structure of all penicillin molecules is the same. It is pharmaceutically termed a *penicillin nucleus*.

AS D.6.2
Explain how penicillins work and discuss the effects of modifying the side-chain.
© IBO 2007

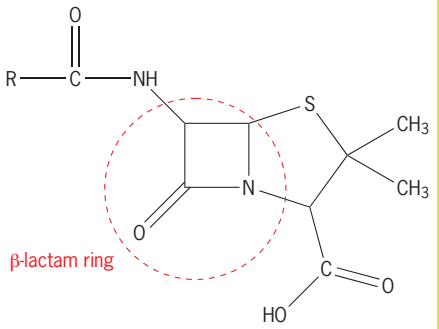


Figure 4.6.3 The basic structure of all penicillin molecules is that of the penicillin nucleus.

The structure of a penicillin affects its ability to work in different environments. The addition of different side-chains (R) changes the environment in which it is active. The first penicillin was the naturally occurring benzylpenicillin. It was found to be deactivated by stomach acid, so it had to be injected. Modification of the side chain produced an acid-resistant penicillin, penicillin V (see figure 4.6.4). Some bacteria avoid the effect of penicillin by the production of penicillinase (also known as β-lactamase), an enzyme that destroys penicillin. Further modification of the side-chain produced a penicillin that was resistant to penicillinase. Amoxicillin is another example of a broad-spectrum penicillin that has a different side-chain and is effective against both Gram-negative bacteria and Gram-positive bacteria. To combat the ability of penicillinase to destroy the structure of penicillins, clavulanic acid is included in the pharmaceutical mixture, because it inhibits many penicillinase enzymes.

Some people are allergic to some or all forms of penicillin. Skin rashes and fever are common. In serious cases anaphylactic shock, which may even lead to death, can occur. This is the reason why any known allergy to penicillins must always be listed when you are filling out medical forms.

When given orally, penicillins, particularly the broad-spectrum type, can alter the bacterial flora in the gut. These bacteria are needed for effective digestion ('good' bacteria) and the imbalance can cause problems such as gastrointestinal disturbances and thrush (a common infection caused by an overgrowth of yeasts in the vagina). In addition, the destroyed bacteria could be replaced by harmful bacteria and further infection could occur.

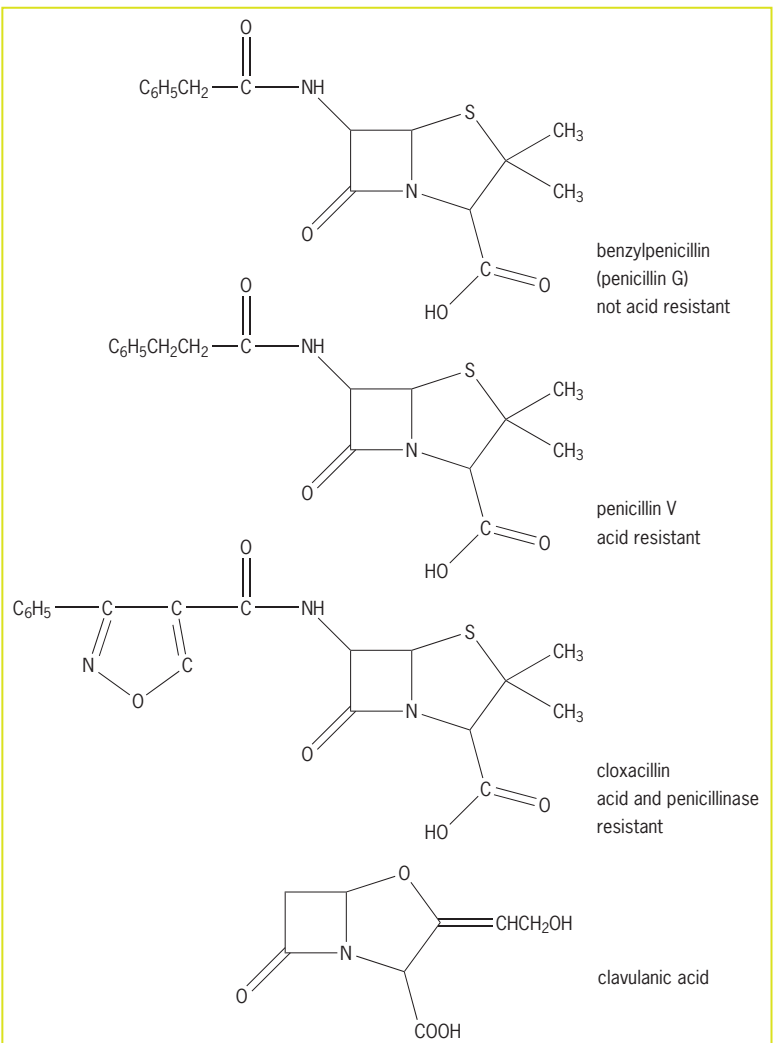


Figure 4.6.4 Structures of three penicillins with different side-chains and also of clavulanic acid, which makes penicillins resistant to penicillinase enzymes (β-lactamase).

Resistance to antibacterials

AS D.6.3
Discuss and explain the importance of patient compliance and the effect of penicillin overprescription.

When a course of antibiotics is prescribed to address an infection, some bacteria, such as those that can produce penicillinase, may survive and pass on their genetic resistance to new generations, thus producing antibiotic-resistant strains of bacteria. Other resistant bacteria may have altered cell walls to which penicillin cannot bind. To prevent this happening you should always finish a course of antibiotics (to kill all the bacteria) and antibiotics should only be used when actually needed. Too often patients go to the doctor with a viral condition such as a 'cold' and expect to be given antibiotics. Doctors who want to please their patients may be tempted to give the unnecessary drugs, and this overprescription can lead to increased opportunity for bacteria to become resistant to penicillins.

Pr
PRAC 4.4
Bacteria and antibiotics

The technique of including antibiotics in animal feedstocks for the control of animal diseases can also result in bacteria becoming resistant to the antibiotic, and so staying in the animal product when it is processed. This may result in humans being exposed to drug-resistant salmonella, *E. coli* etc. In 1983, for example, 18 people in four mid-western states of the USA developed multidrug-resistant salmonella food poisoning after eating beef from cows that had been fed antibiotics. Eleven of the people were hospitalized, and one died.

There are some bacteria that can resist almost all the drugs that we have available. These bacteria are often found in hospitals. The antibiotic vancomycin is literally the drug of 'last resort' for many infections. It is the one used when all else has failed. Some hospital-acquired staphylococcus infections are resistant to all antibiotics except vancomycin. However, vancomycin-resistance has now been observed in another common hospital bug, enterococcus. In fact, vancomycin resistance increased from 0.5% in 1989 to 18% in 1997. It is expected that this resistance will be passed on to the more dangerous *Staphylococcus aureus* bacteria, and we will not have a drug with which to combat it.

To cope with new strains of resistant bacteria, new antibiotics need to be developed. These 'super drugs' are needed to combat the 'super bugs'! While new antibiotic agents are constantly being developed, clues to the identification of additional drugs with new mechanisms for killing the bacteria may come from the genome sequencing (determining the exact genetic makeup) of drug-resistant bacterial strains such as methicillin-resistant *Staphylococcus aureus* (MRSA). Bacterial genes encoding for proteins that might be targets for antibiotic drugs have been identified and give a focus for future research.

CHEM COMPLEMENT

The importance of taking all of your medicine

Tuberculosis (TB) is a disease that in the 1980s we seemed to have beaten. Unfortunately the disease has made resurgence and one in seven cases is the result of a new form of TB that is resistant to the two drugs most commonly used to treat it (isoniazid and rifampin). As a result 5 per cent of these patients die. A study by Stephen Weis and colleagues at the University of North Texas Health Science Center in Fort Worth, USA, was reported in the 28 April 1994, *New England Journal of Medicine*. This study involved the close observance of 581 patients from 1986 to 1992 with nurses supervising them taking their pills. By the end of the study, the relapse rate (which reflects antibiotic resistance) fell from 20.9 to 5.5 per cent. Given that other risk factors for the spread of TB (such as AIDS, intravenous drug use and homelessness) were increasing, this result was particularly important. This study showed that drug resistance can be slowed if patients take medications correctly.

CHEM COMPLEMENT

Crocodile medicine

A major problem facing seriously ill patients in hospitals is the increasing occurrence of bacteria that are resistant to many or all of the forms of penicillin that we have available. The most dangerous of these bacteria is methicillin-resistant *Staphylococcus aureus* (MRSA, also known as multiple resistant *S. aureus*), a strain of the 'golden staph' bacteria that is commonly found in hospitals.

Since 1998 research has been conducted on the immune system of crocodiles, after the discovery of antibodies in the reptile's blood that have been found to kill these penicillin-resistant bacteria. The immune system of crocodiles is necessarily powerful, given their tendency for savage fights with other crocodiles, resulting in gaping wounds and missing limbs. Despite such terrible wounds and their bacteria-filled environment, the crocodiles were found to heal rapidly and almost always without infection.

In addition to its ability to kill bacteria, the serum from crocodile blood also appears to have a greater effect on HIV than human blood serum. The antibodies (a polypeptide) from samples of crocodile blood have been isolated—the obtaining of which is a dangerous task in itself—and are now known as crocodillin. In addition, similar work was undertaken with Indonesia's Komodo dragons, which use a vast array of bacteria in their saliva to kill their prey by a single bite. While it appears that little progress has been made in this area since 2005, we can hope that the potential to produce an antibiotic based on the antibodies of these dangerous reptiles still remains.

Figure 4.6.5 The Australian saltwater crocodile is a fierce creature from which to obtain a blood sample!

Section 4.6 Exercises

1 **a** State the type of micro-organism that is killed by penicillins.

b Describe the method by which penicillins kill this type of micro-organism.

2 Explain why the side-chain of the first penicillin, benzylpenicillin, needed to be modified.

3 In addition to the modification of side-chains for reasons such as that as described in your answer to question **2**, scientists need to continue developing new penicillins for another reason. State that reason.

4 When penicillins are prescribed unnecessarily, there are long-term harmful effects in our community. Describe these effects of overprescription of penicillin.

5 If a patient does not finish a course of antibiotics because he feels better, there is a consequence in the future effectiveness of that antibiotic. Describe this consequence and explain why this occurs.

6 Describe another practice that can result in similar effects to the overprescription of penicillin.

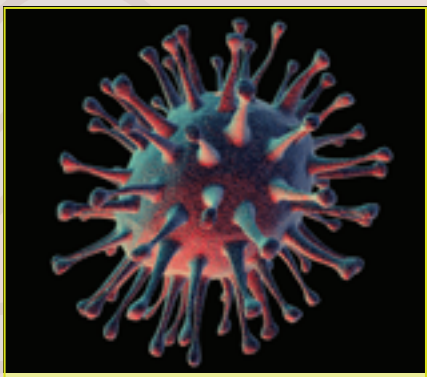


Figure 4.7.1 The protein coat of the avian flu virus, which encloses the genetic material of the virus, has surface proteins that are seen here as spikes.

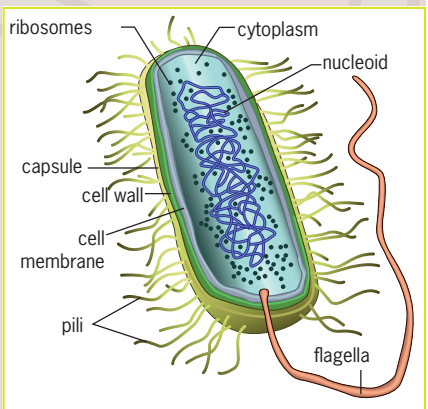


Figure 4.7.2 A simple diagram of a bacterium cell.

AS D.7.1 State how viruses differ from bacteria. © IBO 2007

4.7 ANTIVIRALS

Bacteria are about 20 times the size of viruses. While bacteria are unicellular, viruses are non-cellular in structure. They are tiny particles (**virions**) composed largely of nucleic acid with a protein coating. Bacteria can reproduce on their own (and do so very rapidly), but viruses need to be inside another living cell in order to reproduce. Once inside a host cell (i.e. one that the virus has invaded), the virus takes over the reproductive mechanism (the ribosomes) of the cell and that cell then produces viral proteins and nucleic acids. The host cell is not necessarily destroyed by the virus, but it will be adversely affected, hence the detrimental effects of viral diseases.

TABLE 4.7.1 THE MAIN DIFFERENCES BETWEEN BACTERIA AND VIRUSES

Bacteria	Viruses
Living	Non-living
Single-celled micro-organisms	Non-cellular and sub-microscopic (much smaller)
Bacteria contain: • a single chromosome in the 'nucleoid' (circular strand of DNA) • a rigid cell wall made of polysaccharide molecules • cytoplasm • a cell membrane • ribosomes and enzymes to break down food and build cell parts	Viruses contain: • a central core of DNA or RNA surrounded by a protein coat • no nucleus • no cytoplasm • no cell membrane or cell wall • no ribosomes • the enzymes needed to invade a cell and replicate their nucleic acids
Bacteria: • feed • excrete • reproduce asexually • grow	Viruses: • do no feed • do not excrete • use the ribosomes of the cell they have invaded to reproduce • do not grow

Viruses are controlled most effectively by vaccinations. When a vaccination is given, the material administered can either be a live, but weakened, form of the virus, an inactivated form of the virus, or a purified material such as a protein from the virus. The vaccine stimulates the creation of antibodies that recognize and attack a particular infection. The presence of these antibodies means that when the virus is next encountered, the body already has the antibodies to fight it. The smallpox virus was eradicated across the world by the use of inoculations, and viral diseases such as polio are no longer common. In developed countries, children are routinely vaccinated against diseases such as measles, rubella, and tetanus, whereas in developing countries not all children are as fortunate and these viral diseases are still a threat.

AS D.7.2
Describe the different ways in which antiviral drugs work. © IBO 2007

We have seen that a virus is dependent on the reproductive mechanism (the ribosomes) of the cell that it has invaded in order to carry out its own reproduction. One method by which antiviral drugs can work is to chemically block the use of ribosomes and hence the reproductive enzymes by the virus. This would prevent replication of the virus in host cells. An example of such an antiviral drug is interferon alpha, which is used to treat hepatitis B and C.

Antiviral drugs often mimic one of the nucleosides (deoxyribose sugar bonded to an organic base) from which the DNA of the virus is made. The virus incorporates this fake nucleoside into its genome during replication, and phosphorylation to form a nucleotide continues. The fake nucleotide inhibits the viral enzyme (DNA polymerase) that is catalysing the polymerization reactions and forming the new DNA chain. The newly synthesized DNA becomes inactive and the life cycle of the virus is halted.

An example of an antiviral drug that works in this way is **acyclovir** (also known as aciclovir), which is used to fight herpes virus infections. Acyclovir is one of the oldest and most frequently prescribed antiviral drugs. It is a derivative of the guanosine nucleoside, so is used by the virus when guanosine would have been used. Its structure differs from that of guanosine in that the sugar ring is replaced by an open-chain structure.

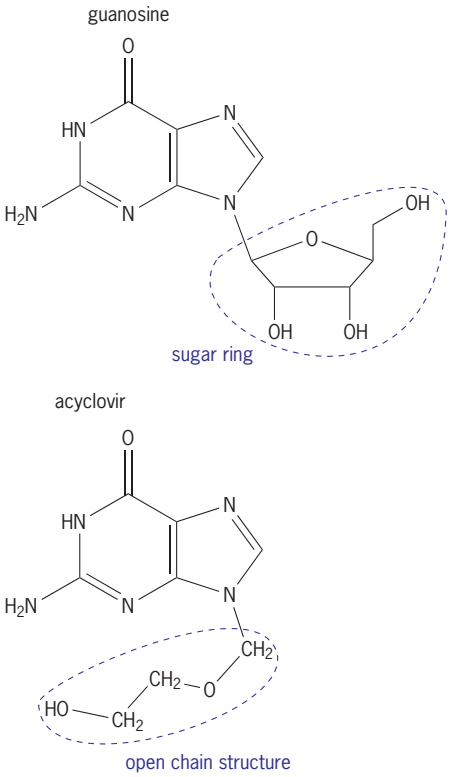


Figure 4.7.3 The structures of guanosine and acyclovir. Note the replacement of the sugar ring in guanosine with an open-chain structure.

Acyclovir inhibits the viral DNA polymerase, terminating the nucleotide chain. In cells that have not been invaded by the virus, acyclovir is not as active. The virus-specific enzyme that carries out phosphorylation of the nucleoside to a nucleotide is much more effective in this process than the host cell's enzyme. Additionally, the acyclovir triphosphate nucleotide is 30 times more potent an inhibitor of the action of the herpes virus enzyme than of the host enzyme.

Acyclovir is used not only for the treatment of *Herpes simplex* infections (cold sores, genital herpes) but also for the treatment of shingles and chickenpox (in patients with a compromised immune system). It can also be given as a **prophylactic** treatment for patients who suffer from frequent recurrences of genital herpes infection. Zovirax®, a commercial preparation for cold sores, contains acyclovir.

Antiviral drugs may also inhibit the attachment to, or the penetration of, the host cell by the virus. Antivirals for use against the influenza virus work in this way.

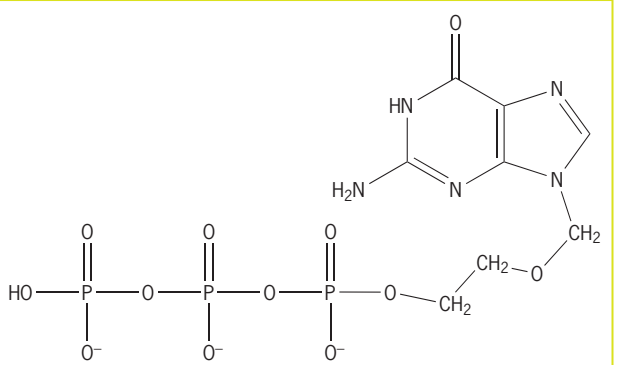
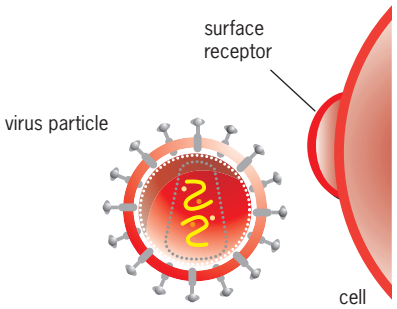
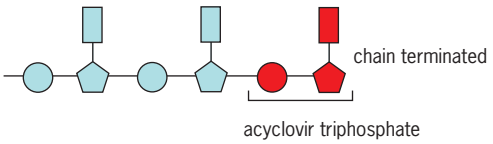
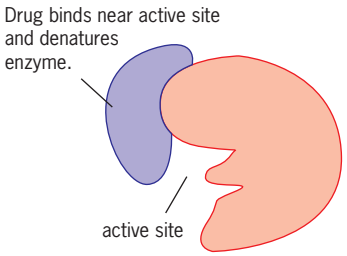
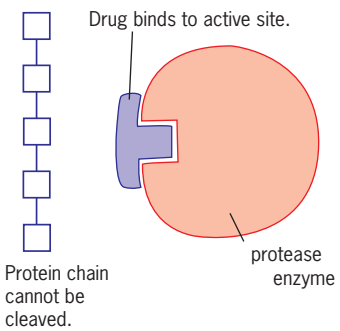


Figure 4.7.4 Acyclovir triphosphate inhibits the action of the DNA polymerase enzyme of the herpes virus.

TABLE 4.7.2 A RANGE OF WAYS IN WHICH ANTIVIRAL DRUGS PREVENT VIRUS REPRODUCTION

Method of action	Diagram	Example
It inhibits attachment to or penetration of the host cell by the virus		Anti-influenza drugs such as amantadine
It is phosphorylated and competes with the host's triphosphates in the formation of the nucleotide chain. The drug triphosphate terminates the chain.		• Acyclovir (herpes drug) • Nucleoside reverse transcriptase inhibitors (see p. 241)
Binds near the active site of the enzyme and denatures it.		Non-nucleoside reverse transcriptase inhibitors (see p. 241), e.g. neverapine, which can prevent the transmission of HIV from mother to a newborn baby
Inhibits virus-specific protease, thus preventing cleavage of translated protein into structural proteins.		Protease inhibitors, e.g. saquinavir (see p. 241)

AIDS

Infection with **human immunodeficiency virus (HIV)** results in acquired immunodeficiency syndrome (AIDS). In 2006 it was estimated by the World Health Organization that about 25 million people had died of AIDS since it was first recognized in 1981, 39.5 million were HIV positive and 4.3 million people had become infected during 2006. In 2006 alone, AIDS claimed an estimated 2.9 million lives, of which 380 000 were children. The epidemic is centred in sub-Saharan Africa, where up to 35% of the adult population is infected with HIV.

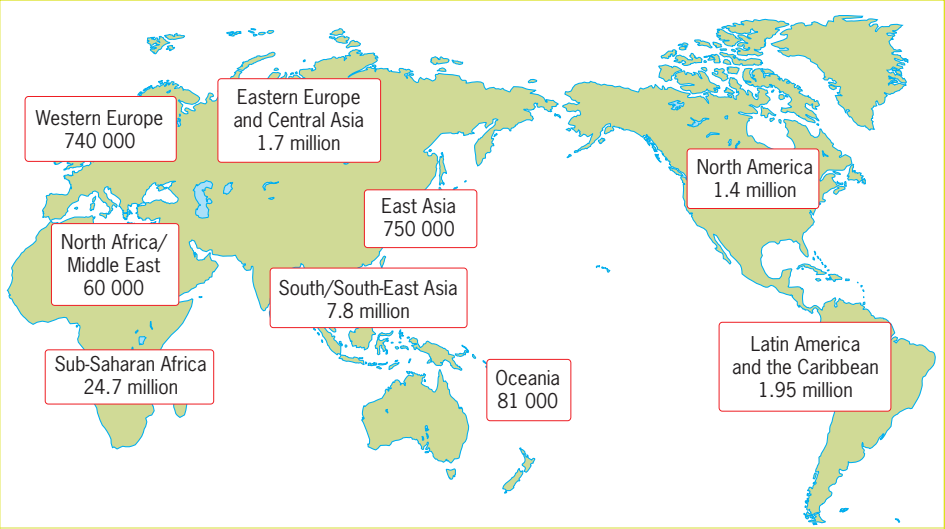


Figure 4.7.5 This map shows the spread of HIV/AIDS infections across the world at the end of 2006 (data from WHO 2006 UNAIDS report).

AS **D.7.3**
Discuss the difficulties associated with solving the AIDS problem. © IBO 2007

HIV is an RNA **retrovirus**. It has an RNA genome and replicates via a DNA intermediate. It contains an enzyme called reverse transcriptase. This enzyme makes a DNA copy of the viral RNA. This viral DNA is then integrated into the genome of the host cell and directs the generation of new virus particles. The interaction of HIV with the host's immune system is complex. It uses the surface proteins of the immune system cells (white blood cells known as T cells) to gain entry to them. These surface proteins are like name badges. The virus is able to recognize the name badges (surface proteins) on some T cells and infects the cells displaying those name badges. The infection that follows this invasion prevents the T cells from signalling to other cells in the immune system to perform their function. HIV progressively destroys the body's ability to fight infections, leading to life-threatening infections such as pneumonia.

Within the cell, HIV is integrated into the host DNA, undergoes transcription (the synthesis of RNA under the direction of DNA) and new virions (virus particles) are made. HIV can mutate, and, as its metabolism is similar to that of the human cell, this makes effective treatment with antiviral drugs and vaccine development very difficult.

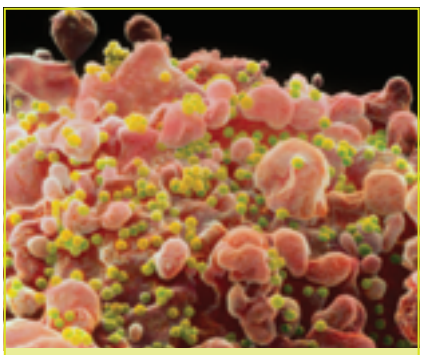


Figure 4.7.6 A coloured scanning electron micrograph of the surface of a T-cell infected with HIV. The small spherical green particles visible on the surface are HIV virus particles in the process of budding away from the cell membrane.

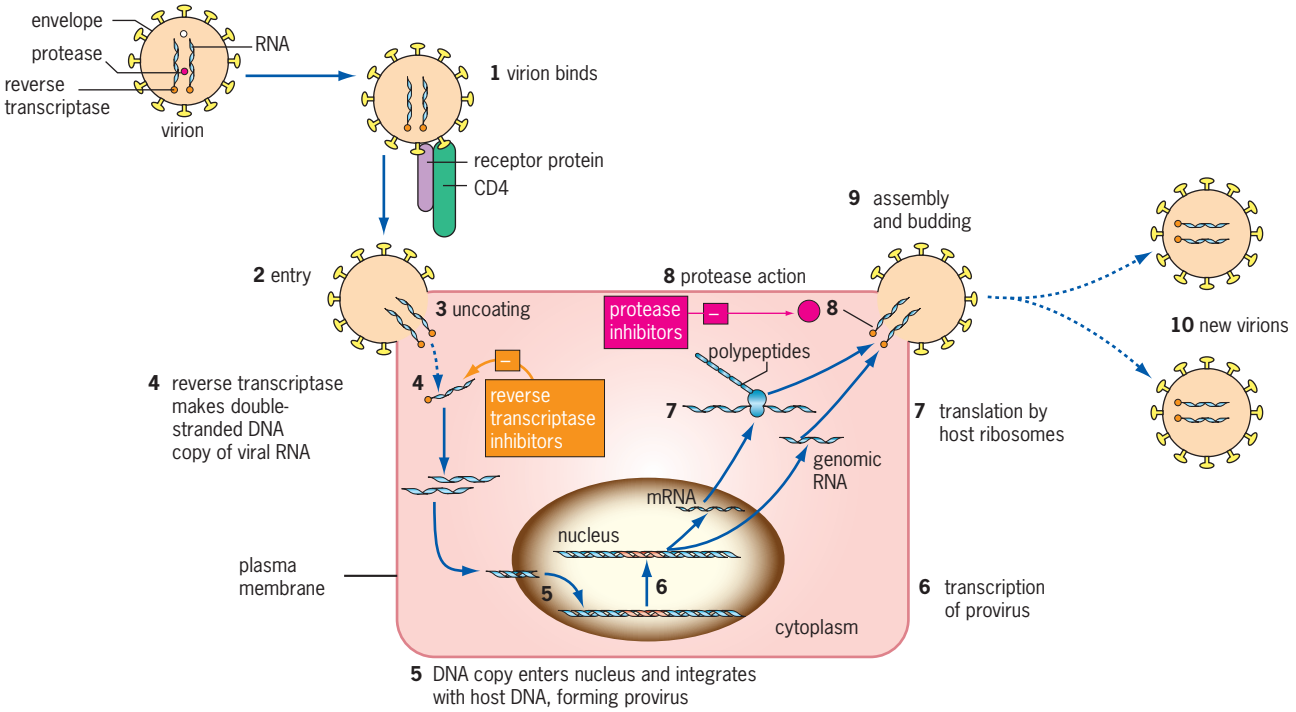


Figure 4.7.7 A schematic diagram of the infection of a T-cell by an HIV virion, showing the sites of action of the two main classes of drug.

A feature of all viruses is that they have a very high reproductive rate and undergo frequent mutations. This allows them to adapt to changing environments quite well. This is important for their reproductive cycle. Because the only way it can reproduce itself is by infecting a cell, a virus must be able to evolve faster than the host cells. If it did not, then the host cells would adapt or evolve to a point where the virus would no longer be able to infect them. The method of adaptation is related to the surface receptors of the cells. In response to the changing of the cell's surface receptors to which the virus needs to attach, the virus changes its surface proteins.

So, while a virus could be treated with an antiviral in one case with success, the same virus may mutate within a population and become immune to the same antiviral in another patient. HIV is particularly adept at mutation, making it very difficult to treat. As an RNA virus, HIV mutates into a slightly different virus and overcomes the animal's 'immune memory'. The surface proteins of HIV are changed so the antibodies produced by the cell no longer attach. HIV mutates so fast that the immune system never clears it from the body and every vaccine that has been developed has failed to prevent infection from this ever-mutating virus. It is like a 'new' virus constantly appearing. RNA retroviruses like HIV have the highest mutation rates ever measured.

In addition, the invasion of the host cell and subsequent use of its reproductive mechanism by the virus creates close links between the metabolism of the virus and that of the host cell. This means that drugs that harm the virus will also harm the host cell, making it difficult to target HIV without damaging the host cell.

There are two main classes of anti-HIV drugs: **reverse transcriptase inhibitors** and **protease inhibitors**. Combinations of these drugs are used in the treatment of HIV–AIDS. The combination treatment is known as HAART (highly active antiretroviral therapy). The effect of HAART is to inhibit HIV replication, so that levels of HIV RNA in blood plasma are reduced to undetectable levels and patient survival is prolonged. HIV is not eradicated and so this treatment, which has many unwanted side-effects, must be kept up for life. These antiretroviral drugs are extremely expensive and may be priced beyond reach of many.

The reverse transcriptase inhibitors can be further divided into two types of drugs: nucleoside based or non-nucleoside based.

- Nucleoside reverse transcriptase inhibitors are phosphorylated and compete with the host cellular triphosphate nucleosides for the formation of proviral DNA using the viral reverse transcriptase enzyme. When this drug version of the nucleoside triphosphate is incorporated into the growing proviral DNA chain, it results in chain termination.
- Non-nucleoside reverse transcriptase inhibitors bind to the reverse transcriptase enzyme of the virus near its active site and denature it. The enzyme is no longer able to catalyse the formation of proviral DNA.

WORKSHEET 4.3
Viruses and antivirals

Other anti-HIV drugs use proteases inhibitors. These act on the enzyme that cleaves polypeptides made in the host cell by the proviral DNA at appropriate positions, to create structural proteins for the new virions.

The mechanism of action of new anti-HIV drugs may involve the modification of the surface protein by which HIV recognizes the cell; as a result HIV cannot bind to cells and enter them. Other possibilities include prevention of the uncoating of the virus after it enters the host cell, prevention of the viral RNA entering the nucleus and so integrating with the host DNA, or prevention of the transcription of viral mRNA from the provirus.

The treatment of HIV is difficult and very expensive, so the need to control the spread of AIDS becomes paramount. The control of AIDS is linked to limiting the transmission of HIV during unprotected sexual intercourse, injection of intravenous drugs using contaminated needles, and from an infected mother to her baby through breast milk and during the baby's birth. Sociocultural issues, such as the lack of education regarding use of condoms, lack of access to condoms and to clean needles for intravenous drug users, and religious reasons for rejecting condoms due to their contraceptive use, can make this limitation very difficult.

Section 4.7 Exercises

1 Draw an annotated diagram that compares the physical structures of bacteria and viruses.

2 Explain why a virus is unable to reproduce on its own.

3 **a** Describe how a virus uses another living cell (a host cell) to reproduce itself.

b Explain how the method by which bacteria multiply differs from your answer to part **a**.

4 Describe two general ways in which antiviral drugs work.

5 Describe the particular way in which the drug acyclovir fights the herpes virus.

6 Explain how HIV identifies the cells that it invades.

7 Explain why HIV is called a retrovirus.

8 Considering the mechanism of action of HIV, economic and sociocultural issues, describe why it is particularly difficult to solve the AIDS problem.

4.8 DRUG ACTION

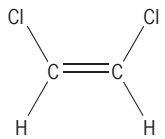
HL

AS D.8.1 Describe the importance of geometrical isomerism in drug action. © IBO 2007

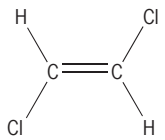
We have seen in many examples so far that the structure of a drug is intrinsically linked to its action. In some cases, a change as simple as the geometric arrangement around a double bond can make the difference between a drug performing its intended function or not.

Geometric isomerism

Stereoisomers have the same molecular formula and the same structural formula, but different arrangements of atoms in space. The presence of a double bond stops rotation around that bond, so the arrangement of the atoms on each carbon can differ. The two forms of the stereoisomer are named according to the arrangement of the functional groups across the double bond (their geometry). Hence these isomers are called **geometric isomers**. The *trans* (meaning 'across') isomer has the two functional groups on opposite sides of the double bond and bonded to the two different carbon atoms. The *cis* isomer has the two functional groups still bonded to the two different carbon atoms, but on the same side of the double bond.



cis-1,2-dichloroethane

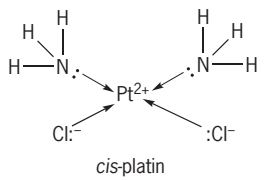


trans-1,2-dichloroethane

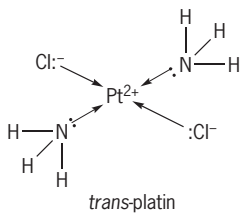
Figure 4.8.1 There are two possible stereoisomers of 1,2-dichloroethene, these are denoted *cis* and *trans*.

Geometric isomers have different physical properties. Their melting and boiling temperatures differ. Usually the *trans* isomer has the higher melting point and the *cis* isomer has the higher boiling point. The difference between the two boiling points is because the *cis* isomer is a polar molecule whereas the *trans* isomer is non-polar, due to its symmetry. In the case of melting points, the ability of the molecules to pack closely together is an even greater influence than polarity. *Trans* isomers pack together more efficiently than *cis* isomers. The poorer packing in the *cis* isomers means that the intermolecular forces aren't as effective as they should be and so less energy is needed to melt the molecule, resulting in a lower melting point.

Geometric isomers can also undergo different chemical reactions. While some chemical properties will be the same because they have the same molecular formula, one way in which they may differ is in their pharmacological effects.



cis-platin



trans-platin

Figure 4.8.2 The *cis* and *trans* isomers *cis*-platin and *trans*-platin. *Cis*-platin is an anti-cancer drug while *trans*-platin is not.

In addition to double-bonded organic compounds, inorganic complexes (see *Chemistry: For use with the IB Diploma Programme Higher Level*, chapter 3) can also form geometric isomers. An example of an inorganic complex with pharmacological effects is the anti-cancer drug cisplatin (*cis*-diamminedichloroplatinium(II)). It is the water-soluble *cis* isomer of the inorganic complex $Pt(NH_3)_2Cl_2$.

The *cis* and *trans* isomers are possible, as there is a square planar arrangement (rather than tetrahedral) around the Pt ion. The *cis*-platin molecule binds to DNA in cells by forming cross-links from one part of the DNA chain to another. The two chlorides on *cis*-platin are replaced by nitrogen atoms of the guanine organic base, a part of the DNA structure. This creates a bend in the DNA chain and prevents replication of the damaged DNA of the cancer cell. The damaged

cell is then destroyed by the body's immune system. Normal cells are affected in the same way by *cis*-platin, but are better able to repair the damage and resume normal functioning. The *trans* isomer would not be able to create such a bend in the DNA chain, so does not have the same effect.

Interactive CW Animation
Isomerism
Geometric isomerism

Chirality

Another form of stereoisomerism is optical isomerism. In *Chemistry: For use with the IB Diploma Programme Higher Level*, chapter 9, **optical isomers** were described as molecules with the same molecular formula, but with structures that are unable to be superimposed on one another. Such molecules are also described as **chiral**. The two different forms of an optical isomer are called **enantiomers**. Molecules in which a carbon atom is bonded to four different substituents will be chiral.

AS **D.8.2 Discuss the importance of chirality in drug action.**

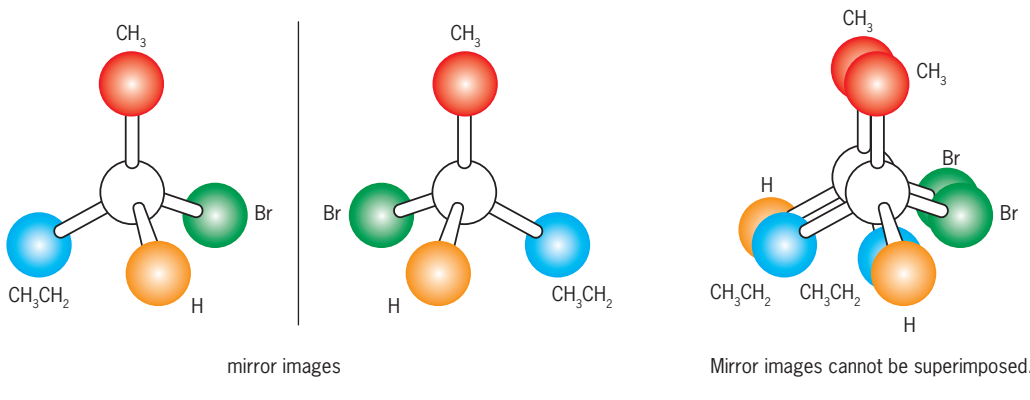


Figure 4.8.3 Optical isomers are mirror images of one another and cannot be superimposed.

A solution of an optical isomer is able to rotate the plane of plane-polarized light to either the left or to the right. The isomers are named according to the direction in which the light is rotated. An enantiomer that rotates plane-polarized light in a clockwise direction (to the right, dextro) is denoted as the + enantiomer; the one that rotates the light in an anticlockwise direction (to the left, levo) is denoted as the – enantiomer. This naming allows us to differentiate between the two different isomers that often behave quite differently in a biological or pharmacological context.

While naturally occurring drugs usually occur as a single enantiomer, optically active substances synthesized in the laboratory often occur as a mixture of the two enantiomers. This is known as a **racemic mixture**, and it has no effect on plane-polarized light. Often in a racemic mixture only one enantiomer has the required pharmacological properties, the other may be totally inactive. It is acceptable for a racemic mixture to be marketed as long as the other enantiomer does not have a detrimental effect on the consumer.

A well-known example of a racemic mixture is the tragedy of the drug thalidomide, which was developed by German pharmaceutical company Grünenthal and was sold from 1957 to 1961 in almost 50 countries under at least 40 names, including Distaval, Talimol, Nibrol, Sedimide, Quietoplex, Contergan, Neurosedyn, and Softenon. Thalidomide was chiefly sold and prescribed to pregnant women, to combat morning sickness and as an aid to help them sleep. Disastrously, one enantiomer alleviated morning sickness, but the other enantiomer caused deformities in the limbs of foetuses and other

birth defects. The drug was sold as a racemic mixture, so both effects were experienced when only the alleviation of morning sickness was sought. Between 1958 and 1962 the births of 10 000 malformed babies were attributed to thalidomide. Today, thalidomide is being used again, but this time as a drug to treat a skin condition associated with leprosy, as well as for an incurable bone marrow cancer called multiple myeloma. The use of thalidomide is monitored carefully to ensure that the user does not become pregnant.

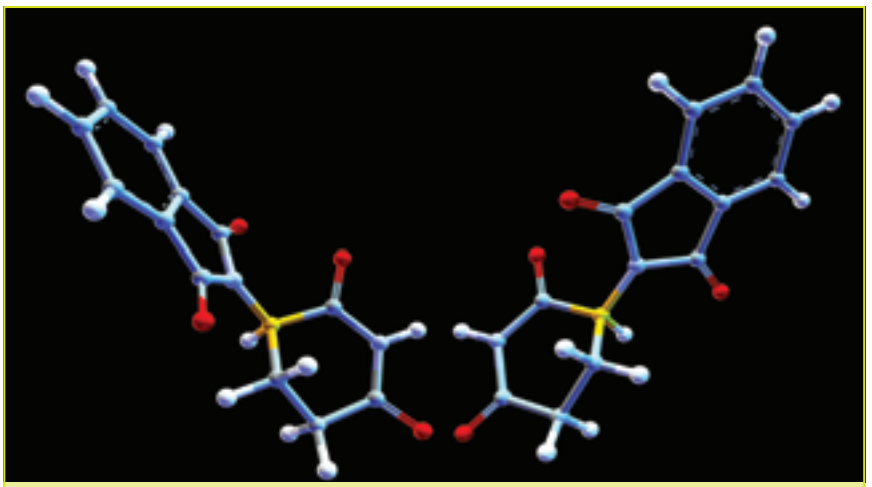


Figure 4.8.4 The drug thalidomide when synthesized in the laboratory is a racemic mixture of two enantiomers, one of which causes birth defects.

Another example of a pair of enantiomers in which one has biological action whereas the other does not is the amino acid DOPA (dihydroxyphenylalanine). Levodopa or L-DOPA (3,4-dihydroxy-L-phenylalanine) is an intermediate in dopamine biosynthesis. In clinical use, levodopa is given to patients to help manage Parkinson's disease. Dextrodopa has no physiological effects on the body.

AS D.8.3
Explain the importance of the beta-lactam ring action of penicillin. © IBO 2007

Another example of the importance of structure in drugs is that of penicillin. Part of its structure is the β-lactam ring (see figure 4.8.5). The strain of the four-membered ring makes the amide group highly reactive. When penicillin comes in contact with bacteria, the ring opens so that the penicillin becomes covalently bonded to the enzyme that synthesizes bacterial cell walls (DD-transpeptidase), thus blocking its action. The cell wall of the bacterium is weakened and the cell eventually bursts.

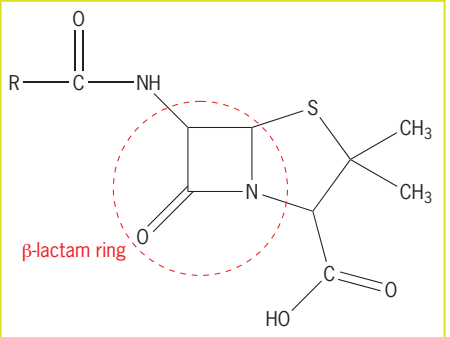


Figure 4.8.5 The β-lactam ring of penicillin is responsible for its action on bacterial cell walls.

AS D.8.4
Explain the increased potency of diamorphine (heroin) compared to morphine. © IBO 2007

In section 4.3 we discussed the differences and similarities in the structure of morphine and diamorphine (heroin). As an analgesic, heroin is much more potent than morphine, and much more addictive. The difference in structure between morphine and heroin is that the polar hydroxyl groups in morphine are replaced by non-polar ester groups in heroin, facilitating its transport into the non-polar environment of the central nervous system.

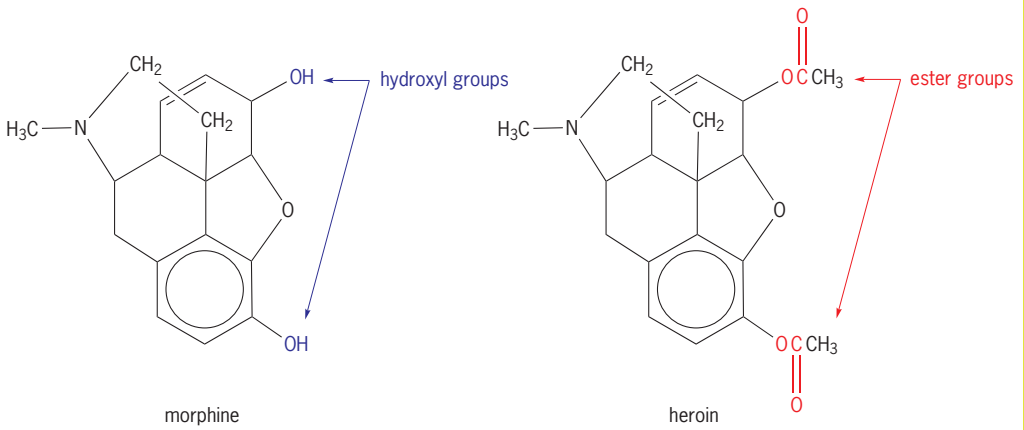


Figure 4.8.6 The structure of morphine and heroin differ in the replacement of hydroxyl groups with ester groups.

Section 4.8 Exercises

1 Explain, using diagrams, the difference between the *cis* and the *trans* geometric isomers of 1,2-dibromopropene.

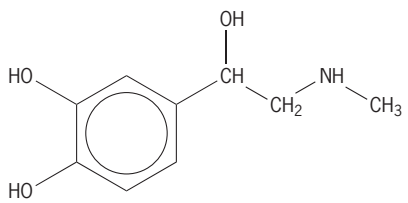
2 Explain, using diagrams, why 1,2-dichloroethene shows *cis* and *trans* isomerism, but the tetrahedral compound dichloromethane does not.

3 Explain why the *cis* isomer of the inorganic complex diamminedichloroplatinium(II) is able to act as an anti-cancer drug, while the *trans* isomer shows no such action.

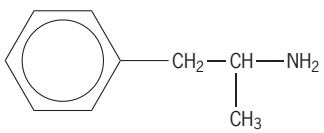
4 Describe, giving an example, what is meant by the term *chirality* when referring to a compound.

5 Identify the chiral centre in each of the structures below and indicate it using an asterisk.

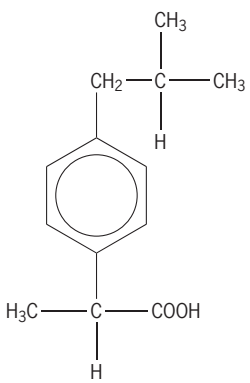
a Adrenaline



b Amphetamine



c Ibuprofen



6 Explain how enantiomers differ from one another.

7 **a** State the number of atoms that are in the β-lactam ring of penicillin.

b Explain why the β-lactam ring is important in the antibacterial action of penicillin.

8 Using diagrams to illustrate your answer, explain why heroin (diamorphine) has a more potent analgesic effect than morphine.

4.9 DRUG DESIGN

HL

Pharmaceutical chemists do not synthesize drugs at random. There are particular features in the function of the drug they are seeking that are linked to the structure. Similarities in structure between a drug that is effective and others that may be effective in treating a disease or condition (perhaps with even better results) are most important in determining the path to follow in order to produce a new drug.

AS **D.9.1 Discuss the use of a compound library in drug design. © IBO 2007**

A compound library is a collection of stored chemicals. It is used for the screening of molecules for a particular use, such as biological activity in a required field (e.g. anti-cancer drugs). Each chemical in a compound library has associated information stored in an electronic database, with information such as the chemical structure, purity, quantity and chemical characteristics of the compound. The more detailed this information is, the more useful the compound library will be.

In seeking a molecule that will perform a particular function, a pharmaceutical chemist will perform a search of the compound library database. Any molecules that appear to have the required traits can then be tested in the laboratory. Traditionally, a large collection of related compounds would have been synthesized and individually evaluated for biological properties. This approach is time-consuming and expensive. The use of the compound library cuts the time wasted on testing molecules that will not meet the requirements and thus saves a lot of money.

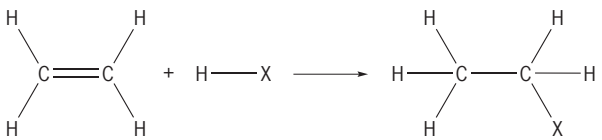
AS **D.9.2 Explain the use of combinatorial and parallel chemistry to synthesize new drugs. © IBO 2007**

When new drugs are being synthesized, the syntheses are based on the initial knowledge of a drug that shows biological activity. This compound, the **lead compound**, is systematically changed to produce more effective drugs. Perhaps the new drugs will have fewer side-effects, or be cheaper to make or easier to take. Such a mechanized mass production of many variations on a drug is known as **combinatorial chemistry**. A set of starting materials is systematically reacted in all possible combinations, with robotics being used to carry out these identical chemical reactions. Computer-controlled syringes carry out the repetitive parts of the synthesis (e.g. adding chemicals) in exactly the same way for all the possible combinations. The two most common methods of combinatorial chemistry are called **parallel synthesis** and **solid-phase synthesis** (or 'mix and split'). In parallel synthesis, the reactions are carried out as many variations on a theme. The products of these reactions (a compound library) may be tested for their potential biological activities in the laboratory (before animal tests).

Example 1 Parallel synthesis

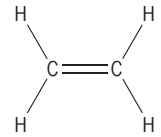
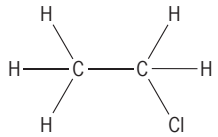
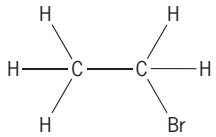
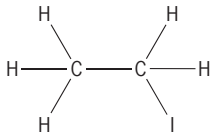
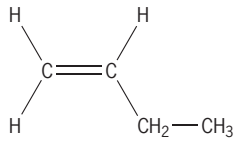
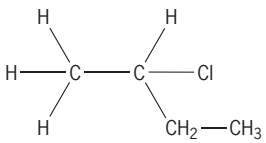
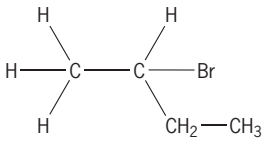
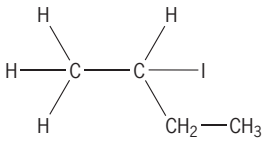
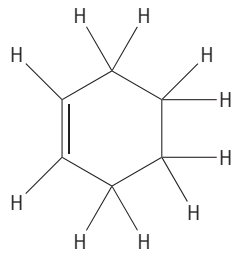
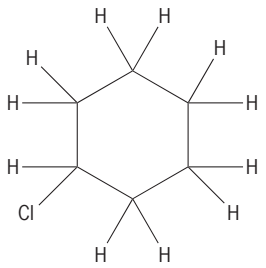
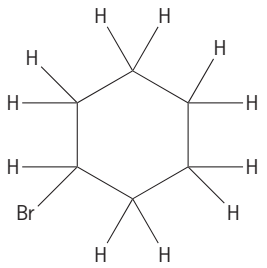
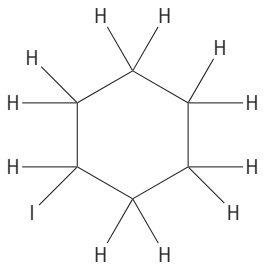
A simple example of how parallel synthesis works can be seen with the reaction between alkenes and hydrogen halides.

Alkenes undergo an addition reaction with hydrogen halides to make halogenoalkanes.



Parallel synthesis could be used to react three different hydrogen halides with three different alkenes.

TABLE 4.9.1 THE COMBINATIONS OF PRODUCTS POSSIBLE BETWEEN THE NINE COMBINATIONS OF REACTANTS

Alkene	Hydrogen chloride, HCl	Hydrogen bromide, HBr	Hydrogen iodide, HI
Ethene 	Chloroethane 	Bromoethane 	Iodoethane 
But-1-ene 	2-Chlorobutane 	2-Bromobutane 	2-Iodobutane 
Cyclohexene 	Chlorocyclohexane 	Bromocyclohexane 	Iodocyclohexane 

Parallel synthesis allows many more combinations of reactants to be reacted than would otherwise be possible. The tedium of performing the simple parts of these preparations is relieved by the use of robotics.

The other method of combinatorial chemistry is called solid-phase synthesis (or 'mix and split'). This involves the use of insoluble, but porous, small solid beads, usually made of polystyrene. This method was developed for the synthesis of polypeptides, but is used these days to make a large number of polymer compounds. The beads are treated with functional units ('linkers') on which peptide chains can be built. The peptide will remain covalently attached to the bead until cleaved from it by a reagent such as trifluoroacetic acid. This means that any unreacted reagents left at the end of any synthetic step can be removed by a simple wash procedure, greatly decreasing the time required for synthesis. Using an automated procedure, polypeptides can be constructed on these beads with every combination of amino acids.

The degree of complication of the mix and split process depends on how many variations are being sought.

Example 2 Solid phase synthesis

Let us use the example of preparing a tripeptide with three different amino acids.

Step 1: The polystyrene beads are split into three groups and a different amino acid (represented as a coloured shape) is attached to the beads in each group.

Step 2: The beads are all mixed together, and then separated randomly into three groups again. Each group will contain beads with each of the three amino acids attached. A second amino acid is attached to the beads and amino acids in each of these new groups.

Step 3: The beads and their two amino acids are all mixed together again. They are split into three groups again, with each group containing (due to the randomness of the procedure) examples of the nine different dipeptides that have been synthesized so far. In each of these new groups a third amino acid (different for each group) is attached to the dipeptide. In this way 27 different tripeptides have been produced.

In this way many combinations of amino acids can be produced with relative ease.

Stage	Reaction vessel 1 (adding ◆)	Reaction vessel 2 (adding ■)	Reaction vessel 3 (adding ▲)	Number of compounds created
1	Bead + ◆	Bead + ■	Bead + ▲	3
Contents of all three vessels is mixed.				
2	Bead + ◆◆ Bead + ■◆ Bead + ▲◆	Bead + ◆■ Bead + ■■ Bead + ▲■	Bead + ◆▲ Bead + ■▲ Bead + ▲▲	9
Contents of all three vessels is mixed.				
3	Bead + ◆◆◆ Bead + ■◆◆ Bead + ▲◆◆ Bead + ◆■◆ Bead + ■■◆ Bead + ▲■◆ Bead + ◆▲◆ Bead + ■▲◆ Bead + ▲▲◆	Bead + ◆◆■ Bead + ■◆■ Bead + ▲◆■ Bead + ◆■■ Bead + ■■■ Bead + ▲■■ Bead + ◆▲■ Bead + ■▲■ Bead + ▲▲■	Bead + ◆◆▲ Bead + ■◆▲ Bead + ▲◆▲ Bead + ◆■▲ Bead + ■■▲ Bead + ▲■▲ Bead + ◆▲▲ Bead + ■▲▲ Bead + ▲▲▲	27
Contents of all three vessels is mixed.				

Figure 4.9.1 In the mix and split process of combinatorial chemistry a huge number of variations can be achieved relatively easily.

AS **D.9.3**
Describe how computers are used in drug design.

While three-dimensional models of molecules used to be something we could only construct from balls and sticks, the computer age has brought a new way to visualize molecules. Computer modelling ('in-silico') has a distinct time advantage over more 'hands-on' approaches, with millions of molecules being modelled by simply running a computer program. Virtual development of new drugs can be achieved by programming the computer to systematically change the structure of a molecule that is already known to have biological activity, then the molecule can be evaluated within set parameters. For example, an inhibitor for a particular enzyme will need to have a structure that will be compatible with the active site of the enzyme, so molecules that are modelled by the computer will need a certain configuration of functional groups to meet this requirement.

CHEM COMPLEMENT

Bruce Merrifield and the synthesis of polypeptides

R. Bruce Merrifield was an American biochemist who graduated with a PhD in 1949. His subsequent work until he died was carried out at the Rockefeller Institute for Medical Research, where he developed an innovative method for synthesizing chains of amino acids in 1959. His method was based on the anchoring of the first amino acid to an insoluble solid. After that, the subsequent amino acids could be joined, one by one, to the fixed terminus. The polypeptide chain could be easily detached from the solid when complete. The process could be carried out by a machine and was highly efficient. Dr. Merrifield's method greatly stimulated progress in biochemistry, pharmacology and medicine, making possible the systematic exploration of the structural basis of the activities of enzymes, hormones and antibodies. Insulin was synthesized in his laboratory for the first time in the mid 1960s. Solid-phase peptide synthesis is now the accepted method for creating peptides and proteins in the laboratory.
Merrifield was awarded the Nobel Prize for Chemistry in 1984 for his work in solid-phase peptide synthesis.

Figure 4.9.2 Bruce Merrifield (1921–2006).

CHEM COMPLEMENT

Using computer down-time to find anti-cancer drugs

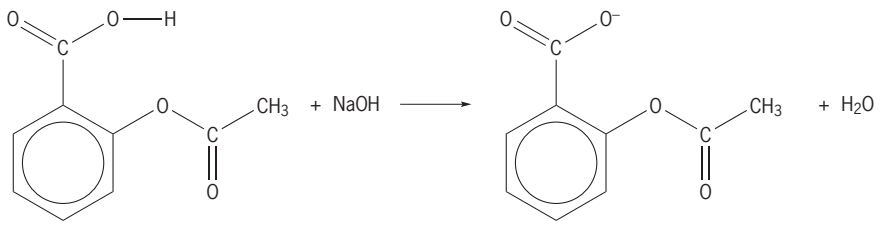
In April 2007 a six-year project at the Oxford University, England, concluded successfully. The project, which included the cooperation of more than 3.5 million computer owners over 200 different countries aimed to screen 3.5 billion different molecular models '*in-silico*'. Oxford University's Centre for Computational Drug Discovery based in the school of Chemistry offered a screen saver for the computers of volunteers which would screen 3.5 billion molecules for cancer-fighting potential while the owner had the computer turned on, but was not using it. The results have been most successful. Of the molecules that were identified as having biological activity, 10% of those already synthesized have shown themselves to be experimentally active. This is a high percentage in drug development circles. Over 450 000 years of computer time was used, making this the world's largest ever computational project.

Figure 4.9.3 Professor Graham Richards, head of the Centre for Drug Discovery at Oxford University, England, with the cancer research screen saver.

AS D.9.4
Discuss how the polarity of a molecule can be modified to increase its aqueous solubility and how this facilitates its distribution around the body. © IBO 2007

Many drugs are essentially non-polar, and so their solubility in aqueous solution is limited. But recall that most transport of solutes within the human body is in aqueous solution. Acidic groups, such as carboxyl groups, and basic groups, amines, that are part of the structure can lose or gain hydrogen ions to form negative or positive ions and then form ionic salts. The positive or negative ion used to make the salt is chosen carefully to ensure that the product is soluble in water; for example, the positive ion could be a group 1 ion and the negative ion could be a nitrate, NO_3^-, ion. Aspirin is an obvious example of this transformation. The acetylsalicylic acid molecule is not soluble in cold water. This property is actually utilized in its preparation (see Practical investigation 4.2 Preparation and analysis of aspirin).

Aspirin tablets are a popular form of the medication, but not everyone is able to swallow a tablet. Children often are quite unable to bring themselves to swallow such an unpleasant tasting pill. Soluble aspirin is produced by forming the sodium salt of acetylsalicylate in an acid–base reaction with sodium hydroxide:



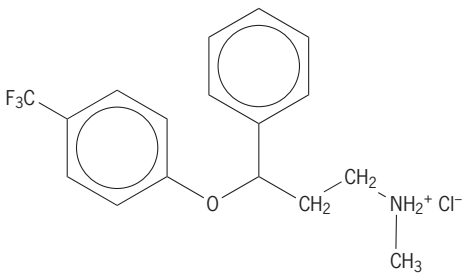


Figure 4.9.4 Fluoxetine hydrochloride is made soluble by the formation of the ammonium ion, which forms an ionic compound with the chloride ion.

The soluble salt formed needs to be crystallized in order to prepare it in tablet form.

Fluoxetine hydrochloride (Prozac®) is the soluble form of fluoxetine. It is the chloride salt formed when the amine group gains a hydrogen ion and forms ionic bonds with chloride ions.

In section 4.8, the danger of a racemic mixture was illustrated by the example of thalidomide. In this and many other cases, it is desirable to be able to synthesize only one enantiomer of a chiral compound. This is called a stereospecific synthesis. It is easier to make the desired enantiomer from a non-chiral molecule than to separate it from a racemic mixture.

AS D.9.5
Discuss the use of chiral auxiliaries to form the desired enantiomer. © IBO 2007

To synthesize just one enantiomer of a compound, a **chiral auxiliary** is used. This is a chiral molecule that is temporarily attached to a non-chiral molecule and directs the synthesis in an asymmetric manner, so that just one enantiomer is formed. Enantiomers have identical chemical properties in relation to non-chiral reagents, but not to other chiral molecules. They may also differ in their biochemical reactions. The auxiliary can be removed (and recycled) when the required enantiomer has been made.

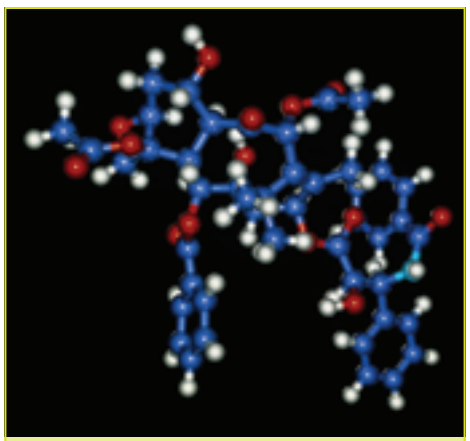


Figure 4.9.5 The structure of taxol, an anti-cancer drug that is prepared by a stereospecific synthetic route.

For example, Taxol, an anti-cancer drug used for the treatment of ovarian cancer and breast cancer is synthesized using a chiral auxiliary.

Section 4.9 Exercises

1 Explain why compound libraries are favoured in the early stages of drug design.

2 Outline the importance of a lead compound in drug design.

3 Describe the two methods of combinatorial chemistry that are commonly used.

4 Describe the advantage of using computer modelling as a preliminary step in drug design.

5 Describe how a previously insoluble compound that has a carboxyl group as part of its structure can be made into a soluble salt.

6 State the functional group (other than carboxyl) that can be used to create a soluble form of a drug and give a general equation for this reaction.

7 Some enantiomers can be prepared by using a chiral auxiliary.

a Explain the purpose of a chiral auxiliary.

b State the advantage of using a chiral auxiliary in synthesizing a drug.

8 **a** Explain the term *racemic mixture*.

b Supporting your answer with an example, describe the possible ill effects of the preparation of a drug as a racemic mixture, rather than a pure sample of one enantiomer only.

4.10 MIND-ALTERING DRUGS

HL

Mind-altering drugs interfere with neurotransmitters in the brain. These 'mind-altering' drugs change the interpretation of the world, behaviour, and mood of the user. Among the more common 'mind-altering' drugs are lysergic acid diethylamide (LSD), mescaline, psilocybin and tetrahydrocannabinol (THC).

AS **D.10.2 Discuss the structural similarities and differences between LSD, mescaline and psilocybin. © IBO 2007**

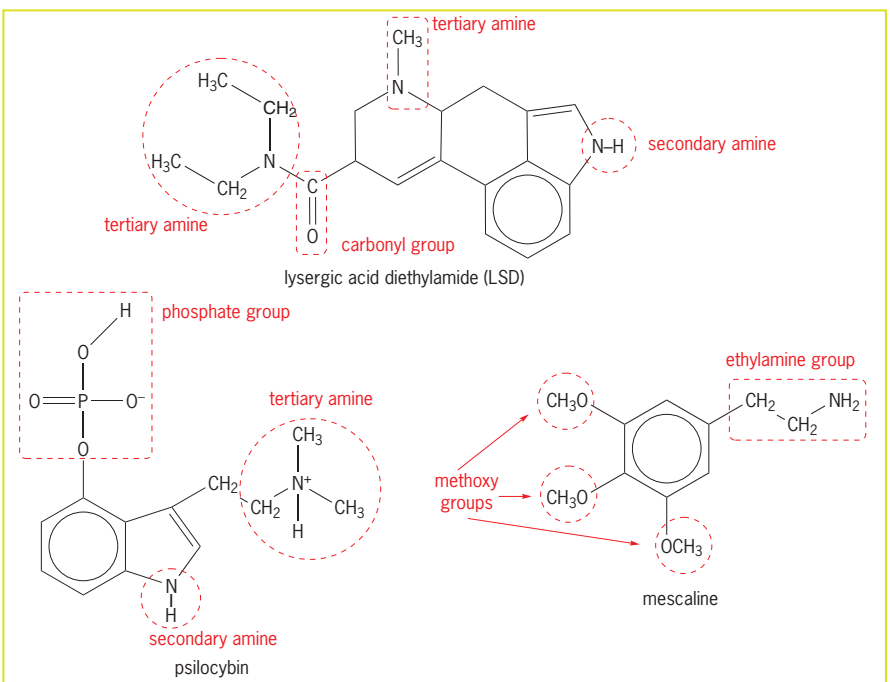


Figure 4.10.1 The structures of LSD, psilocybin and mescaline.

LSD is synthesized from lysergic acid, which is found in a particular fungus (ergot), while mescaline is derived from various cacti. Psilocybin, a psychedelic indole of the tryptamine family, is found in psilocybin mushrooms ('magic mushrooms').

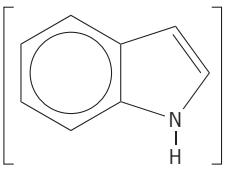


Figure 4.10.2 The indole ring structure.

Psilocybin and LSD differ from mescaline in that they contain an indole ring structure. In psilocybin, the indole ring has a negatively charged phosphate group and a positively charged tertiary amine bonded to it. The amine that is part of the indole ring is a secondary amine. In LSD, the indole ring is part of a fused ring system. A tertiary amine is bonded to the fourth ring in this system and another tertiary amine can be found as part of that ring. The structure of mescaline consists of a benzene ring with three methoxy groups and an ethyl amine group substituted onto the ring.

AS D.10.1
Describe the effects of lysergic acid diethylamide (LSD), mescaline, psilocybin and tetrahydrocannabinol (THC). © IBO 2007

The effects of LSD are the result of the disruption of the transmission of nerve impulses to the brain. This causes changes in perception, especially visual perception, with vivid, often brightly coloured (psychodelic) hallucinations. No physical dependence occurs, but bizarre mental effects and permanent personality changes sometimes occur. LSD mimics serotonin action in the brain, which seems to explain its hallucinogenic effects. Serotonin is a neurotransmitter that is believed to play an important role in the modulation of a range of conditions such as aggression, body temperature, mood and gastrointestinal operation.

Like LSD, mescaline is also a hallucinogenic drug. It is considerably less potent than LSD and slightly less potent than psilocybin. The hallucinations caused by mescaline are different from those caused by LSD. They represent an intensification of actual experience. There are no distortions of form or kaleidoscopic experiences. Mescaline is a sympathomimetic molecule to the neurotransmitter norepinephrine. Together with epinephrine (adrenaline), norepinephrine underlies the fight or flight response, directly increasing heart rate, triggering the release of glucose from energy stores, and increasing blood flow to skeletal muscle. Mescaline also blocks the release of acetylcholine (another neurotransmitter) in muscle.

Psilocybin is converted into psilocin, the pharmacologically active compound in the body, which, like LSD, mimics the action of serotonin in the brain.

THC is a cannabinoid which is the main psychoactive ingredient in cannabis. There are specialized receptors in the brain, known as cannabinoid receptors, to which THC binds. THC affects the release of acetylcholine, noradrenaline and dopamine (which causes happiness) from neurones. THC also has analgesic effects, giving reason for cannabis to be used as a pain-killer. Other effects include relaxation; euphoria; altered space–time perception; alteration of visual, auditory and olfactory senses; disorientation; fatigue and appetite stimulation.

AS D.10.3
Discuss the arguments for and against the legalization of cannabis. © IBO 2007

Marijuana is the dried leaves, flowering parts, stems and seeds of the Indian hemp plant, cannabis.

The stems can also be used to make tough fibres from which rope is made. Cannabis is a hallucinogen. It is not regarded as addictive, but users develop psychological dependence.

Marijuana is smoked in pipes and cigarettes. Like normal cigarettes, the tar from marijuana cigarettes (50% more is present than in normal cigarettes) is carcinogenic.

There is some evidence that long-term users of cannabis are apathetic and sluggish with poor memory. The psychological dependence shows as

restlessness, anxiety, irritability and insomnia. Driving performance, work performance and relationships with other people are all affected by cannabis use. Regular use of cannabis also suppresses the body's immune response and makes the user more susceptible to disease. After a while many users crave a better or different sort of 'high' and turn to other ('harder') drugs (e.g. heroin).

In many countries the use of cannabis is not legal. There has been active debate for many years as to whether it should become legal, supporters suggesting that it has medicinal uses that could benefit many if it were available as a legal medicine. The arguments for and against the legalization of cannabis are summarized in table 4.10.1.

TABLE 4.10.1 ARGUMENTS FOR AND AGAINST THE LEGALIZATION OF CANNABIS

Arguments for the legalization of cannabis	Arguments against the legalization of cannabis
Smoking of cannabis relieves symptoms of diseases such as cancer, AIDS, glaucoma and Parkinson's disease.	Like tobacco, smoking cannabis causes diseases such as respiratory ailments, heart disease and cancer, and it suppresses the immune system.
If it was easily available, there would be no need for it to be sold by criminals, so would result in a reduction in crime.	Once an individual has tried smoking cannabis, the risk of trying more damaging ('harder') drugs may be increased.
The free choice to smoke cannabis should be a matter for the individual.	Like someone under the influence of alcohol, a smoker under the influence of cannabis has a reduced ability to drive or operate machinery safely.

CHEM COMPLEMENT

Prolonged use of marijuana shown to shrink the brain

A paper published in June 2008 reported that long-term heavy use of marijuana may cause parts of the brain to shrink. The particular parts of the brain that were affected were the hippocampus, which regulates memory and emotion, and the amygdala, which plays a critical role in fear and aggression.

The subjects of the study were more likely to exhibit mild signs of psychotic disorders than non-smokers, and they reported experiencing some form of paranoia and social withdrawal. Although the study was based on only 15 men who had been heavy marijuana smokers for an average of 20 years, it still has been recognized as the first concrete evidence that marijuana use does adversely affect the brain.

THEORY OF KNOWLEDGE

Salvador Dali painted the *Slave Market with the Invisible Bust of Voltaire* (1940) and many other works using a style called Surrealism, which portrays a 'paranoiac-critical' way of perceiving reality. Supposedly without using any mind-altering drugs, Dali trained himself to fall into a hallucinatory haze. After returning to a normal state of mind he painted works with multiple images and meanings that came from this altered perception. Dali claimed that this method allowed him to release his innermost thoughts and connect with his subconscious.

Figure 4.10.3 *The Slave Market with the Invisible Bust of Voltaire* by Salvador Dali.

THEORY OF KNOWLEDGE

Albert Einstein said 'reality is merely an illusion.' Imagine you have touched a hot beaker and the stimulus is converted to an electrical signal that is sent to your brain and decoded. Your brain interprets the sensation as an extreme burning pain. Now consider the view that the pain you experienced took place only in your mind. The pain is not real; it doesn't really exist. The philosophical view that nothing actually exists outside of the mind is called solipsism. This book doesn't exist, your chemistry teacher doesn't exist, the physical world you live in doesn't exist. An analogy might be the avatar in a computer game. You're playing a game where you control the avatar, who can move around within the game world. Solipsism is the perspective that says your avatar is the only thing in the game that's real, and the whole simulation is a product of your avatar's mind.

- Is it possible that the inside of your mind is the only thing that exists or is real? Can you offer a counter argument to solipsism? Is there any way of proving if this view is true or not?
- Galileo once said that 'the tickle is not in the feather'. What do you think he meant by this?

Section 4.10 Exercises

1 Explain what is meant by the term *sympathomimetic*.

2 Although both LSD and mescaline are mind-altering drugs, there is a difference in their effects. Describe the difference in the effects of LSD and mescaline.

3 Describe the similarities in action and structure of psilocybin and LSD.

4 Describe the role of serotonin in the brain.

5 Compare the structures of LSD and psilocybin with that of mescaline.

6 Identify the active ingredient present in cannabis (marijuana).

7 Taking the point of view of a person in extreme pain due to illness, state three arguments as to why cannabis should be legalized.

8 Taking the point of view of the mother of a boy who died in a car crash after smoking marijuana, state three arguments as to why cannabis should not be legalized.

Chapter review questions and tests are available on student CD.

Chapter 4 Summary

Terms and definitions

Acyclovir An antiviral drug that is used to fight herpes virus infections (also know as aciclovir).

Alginates Polymers derived from brown seaweeds, which are added to antacid mixtures to form a neutral layer on top of the stomach acid.

Analgesic A pharmaceutical that relieves or prevents pain.

Antacid A base that neutralizes the excess acidity and relieves the pain associated with heartburn and peptic ulcers.

Anti-foaming agents Chemicals that reduce the surface tension of gas bubbles, causing them to coalesce, e.g. dimethicone.

Anxiolytics Drugs that are used to treat anxiety disorders.

Chiral A molecule that does not have a plane of symmetry and its mirror image will not superimpose onto the original molecule.

Chiral auxillary A chiral molecule that directs the synthesis, to produce only one enantiomer.

Combinatorial chemistry The mass production of many variations of a drug. A set of starting materials is reacted in all possible combinations.

Depressant A chemical agent that diminishes the function or activity of a specific part of the body, in particular the central nervous system.

Effective dosage (ED_{50}) The amount of a pharmaceutical that would produce a therapeutic response in 50% of the population.

Enantiomers Two different forms of a chiral molecule.

Gas–liquid chromatography (GLC) A type of chromatography that uses a liquid stationary phase and a gaseous mobile phase or carrier gas.

Geometric isomers Stereoisomers in which the arrangement of the functional groups across the double bond differs.

HIV (Human immunodeficiency virus) The virus that causes AIDS.

Intramuscular Into a muscle.

Intravenous Into a vein.

Lead compound A compound that will lead the way forward in design of a drug; a starting compound in drug design.

Lethal dosage (LD_{50}) The amount of a pharmaceutical that would kill (be lethal to) 50% of the population.

Narcotic An analgesic that produces euphoria as well as causing loss of feeling or paralysis.

Optical isomers Molecules with the same molecular formula, but with structures that are unable to be superimposed on one another.

Parallel synthesis Many identical chemical reactions are carried out on a wide range of compounds at the same time. Robotics are usually used to carry out routine steps.

Pathogenic Disease-causing.

Placebo A tablet that is pharmaceutically inert, i.e. it contains no medicine.

Placebo effect A psychosomatic response to a substance that should be ineffective.

Prophylactic A medication or a treatment designed and used to prevent a disease from occurring.

Protease inhibitors Anti-viral drugs that act on the enzyme responsible for cleaving polypeptides made in the host cell by the proviral DNA and so preventing the synthesis of structural proteins for the new virions.

Racemic mixture A mixture of the two enantiomers of a chiral compound.

Retrovirus A protein-enveloped virus possessing an RNA genome that replicates via a DNA intermediate.

Reverse transcriptase inhibitors anti-HIV drugs which interfere with the process of transcription of proviral DNA by either inhibiting the enzyme or competing with the host cellular triphosphate nucleosides.

Side-effects Unwanted response to a drug; not the main effect of the drug.

Solid-phase synthesis Peptide synthesis using insoluble, but porous, small solid beads on which peptide chains can be built.

Stereoisomers Molecules having the same molecular formula and the same structural formula, but different arrangements of atoms in space.

Stimulant A drug that stimulates the central nervous system, making the user feel less tired and more alert.

Subcutaneous Under the skin.

Sympathomimetic drugs Drugs that mimic the action of hormones such as epinephrine and neurotransmitters such as noradrenaline in the way in which they affect the body.

Synergistic effect The enhancement of the effect of one drug when taken in conjunction with another, e.g. ethanol (alcohol).

Therapeutic window The ratio of the lethal dose (LD_{50}) to the effective dose (ED_{50}), i.e. $\frac{LD_{50}}{ED_{50}}$.

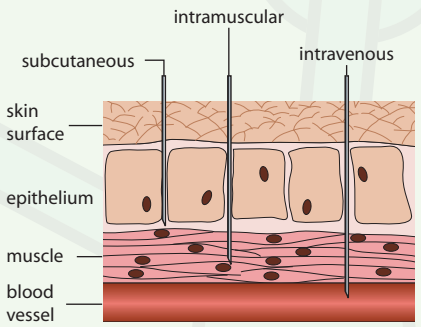
Tolerance The ability of the body to absorb increasing amounts of a drug over time without experiencing the effect of the drug.

Vasoconstrictor A drug that makes the blood vessels narrower.

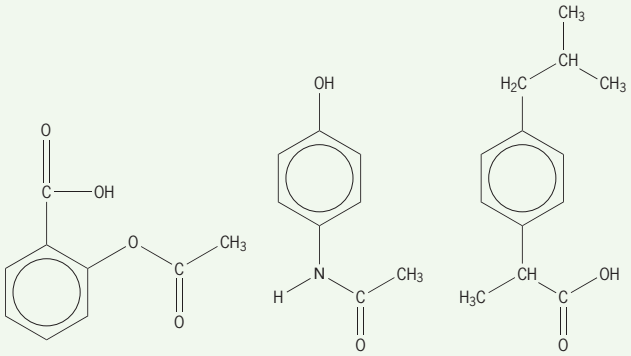
Virion A complete virus particle, consisting of nucleic acid surrounded by a protective coat of protein.

Concepts

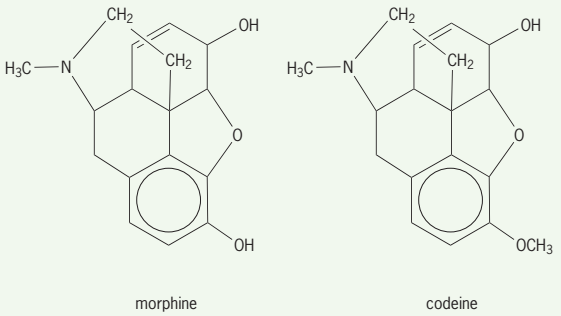
- The method of administration (oral, rectal, inhalation or injection) of a medicine large depends on the effect the medicine is to have and on the condition of the patient.
- Injections may be subcutaneous, intramuscular or intravenous.

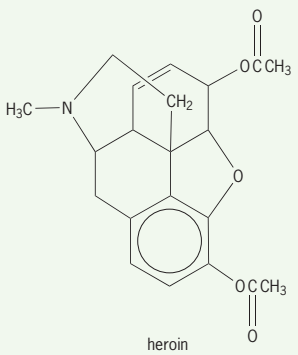


- Research and development of new drugs is an involved, lengthy and costly process that is carefully administered by government organizations.
- Antacids are used to neutralize excess acidity in the stomach and relieve the discomfort caused by this acidity. Antacids commonly are made up of bases such as aluminium and magnesium hydroxides and sodium hydrogencarbonate.
- While mild analgesics intercept the pain stimulus at the source by interfering with the production of prostaglandins that cause pain, swelling or fever, strong analgesics temporarily bond to receptor sites in the brain, thus preventing the transmission of pain impulses without depressing the central nervous system.
- Mild analgesics include aspirin, acetaminophen (paracetamol) and ibuprofen.

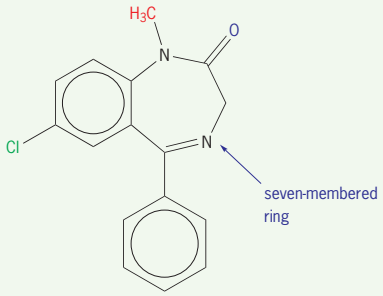


- Although described as mild due to their pain-relieving action, these analgesics still have disadvantages, including stomach bleeding in the case of aspirin and kidney damage in the case of excessive use of acetaminophen.
- Strong analgesics include the naturally occurring morphine and its derivatives codeine and heroin.

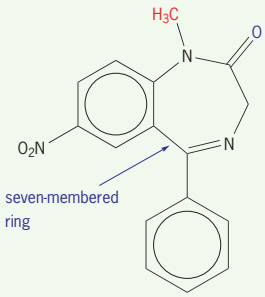




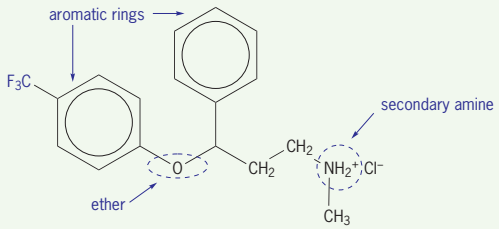
- Depressants are chemical agents that diminish the function or activity of a specific part of the body. Antidepressants are psychiatric medications or other substances (nutrients or herbs) used for alleviating depression. As such they may also be called depressants.
- Ethanol is a common depressant that is consumed by many with varying degrees of effect depending on how much is consumed and for how long.
- Other depressants include diazepam (Valium®) and nitrazepam (Mogadon®) and fluoxetine hydrochloride (Prozac®), which is actually an antidepressant.



diazepam (Valium®)



nitrazepam (Mogadon®)



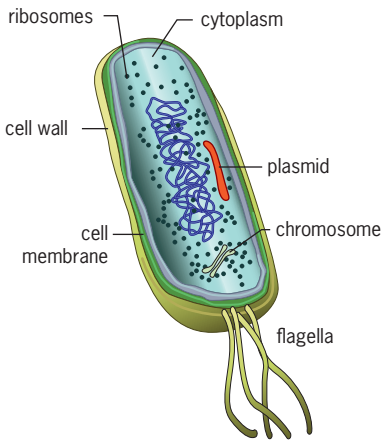
fluoxetine hydrochloride (Prozac®)

- Ethanol may be detected in the breath using a breathalyser or an intoximeter, or in the blood by gas–liquid chromatography.
- Stimulants include nicotine and caffeine. They generally make the user feel less tired and more alert, particularly when the user is already tired. They stimulate the brain and central nervous system.

Effects of depressants	Effects of stimulants
Calm and relax the CNS, slow the heart rate	Increase heart rate and increase blood pressure
Reduce rate of breathing	Increase breathing rate
Dull emotional responses	Dilate the pupils of the eyes
Reduce anxiety	Constrict the arteries

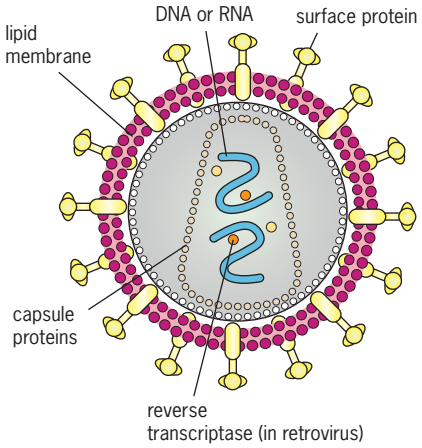
- Penicillin is an antibacterial medicine that kills bacteria by interfering with their ability to synthesize cell walls. The bacteria become unable to divide and as a result the cells burst. Overprescription of penicillins can result in the increasing resistance of bacteria to the penicillin and eventual development of strains of bacteria that are resistant to all antibiotics.
- Antiviral drugs act by:
  - chemically blocking the use by the virus of ribosomes, and hence the reproductive enzymes, thus preventing replication of the virus in the host cell
  - mimicking one of the nucleosides from which the DNA of the virus is made. The virus incorporates this fake nucleoside into its genome during replication and ultimately newly synthesized viral DNA becomes inactive, so the life cycle of the virus is halted.
- Like other viruses, HIV reproduces rapidly and mutates frequently, making it very difficult to treat with antiviral drugs.

Bacteria



- Single-celled micro-organisms
- Single chromosome per cell (circular strand of DNA)
- Rigid cell wall made of polysaccharide molecules
- Cytoplasm, nucleus and cell membrane, as for animal and plant cells
- Cytoplasm inside cell membrane containing ribosomes and enzymes to break down food and build cell parts
- Feed, excrete and grow

Viruses



- Non-cellular and sub-microscopic (much smaller)
- Central core of DNA or RNA surrounded by a protein coat
- No nucleus, cytoplasm, cell membrane or cell wall
- Have no ribosomes of their own and can only reproduce inside a living cell using the ribosomes of that cell
- Have only the enzymes that they need for invading a cell and to replicate their own nucleic acids
- Do not feed excrete or grow

- The action of drugs may depend on their geometric or optical isomerism, with one isomer being biologically active and the other not. A chiral auxiliary may be used to ensure that the correct isomer is synthesized.

- Cisplatin (a geometric isomer) and taxol (an optical isomer) are anti-cancer drugs.
- Modern drug design makes use of computer modelling, compound libraries and techniques of combinatorial chemistry.
- LSD, mescaline, psilocybin and THC are all mind-altering drugs that interfere with neurotransmitters in the brain, and change the interpretation of the world, and the behaviour and mood of the user.

5 ENVIRONMENTAL CHEMISTRY

Chapter overview

This chapter covers the IB Chemistry syllabus Option E: Environmental Chemistry.

By the end of this chapter, you should be able to:

- describe both natural and anthropogenic sources of air pollution and evaluate methods for its reduction
- state what is meant by acid deposition and outline its origins
- describe the greenhouse effect
- describe the formation and depletion of ozone in the stratosphere by both natural and anthropogenic sources
- distinguish between aerobic and anaerobic decomposition of organic matter in water
- describe the processes and the effects of both eutrophication and thermal pollution
- list and give the sources of waste water primary pollutants and outline the three stages of waste water treatment
- evaluate the process to obtain fresh water from sea water using multi-stage distillation and reverse osmosis
- discuss the three causes of soil degradation
- describe the relevance of SOM and outline its physical, biological functions and **(HL)** chemical functions
- list common soil pollutants and their sources
- outline and compare the various methods for waste disposal
- describe the characteristics, storage and disposal of different types of radioactive waste

- describe the mechanism in the catalysis of O_3 depletion by pollutants, and outline the reason for greater ozone depletion in the polar regions HL
- state the source of both primary pollutants, and secondary pollutants and the conditions necessary for the formation of photochemical smog
- describe the mechanism of acid deposition caused by the oxides of nitrogen and oxides of sulfur and explain the role of ammonia
- solve problems relating to the removal of various pollutants from water by chemical precipitation
- explain the importance of CEC.

IBO Assessment statements E.1.1 to E.12.4

An **environment** can be defined as all external conditions that affect an organism (plant or animal, including humans) during its lifetime. All activities in which humans participate influence all aspects of our environment including the air, water and soil. Our influences extend to all organisms living within our habitat, from the tiniest soil nitrogen-fixing bacteria found in Vietnam to larger South African water bugs to the majestic polar bear found in the Canadian Arctic.

Figure 5.0.1 The polar bears' Canadian Arctic habitat is being threatened by human's abuse of the environment.

Figure 5.0.2 Just how safe is our drinking water?

No matter where we live on this planet, we should live by the principle that we have to leave it in as good a shape as we found it, if not better. In the last 150 years, due to the rapid increase in both technology and invention, the human race has caused great destruction to the global environment. What happens in your backyard directly affects your global neighbours. We can see examples of this destruction if we look at newspaper headlines from around the world:

Reuters—Ethiopia's Gebrselassie Misses Marathon due to Pollution

Associated Press—Probe finds Drugs in US Drinking Water

BBC News—European Union Warns of Climate Change Threat

Sydney Morning Herald—Most Aussies Urging Climate Change Action: Poll

Beijing People's Daily—Emissions Make Earth Wetter and Stickier Planet

THEORY OF KNOWLEDGE

The Internet is an important secondary source of information about the environment. As critical thinkers we should take care not to blindly accept what we are told, what we hear or what we read without thinking about whether the source is one we can trust. Information on the Internet exists in large quantities and is continuously being created and revised. It is used to inform, to persuade, to sell, to present a viewpoint, or to change an attitude or belief. In addition, the reliability of information covers the full spectrum from very good to very bad. Before you start searching you need to think about what sort of information you want. Do you want facts, opinions (authoritative or just anyone's), reasoned arguments, statistics, narratives, eyewitness reports, descriptions or new ideas? Do you want to find either factual or reasoned support for a position, to survey opinion, or something else?

- All news media have some sort of bias, written to give one particlar opinion or to influence people's views. Some are definietely more objective than others. Things to consider when using media sources are: Who produced the information?
- What is its purpose? How are the facts conveyed? How are the words chosen and put together? The following headline comes from China Daily Online, 10 September 2007:

  'In pure Arctic air, signs of China's economic boom'

  – Is the headline factual or giving an opinion?

  – Do you think the headline is trying to grab the reader's attention?

  – Is the title honestly stated or is more subtle, value-laden language used?

  – Is it clear where the sympathies of the editor lie?
- Suggest a set of guidelines for evaluating the reliability and objectivity of the most up-to-date research posted on the Internet about climate change. Things to consider are: Who is the author and what are their credentials on the subject? What purpose did the author have in mind? How current is the information? How can you judge the reliability of the information? What do others say about this site? What does your intuition tell you? How well does it fit in with what you already know?

5.1 AIR POLLUTION

Interactive CW Animation
Atmospheric layers

The Earth is the only planet in our solar system in which there is a suitable atmosphere for the maintenance of human life. The atmosphere is a layer of gases that surrounds the Earth and provides the necessary gases for human existence. The air that we breathe contains approximately 78.08% nitrogen (N_2), 20.95% oxygen (O_2), 0.93% argon (Ar), 0.04% carbon dioxide (CO_2) and a fluctuating percentage of water (H_2O) that hovers at about 1%.

Air pollution is the change in the natural atmosphere by the addition of other gases, particulates and volatile organic compounds (**VOCs**). The main gaseous air pollutants are carbon monoxide (CO), oxides of nitrogen (NO_x) and oxides of sulfur (SO_x).

Figure 5.1.1 Our Earth, there is only one so we must prevent its destruction due to pollution.

AS **E.1.1**
Describe the main sources of carbon monoxide (CO), oxides of nitrogen (NO_x), oxides of sulfur (SO_x), particulates and volatile organic compounds (VOCs) in the atmosphere. © IBO 2007

AS **E.1.2**
Evaluate current methods for the reduction of air pollution. © IBO 2007

Figure 5.1.2 Automobile exhaust is the leading causes of anthropogenic carbon monoxide production.

QuickTime CW Video
Action at a catalyst surface

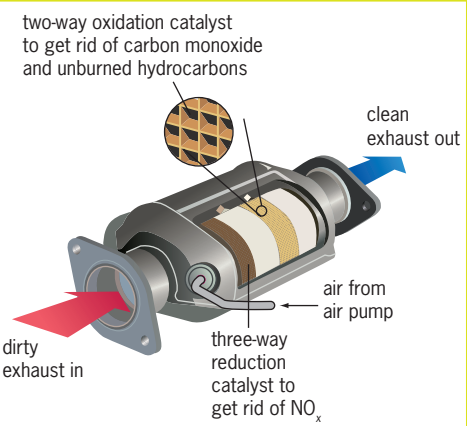


Figure 5.1.4 The inside workings of a typical catalytic converter.

Carbon monoxide (CO)

The largest source of **anthropogenic** (resulting from human activity) CO is the incomplete combustion of fossil fuels in internal combustion engines such as those found in automobiles.

Incomplete combustion occurs when there is not enough oxygen available to the fuel in order to create carbon dioxide as one of the products. Instead carbon monoxide is formed.

$$2C_8H_{18}(l) + 17O_2(g) \rightarrow 16CO(g) + 18H_2O(l)$$

Natural sources of CO include volcanic activity and bushfires. Figure 5.1.3 shows the first image of global carbon monoxide levels that were produced in 2000 by NASA's Terra spacecraft.

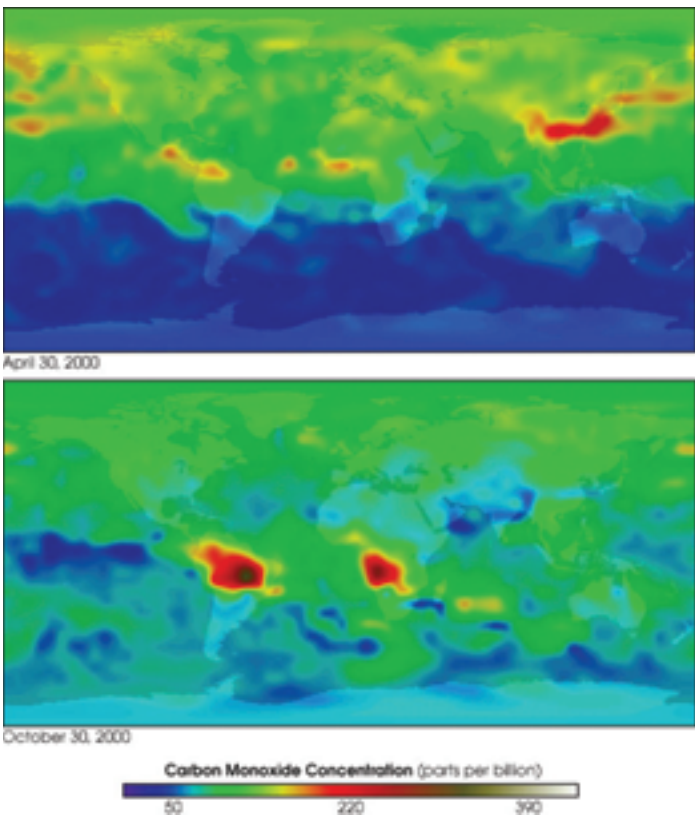


Figure 5.1.3 NASA satellite imagery depicts the global seasonal carbon monoxide concentrations.

To reduce levels of anthropogenic CO, most cars and some buses, trains and industrial machinery are equipped with catalytic converters. Even the newly touted 'people's car', the Tata Nano, selling very cheaply in India has a catalytic converter as a part of its exhaust system. Catalytic converters work to convert the toxic by-products of combustion, CO included, to less toxic substances.

$$2CO(g) + O_2(g) \rightarrow 2CO_2(g)$$

Catalytic converters make use of a platinum, palladium or rhodium catalyst on which the exhaust gases undergo heterogeneous catalysis. Because rare metals are used in the production of catalytic converters, theft of converters from parked cars or junk yards has been on the rise over the past few years.

CHEM COMPLEMENT

Carbon monoxide

Carbon monoxide is sometimes found in coal mines. At one time canaries were taken down in mines to detect this poisonous gas. Canaries would be killed at doses not quite lethal to miners. Today gases are detected more humanely with instruments. Carbon monoxide poisoning can occur when hemoglobin is prevented from carrying vital oxygen in the bloodstream. The oxygen molecule is replaced with a molecule of carbon monoxide and carboxyhemoglobin forms. This means that oxygen can't bind to the hemoglobin and could result in a life-threatening situation. Home owners who burn wood or coal as the main source of heat are encouraged to purchase a carbon monoxide detector, which is similar to a smoke detector but indicates when there is an unsafe level of carbon monoxide in the home.

Figure 5.1.5 Instrumentation used to detect unsafe levels of CO in households. It will emit a high-pitched beep when unsafe levels of CO have been exceeded.

Levels of anthropogenic CO can also be reduced by the use of lean-burn engines. A lean-burn engine has a very high air to fuel ratio, which means that more complete combustion occurs and, as a direct result of this, less CO is produced by the car's engine. Car manufacturers such as Honda and Mitsubishi use lean-burn engines in some of their models.

Oxides of nitrogen

NO_x is the general term used to classify nitrogen oxides (NO and NO_2). These nitrogen oxides are formed at high temperatures inside internal combustion engines when the nitrogen in the air combines with oxygen according to the following equation:

$$N_2(g) + O_2(g) \rightarrow 2NO(g)$$

In areas with large traffic volume, the abundance of NO results in the formation of a red haze, smog (see section 5.10).

Figure 5.1.6 NO is one of the pollutants responsible for smog that forms over large cities, such as New York.

Venturing outside when there are high levels of NO_x can be very harmful to people with lung diseases such as asthma, and can cause damage to healthy lung tissue and reduction in lung function for those exposed to this form of air pollution for any length of time. Leading up to the 2008 Summer Olympics in Beijing, China, there was much discussion about the high levels of pollution in that city.

In nature, lightning strikes produce NO_x within thunderstorm clouds. Additionally, NO and N_2O are produced by the process of denitrification in soils. In the denitrification process highly oxidized forms of nitrogen—nitrate and nitrite—are reduced to NO and N_2O by a microbial process. The sequence of reductions is shown below. The process occurs under anaerobic (low levels of oxygen) conditions, which can be found in very wet soils such as those of wetlands.

$$NO_3^- \rightarrow NO_2^- \rightarrow NO \rightarrow N_2O \rightarrow N_2$$

Figure 5.1.7 Lightning strikes within thunderstorm clouds produce the pollutants NO_x.

Denitrification is costly to the agricultural sector as the nitrogen lost in this process has to be replaced by costly nitrogen-containing fertilizers in order to maintain the high yields of crops.

To reduce the amount of NO_x pollutants in the atmosphere (other than reducing the number of cars on the road) one can make use of the catalytic converter, which is also useful for the lowering the levels of CO. In the catalytic converter the harmful NO is reduced to diatomic nitrogen at the same time as CO is being oxidized to carbon dioxide.

$$2CO(g) + 2NO(g) \rightarrow 2CO_2(g) + N_2(g)$$

Varying the ratio of air to fuel in a lean-burn engine can also lower the levels of NO_x; unfortunately, it counteracts the conditions we saw earlier in which the levels of CO in car exhaust were lowered. By creating a high air to fuel ratio, more complete combustion does indeed occur and less CO is produced. However, this creates a situation in which higher temperatures are produced and, as a result of these higher temperatures, more NO_x^- is produced. At a slightly lower air to fuel ratio, less NO_x^- but more CO is produced. The trick is to find the optimal air to fuel ratio.

Sulfur dioxide SO_2

Figure 5.1.8 During the smelting of metals, large amounts of SO_2 are released to the atmosphere.

Figure 5.1.9 A volcanic eruption of the island of Vanuatu releases hydrogen sulfide gas to the atmosphere.

The main anthropogenic sources of sulfur dioxide are the combustion of sulfur-containing coal ($S(s) + O_2(g) \rightarrow SO_2(g)$) and the smelting of sulfide ores. **Smelting** is a process by which a metal is extracted from its ore. Smelting revolutionized the field of metallurgy and marked the end of the Stone Age and the beginning of the Bronze Age, as workmen were able to make more extensive use of a wider range of metals in their purest form.

Roasting is the first step in smelting in which sulfur-containing ores are converted into oxygen-containing ores. These oxygen-containing ores are then converted to the bare metals. It is the roasting step that produces the pollutant sulfur dioxide from sulfide ores such as iron pyrites, zinc blende and cinnabar as shown below:

Iron pyrite: $4FeS_2(s) + 11O_2(g) \rightarrow 2Fe_2O_3(s) + 8SO_2(g)$
Sphalerite: $2ZnS(s) + 3O_2(g) \rightarrow 2ZnO(s) + 2SO_2(g)$
Cinnabar: $HgS(s) + O_2(g) \rightarrow Hg(l) + SO_2(g)$

Natural sources of sulfur dioxide include volcanic eruptions. When the highly pressurized volcanic gases rise to the Earth's surface there is a huge increase in their molar volume as shown by Boyle's law: PV = constant (see *Chemistry: For use with the IB Diploma Programme Standard Level*, chapter 4). The gas is first released in the form of hydrogen sulfide, the gas that is responsible for the rotten egg smell in stink bombs, and then converted to sulfur dioxide.

$$2H_2S(g) + 3O_2(g) \rightarrow 2SO_2(g) + 2H_2O(g)$$

This same reaction also occurs in 'sour' natural gas. This is natural gas that contains significant amounts of hydrogen sulfide (H_2S) as a result of the decay of organic matter. Sulfate ions (SO_4^{2-}) are found in sea spray and can be reduced to sulfur dioxide.

Sulfur dioxide is harmful to human health when it is inhaled and reacts with the wet mucus linings of our nasal passages and the pleural linings of the lungs to create sulfuric and sulfurous acid. This acts to irritate the entire respiratory system. This is also why your eyes may water when you cut onions. Cutting onions releases sulfur dioxide that mixes with the moisture in your eyes to create sulfuric acid. To dilute this acid, your eyes start to water.

The amount of sulfur in coal can be reduced either by washing the coal before combustion or by using alkaline scrubbers or limestone-based fluidized beds. By crushing and washing the sulfur-containing coal with water, the more dense sulfides sink to the bottom and the less dense, cleaner coals floats to the top and can be skimmed off, dried and then combusted for a cleaner burning fuel.

Alkaline scrubbers make use of a basic mixture of limestone ($CaCO_3$) and lime (CaO) to convert sulfur dioxide into a less harmful slag that can either be used as landfill or to make gypsum. Gypsum is made into plasterboard. This process is shown in figure 5.1.10 and in the equations below:

Limestone: $CaCO_3(s) + SO_2(g) \rightarrow CaSO_3(s) + CO_2(g)$

Lime: $CaO(s) + SO_2(g) \rightarrow CaSO_3(s)$

Gypsum production: $2CaSO_3(s) + O_2(g) + 4H_2O(g) \rightarrow 2CaSO_4.2H_2O(s)$

When performed efficiently, this method will result in the reduction of SO_2 by 95%, a very effective method indeed.

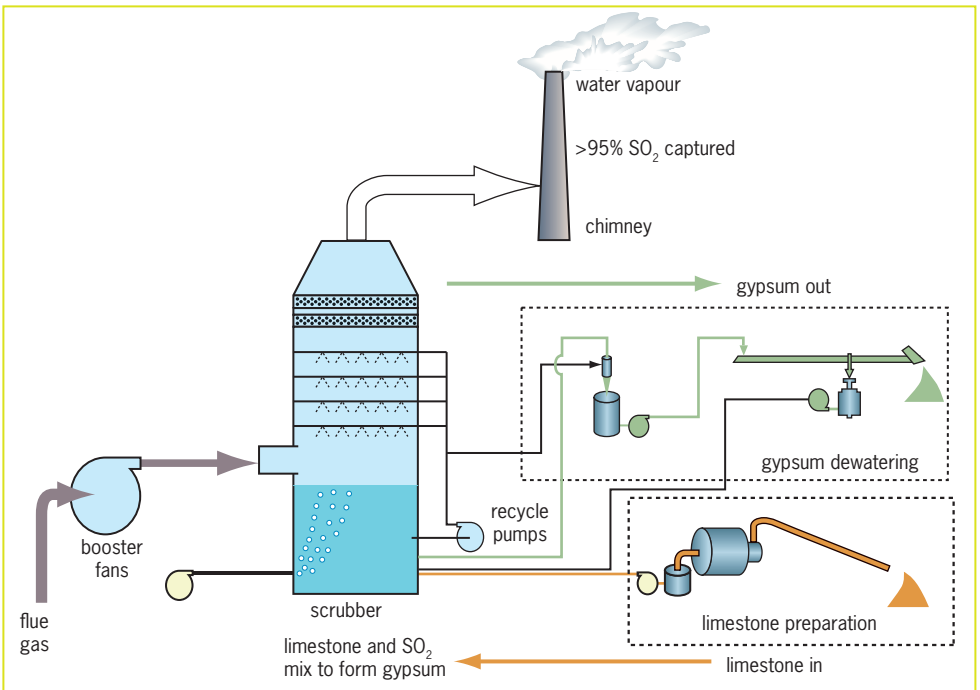


Figure 5.1.10 Limestone is used in alkaline scrubbers to convert harmful SO_2 into useful and harmless gypsum.

A second and more modern method of removing sulfur dioxide from the industrial combustion of sulfur-containing coal is to make use of limestone-based fluidized beds. In this clean-coal technology, a slurry of pressurized limestone and lime is mixed with the solid sulfur-containing coal, creating a 'fluid'. Fluidization allows the mixing of the sulfur-containing coal with limestone and lime, then, in reactions that are similar to those above, the contaminated sludge is removed, a mixture of clean coal and coal gas are produced and the sulfur dioxide pollutant is removed.

Particulates

Particulates are very small solid particles or liquid acid drops that are pollutants in our atmosphere. Particulates can be made up of metals, soot, soil or dust particles. Particulates are formed as the result of forest or bush fires; they are also a by-product of many industrial processes and result from the combustion of fossil fuels in internal combustion engines. Larger particles can be trapped in the nostrils during the inhalation process. The real danger lies when smaller particles enter the lungs and result in major health problems such as decreased lung function, aggravation of asthma and the development of bronchitis, emphysema and various respiratory cancers.

The main way to reduce particulates is by electrostatic precipitation. Electrostatic precipitators can remove 99% of particulate matter, but do not work if the particulate matter has high electrical resistance, as the process is based on these particulates being removed because they can become charged. First the particulates are charged negatively as they pass through pipes that are negatively charged. These negatively charged particles are then moved through positively charged plates that attract the negatively charged particulates. The particulates stick to the positive plates until they are collected. The air that leaves the plates is free from harmful particulate matter.

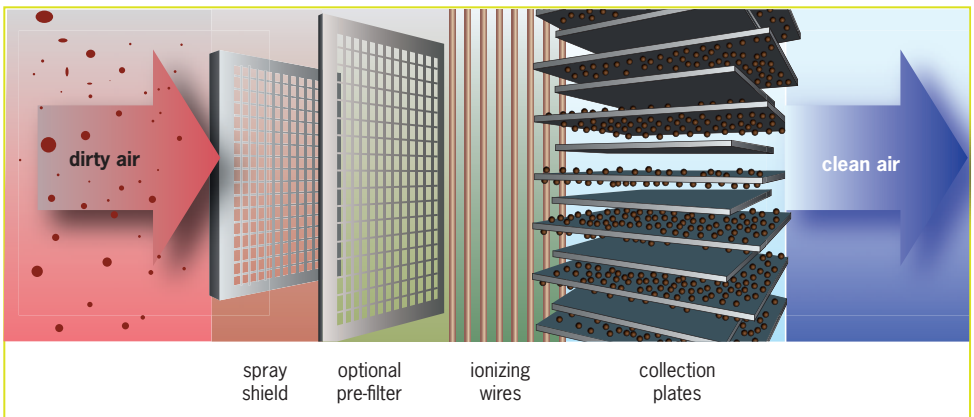


Figure 5.1.11 Electrostatic precipitation in action.

Volatile organic compounds (VOCs)

Volatile organic compounds are another of the by-products of combustion in an internal combustion engine that can be removed by catalytic converters. Like aldehydes and ketones, VOCs are organic compounds with high enough vapour pressures for vapours to be emitted from solid materials. They are emitted from such materials as paints, cleaning supplies, building materials, furnishing, glues and permanent marker pens. VOCs have been linked to 'sick building syndrome' in which occupants of modern office buildings suffer from exposure to these chemicals and show various symptoms. Ways to rid a building of VOCs include placing green plants, which can act as filters, in the building or by reverting back to the natural way of airing buildings and opening windows.

CHEM COMPLEMENT

The principle of minimum wrong and the global perspective

We should pay attention to the principle of minimum wrong, which states:

> When we alter nature to meet what we consider to be basic or non-basic needs, we should choose the method that does the least possible harm to other living things; in minimizing harm it is in general worse to harm a species than an individual organism, and still worse to harm a biotic community.

http://www.klima.ph/public_briefings/pb2003/05_TeacherTraining2/09_PhilQualityOfLife/01_Principles.pdf

People of all cultures, races and ages worldwide have a role to play in preserving our planet and restoring our environment to its once-thriving state. Through the study of this environmental chemistry option, it is hoped that you will find your voice and create an awareness of environmental problems in your community. It only takes one person to make a difference! The problem of pollution is not isolated to one continent or one culture or developed countries. It is a global concern, as shown by the following graph from the World Bank:

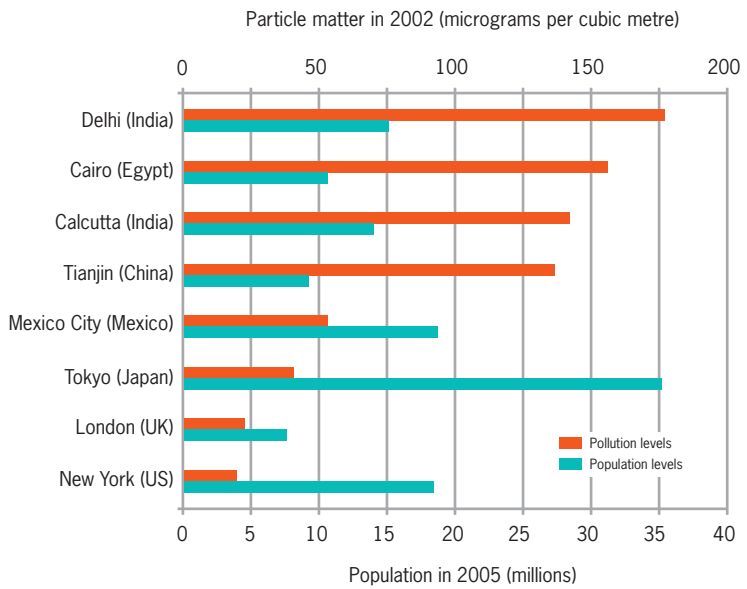


Figure 5.1.12 Graph showing some of the world's most polluted cities.

Summary of air pollutants

TABLE 5.1.1 AIR POLLUTANT SOURCES, HEALTH EFFECTS AND WAYS TO REDUCE EACH POLLUTANT

	Air pollutant				
	CO	**NO_x**	**SO_x**	**VOCs**	**Particulates**
Anthropogenic sources	Incomplete combustion in internal combustion engines	Formed at high temperatures in internal combustion engines	Combustion of sulfur-containing coal, smelting of sulfid ores and roasting processes	By-products of combustion in internal combustion engines; emitted from paints, cleaning supplies, building materials, furnishings	By-products of industrial processes, incomplete combustion
Natural sources	Volcanic activity and forest fires, incomplete oxidation of methane formed by anaerobic decomposition of organic matter	Lightning strikes, denitrification in soils	Volcanic eruptions, sour natural gas, sea spray	Anaerobic decomposition by bacteria	Volcanic eruptions, forest fires
Ways to reduce the pollutant	Catalytic converters, the use of lean-burn engines	Catalytic converters, the use of lean-burn engines	Wash coal prior to burning, use alkaline scrubbers or fluidized beds	Catalytic converters, the use of green plants to improve air quality	Electrostatic precipitation
Effects on human health	Replaces the oxygen carried in hemoglobin in red blood cells, reducing oxygen availability and causing suffocation	Harmful to people already suffering from lung disease; can also damage healthy lung tissue	Irritates the respiratory system	'Sick-building syndrome' due to poor indoor air quality	Decreased lung function, aggravation of asthma, development of bronchitis, emphysema and respiratory cancer

Section 5.1 Exercises

1. Explain why we should be concerned with the relationship between our environment and pollution.
2. State chemical equations that show the difference between complete and incomplete combustion. Use these equations to explain which type of combustion process produced the most harmful air pollutant.
3. List the pollutants that are responsible for air pollution and give both their natural and anthropogenic sources.
4. Evaluate the current methods for the reduction of each of the pollutants that you listed in question **3**.

5 Explain the importance of catalytic converters and state the chemical equation that occurs in a catalytic converter.

6 Outline the chemistry behind alkaline scrubbing, giving appropriate chemical equations for the reactions involved.

7 Explain how electrostatic precipitation removes particulates from the atmosphere.

8 Examine figure 5.1.3 and explain why there is a seasonal difference in the photos. Explain why there is an increased amount of carbon monoxide over South America and Sub-Saharan Africa.

9 Discuss the effect that modern living has on air pollution.

10 Describe the effect of air pollution on human health and on the health of the environment.

11 Apply the principle of minimum wrong to your school community. What can you do to follow this principle?

5.2 ACID DEPOSITION

AS **E.2.1 State what is meant by the term *acid deposition* and outline its origins. © IBO 2007**

Acid deposition refers to the process by which acidic particles, gases (dry deposition) and precipitation (wet deposition) leave the atmosphere. Any form of precipitation, whether it falls as snow, sleet, rain, fog or hail may be contaminated with several different acids. Normal precipitation has a pH level of 5.6. You may think that normal rain is just pure water with a pH level of 7, but this is not the case. Normal wet deposition is subject to reaction with carbon dioxide (CO_2) as shown in this equation:

$$H_2O(l) + CO_2(g) \rightarrow H_2CO_3(l)$$

QuickTime CW Video
Acid behaviour of CO_2 in water

This reaction produces carbonic acid, which naturally makes all forms of wet deposition slightly acidic. Carbon dioxide is produced when combustion occurs, animals respire and micro-organisms undergo fermentation.

Any form of precipitation that has a pH of less than 5.6 is considered acid deposition. The rainfall with the lowest recorded pH level—a pH of 2.4—fell in Scotland in 1974. Eastern Europe, Eastern North America, People's Republic of China and Russia are areas that are most prone to acid deposition. The formation of acid deposition has increased since the Industrial Revolution, primarily due to increased air pollution caused by the combustion of fuels at power plants and in internal combustion engines.

Figure 5.2.1 Rain is only one form of acid deposition.

Sources of acid deposition

Power generating plants and other industrial plants give off large amounts of sulfur dioxide (SO_2), particulate matter and nitrogen oxides (NO_x), all of which are primary pollutants. Wind and other weather patterns transport these pollutants, which eventually form secondary pollutants: nitrogen dioxide (NO_2), nitric acid (HNO_3) and sulfuric acid (H_2SO_4).

Figure 5.2.2 One environmental effect of acid deposition—damage to thousands of acres of forest.

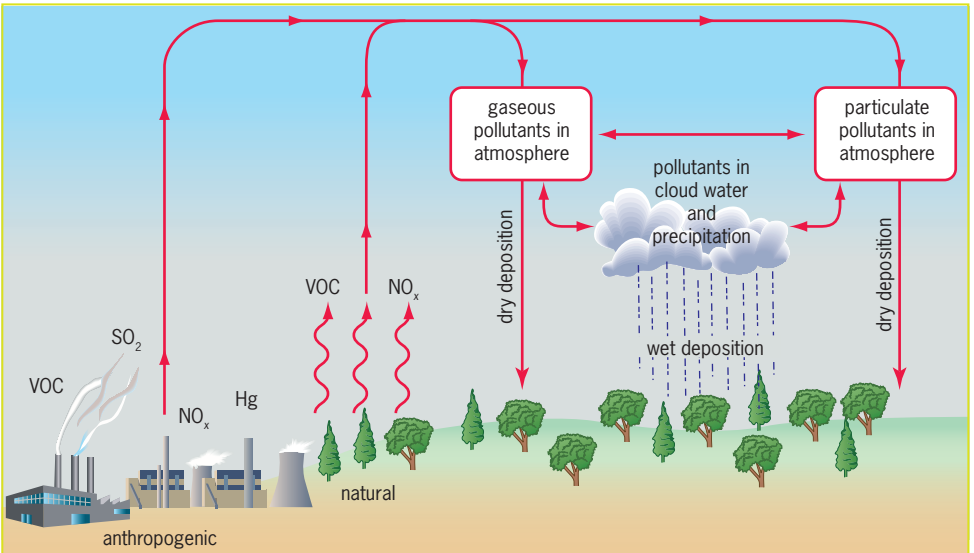


Figure 5.2.3 Main sources of acid deposition.

We are not the only culprits in the formation of acid deposition. SO_2 is also produced naturally from volcanoes, sea spray, plankton, rotting vegetation and sour natural gas. Its natural formation is:

$$2H_2S(g) + 3O_2(g) \rightarrow 2SO_2(g) + 2H_2O(l)$$

Figure 5.2.4 Coal-burning plants, like this one, are one of the biggest sources of SO_2.

The particulate matter that is produced as a by-product of coal burning can act as a heterogeneous catalyst for the direct reaction of SO_2 to SO_3, which then reacts with moisture to produce sulfuric acid:

$$2SO_2(g) + O_2(g) \rightarrow 2SO_3(g)$$
$$SO_3(g) + H_2O(l) \rightarrow H_2SO_4(aq)$$

Another culprit in the formation of acid deposition is the primary pollutant NO_x. The adverse effects of this primary pollutant are some of the leading motivators for the introduction of hybrid cars and other alternatives to the internal combustion engine. Automobile exhaust contains both nitrogen monoxide (NO) and nitrogen dioxide (NO_2), which react via the following mechanism to produce either nitrous (HNO_2) or nitric (HNO_3) acids:

$$N_2 + O_2 + \text{heat} \rightarrow 2NO$$
$$2NO + O_2 \rightarrow 2NO_2$$
$$3NO_2 + H_2O \rightarrow 2HNO_3 + NO$$

or

$$4NO + O_2 + 2H_2O \rightarrow 4HNO_2$$

The formation of nitrogen(II) oxide, NO, results from the combination of nitrogen and oxygen (found naturally in the air) at high temperatures in internal combustion engines. When released to the atmosphere, NO can react further with more oxygen to produce nitrogen(IV) oxide, which in turn reacts with water to form HNO_3. Alternatively, NO can react with oxygen and water to produce HNO_2.

Environmental effects of acid deposition

AS **E.2.2 Discuss the environmental effects of acid deposition and possible methods to counteract them. © IBO 2007**

Acid deposition has many different detrimental effects on both the environment and human health. Acid deposition has directly resulted in the deterioration of many historic marble structures. Marble is primarily made up of limestone ($CaCO_3$), which reacts with sulfuric acid:

$$CaCO_3(s) + H_2SO_4(aq) \rightarrow CaSO_4(aq) + H_2O(l) + CO_2(g)$$

Once this deterioration takes place it cannot be reversed and countless irreplaceable monuments, buildings and statues are forever lost.

Direct evidence that acid deposition is a major problem is seen by its undesirable effects on the Earth's aquatic environments—lakes, rivers, streams etc. Acid deposition will lower the pH levels of the water supplies in which it falls. Only the hardiest of species can survive at low pH levels. The acidity will either directly kill fish or affect their ability to successfully produce eggs or for their young to survive. This in turn limits the food supply that is available to local birds, snakes and insects whose diet is directly linked to the health of water supplies. **Acid shock** occurs during spring run-off into water supplies, when large quantities of snow and ice melt and bombard the water supplies with a huge amount of acid all at once. The vast amount of water is often 100 times more acidic than normal rainfall.

Figure 5.2.5 Acid shock occurs during spring run-off.

Trees are also directly affected when acid deposition occurs and damages the surfaces of leaves and needles. This reduces a tree's ability to withstand cold, and inhibits seed germination and reproduction.

Indirect effects of acid deposition include leaching of valuable nutrients from soils. Acid deposition causes cation exchange to take place in which the hydrogen ions from the acids replace the calcium, magnesium and potassium ions that are found in the soil. Then in a process called **leaching**, these ions are washed deeper into the subsoil or washed out of the topsoil and so are no longer available to be used in the healthy maintenance of the plant. Also, these toxic metal ions can be leached into drinking water supplies.

Figure 5.2.6 Leisure activities such as camping may be at risk due to loss of the environment as a result of acid deposition damage.

Deterioration of plants, wildlife and water supplies will ultimately lead to a decrease in the economies dependent on the environment such as tourism, camping, aquaculture and fishing.

Controlling acid deposition

Steps can be taken to limit the production of acid deposition. In smelting industries where coal is burned and SO_2 is produced, powdered limestone ($CaCO_3$) can be added to convert the SO_2 into $CaSO_3$, which can then be removed as a sludge:

$$CaCO_3(s) \rightarrow CaO(s) + CO_2(g)$$
$$CaO(s) + SO_2(g) \rightarrow CaSO_3(g)$$

To remove any remaining SO_2, an aqueous solution of CaO can be injected into the chamber prior to the gases escaping through a smokestack, a process called scrubbing:

$$2CaO(aq) + 2SO_2(g) + O_2(g) \rightarrow 2CaSO_4(aq)$$

Pr PRAC 5.1 Acid rain

Figure 5.2.7 Sulfuric acid—the most produced chemical in the world.

Alternatively, the SO_2 could be used in the production of sulfuric acid (Contact process), which is the most produced chemical in the world. Proactively, the percentage of sulfur in coal could be reduced before combustion by washing the coal to remove the sulfur or companies could use only low-sulfur coals.

CHEM COMPLEMENT

The chemistry of the Contact process

The Contact process is used in the production of sulfuric acid. It relies on many different chemical properties to be efficiently run: Avogadro's law of combining volumes, reaction rates, catalysis and Le Chatelier's principle. The Contact process has been used since 1881 and is the most widely used and most efficient way of producing large amounts of sulfuric acid.

Step 1: Production of SO_2

$$S(s) + O_2(g) \rightarrow SO_2(g)$$

Step 2: Production of SO_3

This step is closely controlled by equilibrium factors to generate the optimal conditions for SO_3 production.

$$2SO_2(g) + O_2(g) \rightleftharpoons 2SO_3(g) \qquad \Delta H = -196 \text{ kJ mol}^{-1}$$

Vanadium(V) oxide is used as a catalyst to speed up the reaction rate, although the catalyst has no effect on how much SO_3 is produced. Conditions of low temperature, 400–450°C work the best. The reaction is favoured by high pressure, but, due to the high cost of reproducing these high pressures, a pressure close to atmospheric pressure is used and only a small amount of yield is sacrificed. Lastly, increasing the concentration of either one of the reactants will result in a greater amount of SO_3 formed. Since the source of O_2 is the air, this is a cheap way of increasing the yield of SO_3.

Step 3: Production of sulfuric acid from oleum, $H_2S_2O_7$

Sulfuric acid cannot be produced directly by the addition of SO_3 to water, as this process is too exothermic and thus too dangerous.

$$SO_3(g) + H_2SO_4(aq) \rightarrow H_2S_2O_7(l)$$
$$H_2S_2O_7(l) + H_2O(l) \rightarrow 2H_2SO_4(aq)$$

Aquatic life can be helped when some environments have a natural defence mechanism against acid deposition. Some soils and bedrock contain limestone that acts in the same way as that introduced in smelting plants by creating a natural buffer zone. Thus areas that are rich in natural limestone deposits are able to counteract the effects of acid deposition. Lime-based fertilizers can be added to water supplies that are based on granite deposits, which are not natural buffers. This increases the pH of the water supply.

Section 5.2 Exercises

1. Explain why *acid deposition* is a better term to use than *acid rain*.
2. Explain why normal deposition has a pH level of 5.6. Make sure to use the proper chemical equation in your explanation.
3. Describe both the natural and anthropogenic sources of acid deposition and comment on what you think is the main source of acid deposition.
4. State the equations for the formation of sulfuric, sulfurous, nitric and nitrous acids in the atmosphere. Include the equations for the formation of gases responsible for the formation of these acids.
5. State the chemical reaction that occurs when acid rain falls on marble statues.
6. Comment on whether or not modernization is responsible for an increased occurrence of acid deposition.
7. Discuss the environmental effects of acid deposition.
8. Explain why the process of scrubbing is an important means of controlling acid rain. Ensure that you state the appropriate chemical equations.
9. Describe how some environments have natural defence mechanisms towards acid deposition.
10. Comment on the relationship between acid deposition and air pollution.
11. Discuss the relationship between acid deposition and the economy.

5.3 GREENHOUSE EFFECT

The Earth is the third planet in our solar system and is at a distance of roughly 150×10^6 km from the Sun. The Sun is the major source of external energy on Earth and it has been emitting solar radiation since the beginning of our galaxy. The Sun emits energy as electromagnetic radiation (see figure 5.3.1)—energy that does not need to travel through a medium; it can travel through the vacuum of space. It takes a little over 8 minutes for the Sun's energy to reach the Earth. About 30% of the Sun's energy is reflected back by clouds and particulate matter in the atmosphere, 23% is absorbed by the clouds and the remaining 47% makes it to the ground. The 70% of the Sun's energy that makes it to the Earth's surface is responsible for heating our planet and providing us with a comfortable atmosphere to live in.

The greenhouse effect

AS **E.3.1 Describe the greenhouse effect.** © IBO 2007

The Sun's radiation enters the Earth's atmosphere as shorter wavelength UV and visible light. However, when these light waves collide with clouds or the Earth's surface they are reflected back as longer wavelength infrared light. Much of this longer wavelength energy is too long to pass back through the Earth's atmosphere and as a result is trapped by the blanket of gases of the atmosphere—the 'greenhouse gases' carbon dioxide (CO_2), methane (CH_4), water vapour (H_2O), nitrous oxide (N_2O) and chlorofluorocarbons (CFCs). This trapped energy warms the Earth. This is the 'greenhouse effect'.

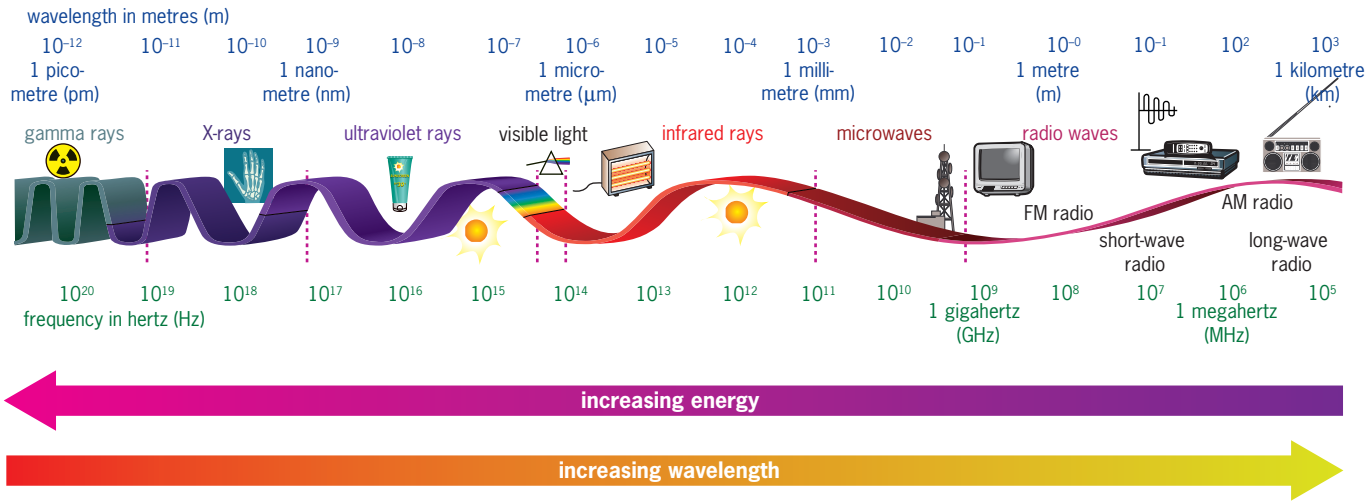


Figure 5.3.1 Electromagnetic spectrum; notice the difference in the wavelengths of visible light and infrared light.

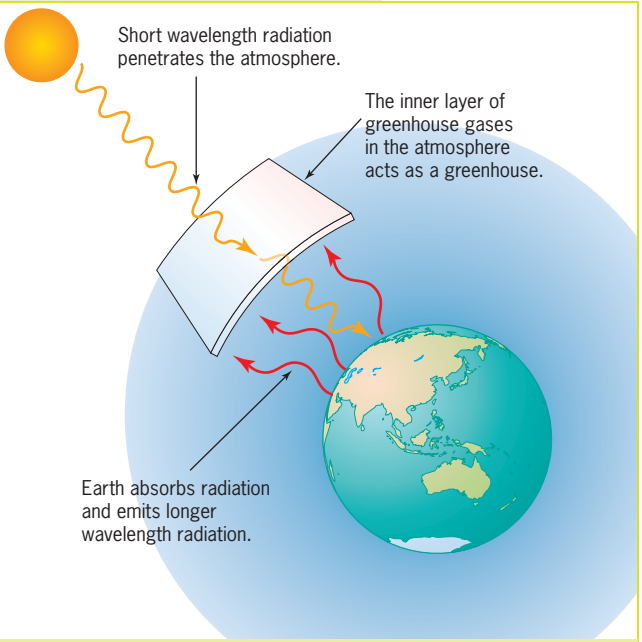


Figure 5.3.2 The Sun emits radiation that keeps the Earth warm and able to sustain life.

The greenhouse effect is a naturally occurring phenomenon and without it the Earth's temperature would be about 45°C lower than it is now. Just how bad is the greenhouse effect? Should it be getting all of this negative attention? For example on Venus, a planet that has an atmosphere comprised of nearly 97% carbon dioxide, compared to Earth's 0.04%, the average surface temperature is 500°C.

Greenhouse gases

Carbon dioxide

With the onset of the Industrial Revolution in the late 1780s, there was a rapid increase in the production of CO_2 as a direct consequence of the increase in the burning of fossil fuels. Additionally, due to a recent surge in tropical deforestation, millions of hectare of tropical forests are burned annually, resulting in increased levels CO_2. Deforestation also means that less CO_2 is removed by the trees during photosynthesis. Ice core samples from the Antarctic show that present CO_2 levels are as high as they have ever been in the Earth's history.

AS **E.3.2**
List the main greenhouse gases and their sources, and discuss their relative effects.

Methane

An increase in human population means more of the Earth has been turned into rice paddies and cow pastures to provide food for the increasing population. A by-product of growing rice in warm, waterlogged soil is methanogenesis—the production of methane gas—one of the major gases contributing to the greenhouse effect. Raising more cows for human consumption means that as the cows digest their food they are also responsible for the production of millions of tonnes of methane each year.

Figure 5.3.3 The destruction of rainforests sush as these in Brazil is reducing the ability of our Earth to absorb CO_2.

Figure 5.3.4 Rice paddies like these have some responsibility for the increase in global methane levels.

Water vapour

With an increase in the Earth's temperature, water is being evaporated from lakes, rivers and oceans, and this means that there is an increased concentration of water vapour in the atmosphere; however, it is ineffective at absorbing longer wavelength IR light and so has little effect as a greenhouse gas.

Nitrous oxide

A greenhouse gas that has much more influence on the increase of the greenhouse effect is nitrous oxide, N_2O. Known in the world of dentistry as 'laughing gas', N_2O plays a large role in trapping radiated energy from the Earth's surface. N_2O enters the atmosphere as a by-product of the nitrification and denitrification process. Nitrogen that has been taken up from the soil by plants or ingested by animals is returned to the soil when these organisms die. Bacteria act upon this biomass (dead plants and animals) and produce N_2O, which is released into the atmosphere.

Chlorofluorocarbons

The greenhouse gases that are most effective at trapping the radiated IR energy are the chlorofluorocarbons, CFCs. CFCs have been widely used as coolants in refrigeration and air conditioners, as solvents in cleaners, as foaming agents in fire extinguishers and as propellants in aerosols. A CFC molecule can exist in the atmosphere for more than 100 years, therefore there is no quick fix for the removal of these gases from the atmosphere. All we can do now is limit the use of CFCs in the future. But not everyone agrees with this, and many newspaper articles worldwide tell the tale of smugglers smuggling not drugs or contraband cigarettes but illegal CFCs across borders. CFCs are still widely used as a cheap coolant chemical in air conditioners and refrigerators.

CHEM COMPLEMENT

Cow power

A popular urban myth is that cows are contributing to global warming through excessive flatulence. Although this is not true, our bovine friends are not entirely innocent. Rather than flatulence, it is actually burping that is the problem. Due to the work of anaerobic bacteria in the rumen on the large amount of plant matter that cows consume each day, the average cow burps about 250 dm^3 of methane into the atmosphere each day. The process is known as *enteric fermentation*. Other ruminant livestock, such as sheep and goats, also contribute to the problem. Things are somewhat better at the other end of the cow's digestive tract. The methane given off by fermenting cow manure can be collected and used to generate electricity.

This is gaining popularity as a 'green' energy source. The leftover manure can be used as fertilizer.

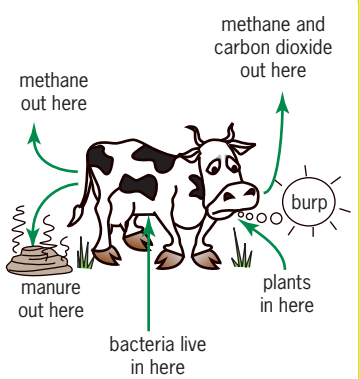


Figure 5.3.5 The cow's contribution to the greenhouse effect.

TABLE 5.3.1 A COMPARISON OF THE NATURAL AND ANTHROPOGENIC CONTRIBUTORS TO THE GREENHOUSE EFFECT. WHO IS REALLY TO BLAME?

Figures are based on concentrations (ppb), which have been adjusted for heat retention characteristics.

Greenhouse gas	% of all greenhouse gases	% Natural	% Anthropogenic
Water vapour	95.000	94.999	0.001
Carbon dioxide	3.618	3.502	0.117
Methane	0.360	0.294	0.066
Nitrous oxide	0.950	0.903	0.047
CFCs and other gases	0.072	0.025	0.047
Total	100.000	99.723	0.278

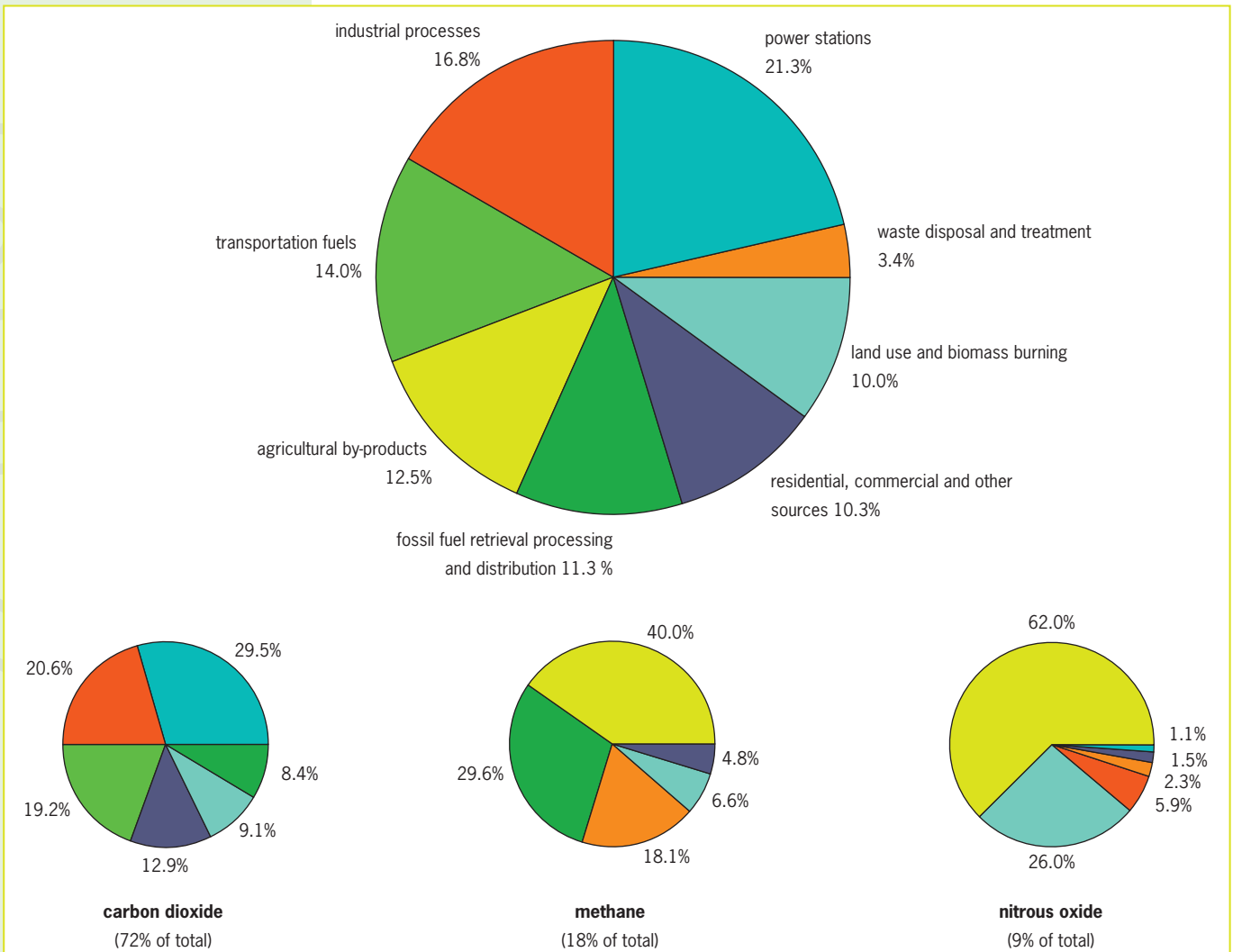


Figure 5.3.6 Annual greenhouse gas emissions by sector.

Atmospheric effects of greenhouse gases

The enhanced greenhouse effect has both beneficial and detrimental effects, some of which are outlined in table 5.3.2.

E.3.3
Discuss the influence of increasing amounts of greenhouse gases on the atmosphere. © IBO 2007

TABLE 5.3.2 POSITIVE AND NEGATIVE INFLUENCES ON THE ATMOSPHERE OF INCREASING AMOUNTS OF GREENHOUSE GASES

Positive	Negative
• Lower heating bills • Longer growing season	• Air conditioning usage is higher and thus more CFCs needed • Changes in precipitation patterns and global temperature • Melting of the polar ice caps • Rising sea levels meaning floods in some areas • Droughts in temperate and desert areas • Increased forest fire (bushfire) hazard • Thermal expansion of the oceans • Changes in the distribution of pests and disease-carrying organisms • Changes in biodiversity

WORKSHEET 5.1
The enhanced greenhouse effect

Section 5.3 Exercises

1 Describe the greenhouse effect and comment on the importance of changes in wavelength on this effect.

2 List the main greenhouse gases and their sources.

3 Describe what you consider to be the main contributor to the greenhouse effect.

4 Discuss the role that industrialization and an expanding global population has on the greenhouse effect.

5 Comment on the relationships shown in figure 5.3.6.

6 Using table 5.3.1 as one source of information, discuss whether or not we should be worried about anthropogenic sources of the greenhouse effect.

7 List the effects that an increasing amount of greenhouse gases has on our environment.

8 Comment on CFC smugglers.

5.4 OZONE

The term *ozone* comes from the Greek *ozien* meaning to smell. Ozone is a gas that has a pungent (strong) odour and is blue in colour. Ozone is an allotrope of oxygen (same atoms but bonded together differently) and has the chemical formula O_3. This triatomic form of oxygen is much more unstable than the diatomic gas oxygen O_2.

Ozone has two resonance structures.

Figure 5.4.1 Resonance structures of the ozone molecule.

AS E.4.1
Describe the formation and depletion of ozone in the stratosphere by natural process. © IBO 2007

QuickTime CW Video
Stratospheric ozone

Ozone formation and depletion

Depending on where ozone is found it can either be good atmospheric ozone or bad ground-level ozone. Ozone is found naturally in the upper atmosphere (stratosphere) where it acts to filter potentially damaging ultraviolet light from reaching the Earth's surface. It is formed naturally in the atmosphere via the following series of reactions:

$$O_2 + \text{light energy} \rightarrow 2O\cdot$$
$$O\cdot + O_2 \rightarrow O_3$$

The oxygen atom produced in the first step is a free radical—it has an unpaired electron (shown by the · symbol) and thus is very unstable and very reactive. This free radical can react with more oxygen in the atmosphere to produce ozone.

When O_3 absorbs ultraviolet light in the upper atmosphere, it denatured by a particular wavelength of light and becomes depleted:

$$O_3 + \text{light energy} \rightarrow O_2 + O\cdot$$
$$O_3 + O\cdot \rightarrow 2O_2$$

Harmful ground-level ozone is produced by reactions of sunlight and other air pollutants (e.g. NO_x and VOCs) and is responsible for the hazy appearance over pavement during warm summer days. This ozone is a constituent of smog (see section 5.10). It has harmful effects on the respiratory systems of humans and animals.

AS E.4.2
List the ozone-depleting pollutants and their sources. © IBO 2007

Ozone-depleting pollutants—CFCs and NO_x

The group of chemicals known as CFCs (see section 5.3) is the main cause of the destruction of atmospheric ozone. CFCs contain chlorine that is weakly bonded to the remainder of the molecule. As a result, the chlorine is readily removed from the CFC molecule in the upper atmosphere. Once the free chlorine is released in the upper atmosphere it is free to catalyse the destruction of ozone molecules. CFCs were first developed in 1930 when they were thought to be a wonder chemical as they were stable, odourless, non-flammable, non-toxic and non-corrosive. We now know that one chlorine atom from a CFC molecule can damage up to 100 000 molecules of ozone.

A second cause of the destruction of atmospheric ozone is the group of chemicals known as nitrogen oxides (NO_x) NO and NO_2. One molecule of NO can stay in the atmosphere for anywhere from 22 to 111 years. This long time translates into the destruction of many, many molecules of ozone.

Scientists have come up with some alternatives that can replace the CFC molecule in its typical uses. They have proposed using various hydrocarbons, fluorocarbons (perfluorocarbons, PFCs) and hydrofluorocarbons (HFCs). Because PFCs and HFCs do not contain chlorine they do not directly affect stratospheric ozone and so have an ozone depletion potential of zero. Table 5.4.1 compares some properties of these alternatives.

Figure 5.4.2 CFCs are used as the foaming agent in chemical fire extinguishers.

AS E.4.3
Discuss the alternatives to CFCs in terms of their properties. © IBO 2007

TABLE 5.4.1 ALTERNATIVES TO CFCS AND THEIR PROPERTIES

	CFC alternative		
	Hydrocarbons	**Fluorocarbons (perfluorocarbons)**	**Hydrofluorocarbons**
Examples	Propane, C_3H_8 Butane, C_4H_{10}	Octafluoropropane, C_3F_8 Perfluorohexane, C_6F_{14}	HFC-23, CHF_3 HFC-152a, CH_3CHF_2
Chemical formula contains	C and H only	C and F only	C, H and F only
Toxicity	Dependent on dose and route of ingestion; hydrocarbon abuse due to sniffing, bagging or huffing	Very low toxicity	Low acute toxicity
Flammability	Quite flammable	Low flammability	Low flammability
Damage to the ozone layer	No damage to ozone layer	F does not catalyse ozone destruction so therefore no damage	F does not catalyse ozone destruction so therefore no damage
Ability to absorb infrared radiation and therefore contribute act as a greenhouse gas	Can absorb	Can absorb	Can absorb

CHEM COMPLEMENT

Effects of the destruction of the ozone layer and what is being done to stop their production

The destruction of ozone in the atmosphere has translated into many problems for the inhabitants of our planet. Since the destruction of the atmospheric ozone began there have been more incidents of skin cancer and of cataracts. As well, there has been a suppression of the immune system, which means our defences against various infectious diseases have been weakened. There has also been a decrease in the yields of important food crops such as corn, rice, wheat and soybeans. Perhaps the hardest hit are oceanic single-celled plankton which have no natural defences. And as plankton are the base of the oceanic food chain, every other organism in that food web is also affected.

Then what can we do to stop the usage of CFCs and the destruction of the ozone layer? The first attempt at a global solution was the Montreal Protocol in 1987. It proposed to freeze CFC production at 1986 levels and a 50% reduction in CFCs by 1999. Initially only 24 countries signed the agreement, although there are now more than 140 signatories. The only problem is that due to the relative indestructibility of CFC molecules even a total stoppage of CFC use today would still mean 100 years of continued depletion.

Don't forget to celebrate World Ozone Day on 16 September each year, the anniversary date of the signing of the Montreal Protocol.

Section 5.4 Exercises

WORKSHEET 5.2
Ozone

1. Draw the Lewis diagrams for the two resonance forms of ozone.
2. Describe the natural formation and depletion of ozone in the atmosphere, making sure to include the relevant chemical equations.
3. Explain the relationship between ozone and air pollution.
4. Explain why one molecule of CFC can be responsible for the depletion of the ozone layer.
5. List the uses of CFCs.
6. List the sources of NO_x.
7. Discuss the effects that the depletion of the ozone layer has on the health of humans and the environment.
8. Discuss the factor that leads to CFCs harming the ozone layer and why the CFC alternatives are classified as having zero ozone destruction capability.
9. Discuss the alternatives to CFCs in terms of their properties.

5.5 DISSOLVED OXYGEN IN WATER

Aquatic plants and animals rely on oxygen gas that has been dissolved in water for survival. Oxygen enters the water from the surrounding air, as a by-product of photosynthesizing bacteria and through aeration (i.e. water flowing through rapids). Aeration rarely occurs in lakes, reservoirs and estuaries, and as a result stratification (layering) of various temperatures of water occurs.

Figure 5.5.1 Aeration occurs where water has a fast flow rate.

Figure 5.5.2 Lake trout require relatively high levels of dissolved oxygen and that is why you will find this species at greater depths and why you have to cast your fishing line a little further.

Typically 5 ppm of dissolved oxygen (DO) is needed for a water source to maintain a healthy environment, although some game fish, for example lake trout, require 8.3 ppm of DO. When the levels of DO fall below 3 ppm matter, life becomes very stressful or even impossible for aquatic life.

Dissolved oxygen levels change and vary according to the seasons and time of day. Solubility of gases decreases with increased temperature (an inverse relationship), therefore dissolved oxygen levels are lowest during the summer months when ambient temperatures are the highest. Additionally, due to global warming, fish stocks are declining or changing their normal habitats. In a quest to find cooler water that has higher levels of DO, fish are moving to greater ocean depths where the more dense cold-temperature waters are found. When DO levels fall below the required value for life, the water is said to be polluted.

Biochemical oxygen demand

AS **E.5.1**
Outline biochemical oxygen demand (BOD) as a measure of oxygen-demanding wastes in water. © IBO 2007

Besides global warming, a second reason for DO depletion is organic-demanding wastes. Oxygen-demanding wastes include bacteria that make use of DO when they decompose dead plant and animal matter and human waste. The amount of DO needed by aerobic decomposers to break down the organic materials in a given volume of water at a certain temperature over a specified time period is called the **biochemical oxygen demand** (BOD). The cleaner the water source, the lower the level of BOD. Clean water has a BOD level of about 1 ppm, whereas any water that is considered polluted has levels above 5 ppm matter. Because they flow, most rivers recover rapidly from an excess of oxygen-demanding wastes.

Pr
PRAC 5.2
Measuring DO

Testing for BOD is now mostly automated; however, it used to be measured using the **Winkler method**. First a sample of water is collected, ensuring that no space is left in the sample bottle for contamination by an outside air source. Second, the amount of oxygen in the water sample is fixed by using manganese sulfate and an alkaline–iodide reagent.

$$2Mn^{2+}(aq) + 4OH^{-}(aq) + O_2(g) \rightarrow 2MnO_2(s) + 2H_2O(l)$$

Formation of a brownish floc of manganese(II) oxide indicates the presence of oxygen in the sample. Then concentrated sulfuric acid is added, removing this floc and fixing the amount of oxygen in the sample.

$$MnO_2(s) + 2I^{-}(aq) + 4H^{+}(aq) \rightarrow Mn^{2+}(aq) + I_2(aq) + 2H_2O(l)$$

At this stage the darker brown the solution is, the higher the concentration of $I_2(aq)$ and therefore the higher the amount of DO. To quantify the amount of DO, a titration is performed, using a standardized solution of sodium thiosulfate and starch as an indicator. The titration is stopped when the blue colour of an iodine–starch complex first disappears.

$$I_2(aq) + 2S_2O_3^{2-}(aq) \rightarrow S_4O_6^{2-}(aq) + 2I^{-}(aq)$$

The greater the volume of sodium thiosulfate used the greater the level of DO in the sample.

Aerobic versus anaerobic decomposition

AS **E.5.2**
Distinguish between aerobic and anaerobic decomposition of organic material in water. © IBO 2007

Aerobic decomposition is the most common decomposition process. The decomposition of organic matter uses oxygen in the process and, as a result, oxides or oxyanions are produced. These processes can be examined in terms of electrochemistry. Elements that are involved in aerobic decomposition are oxidized or lose electrons. For example, in the decomposition of carbon-containing waste, the oxidation number of carbon changes from 0 to +4.

$$C(s) + O_2(g) \rightarrow CO_2(g)$$

Figure 5.5.3 Marshes like this one are often the site of anaerobic decomposition.

In a similar process, the oxidation number of phosphorus changes from 0 to +5.

$$P_4(aq) + 8O_2(g) + 12e^- \rightarrow 4PO_4^{3-}(aq)$$

However, not all decomposition takes place aerobically. When there is not enough oxygen present or when decomposition is done by organisms that do not require oxygen, the decomposition process is called **anaerobic decomposition**. This process involves reduction or gain in electrons. For example, in the anaerobic decomposition of carbon-containing waste, the oxidation number of carbon changes from 0 to –4.

$$C(s) + 2H_2(g) \rightarrow CH_4(g)$$

In a similar process, the oxidation number of sulfur changes from 0 to –2.

$$S_8(s) + 8H_2(g) \rightarrow 8H_2S(g)$$

Dihydrogen sulfide, H_2S, is a foul-smelling gas that has a distinctive rotten egg smell. Because anaerobic decomposition occurs in areas such as marshlands, these areas are often associated with a foul smell.

CHEM COMPLEMENT

Methyl mercaptan—arguably the worst smell in the world

Mercaptans are thiols, organic compounds that are similar in structure to the alcohols but in which the oxygen in the hydroxyl substituent is replaced by a sulfur atom. Mercaptans form as one of the by-products of anaerobic decomposition. It is the sulfur atom that is responsible for the foul odour. Methyl mercaptan has the formula of CH_3SH, which is similar to that of methanol. Methyl mercaptan smells much like rotten boiled cabbage. It may be formed by bacteria in our mouths and results in really bad breath. We can easily detect methyl mercaptan when someone with bad breath speaks to us, however we cannot smell the gas we produce ourselves. Our feet also are home to the bacteria that produce methyl mercaptan. As a beneficial use, methyl mercaptan in low concentrations is often added to propane to make the detection of gas leaks even easier.

Eutrophication

AS **E.5.3 Describe the process of eutrophication and its effects. © IBO 2007**

Eutrophication is the process by which lakes, estuaries and other still bodies of water receive higher than normal levels of nutrients (primarily nitrogen and phosphorus), which results in an excessive growth of plants—an algal bloom. This excessive plant growth reduces the amount of dissolved oxygen in the water source, because when these plants start to decompose they produce a great amount of oxygen-demanding wastes. This process occurs naturally over thousands of years.

Anthropogenic influences can unnecessarily cause eutrophication in a negative manner. Excessive runoff from fertilized agricultural fields and lawns causes an influx of nitrogen-containing compounds into the water source. Untreated household waste water, which has high levels of detergents, allows high levels of phosphates to enter the water system. These nitrogen- and phosphorus-containing compounds result in the increase in plant growth that adversely affects the aquatic ecosystem by reducing the amount of DO and essentially suffocating aquatic life.

Pr PRAC 5.3 Determination of phosphate in washing powder

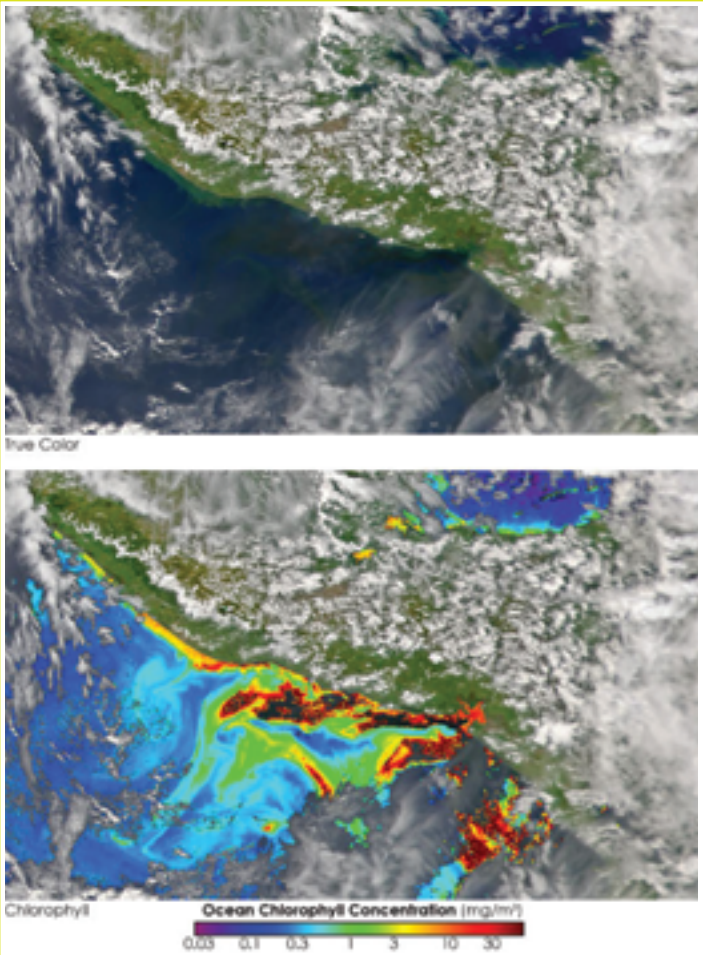


Figure 5.5.4 The brown streak (top) is a large algal bloom off the coast of El Salvador. The bottom photo shows chlorophyll levels that are found in this algal bloom. This type of algal bloom may be indicative of unnaturally high levels of nitrogen- and phosphorus-containing compounds—eutrophication.

AS **E.5.4**
Describe the source and effects of thermal pollution in water. © IBO 2007

Thermal pollution

We saw at the beginning of this section that the level of DO in a water sample is dependent on the water temperature, and that the solubility of gases decreases with increased temperature. As the temperature of water increases, the metabolic rates of fish increase, which means that they have to consume more food, and this can result in a food shortage. As little as a 1–2°C change in temperature can have drastic effects on the ecosystem. This increase in temperature can be caused by anthropogenic means, hence leading to the term **thermal pollution**.

One of the most frequent causes of thermal pollution occurs when water is used as a coolant in industrial processes and then this heated water is released into the environment. Thermal pollution also occurs when unnaturally cold water is released into the environment and results in the death of a large number of fish. Tropical aquarium fish owners must keep the water from getting too warm or risk thermal pollution in their home aquariums.

By cutting down trees, natural shade is removed from the lake, therefore increasing the water temperature, due to the greater amount of direct sunlight now reaching the water's surface.

Figure 5.5.5 Thermal pollution, water pollution and eutrophication are all causes of fish kills.

Section 5.5 Exercises

1. Outline the importance of DO in the water supply and describe how DO naturally enters the water.
2. Explain the relationship between global warming and DO.
3. Describe what is meant by BOD.
4. Describe the Winkler method of detecting BOD.
5. Distinguish between aerobic and anaerobic decomposition. Use appropriate chemical equations to aid in your description.
6. Define the process of eutrophication, describe it and its effects upon the environment.
7. Describe thermal water pollution.

5.6 WATER TREATMENT

All known living organisms depend on water for survival. For instance the human body is made up of 60–78% water. If there is no water our organs will not work properly and dehydration will start. Depending on the person, the longest that you can go without clean water ranges from 8 to 14 days. About 71% of the Earth's surface is covered with water, most of which is found in oceans and other large bodies of water. Only 3% of the Earth's water is fresh water and therefore theoretically suitable for drinking. Unfortunately, of this 3% only 0.3% is surface fresh water that is readily available for human consumption.

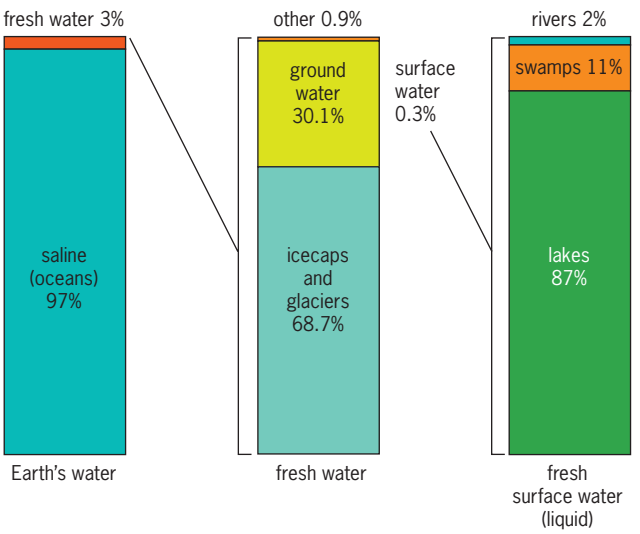


Figure 5.6.1 Distribution of Earth's water.

The World Health Organization (WHO) recognizes that drinking-water quality is an issue of concern for both developing and developed countries. More than 1.1 billion people on the planet still do not have access to safe drinking water.

CHEM COMPLEMENT

Capturing rain water on roofs

The small island nation of Bermuda is only 21 square miles in area. There are no rivers or freshwater lakes on Bermuda, and the only source of drinking water is rainfall. As a result, drinking water is collected on the roofs of all buildings (by law) and in special catchment areas, and stored in underground tanks for each home or property. Roofs are cleaned with a diluted bleach mixture and then whitewashed. As a result, the only purification process for water collected from roofs is the mixing of a small amount of bleach into the underground tanks. During times of drought, water conservation is mandatory.

Figure 5.6.2 Notice the bright white roofs of this school. The roofs have specially designed troughs to funnel the rainwater to underground cisterns

Waste water pollutants

AS **E.6.1**
List the primary pollutants found in waste water and identify their sources.
© IBO 2007

Waste water is any water that has been adversely affected in quality by anthropogenic influences. There are several examples of anthropogenic contaminants that can be found in water supplies. These contaminants enter the water supply by a variety of different means and have a variety of effects on human health and the quality of the water supply.

Before waste water is suitable for human consumption, it must be thoroughly treated and any contaminants removed. Although waste water should be treated before being used for human consumption, water treatment is sporadic or non-existent in some areas of the world. In areas where there is no treatment available, users may have to resort to drinking the same water that is used for sewage runoff, bathing, clothes washing, agricultural and animal tending. In 2006, the World Health Organization produced the third edition of the *Guidelines for Drinking Water Quality*, which describes a 'Framework for Drinking-water Safety' and discusses the roles and responsibilities of national regulators, suppliers, communities and independent 'surveillance' agencies.

AS **E.6.2**
Outline the primary, secondary and tertiary stages of waste water treatment, and state the substance that is removed in each stage. © IBO 2007

In order to make water safe for human consumption it first must be treated. There are three steps to the treatment process—primary, secondary and tertiary—steps involve filtration, sedimentation, **flocculation** and precipitation processes.

TABLE 5.6.1 PRIMARY WATER POLLUTANTS, THEIR SOURCES AND THEIR EFFECTS

Primary pollutant	Source	Effects	Water quality
Pesticides (DDT, herbicides, fungicides)	• Run off from agricultural applications • Run off from municipal applications on lawns, flower gardens, etc.	• Carcinogenic • Birth defects • Neurological disorders	• Polluted • Unsafe to drink at elevated levels
Dioxins	• Natural combustion processes, e.g. forest fires • Volcanic eruptions • Incineration of industrial (hospital) and household waste • As a herbicide, e.g. agent orange, which was used to clear bush during the Vietnam War	• Chloracne, a severe form of skin disease • Reproductive and developmental effects • Liver damage and cancer	• Polluted • Unsafe to drink at elevated levels
Polychlorinated biphenyls (PCBs)	• Environmental cycling of old PCBs (old sources of PCBs include electrical transformers or capacitors) • Many industrial processes such as plasticizers, adhesives, etc.	• Acne-like breakouts • Hearing and vision problems • Irritation of gastrointestinal tract • Affect reproductive efficiency • Liver damage and cancer	• Polluted • Unsafe to drink at elevated levels • Concentrated in fish
Organic matter	• Decaying vegetation • Decaying benthonic (sediment) matter • Decaying organisms	• Retards the photosynthesis process in plants	• Polluted • Ripe for the growth of bacteria • Unsafe to drink at elevated levels
Nitrates	• Fertilizers • Animal wastes • Septic tanks • Decaying plant material • Acid deposition	• Babies can develop methemoglobinemia or 'blue baby syndrome' caused by a lack of oxygen in the blood • Carcinogenic	• Polluted • Unsafe to drink at elevated levels
Phosphates	• Human sewage • Agricultural run-off from crops • Pulp and paper industry • Chemical and fertilizer manufacturing • Detergents	• Eutrophication accelerates plant and algae growth • Kidney damage • Osteoporosis	• Polluted • Unsafe to drink at elevated levels • Necessary for good human health at lower levels

TABLE 5.6.2 HEAVY METAL WATER POLLUTANTS, THEIR SOURCES AND THEIR EFFECTS

Heavy metal pollutant	Source	Effects	Water quality
Copper	• Copper pipes	• Anemia • Liver and kidney damage • Stomach and intestinal irritation	• Polluted • Unsafe to drink at elevated levels
Lead	• Lead pipes • Batteries • Leaded gasoline (petrol) • Paints • Ammunition	• Negatively affects hemoglobin production • Damage to kidneys gastrointestinal tract, joints and reproductive system • Can lower IQ levels in young children	• Polluted • Unsafe to drink at elevated levels
Cadmium	• Smelting • Improper disposal of rechargeable batteries	• Kidney failure • Lung disease • Osteoporosis	• Polluted • Unsafe to drink at elevated levels
Nickel	• Power plants • Waste incinerators • Improper disposal of batteries	• Decreased body weight • Damage to the heart and liver	• Polluted • Unsafe to drink at elevated levels
Zinc	• Mining • Smelting • Steel production	• Anemia • Damage to nervous system and pancreas	• Polluted • Unsafe to drink at elevated levels
Arsenic	• Natural deposits in the Earth • Industry • Agricultural processes	• Carcinogenic • Stomach pain • Numbness • Blindness	• Polluted • Unsafe to drink at elevated levels
Mercury	• Improper disposal of batteries, thermometers and barometers • Various industrial processing	• Tremors • Gingivitis • Spontaneous abortion • Damage to the brain and central nervous system	• Polluted • Unsafe to drink at elevated levels

Water treatment

Primary treatment

Primary treatment is the stage in which filtration, sedimentation and flocculation take place and results in most of the solid materials being removed. Before any chemical treatment of water takes place, waste water is first run through a series of filters and screens to remove large solids such as sticks, dead animals and rubbish. Next it passes through a sand or grit chamber where close control of the flow rate allows for sand, gravel, stones and small grit to settle. These wastes are then disposed of in local landfills or incinerated. The water is then sent to sedimentation tanks where suspended solids settle out as sludge.

Figure 5.6.3 A typical sedimentation tank found in Australia.

Figure 5.6.4 Flocculation circular basins such as this is an important part of water purification.

To speed up this sedimentation process, a process of flocculation is initiated in which smaller suspended particles are coagulated (combined together) into larger particles that can settle in the sedimentation tanks more readily. Flocculation involves adding a mixture of calcium hydroxide and aluminium sulfate to the waste water. This combination results in the precipitation of solid particles of aluminium hydroxide that when formed carry with them the smaller suspended dirt particles, which then sink to the bottom of the sedimentation tank as sludge.

$$3Ca(OH)_2(aq) + 2Al_2(SO_4)_3 \rightarrow 2Al(OH)_3(s) + 3CaSO_4(aq)$$

In the last step in the primary treatment, fats, oils and grease are skimmed off the top of the water. The effluent is then sent for secondary treatment or discharged into the main water supply.

Secondary treatment

Secondary treatment involves a biological treatment process in which dissolved organic matter (human waste, food scraps, soaps and detergents) is removed by the activated sludge process. This process involves the use of bacteria and aerobic (oxygen-filled) conditions to degrade the organic matter. The waste water enters aeration tanks where oxygen is mixed with the water. This creates ideal conditions for aerobic bacteria to grow and flourish. The bacteria then digest the organic matter and any untreated organic matter again settles out as sludge in a second sedimentation or clarifier tank. The water is returned to the aeration tank and the aeration process is repeated until all of the organic matter has been removed.

The effluent is then discharged to be disinfected with either ozone or chlorine, or the water may be forwarded for further processing—tertiary treatment.

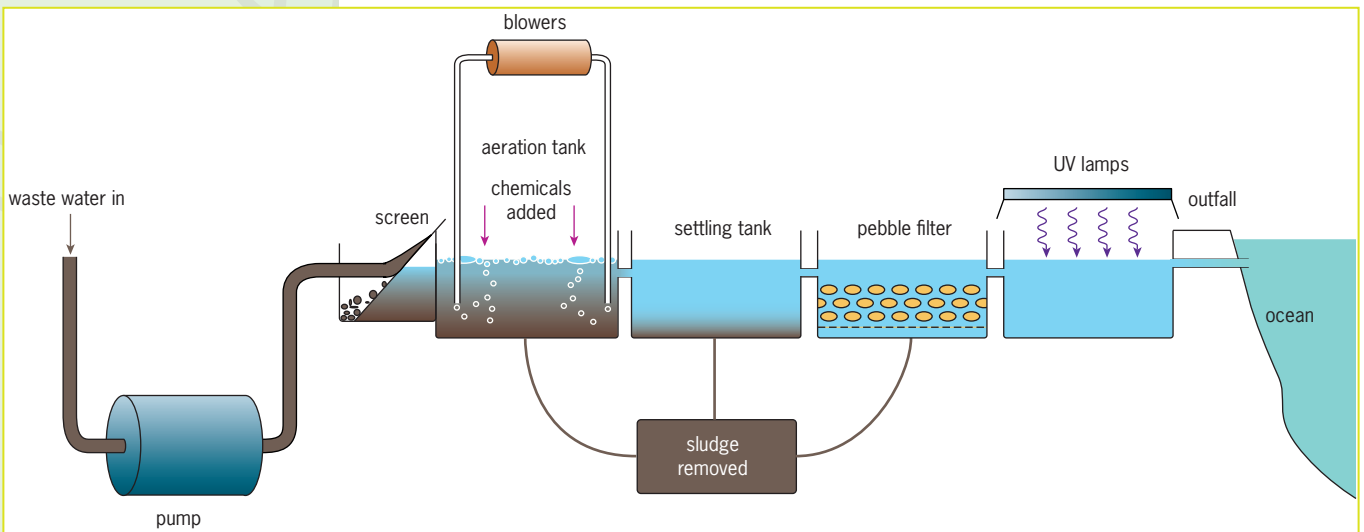


Figure 5.6.5 Schematic of aeration tanks involved in secondary treatment.

Tertiary treatment

Due to higher levels of pollution found in our water supplies, there is a greater and greater need for **tertiary treatment**, which removes heavy metals, phosphates and nitrates. These pollutants are removed via several different chemical or biological processes that we will investigate.

Heavy metals and phosphates can be removed by chemical precipitation reactions that can be both messy and expensive. For example, in removing phosphates (found in the waste water mainly from detergents and soaps) insoluble aluminium phosphate and iron(III) phosphate are produced.

$$PO_4^{3-}(aq) + Al^{3+}(aq) \rightarrow AlPO_4(s)$$
$$PO_4^{3-}(aq) + Fe^{3+}(aq) \rightarrow FePO_4(s)$$

If we examine the removal of arsenic from waste water, we see that it can be precipitated out by the addition of iron(III) salts. During this process arsenic is removed as iron(III) arsenate precipitates. Dangerous heavy metals can be removed via similar precipitation reactions.

Nitrates are harder to remove from water as their compounds are soluble, so precipitation techniques cannot be used. Ion-exchange columns or anaerobic denitrifying bacteria can be used. Ion-exchange columns involve exchanging like-charged ions in large columns filled with resin. The resin contains hydroxide ions (OH^-) that will be exchanged for nitrate ions (NO_3^-) and, as a result, the nitrates are trapped on the ion-exchange resin, while the less harmful hydroxide passes through and out of the column with the effluent. Nitrates can also be removed from the water by the addition of anaerobic denitrifying bacteria that reduce the nitrate ion into nitrogen gas. This biological process is much cheaper and therefore the preferred way to remove nitrates.

The world's water consumption rate is doubling every 20 years, due to increased population as well as greater demand from industry and agriculture. The supply of fresh water on the other hand is dwindling. We learnt at the start of this section that 97% of the water on the Earth's surface is found in the form of salt water and, although it is expensive, many countries are forced to pay for desalination in order to provide adequate and safe water for their citizens.

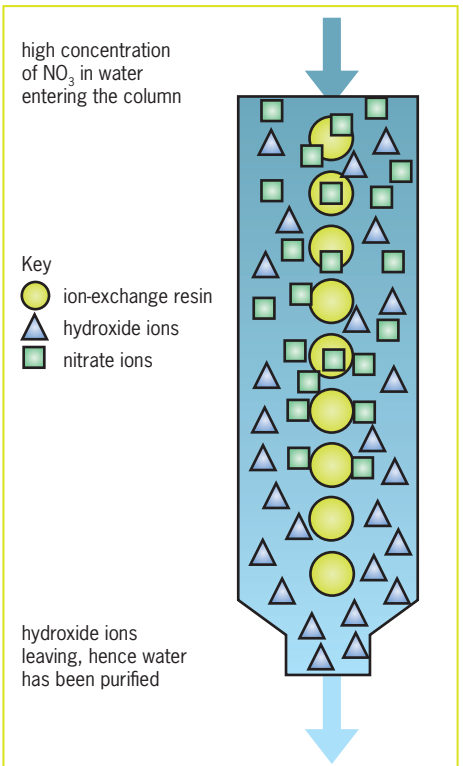


Figure 5.6.6 The process behind ion exchange which is used to remove nitrates from waste water.

TABLE 5.6.3 COMPARISON OF PRIMARY, SECONDARY AND TERTIARY WATER TREATMENT

	Primary	Secondary	Tertiary
Process used	• Filtration • Sedimentation • Flocculation	Activated sludge process	• Precipitation • Ion exchange
Pollutant removed	• Solid material • Oxygen-demanding waste	Dissolved organic matter	• Heavy metals • Phosphates • Nitrate ions

AS E.6.3
Evaluate the process to obtain fresh water from sea water using multistage distillation and reverse osmosis.

Distillation

There are more than 13 000 distillation plants in operation worldwide, most of them located in countries that have warmer climates, such as Jordan, Israel, Bermuda, United Arab Emirates, Australia and the southern US. About 60% of distillation plants are located in the Middle East. The largest one is located in the United Arab Emirates and produces more than 300 billion dm^3 of water each year.

The principle of separation by **distillation** is possible because compounds have different boiling points. Cold salt water enters the distillation chamber and passes into a heating coil. This heated salt water then enters a separate chamber that has a pressure lower than atmospheric pressure. This lower pressure causes the heated water to boil rapidly and be converted into steam. This sudden boiling is called **flashing**. It removes most of the fresh water, and leaves behind a briny solution that has a salt concentration higher than that of normal sea water. The steam that is formed during the flashing process condenses on the cooler salt water coil and fresh distilled water is produced.

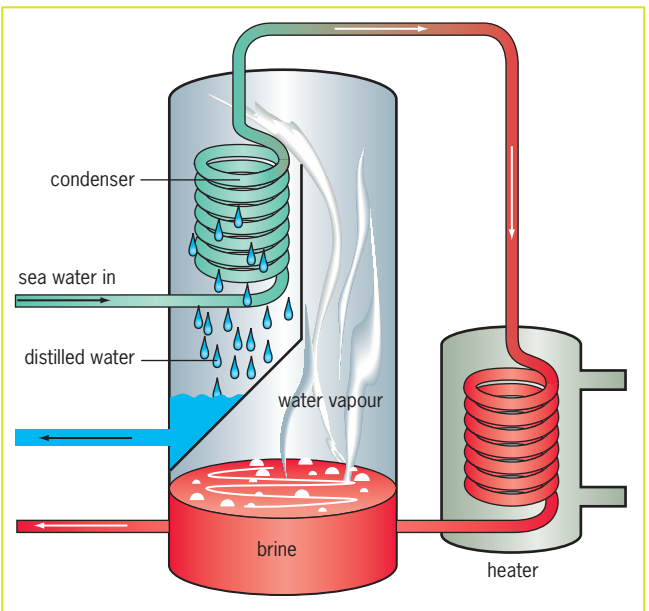


Figure 5.6.7 Flash distillation plants are used to desalinize sea water.

Interactive CW Animation
Semipermeable membrane

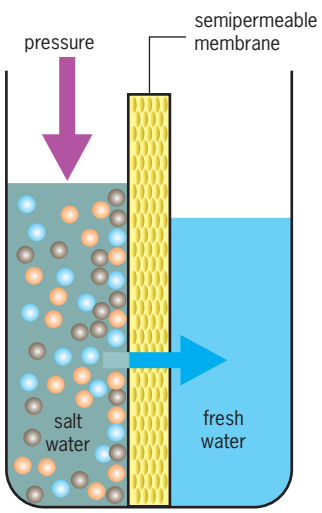


Figure 5.6.8 Diagram of desalination using the method of reverse osmosis.

Reverse osmosis

Osmosis is a natural phenomenon. It is defined as the tendency of a solvent to pass through a **semipermeable membrane** from an area of low solute concentration to an area of higher solute concentration until the concentrations on both sides of the membrane are equal. (A semipermeable membrane is one that allows solvent but not solute to pass through.) Osmotic processes enable plants to absorb nutrients from the ground. Our kidneys purify the blood in our bodies by means of osmosis. It wasn't until the 1960s that the process of osmosis was used in water desalination.

Water desalination uses **reverse osmosis**—the opposite of natural osmosis. During the process of reverse osmosis, a high pressure is applied to the side of the membrane where salt water is located. This high pressure forces the solvent, water, through the membrane, leaving the solute, salt, behind.

Pressures of 6000–8000 kPa are used, depending on the salinity values of the salt water. The major problem with these high pressures, which are roughly 600–800 times atmospheric pressure, is finding a membrane that can withstand these pressures. Typically the membranes that are used in reverse osmosis desalination plants are made from cellulose ethanoate, aromatic polyamides or thin-film polymer composites, which are extremely strong.

Section 5.6 Exercises

1. Explain why fresh water is such an important commodity.
2. List the primary pollutants found in waste water and identify their sources.
3. List the heavy metal water pollutants and identify their sources.
4. Outline why waste water treatment is so important to human health. Comment on what we can do for areas of the world that do not have access to waste water treatment.
5. Discuss the relationship between economics and health in terms of industry construction and water pollution. Which one is more important?
6. Outline the process of primary treatment of waste water including the reasons for, the process itself, the important chemical equations involved and the pollutants removed.
7. Outline the process of secondary treatment of waste water including the reasons for, the process itself, and the pollutants removed.
8. Outline the process of tertiary treatment of waste water including the reasons for, the process itself, and the pollutants removed.
9. Identify the stage of waste water treatment you consider to be the most important.
10. Describe the process of obtaining fresh water from sea water by distillation and reverse osmosis, and evaluate these methods.

CHEM COMPLEMENT

Other uses of reverse osmosis

There are numerous uses of the reverse osmosis process other than in the desalination process. Industry makes use of reverse osmosis to purify water used in the production of semiconductors and various foods (essential oil extraction, maple syrup) and beverages (various citrus juices and wine). The technique is also modified for use in the electroplating process. The dialysis machines used to treat people with kidney failure also make use of the principles of reverse osmosis.

5.7 SOIL

According to the Central Intelligence Agency *World Fact* book, about 40% of the global work force is involved in agriculture, although this contributes to only 4% of the gross domestic product. An important aspect of agriculture is having rich fertile soil in which to grow crops. Crops include cotton in the southern USA, coffee beans in Brazil, tulips in The Netherlands, cassava in Zanzibar and wheat in Australia. Because such a large percentage of the world's population depends upon agriculture to make a living, we must ensure that there will be fertile soil in the years to come.

There is much more to soil than simply calling it dirt. The nutrients in the food you eat come from soil. The crops grown in soil are vital not only as a source of nutrition but also for the production of consumer products—wood, paper and fabrics. Soil also helps to purify the water we drink. Decomposers in soil help to rid us of our waste. Soil also supports our homes, shopping centres, schools and roads. Soil is vital to our existence.

Soil degradation

AS E.7.1 Discuss salinization, nutrient depletion and soil pollution as causes of soil degradation. © IBO 2007

Soil degradation takes place when actions result in the soil being unhealthy or infertile. This degradation can take place naturally as well as a result of human activity. Natural sources of soil degradation include wind and water erosion. The rates of these types of erosion can be increased by improper land-use practices. For example, when small shrubs are removed to make way for farming, the roots that once held soil in place are no longer there, which means that wind and water erosion increase and valuable soil is lost. The three main anthropogenic forms of soil degradation are **salinization**, nutrient depletion and soil pollution.

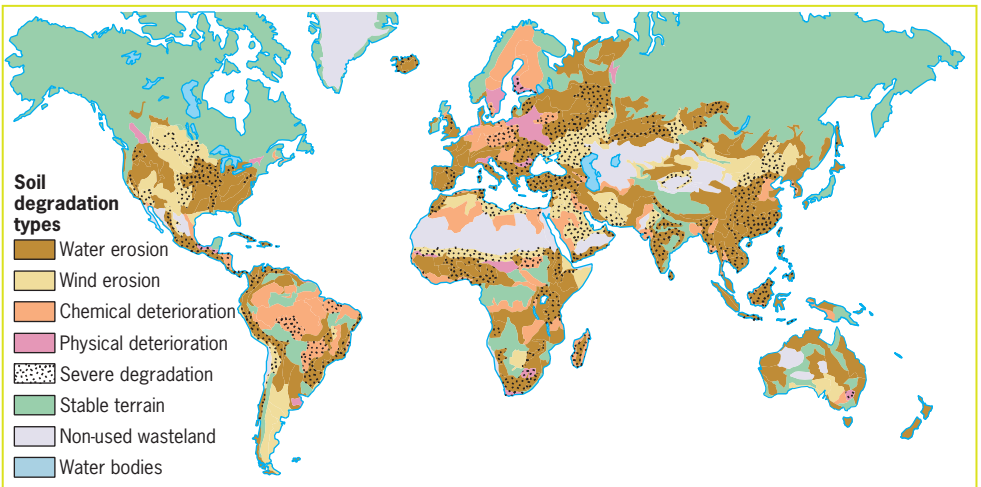


Figure 5.7.1 Map showing both natural and anthropogenic sources of soil degradation.

About 18% of the world's croplands are now irrigated. Irrigation increases crop yield, but it can also have some harmful effects. One of these is salinization, which is the accumulation of salts in soils due to excessive irrigation. As irrigation water flows over and through the ground it dissolves salts, increasing the salinity of the water. Much of the water in saline solution evaporates, leaving a high concentration of salt in the soil. In poorly drained soils, the salts are not washed away and begin to accumulate in the topsoil. Plants cannot grow in soil that is too salty.

Plants, like humans, need nutrients to grow. Nitrates and phosphates are produced naturally by the soil food web (the community of organisms in the soil). However, due to over-farming or bad irrigation practices, the soil can soon be stripped of its nutrients and becomes dead soil. Farmers can mimic the natural soil food web by administering large doses of artificial fertilizers, which, as we have seen, can lead to various forms of water and air pollution. To avoid soil nutrient depletion, farmers observe a practice of crop rotation, cropping only certain areas while other areas remain fallow. The fallow areas would be allowed to be replenished naturally by the soil food web, and could be helped by the addition of various organic materials such as compost or manure, which are rich in the nutrients needed by the soil. These practices can be difficult to maintain on a large scale and when profits are the bottom line. Conservation of soil nutrients must be kept in mind if we want to have a sustainable agricultural economy in the future.

Figure 5.7.2 In crop rotation one field is left fallow while crops are grown in another field.

A third major cause of soil degradation is **soil pollution**, which can be caused by an overuse of chemicals such as pesticides and fertilizers. Because of soil nutrient degradation and unsuitable land-use practices, farmers must supplement the soil by adding fertilizers. These fertilizers can disrupt the soil food web, reduce the soil's biodiversity and ultimately ruin the soil. Pesticides are used to rid the crops of pests and therefore to increase crop yield; however, these pests soon become genetically resistant to the pesticides. Fertilizers and pesticides can also leach into and pollute our water supplies.

Soil organic matter (SOM)

Soil is a complex mixture of inorganic minerals, decaying organic matter, water, air and billions of living organisms. The topsoil layer is usually a porous mixture of partially decomposed organic matter (humus), living organisms and some inorganic minerals particles. Topsoil is usually quite dark in colour and loosely packed. The dark colour of the topsoil helps to absorb heat from the Sun and thus helps to warm cold soil, creating a better environment for plant growth.

AS **E.7.2**
Describe the relevance of the soil organic matter (SOM) in preventing soil degradation, and outline its physical and biological functions. © IBO 2007

The roots of plants and most of the soil's organic matter are concentrated in topsoil. The term **soil organic matter** (SOM) is generally used to represent the organic constituents in the soil, including undecayed plant and animal tissues, their partial decomposition products and the soil biomass. The decomposition of this organic matter results in the formation of high molar mass organic materials, such as polysaccharides and proteins, and simpler substances such as sugars, amino acids and other smaller molecules.

Much of the SOM is insoluble in water but helps to retain water and water-soluble plant nutrients so that they can be taken up by plants roots. Nutrients include inorganic minerals such as phosphorus, nitrogen and sulfur, which are needed for healthy plant growth. As well as acting as a source of nutrients that are important for biological reasons, SOM has many other important physical functions. There are spaces or pores that exist between the SOM which act to hold water (much like a sponge), oxygen and nitrogen. Some of the SOM along with fungi create aggregates that allow for a more stable structure of the soil and are directly linked to better aeration and water filtration. This creates a soil layer that is more resistant to erosion; therefore, soils with greater amounts of SOM have less degradation.

Figure 5.7.3 Healthy soil is a very important aspect of having a vibrant garden.

Soil pollution

Soil pollution can also occur when harmful air pollutants settle onto the topsoil, get turned under and contaminate the soil food web. These contaminants can be released into our water supplies and back into the atmosphere in the form of particulate matter. Soil pollution can also result when hazardous waste from industry is placed in landfills and leaches into the surrounding soil.

Table 5.7.1 lists the most common organic soil pollutants and their sources. While you are examining this table you should keep in mind the earlier sections on air and water pollution and you should see some familiar pollutants. This shows the interdependence that exists between air, water and soil environments.

AS **E.7.3**
List common organic soil pollutants and their sources. © IBO 2007

Figure 5.7.4 As well as being an eyesore, junk yards can also be a source of soil pollution when oil and gasoline from old cars leaks into the soil.

TABLE 5.7.1 ORGANIC SOIL POLLUTANTS	
Organic pollutant	**Source**
Petroleum hydrocarbons	• Leaks from transport vehicles and home oil tanks • Used to reduce dust on unpaved roads
Agrichemicals	• Overuse of pesticides, herbicides and fertilizers
VOCs and semi-volatile organic compounds, solvents	• Internal combustion engines • Emitted from solid materials • Emitted from industrial processes
Polyaromatic hydrocarbons (PAHs)	• Vehicle exhaust • Incomplete combustion of coal, oil and wood • Incineration of waste • Oil spills
Polychlorinated biphenyls (PCBs)	• Electrical transformers or capacitors as insulators • Industrial processes, e.g. plasticizers, adhesives, etc.
Organotin compounds	• Chemical used in the preservation of wood • Insecticides and fungicides

Section 5.7 Exercises

1 Describe the importance of soil in our environment.

2 Define soil degradation.

3 Explain why more traditional methods of farming such as crop rotation and terracing are disappearing from the agricultural sector. Explain why it is important to maintain these traditions despite economic pressure for greater crop yield.

4 Define soil organic matter and explain its link to soil health.

5 Explain the biological and physical function of soil organic matter.

6 Describe why good humus being black in colour is beneficial.

7 List organic soil pollutants and give their sources.

5.8 WASTE

Figure 5.8.1 Garbage truck bringing another load to a landfill.

Anything you throw away in your rubbish bin is called waste, be it paper, a banana peel or old shoes. In 10 000 years your waste may be of interest to an archeologist, but now one of the only people interested in your waste is the sanitary engineer—the garbage man.

As the world population and the standard of living increases, larger and larger amounts of waste are produced—billions and billions of tonnes each year. Many societies are centred on consumerism, and a sofa with a worn cover is replaced as easily as a pair of worn-out jeans. This focus on consumerism means that more and more waste is being generated on a daily basis.

Waste disposal

Once your garbage is driven away in the garbage truck there are two main ways in which it is disposed—buried in landfill or incinerated. Modern landfills are complex engineering feats that have the advantage of being able to deal with massive amounts of waste in an efficient manner. Once filled and covered over they can be used for future building sites. However, until then they have the disadvantage of being an eyesore and a breeding ground for rats and insects. Additionally, as the garbage decomposes anaerobically foul-smelling methane is produced as one of the by-products. Not only is methane unattractive to the nose, it is also unattractive to the environment—it is a greenhouse gas.

E.8.1 Outline and compare the various methods for waste disposal. © IBO 2007

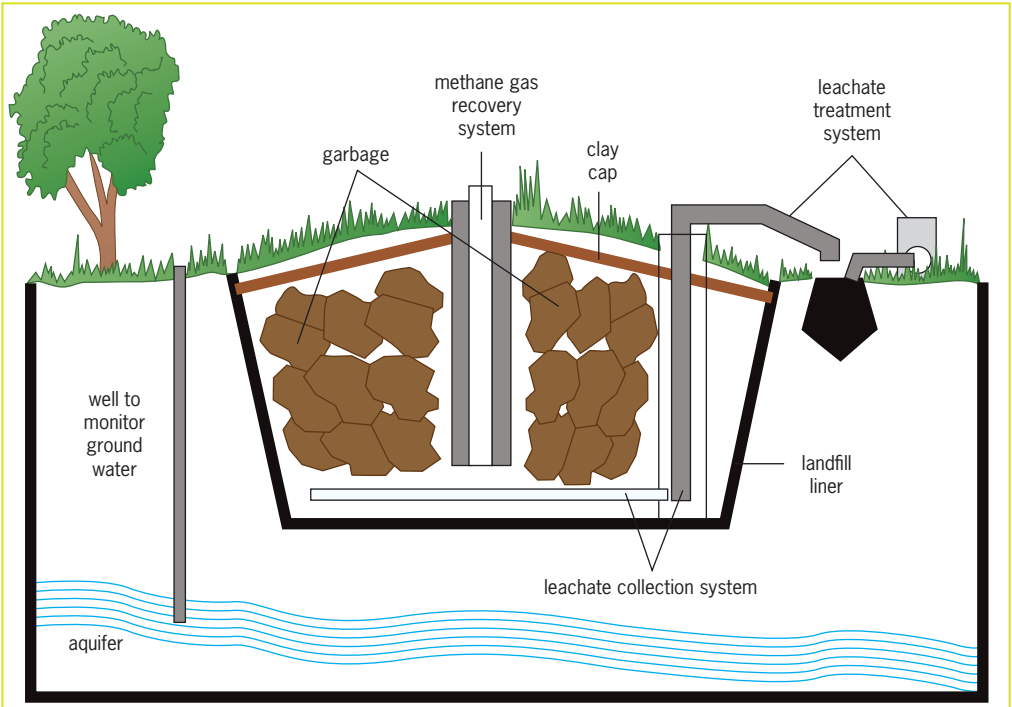


Figure 5.8.2 A diagram of a modern landfill. Garbage is no longer just thrown into a pile and forgotten about.

DEMO 5.1
A biogas generator

Careful engineering is needed so that the leachate, the liquid that drains from the landfill, is collected and treated before being released. Many developing countries still do not have the resources to construct safe and efficient landfills. In some of these countries there are people who survive by sorting through the waste of others in landfills. People are becoming more proactive and ways are being found to use landfills beneficially. In Durban, South Africa, for example, landfill gas is converted into electricity.

The second most common means of disposing of waste is through the process of incineration, which offers a more compact way of dealing with waste. Incineration greatly reduces the volume of waste through high-temperature burning. The initial construction can be quite costly and the remaining ash has to be disposed of in traditional landfills. Burning waste produces large amounts of energy that can be converted to generate electricity. However, incinerators can also produce carbon dioxide, sulfur dioxide and other potentially hazardous air pollutants. We have already seen that these pollutants can be reduced by the use of scrubbers, fluidized beds and filters.

AS E.8.2
Describe the recycling of metal, glass, plastic and paper products, and outline its benefits. © IBO 2007

Recycling

One of the best ways to minimize the influence waste has on our environment and to provide a sustainable environment is to **recycle**. Recycling is one part of resource recovery that involves salvaging usable metals, glass, plastic and paper products from municipal solid waste and selling them to manufacturing industries. This type of waste disposal is seen as a viable economic resource rather than a burden. Recycling can be an expensive way of dealing with problematic waste and not all waste can be recycled; separating the different recyclables can be time consuming and difficult.

TABLE 5.8.1 RECYCLED MATERIALS AND THE ADVANTAGES OF RECYCLING

Resource	Advantages of recycling
Metal	• Metals do not lose their properties when recycled. • The most commonly recycled metals are aluminium and steel. • Recycling aluminium requires only 5% of the energy and produces only 5% of the CO_2 emissions of production from raw materials. • Aluminium metal is so valuable that it makes no sense to landfill it. Once aluminium is sorted, it is shredded and melted and then used in the formation of new aluminium products. • Ferrous metals (made from iron) are separated by the use of magnets. • Steel is melted and turned into new car parts, cans and structural components.
Glass	• Glass is first sorted by colour then washed, crushed, melted and turned into new glass. • To make glass from raw materials takes a massive amount of energy; however, the amount of energy required is greatly reduced as the glass goes through the recycling process. • Glass does not deteriorate during the recycling process, therefore it can be recycled indefinitely.
Plastic	• Many different types of plastic can be recycled; however, due to the large amount of plastic that is found in modern packaging, sorting of these plastics can be a lengthy and costly process. • Recycling of plastics involves a much more in depth process than simply burning them. • Recycling plastics causes much less pollution and requires much less energy than it takes to originally make the plastics. • There are seven main types of plastics although only three are routinely recycled: types 1, 2 and 4. See Table 5.8.2 for the details of the seven types of plastics.
Paper products	• More than 50 different types of paper are sorted and de-inked and then broken down. • These different types of paper can be recycled into newsprint, egg cartons, gift boxes, packaging material, cereal boxes, new paper, tissue paper, paper towels and toilet paper. • Recycling paper reduces the amount of waste in the landfill, saves energy and trees, and reduces the amount of pollution that is involved in producing paper from raw materials.

CHEM COMPLEMENT

Types of plastics

There are seven main types of plastics, not all of which are being recycled, but have the potential to be recycled in the future.

TABLE 5.8.2 MAIN TYPES OF PLASTICS AND EXAMPLES OF EACH

Plastic recyclable symbol	Type of plastic	Examples of uses
 Polyethylene terepthalate	Polyethylene terephthalate	Soft drink and water bottles, oven-ready meal trays
 High density polyethylene	High-density polyethylene	Milk, detergent and juice bottles; yoghurt and margarine containers
 PVC	Polyvinyl chloride	Food trays, shampoo and conditioner bottles, plastic wrap
 Low density polyethylene	Low-density polyethylene	Shrink wrap, freezer bags, rubbish and shopping bags, flexible plastic bottles
 Polypropylene	Polypropylene	Microwave containers, disposable take-away food containers, disposable cups and plates, margarine and yoghurt containers, bottle tops
 Polystyrene	Polystyrene	Meat packaging trays, disposable containers, egg cartons, plastic cutlery, packing for electronic goods
 All other resins and multi-materials	Other	Plastics that do not fall into any of the other six categories

CHEM COMPLEMENT

The 3 Rs—reduce, reuse and recycle

Rather than filling up already full landfills and further adding to air pollution woes, we should follow the three Rs of waste disposal—reduce, reuse and recycle. The following is a top ten list of things you can do to help trim down the amount of waste sent to the local landfill, as provided by the US Environmental Protection Agency.

Reduce

1 Reduce the amount of unnecessary packaging.

2 Adopt practices that reduce waste toxicity.

Reuse

3 Consider reusable products.

4 Maintain and repair durable products.

5 Reuse bags, containers, and other items.

6 Borrow, rent, or share items used infrequently.

7 Sell or donate goods instead of throwing them out.

Recycle

8 Choose recyclable products and containers and recycle them.

9 Select products made from recycled materials.

10 Compost yard trimmings and some food scraps.

Figure 5.8.3 Think of the three Rs before you place something in the rubbish bin. Your decisions can impact the future of our planet.

Radioactive waste

There are about 440 nuclear power plants worldwide, which produce about 16% of the world's electricity.

In the late 20th century, nuclear power was touted as the power source of the future. Rising costs, safety concerns, public outcry, poor management and the problem of radioactive waste disposal have combined to reduce the global usage of nuclear power as an alternative energy supply from its original projections.

Figure 5.8.4 The map shows the locations of nuclear power plants that are in operation as well as those that have been shut down, suspended or decommissioned.

Nuclear power plants have many advantages over other sources of energy. They do not contribute to air pollution or greenhouse gases, and their fuel source, uranium, is abundant enough to supply energy for the next 1000 years.

One of the major problems of nuclear power plants is the disposal of the spent fuel rods that provided the energy for the nuclear fission reactor. These spent fuel rods are sources of **high-level waste** (HLW) because they give off large amounts of ionizing radiation for a long time. These spent fuel rods must be cooled for several years in deep pools inside the plant or in special shielded storage facilities at another site. It is said that this waste must be stored for tens of thousands of years before it can be disposed of safely.

AS **E.8.3**
Describe the characteristics and sources of different types of radioactive waste. © IBO 2007

A second category of radioactive waste is **low-level waste** (LLW), which refers to waste that gives off small amounts of ionizing radiation for a short amount of time. Low-level waste is waste that may have come into contact with radioactive substances directly and therefore has become contaminated. It includes radioactively contaminated industrial or research waste such as paper, rags, plastic bags, rubber gloves, protective clothing and packaging material. Hospitals, medical schools and radiopharmaceutical schools all produce large amounts of LLW. This waste can be placed in steel drums and buried in landfills.

Figure 5.8.5 Radioactive waste must be disposed of by proper means.

Disposal of radioactive waste

There are very different disposal requirements for radioactive waste that depend on whether it is LLW or HLW. The difference in disposal methods is primarily based on the amount of radioactivity remaining in the waste. Low-level waste is the least contaminated and therefore requires less complicated means of disposal. It can simply be stored on site until it has sufficiently decayed and then disposed of with the regular waste. If this is not possible, LLW can be shipped to disposal facilities where it is packaged and buried.

AS **E.8.4**
Compare the storage and disposal methods for different types of radioactive waste. © IBO 2007

Low-level waste was also once disposed of at sea, where its radioactive effects were simply diluted by the sheer volume of the ocean. This practice is now banned by most countries.

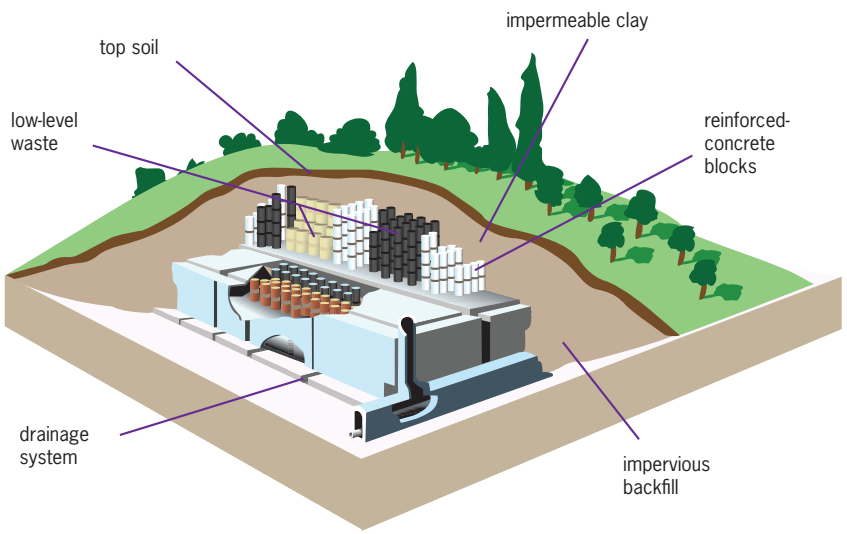


Figure 5.8.6 One type of low-level waste disposal facility.

The spent fuel that is used to fuel a nuclear power plant is initially stored in deep pools of water on the nuclear power plant site. The water in these pools acts to absorb the heat energy that is released by the spent fuel and to protect the workers from the radiation. It is safe to work around the spent fuel rods once they are under water in these pools. These spent fuel rods are then reprocessed to recover the unfissioned uranium ore. This process involves the use of strong chemicals to dissolve the fuel. This liquid waste is then classified as HLW and must be disposed of safely. The US Nuclear Regulatory Commission, for example, has designed a HLW disposal site in the Yucca Mountains in Nevada.

- Canisters of waste, sealed in special casks, are shipped to the site by truck or train.
- Shipping casks are removed, and the inner tube with the waste is placed in a steel, multi-layered storage container.
- An automated system sends the storage containers underground to the tunnels.
- Containers are stored along the tunnels, on their side.

There are concerns that this radioactive material may leach into the water table and then into drinking water.

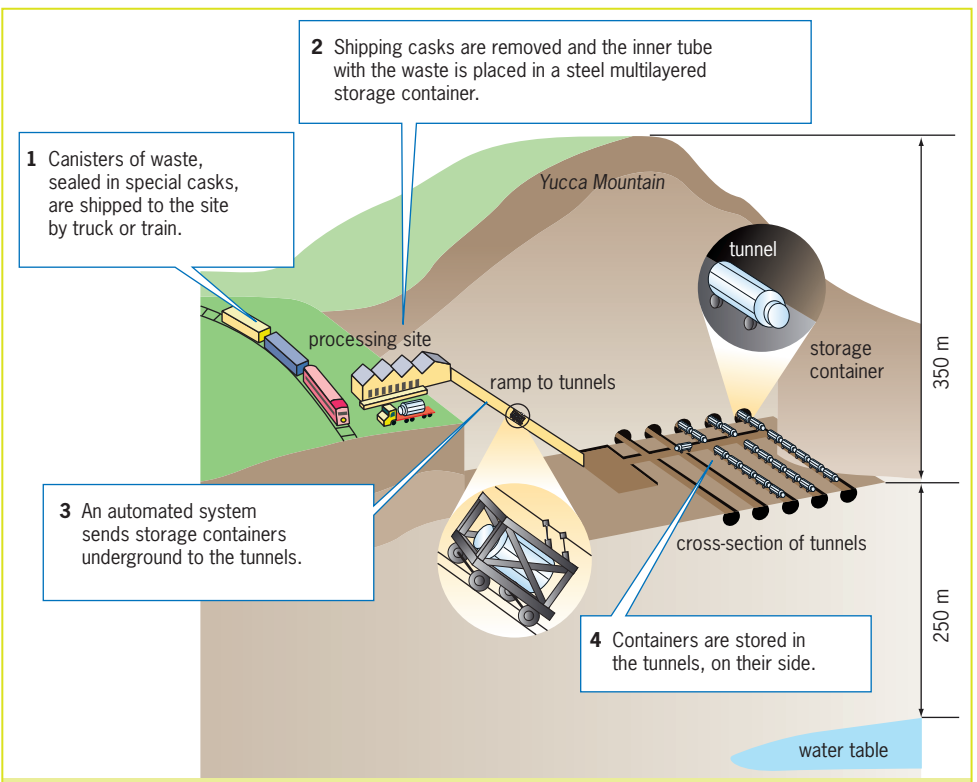


Figure 5.8.7 Conceptual design for a high-level waste disposal location in the Yucca Mountains in Nevada, US.

Section 5.8 Exercises

1. Explain how the phenomenon of consumerism is linked to the global problem of waste.
2. Compare landfill and incineration as two strategies of waste removal.
3. Describe the process of recycling metal, glass, plastic and paper products.
4. As a citizen of the world, explain what aspects of the three Rs program you can incorporate into your daily life.
5. Compare low-level and high-level radioactive wastes in terms of their definitions, sources, storage and disposal methods and effects on the environment.

5.9 OZONE DEPLETION

HL

Structure of oxygen and ozone

Ozone is an allotropic form of oxygen, but its chemistry is very different from that of diatomic oxygen. An oxygen molecule has a relatively strong double bond holding the two oxygen atoms together (figure 5.9.1). An ozone molecule has a much more unstable bonding arrangement of oxygen atoms. It consists of two resonance structures with one double and one single bond (figure 5.9.2).

Figure 5.9.1 Lewis structure of an oxygen molecule.

Figure 5.9.2 Resonance structures of ozone.

Allotropes of oxygen and their relationship to wavelength of light

The bond between the oxygen atoms in O_2 is shorter (121 pm) than that of the oxygen–oxygen bond in ozone (128 pm). The longer oxygen–oxygen bond length indicates that the oxygen–oxygen bond in ozone is weaker than that in O_2. This difference is shown in the chemical equations for the destruction of oxygen and ozone. Light of shorter wavelength, and thus higher energy, is needed to break the bond in a molecule of oxygen than in ozone.

$$O_2 + h\nu\ (242\ \text{nm}) \rightarrow 2O\cdot \qquad O_3 + h\nu\ (330\ \text{nm}) \rightarrow O_2 + O\cdot$$
$$O\cdot + O_2 \rightarrow O_3 \qquad O_3 + O\cdot \rightarrow 2O_2$$

AS **E.9.1**
Explain the dependence of O_2 and O_3 dissociation on the wavelength of light. © IBO 2007

Mechanism of ozone depletion

AS **E.9.2 Describe the mechanism in the catalysis of O_3 depletion by CFCs and NO_x. © IBO 2007**

One atom of chlorine from a CFC molecule can damage up to 100 000 molecules of O_3. The chlorine atom acts as a catalyst in the mechanism for the destruction of ozone, as shown in the reaction mechanism below:

$$CCl_2F_2 \rightarrow CClF_2 + Cl\cdot$$
$$Cl\cdot + O_3 \rightarrow ClO\cdot + O_2$$
$$ClO\cdot + O\cdot \rightarrow Cl\cdot + O_2$$

Overall reaction $O_3 + O\cdot \rightarrow 2O_2$

QuickTime CW Video
CFCs and stratospheric ozone
Catalytic destruction of stratospheric ozone

In the above reaction mechanism the chlorine radical is reproduced in the reaction and so acts as a catalyst. The ClO radical is the reaction intermediate, as it is produced and used up in the reaction. The overall effect is the destruction of a molecule of ozone by the addition of an oxygen radical, to produce two molecules of oxygen.

A second ozone-depleting chemical is NO_x, and a similar reaction occurs between NO_x and ozone. In this reaction mechanism NO acts as a catalyst and NO_2 is the reaction intermediate. The overall effect is the destruction of a molecule of ozone by the addition of an oxygen radical and the production of two molecules of oxygen.

$$NO + O_3 \rightarrow NO_2 + O_2$$
$$NO_2 + O\cdot \rightarrow NO + O_2$$

Overall reaction $O_3 + O\cdot \rightarrow 2O_2$

Polar ozone

AS **E.9.3 Outline the reasons for greater ozone depletion in polar regions. © IBO 2007**

In the polar regions, particularly over Antarctica, there seems to be greater amount of ozone depletion. There are several reasons for this. One is that world weather and wind patterns concentrate the harmful CFCs and NO_x here. Additionally, during the winter when there is 24 hour darkness, the temperature in the stratosphere drops to –80°C, which is cold enough to cause water to condense even in the extremely dry atmosphere to form ice crystal clouds. The ice crystals provide a surface on which the reaction can take place—an example of heterogeneous catalysis.

Most of the chlorine in the polar stratosphere ends up in one of the two compounds: $ClONO_2$ or HCl. These two compounds are usually inert in the stratosphere; however, the ice crystal clouds provide a surface for these compounds to react on, as shown in the reactions below:

$$HCl + ClONO_2 \rightarrow HNO_3 + Cl_2$$
$$H_2O + ClONO_2 \rightarrow HNO_3 + HClO$$

The Cl_2 and HClO are then used in the reaction for the destruction of ozone, as already shown.

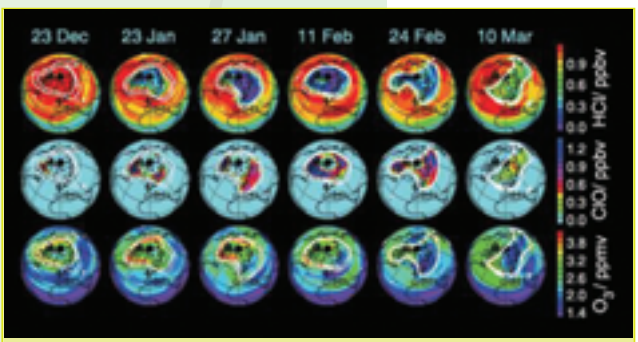


Figure 5.9.3 NASA's Aura microwave limb sounder estimates of ozone loss in the 2004–005 Arctic winter. This figure also shows the relationship between ozone and HCl and ClO.

CHEM COMPLEMENT

How do sunscreens work?

Your skin becomes sunburnt when you have been exposed to too much UV radiation, whereas it becomes suntanned if you have been exposed to much lower amounts of UV radiation. Usually the main source of this UV radiation is the Sun. Over-exposure to UV radiation may lead to one of many different types of skin cancer. The ozone layer acts naturally to filter out UV radiation from the Sun; however, as a result of the thinning of the ozone layer, more and more cases of skin cancer have occurred. Australia is skin-cancer headquarters, with one in two Australians getting the disease in their lifetime.

One method of preventing UV rays reaching the skin is to cover it with clothing. This can be uncomfortably warm in some climates, so another alternative has to be investigated. This is where sunscreens come in. Sunscreens containing zinc oxide or titanium oxide work to reflect the Sun's radiation so that it never reaches the skin. The problem that some people find with these sunscreens is that they appear chalky white and this look may not be desired by all! Other sunscreens have been developed that are transparent and absorb visible light.

Sunscreens have been produced that contain para-aminobenzoic acid (PABA), which blocks the Sun's more penetrating rays. Sunscreens containing PABA absorb UVB but not UVA rays. In the resonance structures of PABA, both the benzene ring and the carboxyl group allow for the bonds of this molecule to vibrate with the exact energy of UVB rays. The UVB rays are therefore absorbed and emitted as harmless IR rays.

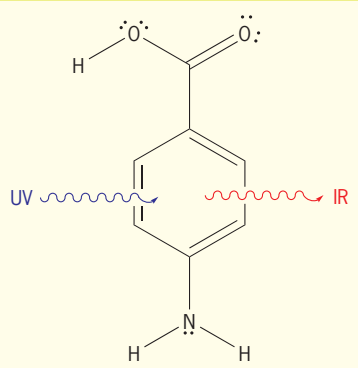


Figure 5.9.4 PABA molecule shown absorbing UV radiation

Section 5.9 Exercises

1. Compare the two allotropes, oxygen and ozone.
2. Explain the dependence of O_2 and O_3 dissociation on the wavelength of light.
3. Describe the mechanisms of ozone depletion by both CFCs and NO_x. Don't forget to mention the role of the catalyst, the reaction intermediate and the radicals.
4. Explain why there is greater ozone depletion in polar regions.
5. State the role of ice crystals in ozone depletion, and give the corresponding equations.

5.10 SMOG

HL

Smog, a word originally termed from a combination of smoke and fog, occurs over large cities such as London, New York, Los Angeles, Mexico City, Houston, Toronto, Athens, Sydney, Buenos Aries, Beijing and Hong Kong, and is a type of air pollution. Smog is formed when a series of primary and secondary pollutants interact under the influence of sunlight; specifically this is called **photochemical smog**.

AS **E.10.1**
State the source of primary pollutants and the conditions necessary for the formation of photochemical smog.

Photochemical smog

Figure 5.10.1 A blanket of smog hangs over a large urban city.

Any city that has a large number of cars on its roads and a sunny, warm and dry climate will have photochemical smog. Smaller rural communities can also be affected by smog that is carried from urban centres by strong wind patterns. Cars are one of the main sources of a primary pollutant that is responsible for smog: NO_x. Recall that nitrogen monoxide, NO, forms in the high temperatures of internal combustion engines as shown:

$$N_2(g) + O_2(g) \rightarrow 2NO(g)$$

The other main primary pollutants responsible for smog formation are the volatile organic compounds (VOCs). VOCs can also form as the result of reactions in high temperature combustion engines or can be released from other anthropogenic sources. High levels of VOCs can be produced in idling cars. In fact, leaving a car idling for 10 minutes a day translates into a waste of about 0.40 litres of petrol (gasoline) a day or 146 litres of petrol a year. It has been suggested that it is better to turn off your car than to leave it idling for more than 10 seconds, in order to save petrol and reduce the amounts of primary pollutants entering our atmosphere.

Optimum conditions for photochemical smog

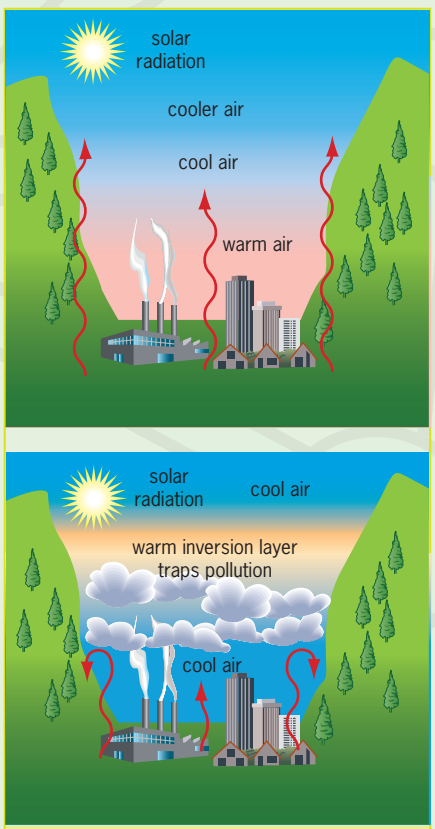


Figure 5.10.2 A thermal inversion and its role in photochemical smog formation is seen in the bottom diagram.

Usually areas with high winds and high annual precipitation of rain or snow, which will naturally cleanse the air, will have a lower occurrence of photochemical smog than cities that have low levels of wind and low levels of annual precipitation. Also, bowl-shaped cities that are surrounded by elevated land do not allow for a free flow of air, and therefore pollutants are allowed to build up at ground level. This problem is compounded in cities that have a large number of tall buildings and skyscrapers, which also reduce the air flow through the city.

During the day the Sun warms the air near the Earth's surface. Normally this heated air expands and rises, carrying the low-lying pollutants with it and in a sense cleansing the air. Cooler, cleaner air then moves down to replace this polluted air. This continual mixing process acts to cleanse air that is closer to the ground of its pollutants.

Sometimes weather conditions trap a layer of dense, cool air beneath a layer of less dense, warm air over a city that is located in a valley. This is a **thermal inversion**. This prevents the lower level air from rising and carrying pollutants with it. As a result, ground-level pollutants become more concentrated and more harmful to human health.

The most deadly one-time incident of smog-related deaths due to thermal inversion occurred in 1948 in Donora, Pennsylvania, US. Donora is a small community situated in a river valley, where thermal inversions can occur. Pollution from a nearby steel mill and zinc smelter was trapped in the lower levels of the atmosphere. This heavily polluted smog lasted for five days and 20 deaths were directly related to this incident.

Secondary pollutants and photochemical smog

The primary pollutant NO can react with oxygen in the atmosphere to create harmful nitrogen dioxide as shown by the following reaction mechanism:

$$2NO + O_2 \rightarrow 2NO_2$$
$$NO_2 + \text{sunlight} \rightarrow NO + O\cdot$$

The oxygen free radical produced in the second step is very unstable and reacts with more oxygen in the atmosphere to produce ozone. This combination requires collision with another molecule, M, to carry away the energy that would otherwise result in the ozone immediately decaying:

$$O\cdot + O_2 + M \rightarrow O_3 + M$$

where M is an inert substance that also acts as a catalyst.

We have already seen that ground-level ozone can be very harmful (section 5.4) to humans as it can aggravate asthma, cause irritation to the respiratory system and cause inflammation and damage to the lungs. It can also cause material damage to substances made of rubber, for example vehicle tyres, and destroy flora (plant life).

This oxygen radical can wreak even more havoc when it reacts with moisture in the atmosphere to form hydroxyl radicals (OH·). The hydroxyl radicals can then react with available hydrocarbons (RH) in the atmosphere to produce aldehydes (RCHO), which further react with more hydroxyl radicals to form peroxyacylnitrates (PANs), $RC(O)O_2NO_2$, which have similar adverse health effects to ozone. The chemistry of this is shown in the reactions below:

$$O\cdot + H_2O \rightarrow 2OH\cdot$$
$$RH + OH\cdot \rightarrow H_2O + R\cdot$$
$$R\cdot + O_2 \rightarrow RO_2\cdot$$
$$RO_2\cdot + NO \rightarrow NO_2 + RO\cdot \qquad \text{(NO from car exhausts)}$$
$$RO\cdot + O_2 \rightarrow RCHO + HO_2\cdot$$
$$RCHO + OH\cdot \rightarrow RCO\cdot + H_2O$$
$$RCO\cdot + O_2 \rightarrow RC(O)O_2\cdot$$
$$RC(O)O_2\cdot + NO_2 \rightarrow RC(O)O_2NO_2$$

It is interesting to see the relationship between all the players involved in the formation of photochemical smog, as shown in figure 5.10.5.

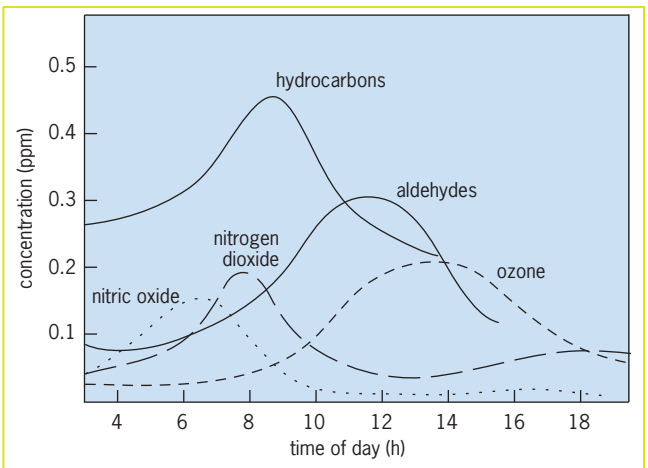


Figure 5.10.5 The components of photochemical smog and their relationships.

AS E.10.2 Outline the formation of secondary pollutants in photochemical smog.

Figure 5.10.3 Sunlight plays an integral part in the creation of ground-level ozone.

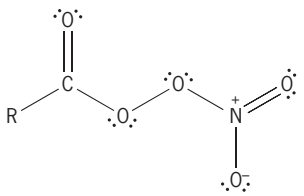


Figure 5.10.4 Structural diagram of a peroxyacylnitrate.

WORKSHEET 5.3
Ozone depletion and photochemical smog

Section 5.10 Exercises

1. Define photochemical smog.
2. List the primary pollutants responsible for the formation of photochemical smog and give the source of each.
3. Describe the optimal conditions necessary for the formation of photochemical smog.
4. Describe the phenomenon of thermal inversion and state the conditions necessary for it to occur.
5. Describe and discuss the role that secondary pollutants play in the formation of photochemical smog, giving all relevant chemical equations.
6. Explain the role that free radicals and sunlight play in the formation of photochemical smog, giving all relevant chemical equations.
7. Explain the relationships that exist in figure 5.10.5 between the primary and secondary pollutants.

5.11 ACID DEPOSITION

HL

AS **E.11.1 Describe the mechanism of acid deposition caused by the oxides of nitrogen and oxides of sulfur.**

We have already seen that power generating plants and other industrial plants give off large amounts of sulfur dioxide (SO_2), particulate matter and nitrogen oxides (NO_x), all of which are primary pollutants.

Mechanism of acid deposition formation

Aging coal-fired power plants burn sulfur-containing coal, which produces SO_2. SO_2 is converted to sulfuric acid as shown below; its production begins with the formation of hydroxyl radicals:

$$H_2O + O_3 \rightarrow 2HO\cdot + O_2$$

or

$$H_2O + O\cdot \rightarrow 2HO\cdot$$
$$HO\cdot + SO_2 \rightarrow HOSO_2\cdot$$
$$HOSO_2\cdot + O_2 \rightarrow HO_2\cdot + SO_3$$
$$SO_3 + H_2O \rightarrow H_2SO_4$$
$$SO_2 + H_2O \rightarrow H_2SO_3$$

SO_2 is also produced in the smelting of metal sulfate ores in the production of iron, nickel and steel. Its formation in the production of nickel is shown:

$$NiS + O_2 \rightarrow SO_2 + Ni$$

The SO_2 is then converted to its acidic form, as shown in the mechanisms above.

One of the other important anthropogenic contributors of acid deposition is the primary pollutant NO_x (NO and NO_2). The mechanism for the formation of nitric acid and nitrous acid, which are both found in acid deposition, is shown:

$$N_2 + O_2 + \text{heat} \rightarrow 2NO$$
$$2NO + O_2 \rightarrow 2NO_2$$
$$HO\cdot + NO_2 \rightarrow HNO_3$$
$$HO\cdot + NO \rightarrow HNO_2$$

or

$$3NO_2 + H_2O \rightarrow 2HNO_3 + NO$$
$$NO_2 + NO + H_2O \rightarrow 2HNO_2$$

or

$$4NO + O_2 + 2H_2O \rightarrow 4HNO_2$$

In the second set of reactions, note that the NO is reproduced, therefore acting as a catalyst in this reaction mechanism.

CHEM COMPLEMENT

Hybrid cars

The adverse effects of primary air pollutants is one of the leading motivators for the introduction of hybrid cars. Most hybrid cars are petrol–electric hybrids, and so are a good fusion of the old technology internal combustion engine and the new technology electric car. Such cars have significantly increased the mileage and reduced the emissions of a petrol-powered car while overcoming the shortcomings of an electric car. A hybrid car has very low emissions of primary pollutants, not only because it runs part of the time on clean electrical power but also because it has a more efficient engine and so relatively low fuel consumption. All of these factors mean that if you have to drive a vehicle, then driving a hybrid car means that you are contributing less to air pollution and therefore less to acid deposition.

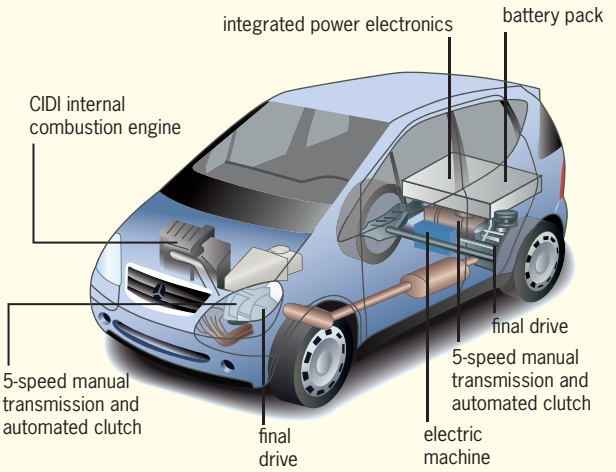


Figure 5.11.1 The insides workings of a hybrid concept vehicle the Mercedes-Benz M-Class HyPer.

The atmosphere contains a natural defence mechanism against acid deposition—ammonia—which neutralizes the acids to form ammonium salts. Ammonia (NH_3) is a weak base that finds its way to the atmosphere via fertilization of agricultural areas or naturally through the nitrogen cycle in which nitrogen gas is converted to NH_4^+.

AS E.11.2
Explain the role of ammonia in acid deposition. © IBO 2007

In the atmosphere, NH_4^+ ions react with nitrates and sulfates to form ammonium salts, which then sink to the ground or are washed out of the atmosphere in rain. The salts then migrate into the soil, where nitrification and acidification can occur:

$$NH_4^+ + 2O_2 \rightarrow 2H^+ + NO_2^- + H_2O$$

Figure 5.11.2 Fertilizers used in agriculture provide a source of ammonia, which helps to neutralize the acids found in acid deposition.

Section 5.11 Exercises

1. List the major pollutants responsible for acid deposition.
2. Describe the mechanism of acid deposition formation from:
   - **a** coal-powered power plants
   - **b** smelters
   - **c** internal combustion engines.
3. Discuss the use of hybrid cars in the battle against acid deposition.
4. Explain the role that ammonia has in acid deposition.
5. Outline how you can contribute to a decrease in acid deposition.

THEORY OF KNOWLEDGE

Ethics, the study of morality, involves thinking critically about *how we should act*, what choices we should make, what everyday decisions and actions should we take. Choices we think are wrong or don't approve of are called immoral and those we approve of are moral.

Thinking about how we should act with respect to our environment can happen at a personal, family, community, business, school, organization and global level. As individuals we need to ask ourselves some difficult questions, such as 'What does it mean to care for the environment? What should I do? What sort of actions show that I care?' and 'How do I justify my choices?'

- Making ethical decisions is difficult because our thoughts and actions are influenced by our emotions, perceptions, culture, expectations, religion, concern for others, obligations and reasoning skills. It can be hard to defend your own ethical beliefs to others. Being value judgments, they reflect on your own personal point of view.

  The United Nations Environment Program's mission statement is:

  > To provide leadership and encourage partnership in caring for the environment by inspiring, informing, and enabling nations and peoples to improve their quality of life without compromising that of future generations.

  What values is the UNEP stating, implying and declaring how it as an organization should act with respect to the environment?

- Reflect on your own environmental awareness and ethic, using the following points as a guide.
  - What are your beliefs, feelings and understanding of how humans are affecting the environment?
  - What are some of the important events and people that have influenced the way you think about the environment? How have they helped you gain a sense of what is the right and wrong?
  - Consider your circle of concern for the environment by thinking about what have you done so far and how you might act in the future.

  Elizabeth Mamo, USA
  —a reflection on my environmental ethic:

  I've always really enjoyed activities in nature such as rafting, hiking and rock climbing. I appreciate how the times I've done these activities, the sites have always been clean and nicely preserved. Despite increasing environmental damage, I want to be able to enjoy nature as natural as it can be.

  All my life I have lived in the United States and I knew there was environmental damage being done daily but it just didn't seem present since in my town I couldn't visibly see it. However since moving to Shanghai it seems that pollution is increasingly present and a situation that is hard to overlook or ignore since it is shown daily in the city. The air feels much thicker, the stars don't show up at night and the sky is constantly a grey colour. None of these signs were present back in Connecticut so coming to Shanghai I really noticed them.

  Something I have done that could be considered a step in the right direction is just noticing and being aware of nature and doing my own little part to recycle, avoid environmentally unfriendly products and care about my surroundings. I have also joined a newly formed Roots and Shoots club. Our projects this year have not been big; we set up a recycling program and put up some posters reminding people to think twice about throwing away reusable paper. Although these are small steps, we are going in the right direction.

5.12 WATER AND SOIL

HL

We have already investigated the importance of both soil and water in our environment, so in this last section of the chapter we will investigate the dual relationship between the two and how their properties influence one another. We have already examined the biological and physical importance of soil organic matter (SOM), but we have yet to examine the chemical importance of SOM.

Chemical function of SOM

AS **E.12.4 Describe the chemical functions of soil organic matter (SOM). © IBO 2007**

Soil organic matter enhances the ability of the soil to **buffer** changes in pH, and binds organic and inorganic materials and keeps them from leaching into the environment. It also reduces the negative environmental effects of pesticides, heavy metals and other pollutants by binding these contaminants. This binding takes place by ion exchange in much the same way as the ion exchange used in the tertiary treatment of waste water (section 5.3). This chemical binding contributes to the overall **cation-exchange capacity (CEC)** of the soil as a whole and means that the SOM contains sites that can bind nutrient cations (positively charged ions). These cations are therefore more readily available to be used by plants and are less likely to be leached into the environment by rain and irrigation.

Cation-exchange capacity of soil

AS **E.12.2 State what is meant by the term cation-exchange capacity (CEC) and outline its importance. © IBO 2007**

Humus and smectites (black swelling clays) have the greatest CEC, as they are both negatively charged, and so can attract positively charged cations. The cations that are exchanged can be acidic—H^+, Fe^{2+}, Mn^{2+} and Al^{3+}—or basic—Ca^{2+}, Mg^{2+} and K^+—in nature. These cations are not removed from the soil with water; instead other cations have to be added in order for an exchange to take place. The CEC of soil is therefore defined as its capacity to exchange cations with the soil solution. It is often used as a measure of potential soil fertility, with the higher CEC indicating a more fertile soil. A soil that has a low percentage of clay can have its CEC increased by the addition of manure, which works to increase the CEC.

Figure 5.12.1 Manure can be added to low percentage clay soil to increase the CEC.

pH and soil

AS **E.12.3 Discuss the effects of soil pH on cation-exchange capacity and availability of nutrients. © IBO 2007**

The quality of a soil is directly connected to its pH level. Soil pH influences how efficiently a crop grows by affecting nutrient availability. The optimal pH level for most plants is in the range of 6.0–8.0 or only slightly acidic or slightly basic, although some plants such as azaleas, rhododendrons, blueberries and conifers thrive best in acidic soils that fall within the pH range of 5.0–5.5.

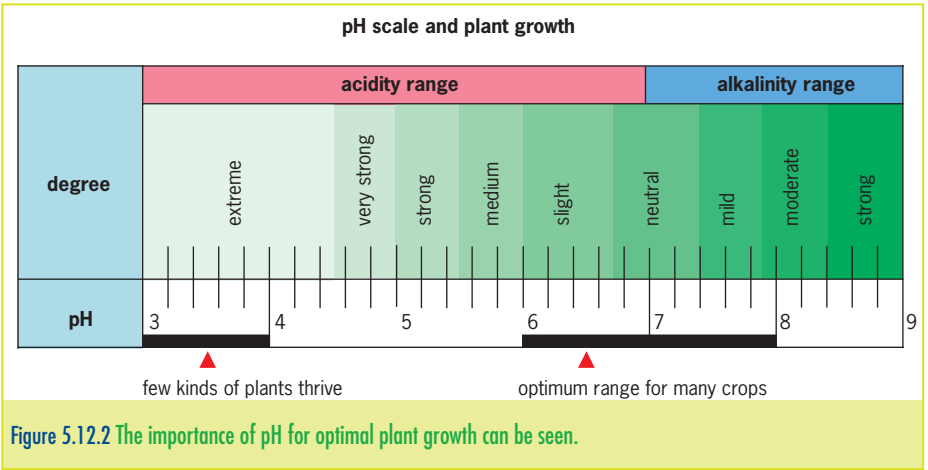


Figure 5.12.2 The importance of pH for optimal plant growth can be seen.

A low pH is an indication of high levels of acidic cations such as aluminium and manganese. Aluminium and manganese are toxic to plants. Acidic soils can be produced by nutrient depletion (section 5.7). Acidic soils can also be produced by acid deposition. The addition of lime to moist acidic soils can raise the pH of the soils to a more suitable pH range.

The acidity or basicity of these cations can be shown using hydrolysis equations. The acidic cations form hydrogen ions (indicating an acid) when mixed with water. The basic cations form hydroxide ions (indicating a base) when mixed with water.

$$Al^{3+}(aq) + 3H_2O(l) \rightarrow Al(OH)_3(aq) + 3H^+(aq)$$

In the hydrolysis of aluminium ions, aluminium hydroxide (a weak base) and hydrogen ions form. The presence of the strong acid and a weak base indicates an acidic solution.

$$Ca^{2+}(aq) + 2H_2O(l) \rightarrow Ca(OH)_2(aq) + 2H^+(aq)$$

In the hydrolysis of calcium ions, calcium hydroxide (a strong base) and hydrogen ions form. The presence of the strong base indicates a basic solution.

AS **E.12.1**
Solve problems relating to the removal of heavy-metal ions, phosphates and nitrates from water by chemical precipitation. © IBO 2007

We can quantitatively determine the concentration of various chemical species, such as heavy-metal ions, phosphates and nitrates, when they form precipitates in water. In a reaction in which a precipitate is formed, the precipitate dissolves to some extent, which allows for a dynamic equilibrium to be established between the precipitate and dissolved ion. Solubilities can therefore be expressed in terms of equilibrium constants. The equilibrium between solid silver chloride and its saturated solution is given:

$$AgCl(s) \rightleftharpoons Ag^+(aq) + Cl^-(aq)$$

The equilibrium constant for this heterogeneous equilibrium is called the **solubility product**, K_{sp}, and can be written for slightly or nearly insoluble ionic compounds:

$$K_{sp} = [Ag^+][Cl^-]$$

Recall from your study of equilibrium that solids do not appear in equilibrium constants.

Generally the solubility product for a salt M_mN_n is shown as:

$$K_{sp} = [M]^m[N]^n$$

Solubility products tend to be very small in magnitude. This is expected as a solid precipitate forms, indicating very low levels of solubility. If the value of the K_{sp} is known, the solubility of the compounds can be calculated.

CHEM COMPLEMENT

Chameleon flowers and mystery colours

Figure 5.12.3 Hydrangeas of two different colours arise from altering the pH levels of the soil.

Hydrangeas are fascinating plants whose flower colour can be changed by changing the pH of the soil. Hydrangeas grown in more basic soil produce pink flowers, whereas hydrangeas that are grown in more acidic soil produce blue flowers. To change the flower colour from pink to blue, aluminium is added to the soil to lower the pH and create acidic conditions. It is harder to change the flower colour from blue to pink, as that would mean having to remove aluminium from the soil.

TABLE 5.12.1 THE RELATIONSHIP BETWEEN pH AND BLOOM COLOUR FOR HYDRANGEAS

pH	Flower colour
4.5	Deep, vivid blue
5.0	Medium blue
5.5	Lavender-purple
6.0	Purplish-pink
6.5	Mauve-pink
6.8	Medium pink
7.0	Deep, vivid pink

Worked example 1

Calculate the solubility of a solution of Ag_2S that has a solubility product of 6.3×10^{-51}.

Solution

Step 1: Write the equation for the equilibrium between the solid and its saturated solution.

$Ag_2S(s) \rightleftharpoons 2Ag^+(aq) + S^{2-}(aq)$

Step 2: Write the solubility product expression.

$K_{sp} = [Ag^+]^2[S^{2-}] = 6.3 \times 10^{-51}$

Step 3: If the assumption is made that $[S^{2-}]$ is x, then $[Ag^+]$ must be twice $[S^{2-}]$, since the mole ratio is 2 : 1, so $[Ag^+] = 2x$.

$K_{sp} = (2x)^2(x) = 6.3 \times 10^{-51}$

Step 4: Factor out the coefficient and combine the pronumerals.

$K_{sp} = 4x^3$

Step 5: Solve for x.

$$x^3 = \frac{K_{sp}}{4}$$
$$= \frac{6.3 \times 10^{-51}}{4}$$
$$= 1.6 \times 10^{-51} \quad \text{(taking the cube root)}$$
$$x = 1.2 \times 10^{-17} \text{ mol dm}^{-3}$$

Step 6: Calculate the mass of silver ions that dissolves in 1 dm^3 of solution.

$m = n \times M$
$= (1.2 \times 10^{-17} \text{ mol})(107.87 \text{ g mol}^{-1})$
$= 1.3 \times 10^{-15}$ g

This relatively small mass of silver ions means that its precipitation is very efficient. The amount of silver ions in solution can be reduced even further by an effect called the **common ion effect**. The common ion effect explains the reduction of the solubility of one salt by the addition of a common ion.

Worked example 2

Calculate the solubility of Ag_2S in a 0.15 mol dm^{-3} solution of $AgNO_3$.

Solution

Step 1: Write the equation for the equilibrium between the solid and its saturated solution. Set up to solve for a traditional equilibrium problem.

$Ag_2S(s) \rightleftharpoons 2Ag^+(aq) + S^{2-}(aq)$

Volume = 1.00 dm^3	Reactants	Products	
	$Ag_2S(s)$	$Ag^+(aq)$	$S^{2-}(aq)$
Molar ratio	1	2	1
$[\ ]_i$ (mol dm^{-3})	(not included as it is a solid)	0.15	0
Change		$+2x$	$+x$
$[\ ]_{eq}$ (mol dm^{-3})		$0.15 + 2x$	$0 + x$

Step 2: Write the solubility product expression.

$$K_{sp} = [Ag^+]^2[S^{2-}] = 6.3 \times 10^{-51}$$

Step 3: Substitute in the known entities.

$$6.3 \times 10^{-51} = (0.15 + 2x)^2(x)$$

Step 4: Rearrange for x.

$$x = \frac{6.3 \times 10^{-51}}{(0.15 + 2x)^2}$$

Step 5: Make assumptions that as Ag_2S is only slightly soluble we would expect $2x$ to be very small compared to 0.15; therefore its value will be negligible compared to 0.15, so we can ignore it.

$$x = \frac{6.3 \times 10^{-51}}{(0.15)^2}$$

Step 6: Solve for x, the solubility of silver ions, remember in regular solution (worked example 1) the value was 1.2×10^{-17} mol dm^{-3}.

$$x = 2.5 \times 10^{-49} \text{ mol dm}^{-3}$$

We can see that the solubility is indeed reduced when a common ion is added to solution.

Section 5.12 Exercises

1 Describe the chemical function of soil organic matter.

2 Define cation-exchange capacity and state its importance in terms of soil fertility.

3 List the acidic and basic ions that can be exchanged in soils.

4 Describe the relationship between pH and soil quality.

5 State the solubility product expressions for the following compounds: $BaSO_4$, $Fe(OH)_3$ and $Ca_3(PO_4)_2$.

6 Calculate how many grams of $PbCl_2$ will dissolve in 0.250 dm^{-3} of water. The K_{sp} value for $PbCl_2$ is 1.60×10^{-5}.

7 If 4.8×10^{-3} g of $Al(OH)_3$ dissolves in 0.600 dm^{-3} of water, calculate the value of K_{sp}.

8 Calculate the minimum concentration of sulfate ion necessary to precipitate $CaSO_4$ from a solution that is 0.0030 mol dm^{-3} in $CaCl_2$. The value of the K_{sp} for $CaSO_4$ is 2.4×10^{-5}.

Chapter review questions and tests are available on student CD.

Chapter 5 Summary

Terms and definitions

Acid deposition The process by which acidic particles and gases (dry deposition) and precipitation (wet deposition) leave the atmosphere.

Acid shock Occurs during spring run-off into water supplies, when large quantities of snow and ice melt and bombard the water supplies with a huge amount of acid all at once.

Aerobic decomposition The decomposition of organic matter that occurs in the presence of oxygen and produces oxides or oxyanions.

Air pollution Is the change in the natural atmosphere by the addition of other gases, particulates and volatile organic compounds (VOCs).

Anaerobic decomposition The decomposition process that occurs when there is not enough oxygen present or when decomposition is carried out by organisms that do not require oxygen.

Anthropogenic The result of human activity.

Biochemical oxygen demand (BOD) The amount of dissolved oxygen needed by aerobic decomposers to break down the organic materials in a given volume of water at a certain temperature over a specified time period.

Buffer (solution) Ability to resist changes in pH when small amounts of either an acid or a base are added.

Cation-exchange capacity (CEC) The capacity of a soil to exchange cations with the soil solution. The higher the value of CEC the more fertile the soil.

Common ion effect The reduction of the solubility of one salt by the addition of a common ion.

Distillation The principle of separation that is dependent on compounds having different boiling points.

Environment All the external conditions that affect an organism (human, plant or animal) during its lifetime.

Eutrophication The process by which lakes, estuaries and other still bodies of water receive higher than normal levels of nutrients (primarily nitrogen and phosphorus), which results in an excessive growth of plants.

Flashing Occurs when heated water enters a chamber that has a pressure lower than atmospheric pressure, causing the water to boil rapidly and be converted into steam.

Flocculation Smaller suspended particles are coagulated (combined together) to form larger particles that can settle.

High-level waste (HLW) Radioactive waste that gives off large amounts of ionizing radiation for a long time.

Incomplete combustion Combustion of a fuel in the absence of enough oxygen produces carbon monoxide rather than carbon dioxide as one of the products.

Leaching Removal of a substance from a solid via a liquid medium. Nutrients can be leached from soils when rain passes through the soil and radioactive waste may leach from waste disposal units.

Low-level waste (LLW) Waste that gives off small amounts of ionizing radiation over a short amount of time.

Osmosis The tendency of a solvent to pass through a semipermeable membrane from an area of low solute concentration to an area of higher solute concentration until the concentrations on both sides of the membrane are equal.

Particulates Very small solid particles such as dust or soil, or liquid drops such as acids that are pollutants in our atmosphere.

Photochemical smog The type of smog that is formed when a series of primary pollutants and secondary pollutants interact under the influence of sunlight.

Primary treatment The stage in waste water treatment in which filtration, sedimentation and flocculation take place and results in most of the solid materials being removed.

Recycle To salvage usable metals, glass, plastic and paper products from municipal solid waste and sell them to manufacturing industries.

Reverse osmosis The opposite of natural osmosis in which a high pressure is applied to the side of the membrane where salt water is located, forcing the solvent (water) through the membrane leaving the solute (salt) behind.

Roasting The first step in smelting in which sulfur-containing ores are converted into oxygen-containing ores.

Salinization The accumulation of salts in soils due to excessive irrigation practices.

Secondary treatment The second stage in the treatment of waste water that involves a biological treatment process to remove dissolved organic matter (activated sludge process).

Semipermeable membrane Membrane that allows solvent but not solute through.

Smelting The process by which a metal is extracted from its ore.

Soil degradation The process that results in the soil being unhealthy or infertile; can occur naturally or anthropogenically.

Soil organic matter (SOM) The organic constituents in the soil; include undecayed plant and animal tissues, their partial decomposition products and the soil biomass.

Soil pollution The overuse of chemicals such as pesticides and fertilizers which remain in the soil.

Solubility product (K_{sp}) Equilibrium constant used to express solubility equilibrium.

Tertiary treatment The third step is the treatment of waste water in which heavy metals, phosphates and nitrates are removed.

Thermal inversion The weather conditions that trap a layer of dense, cool air beneath a layer of less dense, warm air over a city that is located in a valley.

Thermal pollution The addition of abnormally warm or cool water to a body of water, disrupting the ecosystem.

VOCs Volatile organic compounds.

Waste water Water that has been adversely affected in quality by anthropogenic influences.

Winkler method The chemical method for measuring the level of dissolved oxygen in a water sample.

Concepts

- Pollutants, their sources and how to reduce them are summarized in the table on the next page.
- Normal precipitation has a pH level of 5.6 due to dissolved CO_2.
- Acid deposition consists of both wet and dry forms of precipitation.
- Electric power plants and other industrial plants give off large amounts of sulfur dioxide (SO_2), particulate matter and nitrogen oxides (NO_x), all of which are primary pollutants. These primary pollutants are converted into the main components of acid deposition, the secondary pollutants—nitrogen dioxide (NO_2), nitric acid (HNO_3) and sulfuric acid (H_2SO_4).
- (NO_x) forms in internal combustion engines and can be reduced by the use of catalytic converters.
- Acid deposition destroys marble (limestone) statues via the following reaction:
  $CaCO_3 + H_2SO_4 \rightarrow CaSO_4 + H_2O + CO_2$
- Acid deposition is extremely harmful to the Earth's aquatic environment and causes leaching in soils.
- The Sun's radiation enters the Earth's atmosphere as shorter wavelength visible light. When these light waves collide with clouds or the Earth's surface they are reflected back as longer wavelength infrared light. Most of this longer wavelength energy is too long to pass back through the Earth's atmosphere and is trapped by the 'greenhouse' gases.
- The greenhouse gases include carbon dioxide (CO_2), methane (CH_4), water vapour (H_2O), nitrous oxide (N_2O) and chlorofluorocarbons (CFCs).
  See table on the next page.
- Ozone is formed naturally in the atmosphere via the following series of reactions:
  O_2 + light energy $\rightarrow 2O\cdot$
  $O\cdot + O_2 \rightarrow O_3$
  O_3 + light energy $\rightarrow O_2 + O\cdot$
  $O_3 + O\cdot \rightarrow 2O_2$
- Both CFCs and NO_x act as ozone depleters. CFCs are used as refrigerants and propellants.
- NO_x forms in internal combustion engines as a by-product of high temperature combustion. Alternatives to CFCs are given in the table on the next page.
- The amount of dissolved oxygen needed by aerobic decomposers to break down the organic materials in a given volume of water at a certain temperature over a specified time period is called the biochemical oxygen demand (BOD). The cleaner the water source, the lower the levels of BOD.
- The amount of dissolved oxygen can be measured using the Winkler titration method.
- Aerobic decomposition is the most common decomposition process. During this process the decomposition of organic matter uses oxygen and produces oxides or oxyanions.
- Anaerobic decomposition occurs when there is not enough oxygen present or when decomposition is carried out by organisms that do not require oxygen. This process involves the loss or gain of electrons.

Summary of air pollutants and their sources, ways to reduce them and their effects on human health					
	CO	**NO_x**	**SO_x**	**VOCs**	**Particulates**
Anthropogenic sources	Incomplete combustion in internal combustion engines	Formed at high temperatures in internal combustion engines	Combustion of sulfur-containing coal, smelting of sulfide ores and roasting processes	By-products of combustion in internal combustion engines, emitted from paints, cleaning supplies, building materials, furnishings	By-products of industrial processes, and incomplete combustion
Natural sources	Volcanic activity and forest fires, incomplete oxidation of methane formed by anaerobic decomposition of organic matter	Lightening strikes, denitrification in soils	Volcanic eruptions, sour natural gas, sea spray	Anaerobic decomposition by bacteria	Volcanic eruptions, forest fires
Ways to reduce the pollutant	Catalytic converters, the use of lean-burn engines	Catalytic converters, the use of lean-burn engines	Wash coal prior to burning, usage of alkaline scrubbers or fluidized beds	Catalytic converters, the use of green plants	Electrostatic precipitation
Effects on human health	Replaces the oxygen carried in hemoglobin in red blood cells, reducing oxygen availability and causing suffocation	Harmful to people already suffering from lung disease, can damage healthy lung tissue	Irritates the respiratory system.	'Sick-building syndrome due to poor air quality	Decreased lung function, aggravation of asthma, development of bronchitis, emphysema and respiratory cancer

Positive and negative influences on the atmosphere of increasing amounts of greenhouse gases	
Positive	**Negative**
• Lower heating bills • Longer growing season	• Air conditioning usage is higher and thus more CFCs needed • Changes in precipitation patterns and global temperature • Melting of the polar ice caps • Rising sea levels causing flooding in some areas • Droughts in temperate and desert areas • Increased forest fire hazard • Thermal expansion of the oceans • Changes in the distribution of pests and disease-carrying organisms

- Eutrophication is the process by which lakes, estuaries and other still bodies of water receive higher than normal levels of nutrients (primarily nitrogen and phosphorus), which results in an excessive growth of plants—an algal bloom. This process results in reduced levels of dissolved oxygen.
- Thermal pollution occurs when anthropogenic sources, such as factories or other industries, release large amounts of water that has either a lower or higher temperature than the natural water source. Thermal pollution can result in the killing of a large number of fish.

CFC alternatives and their properties			
	Hydrocarbons	**Fluorocarbons (perfluorocarbons)**	**Hydrofluorocarbons**
Examples	Propane, C_3H_8 Butane, C_4H_{10}	Octafluoropropane, C_3H_8 Perfluorohexane, C_6F_{14}	HFC-23, CHF_3 HFC-152a, CH_3CHF_2
Molecular formula	C and H only	C and F only	C, H and F only
Toxicity	Dependent on dose and route of ingestion; hydrocarbon abuse due to sniffing, bagging or huffing	Very low toxicity	Low acute toxicity
Flammability	Quite flammable	Low flammability	Low flammability
Damage to the ozone layer	No damage to ozone layer	F does not catalyse ozone destruction so therefore no damage	F does not catalyse ozone destruction so therefore no damage
Ability to absorb infrared radiation and therefore contribute act as a greenhouse gas	Can absorb	Can absorb	Can absorb

The sources and effects of water pollutants are listed in the tables on pages 392 and 393.

- Primary treatment of waste water is the stage in which filtration, sedimentation and flocculation take place and results in most of the solid materials being removed.
- Secondary treatment of waste water involves biological treatment process whereby dissolved organic matter is removed using the activated sludge process.
- Tertiary treatment of waste water removes heavy metals, phosphates and nitrates through precipitation and ion-exchange methods.
- Fresh water can be obtained from sea water through the processes of distillation and reverse osmosis.
- Soil degradation can occur naturally as a result of wind and water erosion or anthropogenically by salinization, nutrient depletion or soil pollution.
- Soil organic matter (SOM) is defined as the organic constituents of a soil—undecayed plant material and animal tissue, their partial decomposition products and soil biomass. A high level of SOM is an indicator of good soil health and low soil degradation.
- Waste can be disposed of in landfills or it can be incinerated.
- Radioactive waste can be classified as either high-level or low-level waste on the basis of the ionization energies, which they release over a period of time.
- The double bond between the oxygen atoms in O_2 is a stronger and shorter (121 pm) than that of ozone (128 pm).
- CFC and NO_x are ozone depleters, their mechanisms are shown below:

$CCl_2F_2 \rightarrow CClF_2 + Cl\cdot$
$Cl\cdot + O_3 \rightarrow ClO\cdot + O_2$
$ClO\cdot + O\cdot \rightarrow Cl\cdot + O_2$
$O_3 + O\cdot \rightarrow 2O_2$ Overall effect

$NO + O_3 \rightarrow NO_2 + O_2$
$NO_2 + O\cdot \rightarrow NO + O_2$
$O_3 + O\cdot \rightarrow 2O_2$ Overall effect

- There is greater ozone depletion at the polar regions due to wind and weather patterns that concentrate CFCs and NO_x there. Destruction of the ozone layer is sped up by surface adsorption on ice crystals in the atmosphere.
- Photochemical smog occurs when primary pollutants—NO_x and VOCs—and secondary pollutants—NO_2, ozone, aldehydes, and PANs—interact under the influence of sunlight. Photochemical smog is influenced by both topography and climate.
- Thermal inversion occurs when a layer of cool air is trapped below a layer of warm air, and aids in the formation of photochemical smog.

Primary water pollutants, their sources and their effects			
Primary pollutant	**Source**	**Effects**	**Water quality**
Pesticides (DDT, herbicides, fungicides)	• Run off from agricultural applications • Run off from municipal applications on lawns, flower gardens, etc.	• Carcinogenic • Birth defects • Neurological disorders	• Polluted • Unsafe to drink at elevated levels
Dioxins	• Natural combustion processes, e.g. forest fires • Volcanic eruptions • Incineration of industrial (hospital) and household waste • As a herbicide, such as agent orange, which was used to clear bush during the Vietnam War	• Chloracne, a severe form of skin disease • Reproductive and developmental effects • Liver damage and cancer	• Polluted • Unsafe to drink at elevated levels
Polychlorinated biphenyls (PCBs)	• Environmentally cycling of old PCBs (old source of PCBs include electrical transformers or capacitors) • Many industrial processes such as plasticizers, adhesives, etc.	• Acne-like breakouts • Hearing and vision problems • Irritation of gastrointestinal tract • Affect reproductive efficiency • Liver damage and cancer	• Polluted • Unsafe to drink at elevated levels • Concentrated in fish
Organic matter	• Decaying vegetation • Decaying benthonic matter • Decaying organisms	• Retards photosynthesis in plants	• Polluted • Ripe for the growth of bacteria • Unsafe to drink at elevated levels
Nitrates	• Fertilizers • Animal wastes • Septic tanks • Decaying plant material • Acid deposition	• Babies can develop methemoglobinemia or 'blue baby syndrome' caused by a lack of oxygen in the blood. • Carcinogenic	• Polluted • Unsafe to drink at elevated levels
Phosphates	• Human sewage • Agricultural runoff from crops • Pulp and paper industry • Chemical and fertilizer manufacturing • Detergents	• Eutrophication—accelerates plant and algae growth • Kidney damage • Osteoporosis	• Polluted • Unsafe to drink at elevated levels • Necessary for good human health at lower levels

Heavy metal water pollutants, their sources and their effects			
Heavy metal pollutant	**Source**	**Effects**	**Water quality**
Copper	• Copper pipes	• Anemia • Liver and kidney damage • Stomach and intestinal irritation	• Polluted • Unsafe to drink at elevated levels
Lead	• Lead pipes • Batteries • Leaded gasoline (petrol) • Paints • Ammunition	• Negatively affects hemoglobin production • Damage to kidneys gastrointestinal tract, joints and reproductive system • Can lower IQ levels in young children	• Polluted • Unsafe to drink at elevated levels
Cadmium	• Smelting • Improper disposal of rechargeable batteries	• Kidney failure • Lung disease • Osteoporosis	• Polluted • Unsafe to drink at elevated levels
Nickel	• Power plants • Waste incinerators • Improper disposal of batteries	• Decreased body weight • Damage to the heart and liver	• Polluted • Unsafe to drink at elevated levels
Zinc	• Mining • Smelting • Steel production	• Anemia • Damage to nervous system and pancreas	• Polluted • Unsafe to drink at elevated levels
Arsenic	• Natural deposits in the Earth • Industry • Agricultural processes	• Carcinogenic • Stomach pain • Numbness • Blindness	• Polluted • Unsafe to drink at elevated levels
Mercury	• Improper disposal of batteries, thermometers and barometers • Various industrial processing	• Tremors • Gingivitis • Spontaneous abortion • Damage to the brain and central nervous system	• Polluted • Unsafe to drink at elevated levels

Organic soil pollutants	
Organic pollutant	**Source**
Petroleum hydrocarbons	• Leaks from transport vehicles and home oil tanks • Used to reduce dust on unpaved roads
Agrichemicals	• Overuse of pesticides, herbicides and fertilizers
VOCs and semivolatile organic compounds, solvents	• Internal combustion engines • Emitted from solid materials • Emitted from industrial processes
Polyaromatic hydrocarbons (PAHs)	• Vehicle exhaust • Incomplete combustion of coal, oil and wood • Incineration of waste • Oil spills
Polychlorinated biphenyls (PCBs)	• Electrical transformers or capacitors as insulators • Industrial processes, e.g. plasticizers, adhesives, etc.
Organotin compounds	• Chemical used in the preservation of wood • Insecticides and fungicides

Recycled materials and their method of recycling	
Resource	**Method of recycling**
Metal	• Metals do not lose their properties when recycled. • The most common metals that are recycled are aluminium and steel. • Recycling aluminium requires only 5% of the energy and produces only 5% of the CO_2 emissions of production from raw materials. • Ferrous metals (made from iron) are separated by the use of magnets. • Steel is melted and turned into new car parts, cans and structural components. • Aluminium metal is so valuable that it makes no sense to landfill it. Once aluminium is sorted it is shredded and melted and then used to produce new aluminium products.
Glass	• Glass is first sorted by colour then washed, crushed and melted and turned into new glass. • To make glass from raw materials takes a massive amount of energy. The amount of energy required is greatly reduced as the glass goes through the recycling process. • Glass does not deteriorate during the recycling process, therefore it can be recycled indefinitely.
Plastic	• Many different types of plastic can be recycled; however, due to the large amount of plastic that is found in modern packaging, sorting of these plastics can be a lengthy and costly process. • Recycling of plastics involves a much more in-depth process than simply burning them. • Recycling plastics causes much less pollution and requires much less energy than that required to make the plastics from raw materials. • There are seven main types of plastics although only types 1, 2 and 4 are routinely recycled.
Paper products	• The over 50 different types of paper are sorted and de-inked and then broken down. • These different types of paper can be recycled into newsprint, egg cartons, gift boxes, packaging material, cereal boxes, new paper, tissue paper, paper towels and toilet paper. • Recycling paper reduces the amount of waste in the landfill, saves energy and trees and reduces the amount of pollution that is involved in producing paper from raw materials.

- The mechanism for acid formation deposition is:

  $H_2O + O_3 \rightarrow 2HO\cdot + O_2$

  or

  $H_2O + O\cdot \rightarrow 2HO\cdot$
  $HO\cdot + SO_2 \rightarrow HOSO_2\cdot$
  $HOSO_2\cdot + O_2 \rightarrow HO_2\cdot + SO_3$
  $SO_3 + H_2O \rightarrow H_2SO_4$
  $SO_2 + H_2O \rightarrow H_2SO_3$
  $N_2 + O_2 + \text{heat} \rightarrow 2NO$ (first forms in the engine)
  $2NO + O_2 \rightarrow 2NO_2$
  $HO\cdot + NO_2 \rightarrow HNO_3$
  $HO\cdot + NO \rightarrow HNO_2$

  or

  $3NO_2 + H_2O \rightarrow 2HNO_3 + NO$
  $NO_2 + NO + H_2O \rightarrow 2HNO_2$

- In the atmosphere, ammonia neutralizes the acids formed to a large extent, and forms ammonium salts.
- CEC can be defined as the capacity of soil to exchange cations with the soil solution. A higher CEC is indicative of a more fertile soil.
- The optimal pH level for most plants is in the range of 6.0—8.0. Low pH levels indicate high levels of aluminium or manganese, which are toxic to plants. Lime can be added to soil to raise the pH levels.
- Solubility product expressions can be written to calculate the concentration of precipitates in water.

6 FOOD CHEMISTRY

Chapter overview

This chapter covers the IB Chemistry syllabus Option F: Food Chemistry.

By the end of this chapter, you should be able to:

- distinguish between foods and nutrients
- describe the chemical composition of lipids, carbohydrates and proteins
- describe structural differences of saturated and unsaturated fatty acids
- predict the stability and melting points of fats and oils using their structure
- describe the hydrogenation of unsaturated fats and discuss the advantages and disadvantages of this process
- explain what it meant by the term *shelf life*, and discuss factors affecting it
- describe fat rancidity, and compare oxidative and hydrolytic rancidity
- describe ways to minimize rancidity and prolong shelf life
- define the term *antioxidant* and list examples of naturally occurring antioxidants
- compare advantages, disadvantages and structural features of synthetic antioxidants
- distinguish between a dye and a pigment
- explain the occurrence of colour in natural pigments such as anthocyanins and heme
- describe factors that affect colour stability in natural pigments
- discuss the safety of synthetic colourants
- compare browning processes
- define what is meant by *genetically modified foods* and discuss their benefits and concerns
- describe dispersed food systems, such as emulsions, suspensions and foams
- describe how emulsifiers work

- describe the occurrence of oxidative rancidity via a free-radical chain mechanism HL
- explain differences between the three major antioxidant types
- explain three different conventions for naming enantiomers
- describe the different properties of the enantiomeric forms of stereoisomers found in food
- compare the structures of anthocyanins, carotenoids, chlorophyll and heme and explain why they are coloured while many other organics are not
- deduce the water or fat solubility of anthocyanins and carotenoids from their structures.

IBO Assessment statements F.1.1 to F.10.3

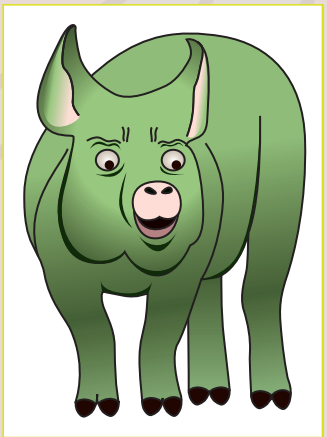


Figure 6.0.1 Pigach anyone? Is this what happens when spinach genes are inserted into pigs?

If, as the saying goes, we are what we eat, then what we are is a complex mixture of minerals and molecules. Eating can be an enormously pleasurable experience and we must have a certain quantity of food every day to survive. Unfortunately, not everyone in the world has access to a steady supply of nutritious food. One proposed solution for world hunger is genetically modified (GM) foods, which are discussed in this chapter. Also in this chapter you will discover:

- why salmon is pink
- what blueberries and dark chocolate have in common
- how oils are turned into margarine
- why fats smell so bad when they turn rancid
- how the shelf life of food can be extended
- what is produced when spinach genes are inserted into pigs.

6.1 FOOD GROUPS

AS F.1.1 Distinguish between a food and a nutrient. © IBO 2007

Are you hungry, or perhaps just in the mood for a snack? Our days are somewhat punctuated by our need and/or desire for food. **Food** may be defined as "any substance, whether processed, semi-processed or raw, that is intended for human consumption, and includes drink, chewing gum and any substance which has been used in the manufacture, preparation or treatment of 'food', but does not include cosmetics or tobacco or substances used only as drugs" (Codex Alimentarius (FAO/WHO) definition 2005). What we obtain from food (in addition to the pleasure of eating it!) are nutrients. A **nutrient** may be defined as any substance obtained from food and used by the body to provide energy, regulate growth, maintenance and repair of the body's tissues. Proteins, lipids, carbohydrates, vitamins, minerals and water are considered nutrients.

Nutrients have a number of functions in the body, with different nutrients having different primary functions. Carbohydrates, lipids and proteins are the energy providers. Protein also has an important role in growth and tissue repair. We require these nutrients in relatively large quantities, and they are classed as **macronutrients**. Also included in this class are minerals such as sodium, potassium, magnesium, phosphorus, sulfur and chlorine. Many other substances are needed in our diet in much smaller amounts (as little as milligrams or micrograms). These **micronutrients** are essential for normal functioning of the body, and include vitamins and trace minerals such as iron, copper, fluorine, zinc, iodine, selenium, manganese, molybdenum, chromium, cobalt and boron.

Good health requires a balanced diet, one that contains all the required nutrients in the required amounts. While the actual amounts required vary with such factors as age, gender, level of activity and weight, a balanced diet should have approximately 60% carbohydrate, 25% protein and 15% fats and oils. A balanced eating plan can be represented in many ways, but is often shown in the shape of a pyramid. The pyramid shows which foods should be eaten in the largest amounts, and which in the smallest amounts. While there are a variety of pyramids available, most emphasize certain principles:

- Eat a wide variety of foods.
- Include plenty of fruit, vegetables and grains.
- Limit intake of sugar, salt and alcohol.
- Reduce intake of saturated fat, trans fat and cholesterol.

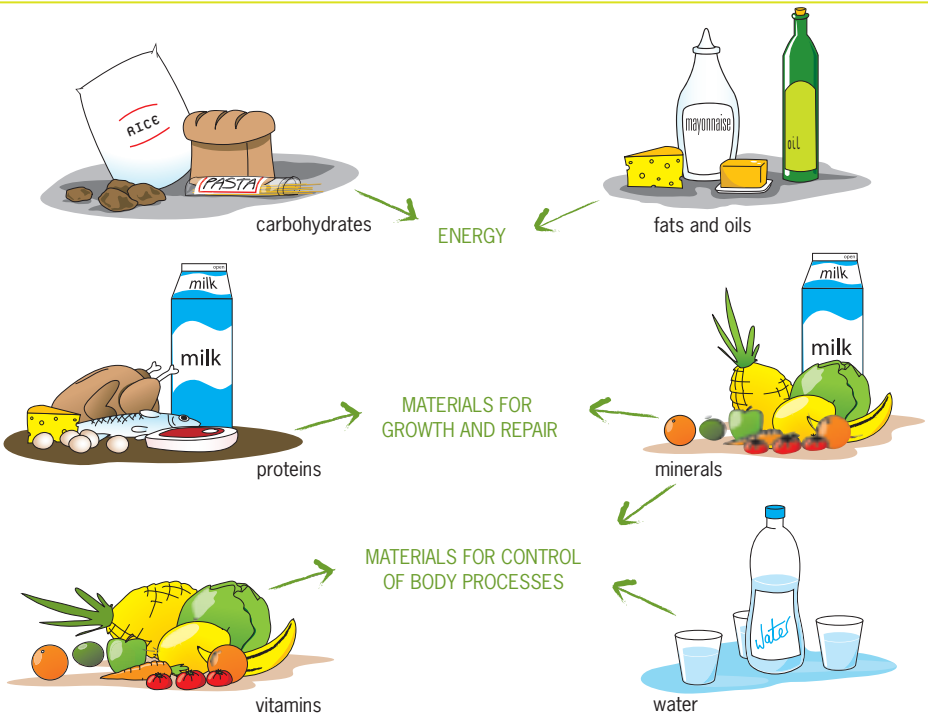


Figure 6.1.1 Nutrients have a number of different functions.

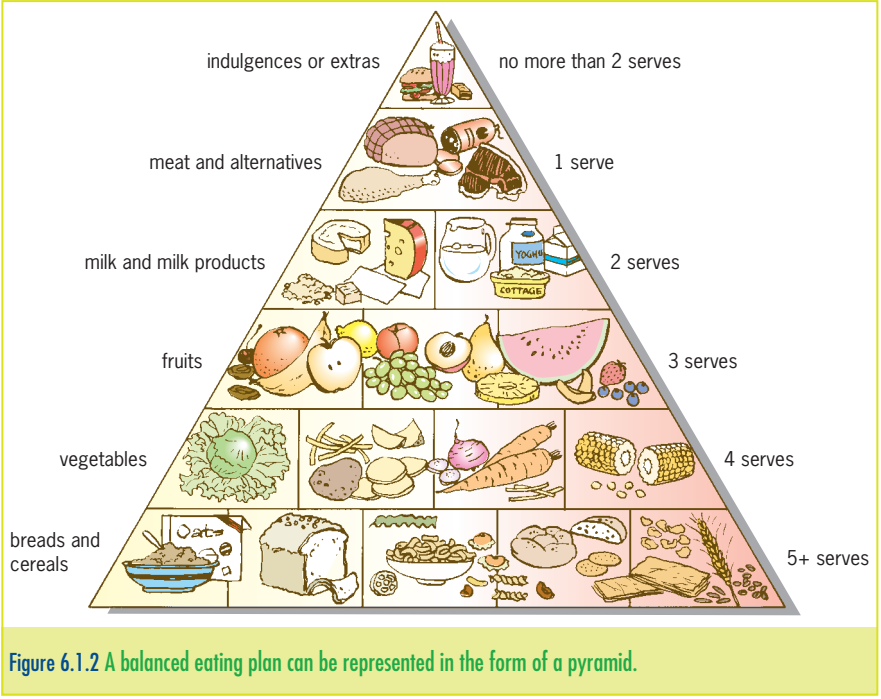


Figure 6.1.2 A balanced eating plan can be represented in the form of a pyramid.

Many of the foods we eat consist of large, complex and insoluble molecules that must be broken down into smaller, more soluble molecules before they can be used in the body. This breakdown, by both chemical and physical processes, is the work of your digestive system. The breakdown and subsequent rebuilding of energy-providing nutrients—the carbohydrates, proteins and lipids—involves two main chemical reaction types: hydrolysis and condensation. **Hydrolysis** is

the splitting up of large molecules by the addition of water molecules, and **condensation** is the joining of small molecules by a reaction involving the removal of atoms, which in turn form a water molecule. Further reactions, such as respiration, are involved to release the energy from the nutrients.

Each of these reactions occurring in the body is characterized by involving a number of small steps. Substances are gradually transformed in a series of reactions, each catalysed by an enzyme (a biological catalyst that is itself a protein). The processes involved in the digestion and utilization of nutrients are summarized in figure 6.1.3. We will now consider the molecular structures of these nutrients.

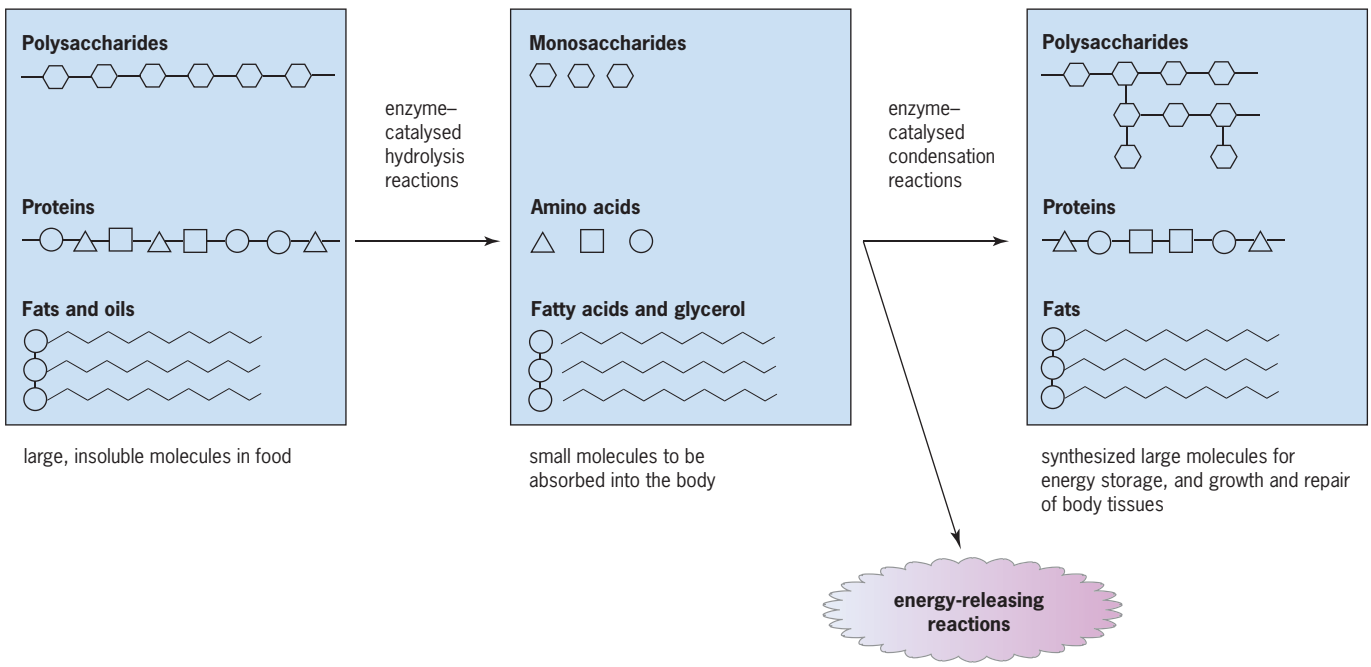


Figure 6.1.3 The breakdown and use of nutrients in the body involves several hydrolysis and condensation reactions.

Proteins

AS F.1.2 Describe the chemical composition of lipids (fats and oils), carbohydrates and proteins. © IBO 2007

Proteins are natural polymers that are essential to life. All organisms, even the tiniest viruses, contain proteins. Proteins have many different functions in living things. For example, they act as biological catalysts (enzymes), they give structure (hair, muscle, feathers and nails), they provide energy and, in some cases, they are hormones.

All proteins contain the elements carbon, hydrogen, oxygen and nitrogen, while many also contain sulfur. Proteins are polymers of **2-amino acids**, the general structure of which is shown in figure 6.1.4. Proteins differ from the synthetic polymers discussed in our studies of organic chemistry in that twenty different monomers are used to produce proteins. The amino acids differ according to their side chains, which are represented in the general structure as 'R'. Table 2.2.1 (p. 82) shows the side chains that distinguish the 20 amino acids. Plants are able to synthesize these amino acid monomers from CO_2, H_2O and minerals such as NO_3^- and SO_4^{2-}. Animals obtain many of their amino acids through their diet. Humans can synthesize 10 amino acids in our bodies.

These are termed non-essential amino acids. However, the remaining 10, termed essential amino acids, must come from our diet.

Amino acids undergo condensation reactions with other amino acids. A condensation reaction will occur between the $–NH_2$ group of one amino acid and the –COOH group of another. One example is shown in figure 6.1.5. Notice that when two amino acids react together, two products, called **dipeptides**, are possible. The link formed (–CONH–) is an amide bond, but it may be called a **peptide link** when it joins amino acids. For each amino acid that bonds to another amino acid by a condensation reaction, a water molecule is produced.

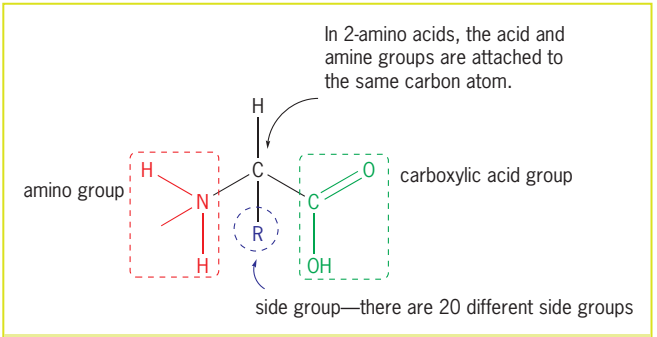


Figure 6.1.4 The general structure of a 2-amino acid.

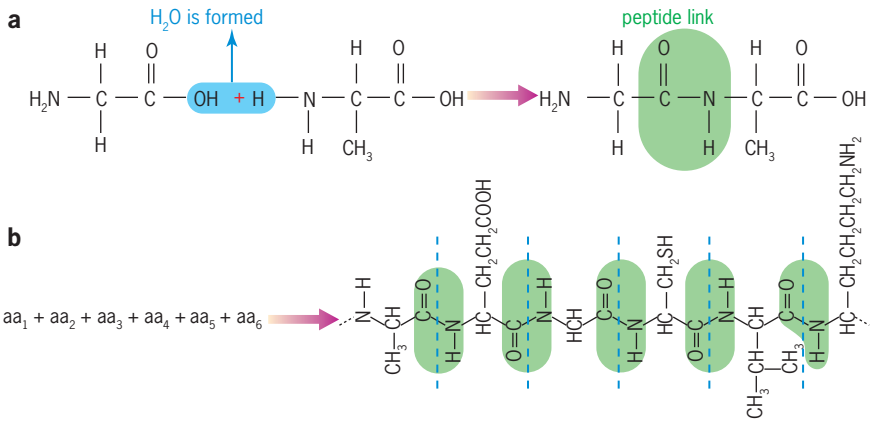


Figure 6.1.5 (a) Two amino acids undergo a condensation reaction to form a dipeptide. (b) Many amino acids undergo condensation reactions to form a polypeptide.

Interactive Animation
Peptide bond

QuickTime Video
Proteins and amino acids

The reaction of many amino acids produces a condensation polymer called a **polypeptide**, in which hundreds or even thousands of amino acids are joined. There are many millions of different combinations of sequences of the twenty different amino acids. If there are a large enough number of amino acids involved, the polymer is called a protein. The actual size distinction is somewhat arbitrary, but the molar mass for a protein is usually above 10 000 g mol^{-1}. The polypeptide chain may fold, coil and in other ways take up a complex, three-dimensional shape. This shape is essential to the functioning of many proteins, and involves various types of bonding between the side groups of amino acid residues at different points along the polypeptide chain.

Digestion of consumed protein involves the enzyme-catalysed hydrolysis of the bonds between the amino acid residues. Reactions begin in the stomach, where dipeptides are produced, and continue in the small intestine, where amino acids are produced. These small, soluble molecules are absorbed into the bloodstream. They are then transported to various parts of the body to be reassembled into proteins and so to perform their various functions.

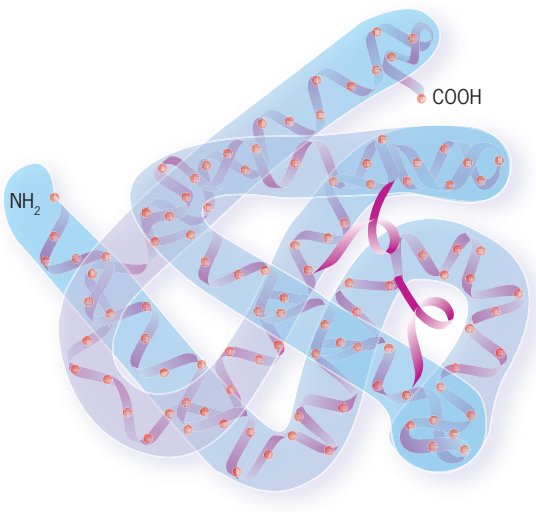


Figure 6.1.6 Protein chains may fold, coil and in other ways take up a complex, three-dimensional shape.

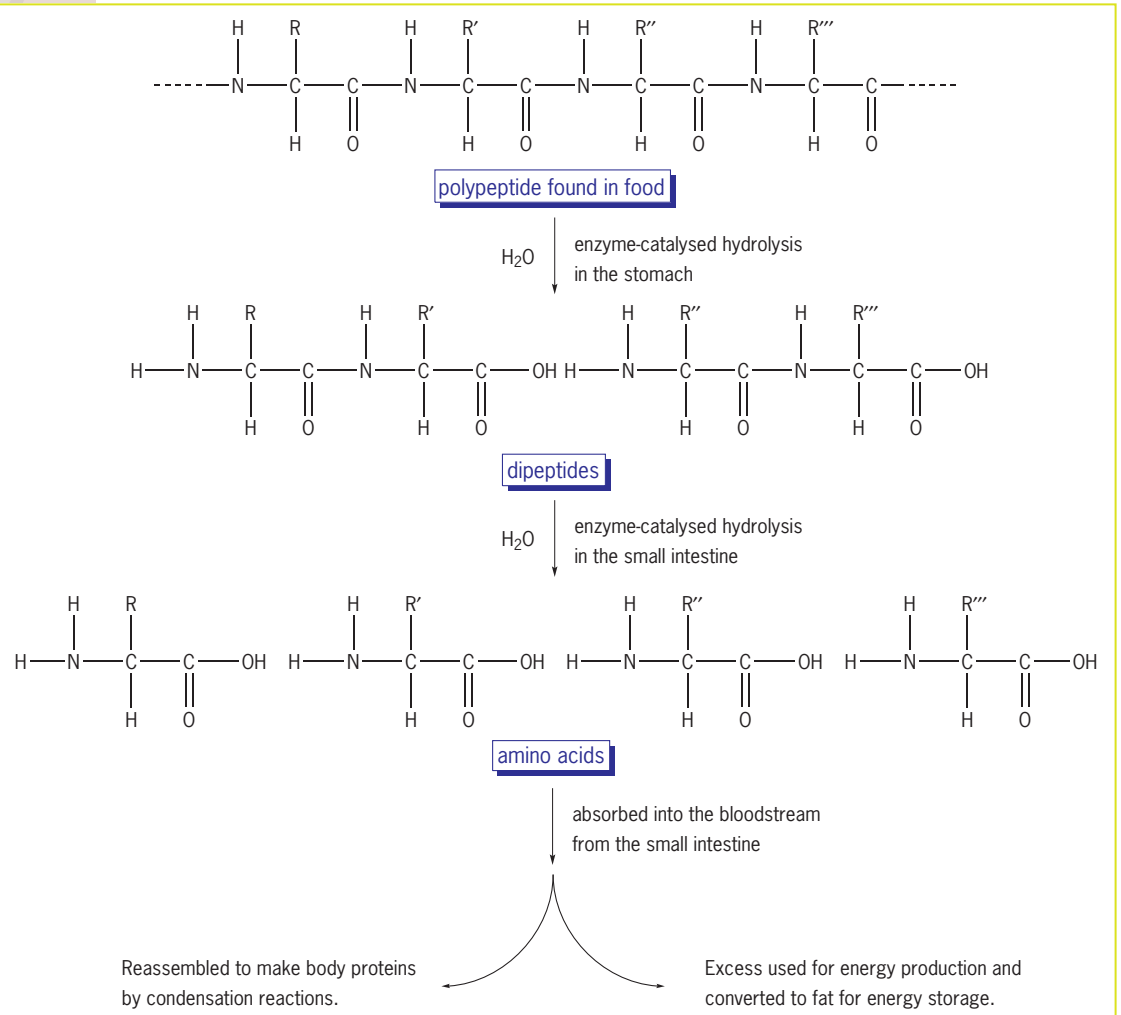


Figure 6.1.7 A summary of protein structure, digestion and use.

Carbohydrates

The name '**carbohydrate**' is derived from the observation that many members of this group have the empirical formula $C_x(H_2O)_y$, where x and y are whole numbers. Examples include the sugar found in grapes (glucose, $C_6H_{12}O_6$) and cane sugar (sucrose, $C_{12}H_{22}O_{11}$). Carbohydrates are often called saccharides (from the Latin, *saccharum*, for sugar) because of the sweet taste of these simple members of the group. Carbohydrates are among the most abundant components of living things. They serve several functions: as an energy source and store (starch, glycogen and glucose), as a structural material (cellulose), as an essential component of the genetic material (ribose and deoxyribose in nucleic acids) and as a precursor of other biologically important molecules.

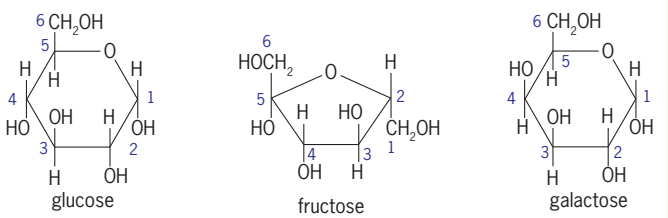


Figure 6.1.8 Structures of the monosaccharides glucose, fructose and galactose.

Glucose is one of the simplest of carbohydrate types, known as **monosaccharides** (simple sugars). All monosaccharides have the empirical formula CH_2O. In their linear form they contain a carbonyl (C=O) group and have at least two hydroxyl (–OH) groups. Fructose, found in fruits and honey, is also a monosaccharide, and galactose, found in dairy products, is another. All are

white, crystalline solids with a sweet taste, and all have the formula $C_6H_{12}O_6$. They are structural isomers of each other and are known as hexoses because they have six carbon atoms in their formula. Note that in figure 6.1.8 the majority of the carbon atoms are not drawn in the simplified structures shown. In each case, a carbon atom is to be found where four lines intersect, that is, at the vertices of the hexagon or pentagon. Notice the very slight difference in structure between glucose and galactose, in the arrangement of the atoms around carbon number 4.

PRAC 6.1
Detecting starch and sugars in food

In solution, glucose is in equilibrium between three isomeric forms—two with ring structures and a straight-chain molecule. While these ring structures of glucose appear to be very similar, there is a fundamental difference. Consider the orientation of the hydroxyl groups on carbons 1 and 4. In α-glucose, both hydroxyl groups are pointing 'downwards' whereas in β-glucose, the hydroxyl group on carbon 1 points upwards, and that on carbon 4 points 'downwards'. This difference is most important when glucose is polymerized. Although free rotation around all the bonds, except the C=O bond, is possible in the linear glucose molecule, the hydroxyl groups are fixed in their positions (pointing up or down) once the ring structure is formed.

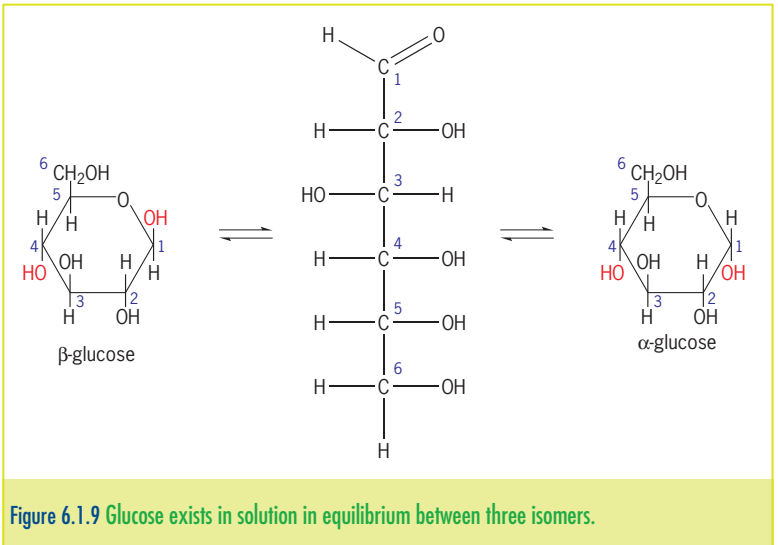


Figure 6.1.9 Glucose exists in solution in equilibrium between three isomers.

Two monosaccharides undergo a condensation reaction to produce a **disaccharide** (di- meaning two), eliminating a water molecule (figure 6.1.10). The linkage formed is a **glycosidic** or **ether linkage**, and the product formula is $C_{12}H_{22}O_{11}$. A condensation reaction between α-glucose and β-fructose produces sucrose. Sucrose, the most common disaccharide, is added as a sweetener to foods.

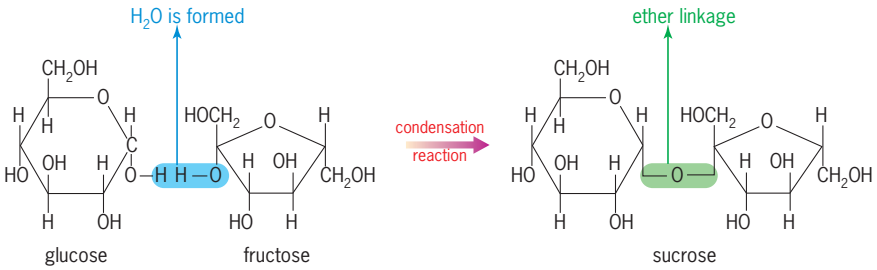


Figure 6.1.10 Condensation reactions form disaccharides.

A condensation reaction between α-glucose and β-galactose produces lactose, the disaccharide present in the milk of mammals and makes milk an energy source for the young. It is less sweet than sucrose, possibly a design of nature to protect the taste buds of the young.

Condensation reactions between many monosaccharides produce a polymer, a **polysaccharide** containing thousands of glucose units. Polysaccharides are less soluble than the smaller saccharides and do not have a sweet taste. Polymerization of glucose in living things can produce three different polysaccharides: starch, cellulose and glycogen. When α-glucose is polymerized by plants, the product is **starch**, a food source for animals and a food store for plants. Foods such as potato and sago are well known for their high starch content.

Interactive Animation
Carbohydrate structures

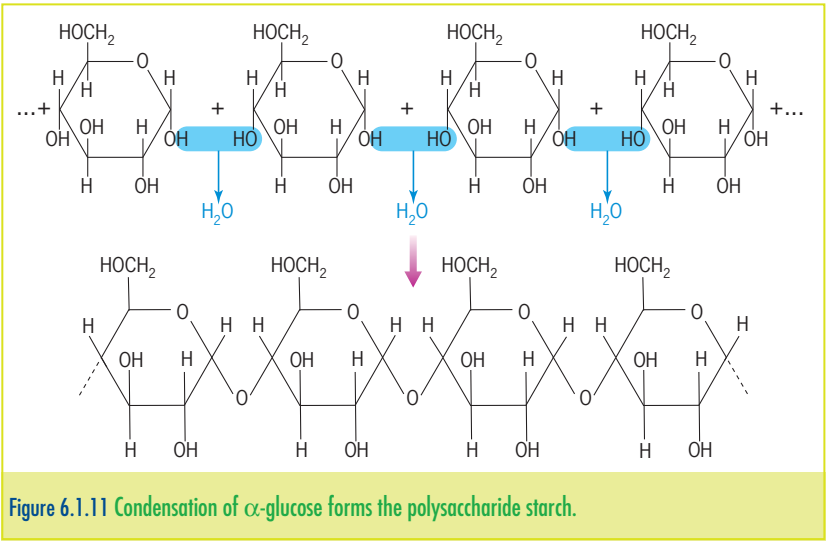


Figure 6.1.11 Condensation of α-glucose forms the polysaccharide starch.

A second polysaccharide, **cellulose**, forms when β-glucose is polymerized. Cellulose is the most abundant molecule in living tissue, making up about fifty per cent of the total organic carbon in the biosphere. It is a structural polysaccharide, found in plant cell walls. The cellulose polymer is a straight-chain molecule, with these molecules acting rather like stiff rods. Extensive hydrogen bonding between the molecules produces a strong material.

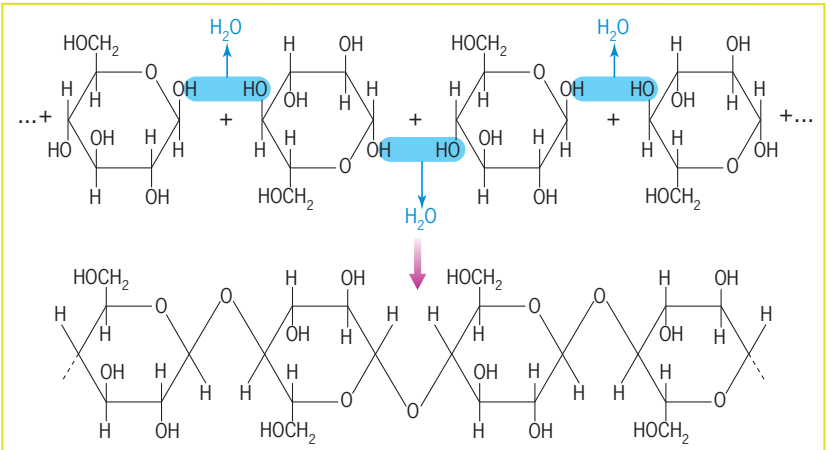


Figure 6.1.12 Condensation of β-glucose forms the polysaccharide cellulose. Note that alternate molecules are inverted to show the formation of the glycoside linkages.

The third polysaccharide, formed by the polymerization of α-glucose, is **glycogen**, a highly branched polymer. It is stored in muscles and the liver, and serves as a ready, short-term store of energy. The branched structure of glycogen makes it a more readily available source of energy than starch, since it can be hydrolysed more rapidly than the long chains of starch.

Digestion of polysaccharides involves the enzyme-catalysed hydrolysis of the bonds between the monosaccharide residues. Reactions begin in the mouth, where disaccharides are produced, and continue in the small intestines where monosaccharides are produced. These small, soluble monosaccharides are absorbed into the bloodstream. They are then transported to various parts of the body for use in energy release, or for the synthesis of energy storage molecules.

CHEM COMPLEMENT

Indigestible

Humans and most other animals lack the enzyme cellulase that is required to hydrolyse the β-1,4 linkage, so cellulose cannot be digested. Animals such as sheep, cows and koalas can digest cellulose because their stomach and other digestive system parts contain bacteria that provide the necessary enzyme for cellulose breakdown. Cellulose forms a significant component of human dietary fibre. This is material that is not hydrolysed by enzymes secreted by the human digestive tract but may be digested by microflora (bacteria) in the gut. A diet high in cellulose (dietary fibre) provides 'bulk' to aid the passage of food through the digestive system. Such 'bulk' helps prevent constipation and hemorrhoids, and reduces the calorific intake of a diet, thus helping to prevent obesity and diabetes mellitus (type 2 diabetes). Dietary fibre also helps reduce the risk of diseases such as colon cancer, diverticulosis and irritable bowel syndrome.

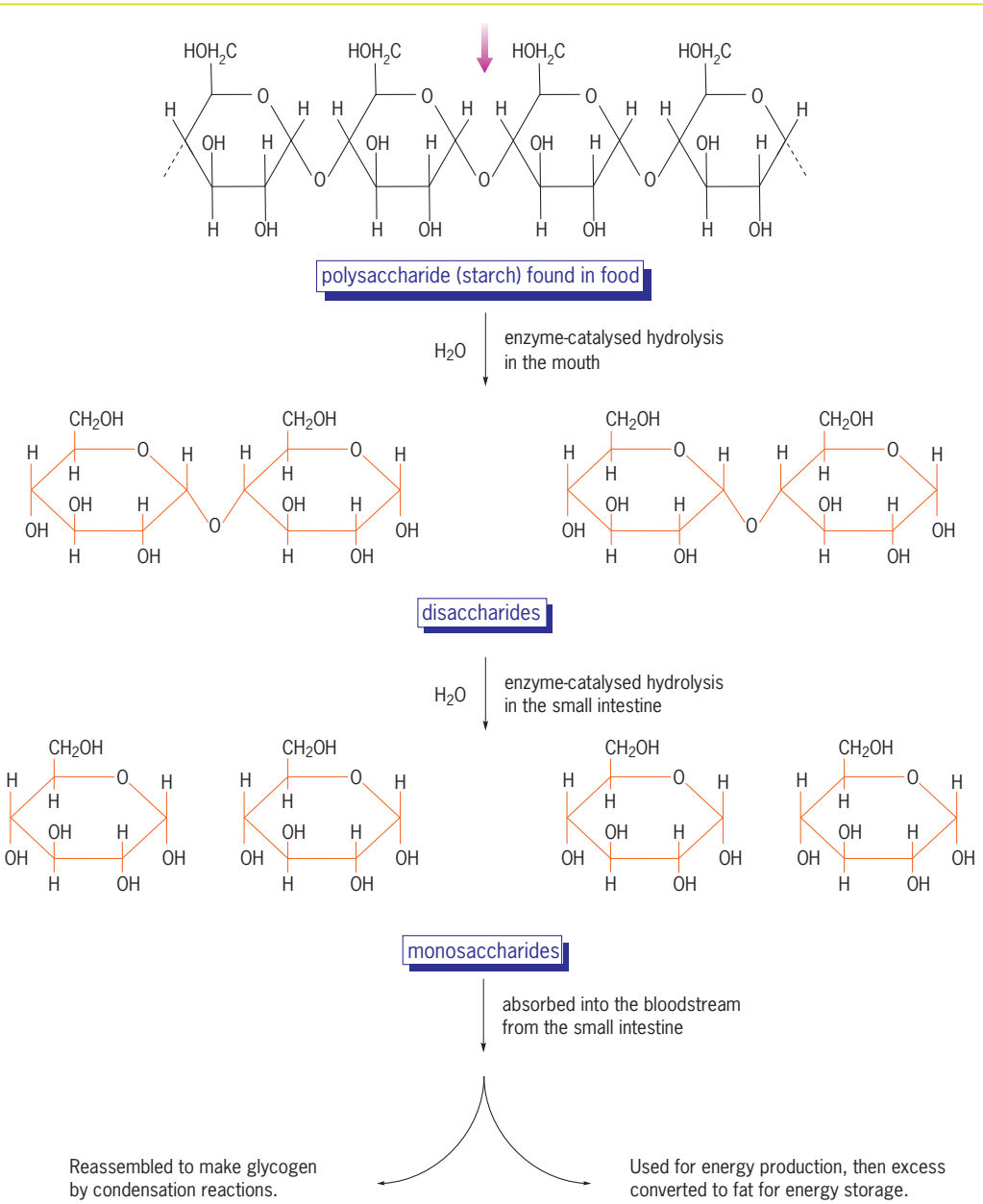


Figure 6.1.13 A summary of carbohydrate structure, digestion and use.

Lipids

Lipids consist of a broad group of compounds that are generally soluble in organic solvents but only sparingly soluble in water. They are major components of adipose tissue, and together with proteins and carbohydrates are the principal structural components of cells. Ninety-nine per cent of the lipids of plant and animal origin are glycerol esters of fatty acids, known as triglycerides, and traditionally called fats and oils. Other lipids found in the human body include phospholipids and steroids (discussed on pages 98–99). Dietary lipids play an important role in nutrition; they supply calories and essential fatty acids, act as vitamin carriers, and increase the palatability of food.

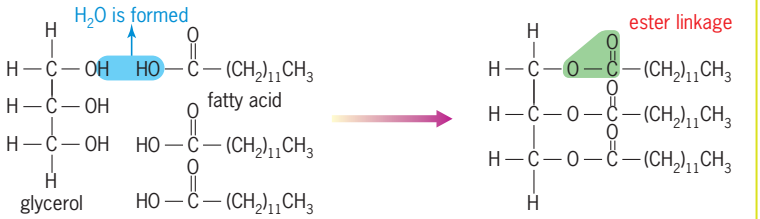


Figure 6.1.14 Triglycerides form by condensation reactions.

Triglycerides are formed by the condensation reaction between three **fatty acid** molecules (carboxylic acids with very long hydrocarbon chains) and one **glycerol** molecule. Glycerol (1,2,3-propantriol), $C_3H_8O_3$, is a polyalcohol containing three hydroxyl groups per molecule. Reaction between a hydroxyl group of glycerol and the carboxyl group of the fatty acid forms an ester linkage. Three fatty acid molecules react, so the triglyceride contains three ester linkages. Triglycerides vary according to the type of fatty acids involved in their formation. They are either fats (solids at room temperature) or oils (liquids at room temperature). We will consider these differences in section 6.2.

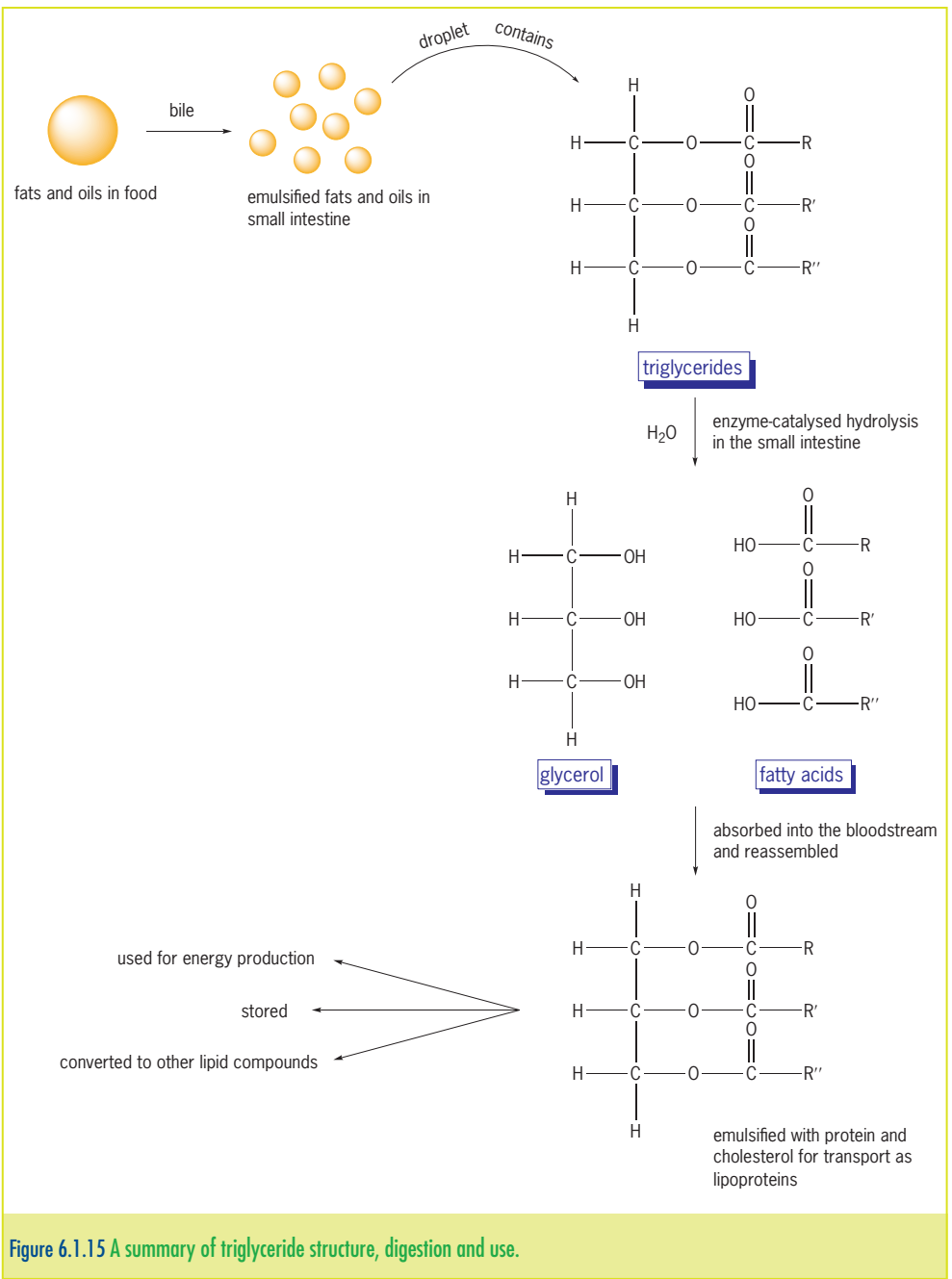


Figure 6.1.15 A summary of triglyceride structure, digestion and use.

Like proteins and polysaccharides, triglycerides undergo enzyme-catalysed hydrolysis during digestion. Triglycerides, however, are insoluble in water, due to their non-polar nature. They are therefore not easily broken down by enzymes (lipases) in the watery content of the digestive tract; thus fats tend to take longer to digest. Digestion occurs in the small intestine. Here bile, produced in the liver, emulsifies fats, dispersing them into small droplets that then become suspended in the aqueous contents of the digestive tract. This process of emulsification increases the surface area of the fats and allows the lipases to gain easier access to the triglyceride molecules, thus accelerating their hydrolysis to form glycerol and fatty acids. Following absorption from the intestines, the glycerol and fatty acids are reassembled into triglycerides. Being large and insoluble, these molecules are not easily transported in the watery blood. For transport they are converted to an emulsion with proteins and cholesterol. The **lipoproteins** formed transport the triglycerides to various locations in the body.

Section 6.1 Exercises

1 **a** Distinguish between a food and a nutrient.

b List the six major groups of nutrients.

c Three major nutrients required by humans are proteins, carbohydrates and fats. State one function that is common to these three nutrients in the human body.

2 Discuss the role that fat plays in a healthy diet and explain why you should not have a 'fat-free' diet.

3 State the major nutrient groups in each of the following foods.

a Spaghetti bolognaise

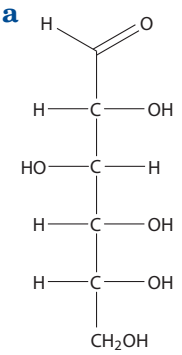
b Steak and potato chips

c Stir-fried vegetables with steamed rice

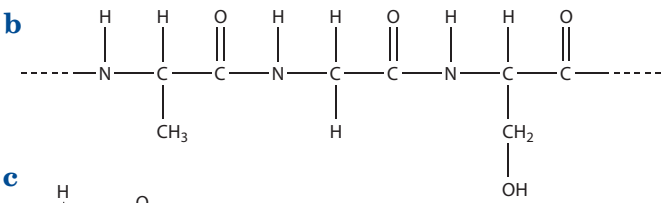
d Ham and cheese sandwich

4 Identify each of the following compounds as either a lipid, a carbohydrate or a protein. For each compound, circle and name the functional group that you used to help identify the nutrient.

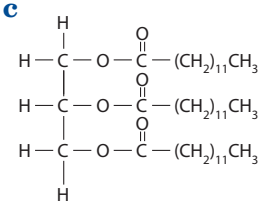
a



b



c



5 Explain why amino acids and monosaccharides are soluble in water, while fatty acids are not.

6 **a** Compare the straight chain and ring forms of glucose.

b Describe the difference between the α and β isomers of glucose.

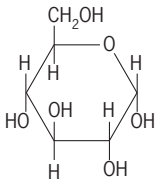
CHEM COMPLEMENT

Good and bad cholesterol

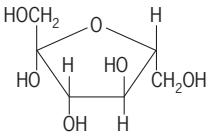
Cholesterol is produced by the liver and is found in all body tissues, where it helps to control the permeability of cell membranes. Cholesterol derivatives in the skin are converted to vitamin D when exposed to sunlight.

Cholesterol is insoluble in blood, and so is transported in the circulatory system within lipoproteins. There is a large range of lipoproteins within blood, of which two are **low density lipoproteins (LDL)** and **high density lipoproteins (HDL)**. There is no difference between the cholesterol carried by the lipoproteins; however, HDL molecules are smaller than LDL molecules, and are denser due to their larger proportion of protein. LDL molecules contain much more cholesterol than HDL molecules. High concentrations of LDL are strongly associated with cardiovascular disease. LDL promotes the narrowing of arteries (atherosclerosis) by accumulating beneath the inner elastic wall of the artery and the smooth muscle surrounding it. This disease process leads to heart attack, stroke and other diseases caused by the blockage of large peripheral arteries. As a result, cholesterol bound up in LDL is known as 'bad cholesterol'. It is hypothesized that high concentrations of HDL can remove cholesterol from cells and reduce atherosclerosis by removing cholesterol from blockages within arteries and transporting it back to the liver for excretion or re-utilization. For this reason, HDL-bound cholesterol is sometimes called 'good cholesterol'.

7 Shown below are the molecules of the structural isomers, glucose and fructose.



glucose



fructose

a Draw the disaccharide formed if these two molecules combine with each other.

b State the molecular formulas of the products of this reaction.

8 **a** Draw the general structure of an amino acid, labelling the functional groups that are common to all amino acids.

b Using table 2.2.1 (p. 81) to identify the correct side groups, draw and name an amino acid with molecular formula:

i $C_5H_9O_4N$

ii $C_4H_8O_3N_2$

9 Amino acids combine to form dipeptides and polypeptides by means of condensation reactions.

a State the name of the functional group formed by the condensation reaction between two amino acids.

b Draw the structure of a dipeptide formed when lysine and arginine react together.

c State the name of the inorganic product formed in the reaction.

10 **a** Draw the lipid structure that results when a glycerol molecule combines with a molecule of decanoic acid ($CH_3(CH_2)_8COOH$) and two molecules of stearic acid ($CH_3(CH_2)_{16}COOH$).

b State the name of the functional group formed during the reaction.

6.2 FATS AND OILS

Saturated and unsaturated fatty acids

AS F.2.1 Describe the difference in structure between saturated and unsaturated (mono- and polyunsaturated) fatty acids.

Fats and oils vary according to the type of fatty acids involved in their formation. Most naturally occurring fats and oils contain a mixture of saturated, mono-unsaturated and polyunsaturated fatty acids and are classified according to the predominant type of unsaturation present.

Saturated fatty acids have a hydrocarbon chain containing no carbon–carbon double bonds. **Mono-unsaturated fatty acids** have one carbon–carbon double bond and **polyunsaturated fatty acids** have more than one carbon–carbon double bond in the hydrocarbon chain. The formula of a fatty acid can be used to determine the number of carbon–carbon double bonds. A saturated fatty acid will have the general formula $C_nH_{2n+1}COOH$. For every double bond present in a fatty acid (the degree of unsaturation), two hydrogen atoms will be lost from the formula. Consequently the general formula of a mono-unsaturated fatty acid will be $C_nH_{2n-1}COOH$, and of a polyunsaturated

fatty acid with two carbon–carbon double bonds the formula will be $C_nH_{2n-3}COOH$. The pattern continues for other polyunsaturated fatty acids.

Fatty acids can be systematically named after their parent hydrocarbons. However, common names are still frequently used. Thus, the fatty acid $CH_3(CH_2)_{10}COOH$ is named dodecanoic acid (12 carbons) or, commonly, lauric acid. Unsaturated acids can also be named after their parent hydrocarbons. The suffix -anoic is replaced by -enoic to indicate unsaturation, and di-, tri-, and so on are used to represent the number of double bonds present. For example, $C_{17}H_{33}COOH$, is octadecenoic or, commonly, oleic acid. Location of double bonds is shown by a number (given as a prefix) for each carbon–carbon double bond. Oleic acid, for example, with one double bond between carbons 9 and 10, is named 9-octadecenoic acid. In some cases it is convenient to distinguish unsaturated fatty acids by the location of the first double bond from the methyl end of the molecule, that is, the omega carbon. Linoleic acid ($C_{17}H_{31}COOH$, 9,12-octadecadienoic acid) is therefore an omega-6 fatty acid, while linolenic acid ($C_{17}H_{29}COOH$, 9,12,15-octadecatrienoic acid) is an omega-3 fatty acid.

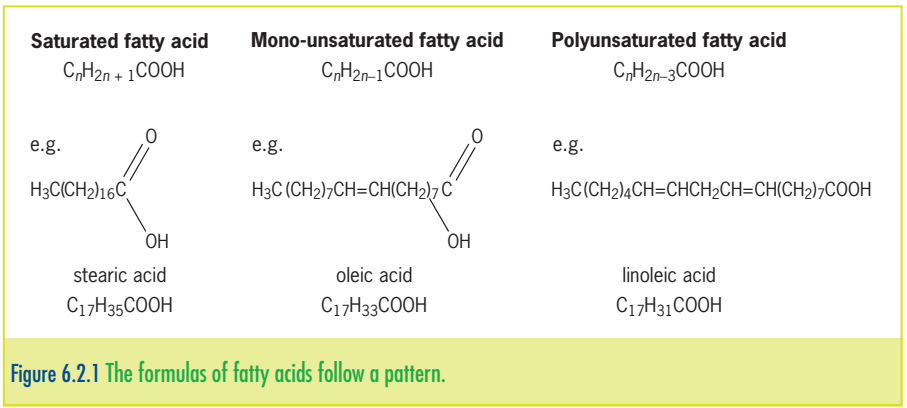


Figure 6.2.1 The formulas of fatty acids follow a pattern.

CHEM COMPLEMENT

Omega-6 and omega-3 fatty acids … which is better?

Two important fatty acids are linoleic acid, $C_{17}H_{31}COOH$, and linolenic acid, $C_{17}H_{29}COOH$. Both are essential fatty acids, as they must be included in the diet of all mammals. Their structures are very similar, differing only by one carbon–carbon double bond and the consequent reduction in number of hydrogen atoms. These two substances work together in the body. Linoleic acid is used in the biosynthesis of prostaglandins, while linolenic acid is converted into two other omega-3 fatty acids: eicosapentaenoic acid and docosahexaenoic acid. These omega-3 fatty acids help reduce inflammation, while most omega-6 fatty acids tend to promote inflammation. An inappropriate balance of these essential fatty acids contributes to the development of disease, while a proper balance helps maintain and even improve health.
A healthy diet should consist of roughly two to four times more omega-6 fatty acids than omega-3 fatty acids.

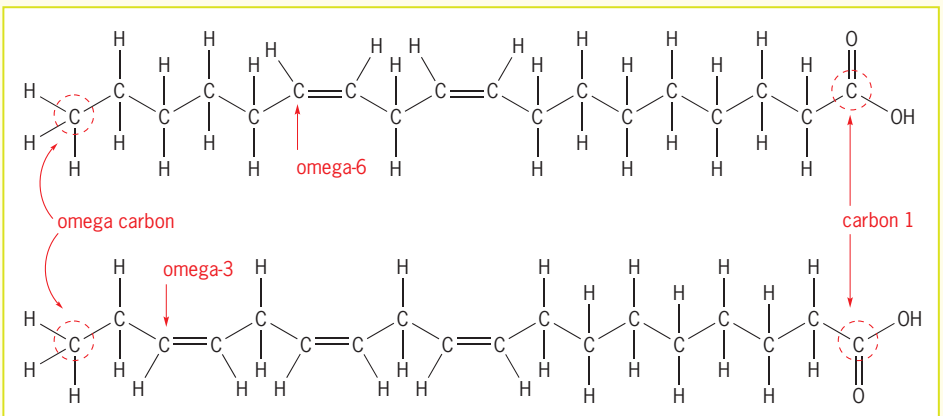


Figure 6.2.2 The structures of linoleic acid, an omega-6 fatty acid, and linolenic acid, an omega-3 fatty acid.

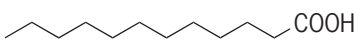
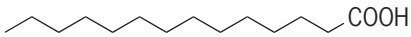
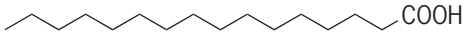
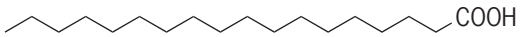
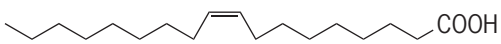
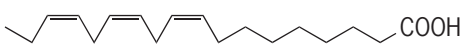
AS F.2.2
Predict the degree of crystallization (solidification) and melting point of fats and oils from their structure, and explain the relevance of this property in the home and in industry. © IBO 2007

Pr
PRAC 6.2
Detection of unsaturated fats

Unsaturation and physical properties

The presence of double bonds in the fatty acid chain influences the physical properties of the fatty acid. For example, unsaturation in fatty acids leads to 'kinks' in the chain. The unsaturated molecules therefore do not pack closely together. This leads to weaker van der Waals' forces between the chains and a lower melting point results. Saturated molecules have a regular tetrahedral arrangement around each carbon atom along the carbon skeleton. The molecules therefore pack together neatly, enabling relatively strong van der Waals' forces between the chains. The melting point of a saturated fatty acid is therefore higher than that of the unsaturated fatty acid of corresponding molecular mass. The melting points of fatty acids increase with increasing relative molecular mass (due to stronger van der Waals' forces) and increasing degree of saturation. Table 6.2.1 shows this trend for a number of important fatty acids.

TABLE 6.2.1 SOME IMPORTANT FATTY ACIDS AND THEIR PROPERTIES

Fatty acid	Number of carbon atoms	Number of double bonds	Semi-structural and schematic structural formula	Melting point (°C)
Lauric	12	0	$CH_3(CH_2)_{10}COOH$ 	44
Myristic	14	0	$CH_3(CH_2)_{12}COOH$ 	59
Palmitic	16	0	$CH_3(CH_2)_{14}COOH$ 	63
Stearic	18	0	$CH_3(CH_2)_{16}COOH$ 	70
Oleic	18	1	$CH_3(CH_2)_7CH{=}CH(CH_2)_7COOH$ 	13
Linoleic	18	2	$CH_3(CH_2)_4CH{=}CHCH_2CH{=}CH(CH_2)_7COOH$ 	–5
Linolenic	18	3	$CH_3CH_2CH{=}CHCH_2CH{=}CHCH_2CH{=}CH(CH_2)_7COOH$ 	–11

QuickTime CW Video
Testing for unsaturation with bromine

It is not only the molecular length and degree of unsaturation that determines fatty acid melting points. The geometric configuration of double bonds is also important. In the *cis* (Latin for 'on this side') configuration, the alkyl groups are on the same side of the molecule, while in the *trans* (Latin for 'across') configuration the alkyl groups are on opposite sides of the molecule. The *cis* configuration is the naturally occurring form, but the *trans* configuration is the

more thermodynamically favoured (due to less steric hindrance by the alkyl groups, as discussed in *Chemistry: For use with the IB Diploma Programme Higher Level*, p. 297). In the *trans* configuration, the fatty acid chain straightens, and becomes more like that of a saturated fatty acid, rather than kinked as in the case of naturally occurring *cis* fatty acids. This straightening raises the melting point of the fatty acid. Compare, for example, the melting points of *cis* and *trans* $C_{17}H_{33}COOH$, shown in figure 6.2.3.

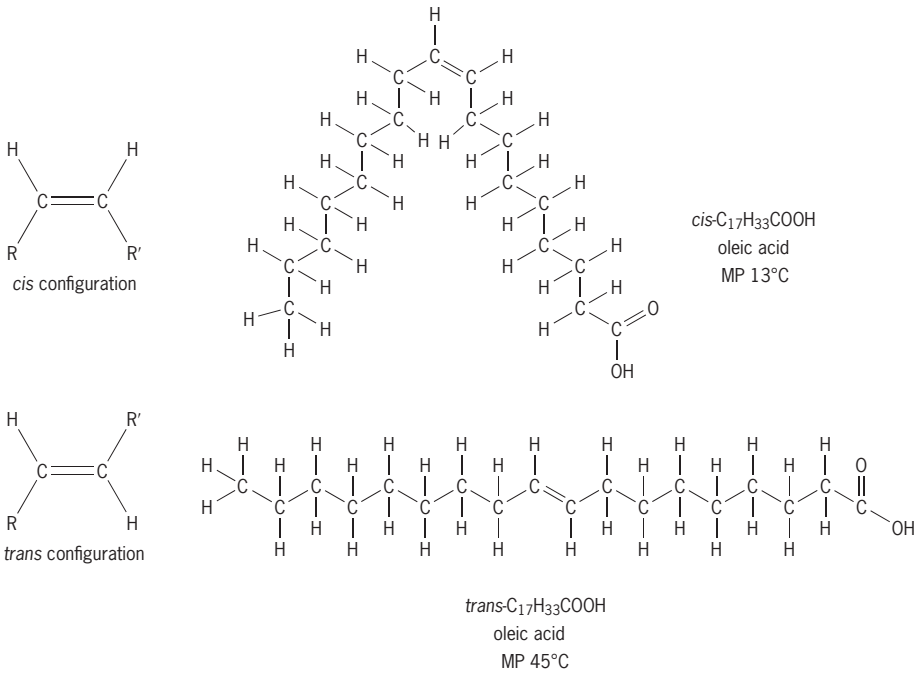


Figure 6.2.3 *Cis* unsaturated fatty acids have a lower melting point than their *trans* isomers.

Generally we classify fats as triglycerides that are solid at room temperature, and oils as those that are liquids at room temperature. Solid fats, those with the higher melting points, are more likely to be crystalline, are more saturated and/or have longer fatty acid hydrocarbon chains. Saturated fats include palm oil, coconut oil, lard, butter and shortening, all solids at room temperature. Mono-unsaturated fats such as olive, canola and peanut oils and polyunsaturated fats and fatty acids (safflower, sunflower, corn and fish oils, linoleic and linolenic acids) are liquids at room temperature.

Fats in foods are either consumed in the form of visible fats, that have been separated from their original sources, such as butter and lard, or as constituents of foods such as milk, cheese and meat. Vegetable oils come from seeds, such as those of soybeans and peanuts, and the oil-bearing trees such as palms and olives. These fats and oils in food exhibit unique physical properties. Their composition, crystalline structure, melting properties, and ability to associate with water are especially important to their functioning in many foods. For example, fats and oils are chosen for cooking on the basis of their melting temperature. Cocoa butter melts at close to body temperature and is used in confectionary. Fats chosen for cake-making melt over a wide range of temperatures. Semi-solid fats are preferred for baking because these fats mix well with flour to produce the desired texture of baked goods.

Figure 6.2.4 Differences between these fats and oils are related to their fatty acid composition.

In fats and oils that are mixtures of triglycerides, each triglyceride will have its own melting point, due to the combination of saturated and unsaturated fatty acids it contains. At a given temperature some of the triglycerides will be liquid, and some will form a crystalline solid. This leads to plasticity at certain temperatures; that is, the fat will be soft and spreadable at those temperatures. Margarines contain a mix of saturated and unsaturated fatty acids and have a wide plastic range. They are therefore spreadable when taken straight from the refrigerator. Most animal fats, with high proportions of saturated fats, have a narrow plastic range and so are hard and difficult to spread.

Fats and oils in foods may be grouped according to their origin and composition. These factors in turn influence their properties and so their uses.

- ***Milk fats*** are derived from the milk of ruminants, particularly cows. The major fatty acids of milk fat are palmitic, oleic, and stearic. This fat is unique in that it contains considerable amounts of the shorter chain acids (C_4 to C_{12}), small amounts of branched and odd-numbered acids, and *trans* double bonds.
- ***Lauric acids*** are derived from certain species of palm, such as coconut, and have low melting temperatures. These fats have high lauric acid content (40–50%), moderate amounts of the shorter C_6, C_8, and C_{10} fatty acids and small amounts of unsaturated acids.
- ***Vegetable butters*** are derived from the seeds of various tropical trees and have narrow melting ranges, due mainly to the arrangement of fatty acids in the triglyceride molecules. Cocoa butter is an important member of this group. Vegetable butters have a large ratio of saturated to unsaturated fatty acids, and are extensively used in the manufacture of confectionary.
- ***Oleic–linoleic acids*** are the most abundant. All vegetable oils contain large amounts of oleic and linoleic acids, and less than 20% saturated fatty acids. Important members of this group are corn, peanut, sunflower, safflower, olive, palm and sesame oils.
- ***Linolenic acids*** contain substantial amounts of linolenic acid. Examples include soybean and wheat germ.
- ***Animal fats*** consists of fats from domestic land animals (e.g. lard). All contain large amounts of C_{16} and C_{18} fatty acids, medium amounts of unsaturated acids, mostly oleic and linoleic, and small amounts of odd-numbered acids. These fats have relatively high melting points. Egg lipids are an important member of this group because of their emulsifying properties and their high cholesterol content.
- ***Marine oils*** are made by lower plant forms such as marine microalgae and accumulate in fish via the marine food chain. They typically contain large amounts of long-chain omega-3-polyunsaturated fatty acids, with up to six double bonds, and they are usually rich in vitamins A and D.

Figure 6.2.5 Why is margarine more spreadable than butter?

Unsaturation and chemical reactions

AS F.2.3
Deduce the stability of fats and oils from their structure.

In addition to influencing melting temperatures, the degree of unsaturation of fats and oils also influences their chemical properties. Saturated fats are more chemically stable than unsaturated fats. The carbon–carbon double bonds in unsaturated fats react in a number of ways: with enzymes, heat and water (hydrolysis); with oxygen (**auto-oxidation**); with light (**photo-oxidation**) and with hydrogen (**hydrogenation**).

Hydrolysis may occur by the addition of water across the double carbon–carbon bond in unsaturated fatty acids to produce products containing hydroxyl groups. In addition to the hydrolysis of the carbon–carbon double bond, hydrolysis of ester bonds in all fats and oils may occur by enzyme action, especially aided by heat and moisture, resulting in the liberation of free fatty acids. These free fatty acids are more susceptible to oxidation than are esterified fatty acids. The release of fatty acids by hydrolysis is sometimes responsible for the development of an undesirable rancid flavour. This **hydrolytic rancidity** will be considered in section 6.3.

Auto-oxidation is one of the major causes of food spoilage. While both saturated and unsaturated fatty acids undergo chemical decomposition when exposed to heat in the presence of oxygen, the reaction of fats and oils with oxygen typically occurs by the addition of oxygen across the carbon–carbon double bond. A variety of compounds are formed, including **hydroperoxides** (ROOH), aldehydes and ketones, which give rise to an unpleasant taste. Oxidation is therefore of great economic concern to the food industry because it leads to the development of various 'off' flavours and odours that make the foods less acceptable. This oxidative rancidity will be considered in sections 6.3 and 6.7. In addition to causing rancidity, oxidation reactions can decrease the nutritional quality of food, and certain oxidation products are potentially toxic. On the other hand, under certain conditions, a limited degree of oxidation may be desirable, as in aged cheeses.

Photo-oxidation (oxidation catalysed by light) of fatty acids leads to the formation of products that are similar but not identical to those of auto-oxidation. Hydroperoxides are formed, and later transformed to produce aldehyde derivatives. Both photo-oxidation and auto-oxidation produce complex mixtures of products.

Antioxidants are substances that can delay onset, or slow the rate, of oxidation of auto-oxidizable materials. Literally hundreds of compounds, both natural and synthesized, have been reported to have antioxidant properties. Antioxidants will be considered in sections 6.3 and 6.8.

AS **F.2.4**
Describe the process of hydrogenation of unsaturated fats. ©IBO 2007

Hydrogenation of fats and oils involves the addition of hydrogen to double bonds in the fatty acid chains. The process is very important in the food industry, as it accomplishes two major objectives:

- it allows the conversion of liquid oils into semi-solid fats more suitable for specific applications, such as in margarine, and
- it improves the oxidative stability of the oil.

The oil is first mixed with a suitable catalyst (finely divided metal (Zn, Cu, Ni)), heated to the desired temperature (140–225°C), then exposed to hydrogen under pressure. Stirring aids in dissolving the hydrogen, creating a uniform mix of catalyst and oil, and in dissipating the heat generated. The progress of the reaction is usually monitored by determining the change in **refractive index** of the oil, as this index is related to the degree of saturation of the oil. When the desired end point is reached, the hydrogenated oil is cooled and the catalyst is removed by filtration.

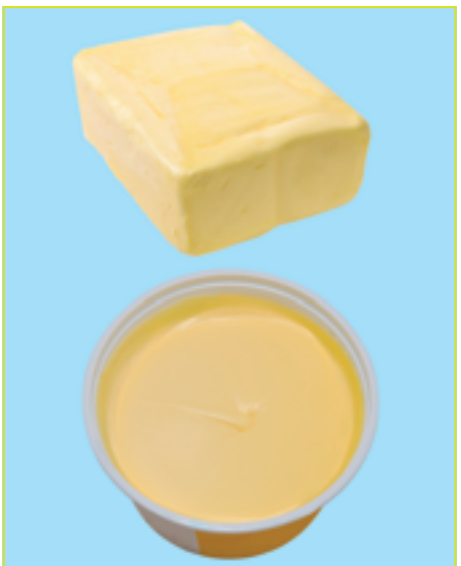


Figure 6.2.6 Butter or margarine … which is better?

Hydrogenation, first developed in the 1890s by Paul Sabatier and further improved by the German chemist Wilhelm Normann in 1901, was found to be the way to decrease the number of double bonds in a triglyceride and consequently increase its melting point. With butter not always as available as consumers would like, the possibility of a mass-produced vegetable-based substitute was attractive. The difference between butter and vegetable oils is the degree of saturation. Butter is made up of saturated triglycerides, so has a higher melting point than polyunsaturated vegetable oils. Hydrogenation enabled the conversion of the polyunsaturated vegetable oils to margarine, a more solid butter substitute.

During hydrogenation, not only are some of the double bonds saturated, but some may also be relocated and/or transformed from the usual *cis* to the *trans* configuration. This further raises the melting point of the oil, making it more convenient to use. *Trans* fatty acids may constitute 20–40% of the total acids of some margarines. However, *trans* fatty acids are not biologically equivalent to their *cis* isomers. Although precise knowledge of

WORKSHEET 6.1
Proteins, carbohydrates and lipids

F.2.5
Discuss the advantages and disadvantages of hydrogenating fats and oils. ©IBO 2007

their physiological properties, metabolism, and long-term effects on health remain controversial, *trans* fats were suggested as early as 1988 to be the cause of increases in coronary artery disease. It is generally agreed that the consumption of *trans* fatty acids is detrimental to our health. It increases the risk of coronary heart disease by raising the levels of LDL cholesterol in the blood, and also by lowering the levels of HDL cholesterol. In recent years health organizations have lobbied governments for the removal of *trans* fats from foods, and currently many countries require that *trans* fats be included on the labelling of foods so that consumers can avoid them, as their presence may be just as bad for the health as the saturated fats they have replaced!

Partial hydrogenation of oils may thus result in the formation of a relatively complex mixture of reaction products, depending on which of the double bonds are hydrogenated, and the type and degree of isomerization. The situation is further complicated because the starting materials already contain an extremely complex mixture of molecules. Additional chemicals may also be added to the hydrogenated oil. These include artificial colourings to make it appear yellow and buttery, additives to make it smell like butter, vitamins, antioxidants, emulsifiers and binding agents to hold it all together. Table 6.2.2 summarizes the advantages and disadvantages of the hydrogenation process.

TABLE 6.2.2 ADVANTAGES AND DISADVANTAGES OF HYDROGENATION OF FATS AND OILS

Advantages	Disadvantages
Changes a liquid oil to a semi-solid or solid, to make the melting point of an unsaturated fat more like that of a saturated fat	Mono- and polyunsaturated fats are healthier for the heart than saturated fats
Decreases the rate of oxidation of the fat or oil (stability increases with increasing saturation)	In partial hydrogenation, *trans* fatty acids can form. *Trans* fatty acids are hard to metabolize, accumulate in fatty tissue, are difficult to excrete from the body, increase levels of LDL (bad) cholesterol and are a low-quality energy source.
Increases hardness	
Controls the feel and plasticity (stiffness)	

For decades fats and oils have been at the subject of debate regarding toxicity, obesity and disease. The type and amount of fats in the diet is one of several factors believed to have an influence on the incidence of coronary heart disease. Fatty acid features of special attention have been chain length, degree and position (omega-3 and omega-6) of unsaturation, and geometric isomerism (*cis* versus *trans*). Some studies suggest that a high intake of saturated fatty acids is also associated with increased risk of colon, prostate and breast cancers. These effects of fats on human health are the subject of debate, and the relationship between dietary fats and health is an active area of nutritional and medical research. Consumers are bombarded with information and advice that can be contradictory and confusing. Complex metabolic interrelationships are involved, so experimental work in this area is complicated. Although an appropriate balance of mono-unsaturated, omega-3 and omega-6 fatty acids is advisable, the role of dietary fat in health and disease remains controversial. The best advice for most people is the same as it has been for years: have a well-balanced diet that includes plenty of all essential nutrients, fresh fruits and vegetables, and regular exercise.

Section 6.2 Exercises

1 **a** Name two chemicals which combine to form a fat.

b Describe the difference between a fat and a fatty acid.

2 **a** Explain what is meant by the terms:

i saturated fatty acids

ii mono-unsaturated fatty acids

iii polyunsaturated fatty acids

iv essential fatty acid.

b Give the name and formula for a common example of each type of fatty acid listed in part **a**.

3 **a** Explain why palmitic acid, $CH_3(CH_2)_{14}COOH$, has a higher melting point than myristic acid, $CH_3(CH_2)_{12}COOH$.

b Explain why palmitic acid has a higher melting point than linoleic acid, $CH_3(CH_2)_4CH{=}CHCH_2CH{=}CH(CH_2)_7COOH$.

c Explain why *trans*-oleic acid, $CH_3(CH_2)_7CH{=}CH(CH_2)_7COOH$, has a higher melting point than *cis*-oleic acid.

4 Explain, with reference to the molecular level, why olive oil exists as a liquid and lard exists as a solid.

5 Describe the reactions occurring when unsaturated fatty acids undergo:

a auto-oxidation

b photo-oxidation

c hydrolysis.

6 Describe the process of hydrogenation of fats and oils.

7 Discuss the advantages and disadvantages of hydrogenating fats and oils.

6.3 SHELF LIFE

How do we know when a food has gone off? Do you smell it, look at the use-by date, or are there other clues? Sometimes it is obvious that a food is no longer safe to eat, such as when it is covered by colourful fungi or bacterial colonies. At other times, the indications are more subtle and easy to miss.

Figure 6.3.1 The appearance of microbes indicates that this food should not be eaten.

Shelf life

Shelf life is the length of time a food can be stored under specified conditions before its quality begins to diminish. This differs from the 'expiration' date because a food may be past its shelf life before the expiration date has been reached. The expiration date exists to indicate when a food is no longer safe to eat. Once a food is past its shelf life, the quality expected of that food by the consumer will not be present; the appearance (colour and mass), flavour, texture and smell may have changed, or microbial spoilage may have occurred. In 1993, the Institute of Food Science & Technology (IFST) defined shelf life as the period of time during which the food product will:

1 remain safe

2 be certain to retain desired sensory, chemical, physical and microbial characteristics

3 comply with any label declaration of nutritional data when stored under the recommended conditions.

AS **F.3.1**
Explain the meaning of the term *shelf life*. © IBO 2007

CHEM COMPLEMENT

Labels—what do they mean?

We all know that foods have to be date-stamped, but do you know what the different labels mean? A 'use-by' date indicates the end date that the food is safe to eat, if it has been stored correctly. Any retailer selling food after its use-by date will face hefty fines. A 'best-before' date indicates the end of the period that the food will retain its quality, and therefore is a good indicator of shelf life. After the best-before date, the food is still safe, but might not be pleasant to eat if the quality has changed dramatically.

AS **F.3.2 Discuss the factors that affect the shelf life and quality of food. © IBO 2007**

The shelf life of foods will vary widely; a potato will not have the same shelf life as a bottle of wine. The estimated shelf lives of four different foods are shown in table 6.3.1.

TABLE 6.3.1 ESTIMATED SHELF LIFE

Food	Shelf life
Pasta	2 years
Ice cream	2 months
Sweet potatoes	2 weeks
Asparagus	2 days

Factors affecting shelf life

Storage conditions affect the shelf life of foods. Think about how long foods last without refrigeration in summer compared to winter; temperature is a major factor affecting shelf life. Factors that affect shelf life fall within two major groups: compositional factors and environmental factors.

Compositional factors are related to the chemical make-up of the food. These include:

- ***pH***—The pH has a major effect on micro-organism (microbe) growth. If the pH of the food is within that needed for optimal growth of one of the microbes present, the food will spoil more quickly. Likewise, in high acid foods (pH < 4.6), bacteria tend not to grow well. *Salmonella* bacteria, which cause food poisoning, don't grow below pH 4.0.
- ***Water activity*** **(a_w)**—This refers to the amount of water in the food not bound to food molecules. Free water can support microbial growth. The water activity scale goes from 0 to 1. Foods with low water activity, such as biscuits ($a_w = 0.3$), are less likely to grow microbes than foods with higher water activity, such as fresh meat ($a_w = 0.99$). Note that water activity is not the same as water or moisture content. Although moist foods are likely to have greater water activity than dry foods, this is not always so; in fact a variety of foods may have exactly the same moisture content and yet have quite different water activities.

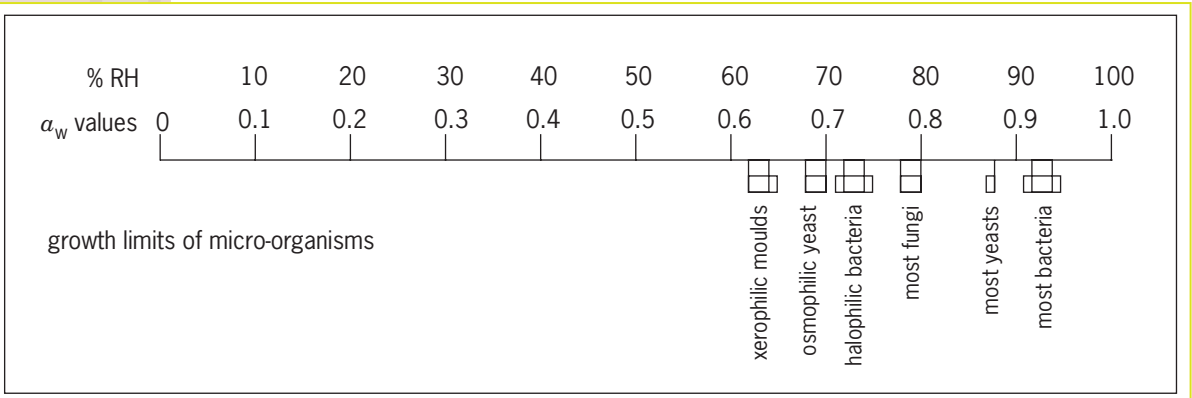


Figure 6.3.2 Water activity and microbial growth chart (RH = relative humidity).

- ***Presence of food additives***—Additives such as preservatives can be used to prolong the shelf life of food and retard microbial growth. One example is sodium benzoate (E211), a preservative found in many soft drinks.
- ***Quantity of reactive substances***—If the food contains substances prone to oxidation or other reactions, it will tend to have a relatively short shelf life, especially if exposed to air. Asparagus only has a shelf life of two days because, once picked, chemical reactions (in particular respiration) occur that result in stem wrinkling and hardening, as well as a loss of green colour. Enzyme action also contributes to a reduced shelf life.
- ***Microbe populations***—Food spoilage will also depend on the natural microbe populations present in the food. For example, *salmonella* bacteria are common in chicken and *Escherichia coli (E. coli)* bacteria are common in beef.

Environmental factors are most significant because they affect the rate of food spoilage reactions. Some factors are listed below.

- ***Temperature***—Temperature directly affects the rate of chemical reactions. In general, the rates of chemicals reactions rise with increasing temperature. This does not hold true for enzymatic reactions because the enzymes denature at high temperatures. Moderate increases in temperature will tend to reduce food shelf life. Heat can also cause the loss of volatile organic compounds that contribute to smell and flavour.
- ***Light***—Both artificial and natural light can adversely affect shelf life. Photodegradation reactions are common. Effects of light include the development of off-flavours in milk, rancidity in fats and oils, loss of vitamins A and C, and the loss of red colour from meat due to photo-oxidation of the pigment.
- ***Contact with air***—When exposed to air, some foods can be oxidized. Vitamin C is easily oxidized, as are many of the pigments that give foods their colour. Fats and oils are oxidized, even during freezing, and become rancid.
- ***Humidity (water content)***—Foods can take on or lose moisture, and humidity will influence this. Dehydration is a commonly used preservation method, but slow water loss can lead to Maillard browning (particularly at 15–20% moisture) and nutrient loss. Some foods absorb atmospheric moisture. This may cause loss of nutrients, browning and rancidity, while dry foods will become more vulnerable to microbial spoilage as their water content increases. Texture is also greatly affected by water content.

Many of the changes that occur in foods during spoilage are important indicators that the food is unsafe to eat. **Rancidity** produces volatile organic compounds with unpleasant odours and the colours produced by oxidized pigments are not enticing. If you go so far as to taste the food, the off flavours produced by compounds created during spoilage should tell you not to eat any more. If you think a food is not quite right, your instincts are probably correct.

Figure 6.3.3 Examples of foods that have exceeded their shelf life.

Rancidity

AS F.3.3 Describe the rancidity of fats. © IBO 2007

Unsaturated fats and oils are susceptible to oxidation across the carbon–carbon double bond. The carbon chain can split at the double bond and then react further. The sequence of reactions produces smaller volatile organic compounds with noxious odours and flavours. When this happens, the foodstuff containing the oxidized lipids is described as rancid. Lipids can also become rancid when the triglyceride molecule hydrolyses into glycerol and fatty acids. Some of the fatty acids have unpleasant odours and flavours. Hence, rancidity

Figure 6.3.4 Hákarl is considered a delicacy in Iceland, where it is prepared by allowing shark meat to hang until it is rancid.

is a noticeable deterioration in sensory quality. This particularly refers to odours and flavours, but texture and appearance can also be affected.

Many different compounds can be produced in these complex processes, such as carboxylic acids, aldehydes, ketones and peroxide derivatives. For example, butyric (butanoic) acid is the substance chiefly responsible for the sharp, unpleasant odour of rancid butter. You probably wouldn't be tempted to eat a food that smells rancid, but if you did, it is interesting to note that rancidity also reduces the nutrient value of the food by depleting fat-soluble vitamins.

Two different types of reactions that can lead to rancidity are hydrolysis and oxidation. The pathway taken to rancidity depends very much upon reaction conditions and triglyceride structure. **Hydrolytic rancidity** (also called lipolytic rancidity) is not of much concern in food safety; it is **oxidative rancidity** that leads to the formation of potentially harmful compounds. However, sometimes rancidity is desirable, as in the production of strong cheeses or the Icelandic specialty *Hákarl*—rancid shark meat.

Hydrolytic and oxidative rancidity

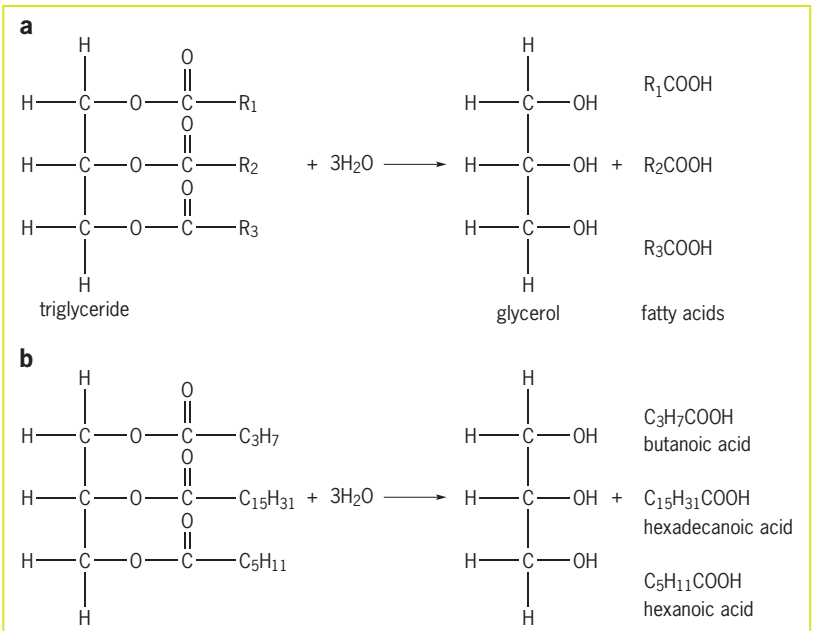


Figure 6.3.5 (a) General lipid hydrolysis and (b) the hydrolysis of a milk fat triglyceride.

AS F.3.4 Compare the processes of hydrolytic and oxidative rancidity in lipids. © IBO 2007

Hydrolytic rancidity is the breaking down (hydrolysis) of a lipid, producing glycerol and its component fatty acids, as shown in figure 6.3.5. This type of rancidity mainly occurs due to the action of enzymes called lipases, and it is particularly a problem in dairy products such as milk and butter. Long-chain free fatty acids are not detectable to the human palate, except to impart an oily or waxy sensation. Milk contains lipids with shorter hydrocarbon chains, which produce short chain fatty acids upon hydrolysis. This is helped along by lipoprotein lipase (LPL), an enzyme which may be **endogenous** (naturally present) or secreted by bacteria.

It is the shorter chain fatty acids that produce the flavours and smells typical of rancid dairy products. Any agitation will hasten lipolysis, but the enzyme denatures at 55–60°C, so is destroyed during pasteurization. Heat and moisture can also hasten hydrolytic rancidity. For example, when potato chips are fried, water leaves the chips. The water then causes some of the oil to hydrolyse. Table 6.3.2 shows some of the shorter chain acid products of lipid hydrolysis that create rancidity in dairy foods and oil.

Oxidative rancidity produces different compounds to those produced by hydrolytic rancidity. Any food containing a high proportion of unsaturated fatty acids, such as mackerel, herring and other oily fish, will be vulnerable to this type of rancidity. The oxidation can occur whether the fatty acids are free or esterified in triglycerides, and typically involves the addition of oxygen across the carbon–carbon double bond of the unsaturated fatty acid. The process can be catalysed by light (photo-oxidation), enzymes (such as lipoxygenases), trace metals and free radical species.

TABLE 6.3.2 LIPID HYDROLYSIS PRODUCTS

Food	Fatty acid name	Semi-structural formula	Characteristics
Milk and butter	Butanoic	$CH_3(CH_2)_2COOH$	Unpleasant odour, acrid taste
Milk only	Hexanoic Octanoic	$CH_3(CH_2)_4COOH$ $CH_3(CH_2)_6COOH$	Goat odour Sweat/body-odour smell
Palm and coconut oil in cocoa butter	Hexanoic Octanoic Decanoic Lauric (dodecanoic)	$CH_3(CH_2)_4COOH$ $CH_3(CH_2)_6COOH$ $CH_3(CH_2)_8COOH$ $CH_3(CH_2)_{10}COOH$	Goat odour Sweat/body odour smell Sweaty smell, soap flavour Soapy smell and flavour
Chocolate	Palmitic Stearic oleic	$CH_3(CH_2)_{14}COOH$ $CH_3(CH_2)_{16}COOH$ $CH_3(CH_2)_7CH{=}CH(CH_2)_7COOH$	Oily flavour Soapy flavour Oily flavour

Enzymatic oxidation occurs when lipoxygenase enzymes are present. These enzymes are found in all legume seeds. They go to work when cell walls are ruptured and can cause unpleasant flavours, even in foods with relatively low fat content. Photo-oxidation is initiated by light. It is usually of little significance if foods are handled in bulk. One example is the metallic taste acquired by milk left in direct light, which is due to the formation of the ketone 1-octene-3-one.

The most significant type of oxidative rancidity is auto-oxidation, which occurs via a free radical chain mechanism (see section 6.7, p. 370 for Higher Level students only). Figure 6.3.6 shows four compounds formed in rancid meat and fish.

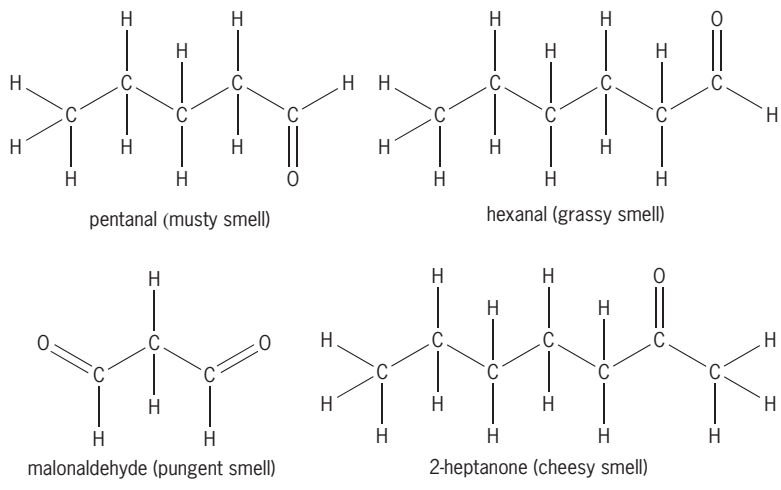


Figure 6.3.6 Oxidation products in meat and fish.

AS F.3.5
Describe ways to minimize the rate of rancidity and prolong the shelf life of foods.

Prolonging shelf life

There are many ways in which we can slow rancidity and make food last longer. Some involve the use of additives, while others aim to reduce exposure to reactive species or conditions.

Refrigeration and freezing are effective at slowing reactions, but not stopping them altogether. All foods last longer at lower temperatures, particularly dairy products. Rancidity will still occur, but it will take longer than if the foods were left at room temperature. It is important not to let the food reach room temperature before putting it back in the refrigerator because a phenomenon called temperature activation can occur, promoting lipolysis of dairy products.

Reducing exposure to light limits photo-oxidation. Storing foods in dark places or in opaque containers can extend their life. Moisture content is also significant. As previously mentioned, a high water activity (a_w) promotes microbial growth. The presence of moisture also makes fat and oil hydrolysis more likely. Sun or air drying is an ancient method of food preservation. Water activity can be reduced by adding **humectants**—water-absorbing chemicals such as glycerol. Water can also be drawn from the food using salt or sugar in a **curing** process. Smoking is often used to preserve meat. Heat from a smoke fire removes moisture, while aromatic hydrocarbons add flavour, aroma and preserving chemicals.

Irradiation is a method to deactivate enzymes, kill bacteria and sterilize food. Milk is **pasteurized** to reduce microbe populations significantly. In the ultra-high temperature (UHT) treatment, the milk is heated to 138°C for a fraction of a second. This treatment also denatures approximately 80% of the lipoprotein lipases, thus inhibiting hydrolytic rancidity.

Methods of packaging aim to reduce or avoid contact with moisture and atmospheric gases in order to considerably increase shelf life. Foods such as potato crisps are kept fresh using an inert, dry nitrogen atmosphere inside the packet. Vacuum-packing, in which the air is sucked out of the container, achieves the same aim. For longer-term storage in vacuum-packaging, gas-impermeable packaging is required. Problems can sometimes occur when molecules seep into or out of packages. Lettuce leaves and other fresh produce are kept fresh in an atmosphere of increased carbon dioxide and reduced oxygen. In cans and jars, foods are best preserved if the container is filled as much as possible, thus leaving little room for air. **Hermetic sealing** (airtight seal) is also important.

Figure 6.3.7 Meat packed in airtight vacuum-packaging will have a lower rate of oxidation than meat exposed to air.

Additives may be used to prolong the shelf life of foods. The addition of a weak acid is an effective preservation method because few bacteria can grow when the pH falls below 4. Another possibility is to add a substance that will inhibit oxidation, or will be preferentially oxidized. Food additives are identified by numbers. For example, additive 220 indicates sulfur dioxide, a common antioxidant. Table 6.3.3 lists the major food additives used to prevent food spoilage.

TABLE 6.3.3 FOOD PRESERVATION ADDITIVES

Additive type	Function	Examples
Nitrites and nitrates	Stabilize flavour, give colour to meat, prevent rancidity, protect against botulism pathogens	• Sodium and potassium nitrite and nitrate are used in curing ham and in other processed meats such as salami, hot dogs, pepperoni salami, bologna, bacon and SPAM.
Antioxidants	Protect against oxidation by being preferentially oxidized	• Butylated hydroxyanisole (BHA) is used to delay oxidative rancidity in lipids. • Citric acid is used as a flavouring and preservative in foods.
Non-acid antimicrobials	Prevent the growth of micro-organisms	• Sodium propanoate prevents mould and bacterial growth in cheeses. • Calcium propanoate is used as a preservative in bread.
Anti-browning agents (sulfites and hydrogen sulfites)	Stop food, especially fruit and vegetables, from non-enzymatic browning	• Sodium hydrogensulfite and sodium sulfite are added to wine and dried fruits to maintain colour.
Acids	Both organic and inorganic acids may be added to lower the pH and reduce microbe growth; may add to the flavour	• Propanoic (propionic) acid inhibits mould growth in cheese. • Ethanoic acid and benzoic acid delay mould and bacterial growth in pickled meats and fish products. • Sorbic acid is used in a range of foods.

Traditional methods

AS F.3.6 Describe the traditional methods used by different cultures to extend the shelf life of foods. © IBO 2007

Food poisoning is nothing new—people have been dying from it for thousands of years—so an awareness that food can kill if not properly treated has also been around for a long time. In days gone by, food preservation was essential so that people could be fed on long sea voyages, in battles, and during periods of cold or drought. Without refrigeration or any of the modern methods such as irradiation and vacuum-packing, people came up with often ingenious ways to prolong the shelf life of their food.

As mentioned in the previous section, drying is the most common and oldest known method. It is also very effective, since microbes cannot grow without unbound water. Dried meat was the staple diet of nomads and the ancient Egyptians used to store dried grain. Smoked meats and fish have also been around for hundreds of years.

Freezing and cooling were only available in certain places. In Scandinavian regions, it was common to freeze-dry fish, and Romans used mountain snow in icehouses to keep food fresh. Salting is known to have been used in ancient China. Through osmosis, adding salt causes water to leach out of the food. There was much experimentation with substances other than salt. For hundreds of years, spices such as cinnamon and cloves have been used to preserve food.

Fermentation has been around since the Sumerians accidentally discovered it about 6000 years ago via a piece of soggy bread. Fermentation produces acids and/or ethanol, both of which act as antimicrobials. Wine, beer, cheese and yoghurt were common to ancient civilizations. Pickling in weak acids such as vinegar dates back to the Mesopotamians in about 2400 BC. Pickle lovers in history include Julius Caesar and Aristotle.

In 1795, Emperor Napoleon Bonaparte of France offered a 12 000 franc reward to anyone who could come up with a method of food preservation suitable for his troops. In 1809, Nicolas Appert won the prize by putting food in jars, which he then stoppered and heated. A year later, Englishman Peter Durand patented a tin canister, from which we get the word 'can'. Interestingly, the can opener wasn't invented until forty-eight years after the tin can. In 1938, a can of veal packed in 1824 was opened and found to be in good condition.

Antioxidants

AS F.3.7 Define the term *antioxidant*. © IBO 2007

Antioxidants are present in many foods, and are common food additives. If the popular press is to be believed, they are also a potent antidote to the less pleasant effects of ageing. An antioxidant is a substance that slows the rate at which another substance is oxidized, or simply delays its oxidation. Antioxidants do this by being preferentially oxidized themselves. The presence or addition of antioxidants can significantly increase a food's shelf life, preventing oxidative reactions that produce off flavours, aromas and colours.

AS F.3.8 List examples of common naturally occurring antioxidants and their sources. © IBO 2007

Many foods naturally contain antioxidants. A diet high in these foods is rumoured to fight disease and age-related degeneration. Table 6.3.4 lists a selection of foods containing the four most common natural antioxidants.

TABLE 6.3.4 SOURCES OF NATURAL ANTIOXIDANTS

Antioxidant	Food sources
Selenium	Fish, red meat, eggs, chicken, shellfish, whole grains, brazil nuts, garlic
α-carotene	Carrots, sweet potatoes, squash, broccoli, kale, spinach, red palm oil, peaches, tomatoes, apricots, mango, papaya, cantaloupe (rockmelon), melon
Ascorbic acid (vitamin C)	Citrus fruits, green leafy vegetables, green peppers, broccoli, strawberries, kiwifruit, rosehip, raw cabbage
Vitamin E (contains the antioxidants called tocopherols)	Soybean, sunflower, olive and canola oils, wheat germ, nuts, seeds, whole grains, green leafy vegetables

Pr PRAC 6.3 Investigating antioxidents

Antioxidant structure

AS F.3.9 Compare the structural features of the major synthetic antioxidants in foods. © IBO 2007

Most synthetic antioxidants are **phenolic** compounds (they have a hydroxyl group bonded to an aromatic ring). By far, the two most common synthetic antioxidants added to foods are BHA (butylated hydroxyanisole or 2- and 3-*tert*-butyl-4-hydroxyanisole) and BHT (butylated hydroxytoluene or 3,5-di-*tert*-butyl-4-hydroxytoluene). Both of these compounds are steam volatile, and so can be lost during processes such as vacuum drying. BHA exists in two isomeric forms, as shown in figure 6.3.8.

Both these compounds are fat-soluble and are used to help delay oxidative fat rancidity. In addition, BHA is used as a yeast-defoaming agent and BHT is used in packaging materials. They are most commonly used at levels of 100–200 ppm (parts per million or mg per kg). These phenolic antioxidants use their hydroxyl groups to react with oxygen in preference to the food they are contained in.

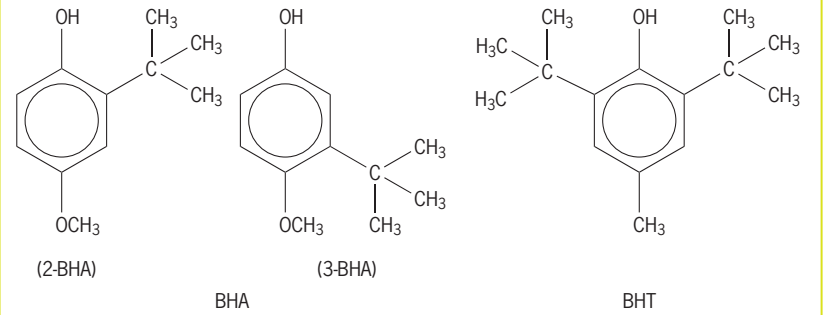


Figure 6.3.8 BHA (butylated hydroxyanisole) and BHT (butylated hydroxytoluene), two synthetic antioxdants.

Two other common synthetic antioxidants are mono-*tert*-butylyhydroquinone (TBHQ) and propyl gallate (PG). TBHQ has moderate oil solubility and slight solubility in water. It is very effective at stabilizing oils. Propyl gallate (PG) is particularly widely used in the United States to inhibit oxidative rancidity of lipids. Like BHA and BHT, TBHQ and PG are phenolic and contain reactive hydroxyl groups. Their structures are shown in figure 6.3.9. Another phenolic antioxidant is 2,4,5-trihydroxybutyrophenone (THBP), also shown in figure 6.3.9.

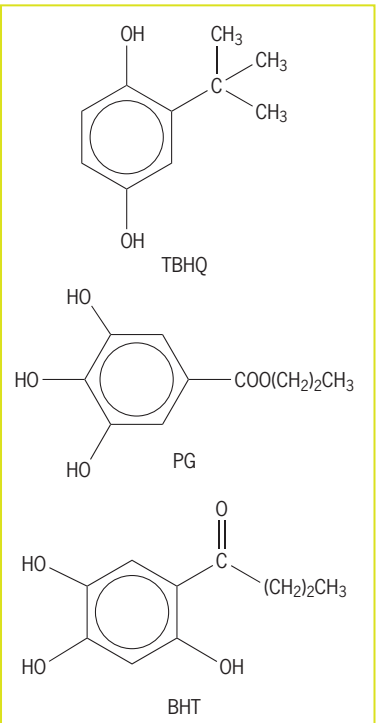


Figure 6.3.9 TBHQ, PG and THBP.

It seems that nature has taught us which antioxidants work best; the five synthetic antioxidants bear a striking resemblance to the two natural antioxidants shown in figure 6.3.10—ascorbic acid (vitamin C) and vitamin E—which also feature reactive hydroxyl groups.

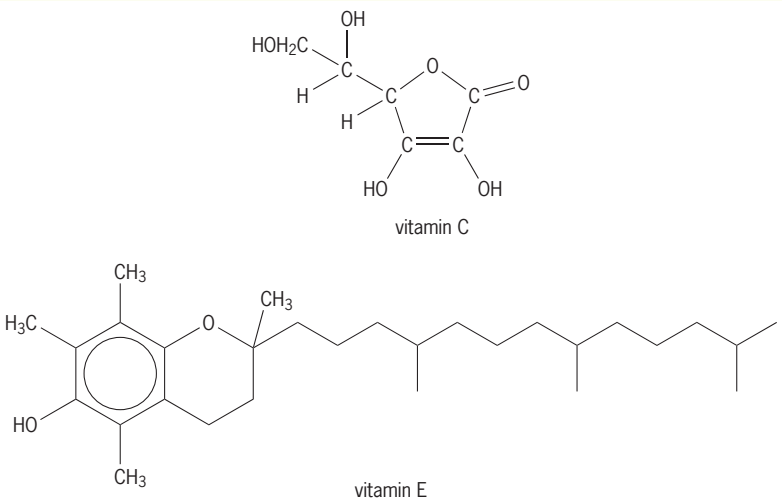


Figure 6.3.10 Vitamins C and E are natural antioxidants.

Advantages and disadvantages of antioxidants

Both synthetic and natural antioxidants have their advantages and disadvantages. The major disadvantage of synthetic antioxidants is their perception by consumers. There has been a growing trend towards 'natural' and 'non-chemical' products in the past 30 years. People are wary of synthetic antioxidants because they fall into the dreaded food additive class. Additives have been identified as being the cause of many modern health problems. As additives, synthetic antioxidants are strictly controlled, and there are clear

AS **F.3.10**
Discuss the advantages and disadvantages associated with natural and synthetic antioxidants. © IBO 2007

rules governing the maximum allowable levels and which foods they can go into. There may be toxicological effects at higher levels, but they have been proven safe if used within guidelines.

An advantage of synthetic antioxidants is their purity; any impurities would be present in trace amounts. Because of their purity, their properties are constant. Many preparations of natural antioxidants are not well-purified because that would increase their price. They are more expensive than synthetics even in impure form. Since the preparations are mixtures of unspecified composition, they have variant properties and questionable safety. In addition, the dosages cannot be controlled; safety tests are not required for natural products. Despite consumer perception, natural does not equal benign, and toxicity is highly dose-dependent. Natural antioxidants have also been shown to be less efficient as antioxidants than their synthetic counterparts. They may also impart undesirable colours, flavours or smells to food. This contrasts with synthetic antioxidants, which are not detectable to humans at the recommended dosages.

Pr

PRAC 6.4
Determination of vitamin C in fruit juices

Despite the lack of dosage control, many natural antioxidants have been safely used for centuries. An advantage of natural antioxidants is their multi-faceted capabilities. Not only do they act as antioxidants, they also have other important roles. Vitamin C is important to immune system functioning and collagen synthesis. Vitamin E is also thought to play a role in the immune system and can increase the uptake of other vitamins. β-carotene is a precursor to vitamin A, needed for healthy eyes. It can also be used as a natural food colour. BHT has been shown to have antiviral properties, but most synthetic antioxidants do not have the multiple benefits of natural antioxidants.

Figure 6.3.11 The name 'oregano' comes from a Greek phrase meaning 'joy of the mountains'.

TABLE 6.3.5 ADVANTAGES AND DISADVANTAGES ASSOCIATED WITH NATURAL AND SYNTHETIC ANTIOXIDANTS

Natural antioxidants	
Advantages	**Disadvantages**
• Multiple roles (e.g. vitamin C has multiple roles in the body, not just as an antioxidant) • Safely used for centuries	• Impure (variable properties; hard to control dose) • Use is usually not regulated • Change taste, colour and odour of food • More expensive than synthetics • Less effective than synthetics
Synthetic antioxidants	
Advantages	**Disadvantages**
• High purity (constant properties; easy dosage control) • Use is closely regulated by authorities and therefore safer • Can't detect within food • Cheaper than natural antioxidants • Greater efficacy	• Poor consumer perception (consumers prefer 'natural' substances)

Antioxidants in traditional foods

Nutritionists agree that it is always better to get your nutrients from foods rather than supplements. Foods traditionally eaten for their health benefits often contain high concentrations of antioxidants. Table 6.3.6 lists some traditional foods and the antioxidants they contain.

AS F.3.11
List some antioxidants found in the traditional foods of different cultures that may have health benefits.

TABLE 6.3.6 FOODS AND ANTIOXIDANTS

Food	Found in	Antioxidant	Health benefits
Green tea	Asian cultures	Catechins	Protect against degenerative diseases, including heart problems
Blueberries	Pemmican, a North American Indian dish	Anthocyanidins	Enhance vitamin C functioning and strengthen collagen, cancer protection
Oregano	First used as healing herb in ancient Greece	Rosmarinic acid	Antioxidant, anti-inflammatory, anti-microbial
Dark chocolate	Xocoatyl, a bitter drink imbibed by the Aztecs	Gallic acid equivalents, flavonoids	Lower blood pressure, degenerative disease protection
Sage	Used as a healing herb in the middle ages	Rosmarinic acid	Antioxidant, anti-inflammatory, anti-microbial
Cranberries	Native North American fruit used by native Indians as a food and dye	Anthocyanins	Protect against cancer and heart disease, lower LDL, urinary tract anti-microbial
Turmeric	Deep orange-yellow powder commonly used as a spice in curries and other South Asian cuisine	Curcumin	Antioxidant, anti-arthritic and anti-inflammatory
Kidney beans	South American cuisine	Flavonoids	Degenerative disease protection

Section 6.3 Exercises

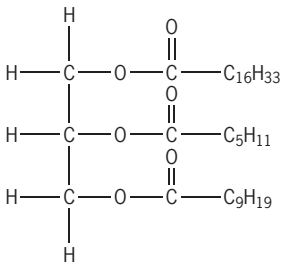
1 Explain what is meant by the term *shelf life*.

2 Explain why it is important that consumers should be aware of the shelf life of foods.

3 Discuss the chemical factors that cause a decrease in shelf life.

4 Define the term *rancidity*.

5 Draw the reaction products when the triglyceride shown at right undergoes hydrolytic rancidity.



6 Compare the processes of hydrolytic and oxidative rancidity in lipids.

7 Describe three methods that you can use to delay rancidity and prolong the shelf life of food.

8 **a** Define the term *antioxidant*.

b Vitamin C is an example of a natural antioxidant. Vitamin C is found in citrus fruits. Name one other food that also contains vitamin C.

c List three other naturally occurring antioxidants and state the foods in which they are found.

9 Describe the structural features common to most antioxidants.

10 Discuss the advantages and disadvantages of natural and synthetic antioxidants.

6.4 COLOUR

Throughout the early 21st century the HJ Heinz Co, one of the world's leading manufacturers of tomato sauce (ketchup) had been experimenting with different coloured products besides the more traditional red tomato sauce. We have all heard of blue cheese, blueberries and chicken cordon bleu … so why not blue tomato sauce? As well as the colour blue, Heinz has also produced its tomato ketchup in green, purple, pink, orange and teal colours. Some consumers bought this product as a novelty item but soon returned to their beloved red tomato ketchup as it was too weird serving blue tomato sauce on French fries or green tomato sauce on barbequed hamburgers. The taste of the coloured tomato sauce was identical to that of the traditional red tomato ketchup, but the colour threw consumers off. As a result, Heinz has stopped producing these coloured tomato sauces.

Figure 6.4.1 Colour plays a key role in the consumers' view of food and also in the nutritional content of foods.

DEMO 6.1
Activated charcoal

There was a similar short-lived phenomenon during the early 1990s when the Pepsi-Cola Co. produced a clear cola instead of the more traditional dark soda. The unique feature of this soda was that unlike other colourless soft drinks, which usually have a lemon-lime flavour, Pepsi Clear tasted very similar to regular Pepsi. Consumers did not respond to this new product, as they either wanted a dark Pepsi or a lemon-lime flavoured clear soda, and this product was soon withdrawn from sale.

Both of these experiments, which centred on the alteration of the colour of well-known products, show the importance of food colour on the marketability of a consumer product. Humans have always used the colour of a food to form judgments about its desirability. The act of eating is a multi-sensory experience, involving the perceptions of sight, taste, smell and touch. Colour provides visual information about a food's quality and condition, and influences the perception of its flavour. The ancient Egyptians were the first to add colour to their food through the use of the yellow-coloured plant part—saffron. Rich Romans ate white bread that had been coloured by the addition of alum, $KAl(SO_4)_2.12H_2O$. The colour of food is not only an indication of its nutritional value but also of its inherent medicinal powers.

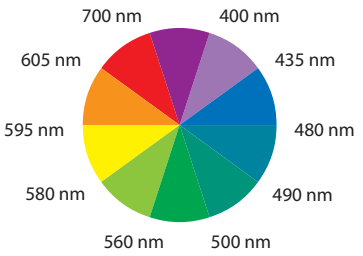


Figure 6.4.2 A colour wheel can be used to explain what colour of visible light is absorbed by a food and what complementary colour we see when we look at a certain food.

The different colours that we see in foods are the result of some wavelengths of visible light (400–700 nm) being absorbed by the chemical bonds and substituents of the substances in food and other wavelengths being reflected.

Interactive CW Animation
Colour wheel

The light that is reflected is called a complementary colour. Whatever wavelength of light is not absorbed becomes emitted as the colour we see. For example a carrot absorbs blue light and therefore appears orange, and a green pepper absorbs red light and therefore appears green.

AS F.4.2 Explain the occurrence of colour in naturally occurring pigments. © IBO 2007

We will examine the relationship between colour and food chemistry in the following paragraphs.

There are two general substances that are responsible for the colour in foods—**pigments**, which are natural, and **dyes**, which are synthetic. Pigments are natural colours are found as suspensions, as they are insoluble, in the cells of plants and animals. Dyes are water-soluble substances that can be added to food to change or enhance its natural colour. Food dyes are manufactured to a higher standard than dyes used to colour fabrics.

AS F.4.1 Distinguish between a dye and a pigment. © IBO 2007

CHEM COMPLEMENT

The 'bug' colour

One way to naturally achieve the red colouring of candy, dairy products, drinks, fruit fillings and surimi is to add the pigment carmine, which comes from the insect *Dactylopius coccus* that lives on cacti from the genus *Opuntia*. The insect is native to tropical South and Central America and produces the pigment as a deterrent against other insects.

The pigment is obtained from the body and eggs of the insect.

The cochineal dye was used by the Aztec and Maya peoples of Central and North America. In 1788, Captain Arthur Phillip (founder of Botany Bay) brought infested cacti to Australia to ensure that there was a reliable source of red pigments. The actual colour in carmine comes from carminic acid ($C_{22}H_{20}O_{13}$); it is the double bonds that are responsible for the absorption of a specific wavelength of light.

Figure 6.4.3 Cochineal insects cover cacti. Their bodies and eggs are crushed to extract the carmine pigment.

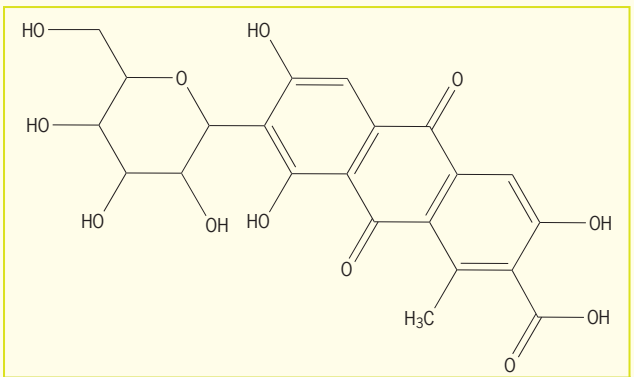


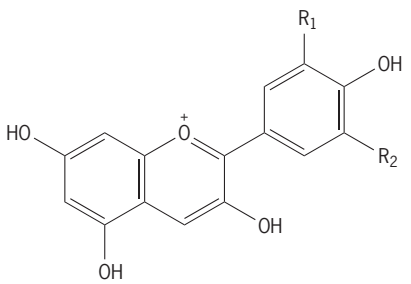
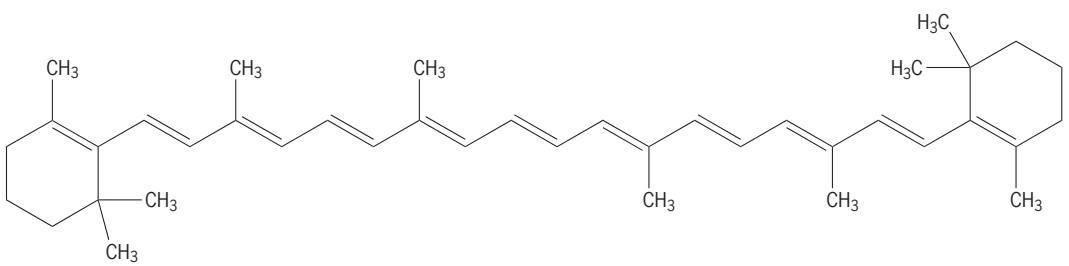
Figure 6.4.4 Carminic acid.

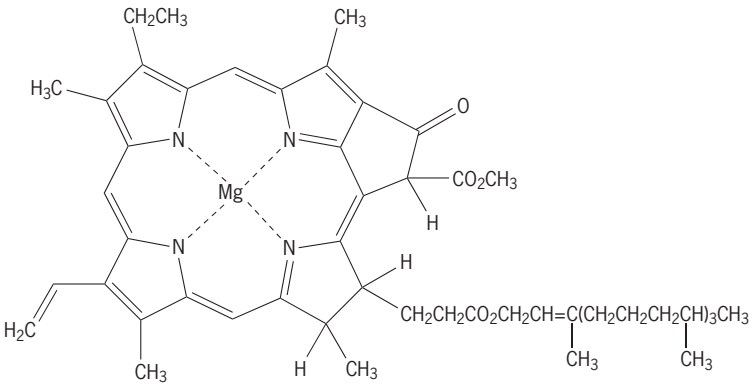
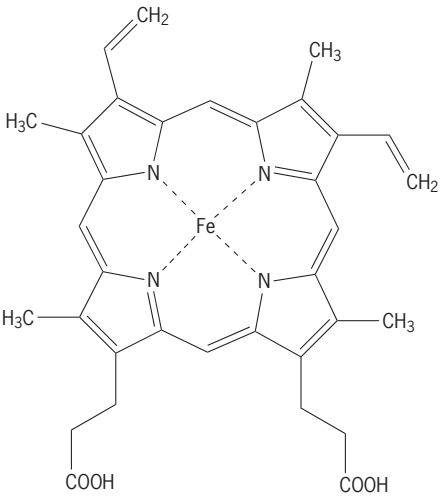
AS F.4.3
Describe the range of colours and sources of naturally occurring pigments anthocyanins, carotenoids, chlorophyll and heme.

Pigments

The reason for the addition of colourings to food has changed little since the 18th and 19th centuries when colouring was added to food to disguise inferior food. The first food colours used were natural pigments, many of which remain in use today. Table 6.4.1 lists some of the most common pigments.

TABLE 6.4.1 NATURALLY OCCURRING PIGMENTS

Pigment	
Anthocyanins	
Colours	Orange-red to red-blue
Sources	Blackcurrant, raspberry, eggplant, orange, blueberry, cherry, red grapes, strawberry, cranberry, redcurrant and avocado
Functions	• Attracting pollinators and animals which will disperse the seeds • Plant sunscreen
Foods used in	Candy, fruit beverages, ice cream, yoghurt and jams
Interesting additional facts	• Colour is dependent on pH level of the food; acidic foods are the reddest and basic foods are more bluish • Most widely occurring pigment in plants • Water-soluble
Structure	
Carotenoids	
Colours	Yellow to orange to red
Sources	Oranges, carrots, tomatoes, bananas, watermelon, red/yellow peppers, saffron, various bacteria, fungi and algae and seaweed
Functions	Colouration in plants and animals
Foods used in	Meat products, cheese, butter, spice mixes, salad dressings and soft drinks
Interesting additional facts	• The most widespread pigment in nature, more than 1000 million tonnes are produced each year • Fat-soluble • Red **astaxanthin**, when present as a complex with protein, gives the blue or green hue found in live lobsters and crabs and the pink colour of salmon and flamingos
Structure	

Chlorophyll	
Colours	Green to olive-green
Sources	Green leafy plants, algae and bacteria
Functions	Photosynthesis in plants
Foods used in	Green pasta and dehydrated spinach
Interesting additional facts	• There are five different naturally occurring variations of chlorophyll • Fat-soluble
Structure	
Heme and myoglobin	
Colours	Red to red-purple
Sources	Red blood cells and muscle tissue from meat, poultry and fish
Functions	Oxygen transport in animals
Foods used in	Blood pudding, sausage, pepperoni, salami and other processed meats
Interesting additional facts	• Myoglobin is responsible for the purple-red colour of meat • Myoglobin structure contains a heme molecule as well as 153 different amino acids
Structure	

AS F.4.4
Describe the factors that affect the colour stability of anthocyanins, carotenoids, chlorophyll and heme.

QuickTime CW Video
Anthocyanin indicators

Pr
PRAC 6.5
Using an anthocyanin indicator

Anthocyanins stability

There are more than 500 naturally occurring **anthocyanins**. They are water-soluble due to their phenolic groups (–OH attached to a benzene ring), a property that restricts their usage. Because of their water-soluble nature, they are sensitive to changes in pH. When the pH of the food changes, their colour changes. For example the colour of red cabbage is enhanced with the addition of vinegar (ethanoic acid). On the other hand, when cooked in aluminium pans, which cause a more alkaline environment, the colour changes to purple and blue. The chemistry behind this change can be explained using a series of complex equilibria.

quinoid (A) ⇌ flavylium (AH^+) ⇌ carbinol (B) ⇌ chalcone (C)
blue red colourless yellow

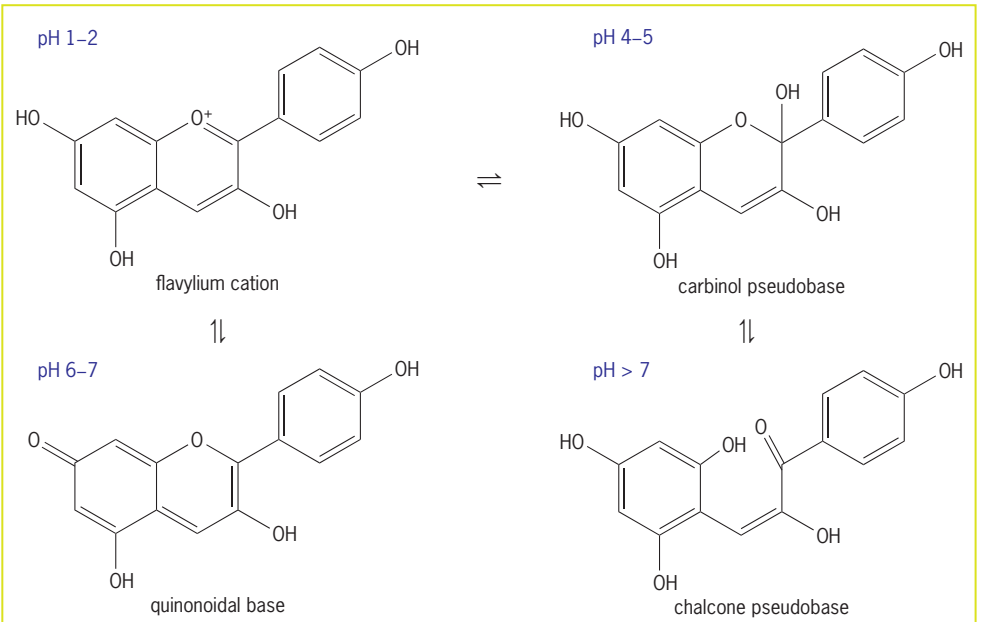


Figure 6.4.5 Structural transformations and colour changes in anthocyanins at various pH levels.

The various colours found in anthocyanins are also susceptible to temperature, oxygen, UV light and different co-factors such as metal ions. They will form deeply coloured coordination complexes with Fe^{3+} and Al^{3+}. This explains why foods that are rich in red anthocyanins when cooked in aluminium pans or cast iron frying pans are often discoloured into unappetizing shades of blue. Additionally, fruits and vegetables that are manufactured in cans containing ferric or aluminium ions can also become discoloured, when the ions leach from the can into the food. The high temperature necessary for pasteurization of canned food may also be responsible for the discolouration and browning of canned foods that contain anthocyanins, as anthocyanins are denatured at high temperatures.

Carotenoid stability

Carotenoids are fat-soluble and therefore do not readily dissolve in water. They are not very susceptible to changes in pH and are quite stable in the pH range 2–7. In a very acidic medium, for example if you cook carrots in vinegar-infused water, the bright orange-red colour will be replaced with a dull orange-yellow colour. They are stable up to 50ºC and therefore are not degraded by most forms of processing.

When you refer back to the structure of β-carotene that is found in table 6.4.1, a common carotenoid, which functional group occurs most frequently? The presence of these carbon–carbon double bonds makes carotenoids susceptible to **oxidation** (addition of oxygen to the double bonds), and they can then be catalysed by light, metals, hydroperoxides and even the chemicals found in cigarette smoke. Oxidation results in changes to the bonds in carotenoids which in turn causes bleaching of colour, loss of vitamin A activity and off odours. This is one reason to keep your fresh fruit and vegetables in dark cupboards or in the refrigerator to protect their pigments from exposure to light, which can degrade them. Also, when preparing your fruit and vegetables you should use a plastic or wooden chopping board, for if you use a metal chopping board it will catalyse the destruction of these beneficial pigments.

Chlorophyll stability

Chlorophyll is a mixture of several highly complex molecules, which consist of a ring structure (the porphyrin structure) with a central magnesium ion, and a long hydrophobic side chain, which means that they are not water-soluble. Chlorophyll can be added unlimitedly to nearly all foods since it is a relatively safe pigment. However, as the colour is rather unstable, the applications are limited. Air, light, heat and pH all negatively influence its stability. Chlorophyll-coloured products should be dry and not exposed to light, air or high temperatures. In many cases the products are packed in dark packaging with a modified atmosphere to prevent chlorophyll degradation. When heated or exposed to light, the cell membrane of the plant deteriorates, releasing acids and therefore decreasing the pH. At lower pH levels, the magnesium atom at the centre of the chlorophyll molecule (refer to the structure in table 6.4.1) is replaced with two hydrogen ions, resulting in the formation of an olive-brown pheophytin complex. During the canning of green vegetables the amount of chlorophyll content can be reduced by as much as 80–100%. An olive-brown coloured vegetable does not appear as pleasing to the eye as a bright green one.

Figure 6.4.6 Notice the difference in colour between canned peas and fresh peas.

In chlorophyll-containing foods, one way to stop the replacement of the magnesium ion by hydrogen ions is by the addition of zinc or copper ions. The zinc and copper complexes have a green colour similar to that of chlorophyll, and are considerably more stable in acidic solutions. Thus, hydrogen ions will not replace the zinc or copper present in these complexes, and the olive-brown pheophytin will not be formed. This 'regreening' effect of copper and zinc ions has been known since the 1940s; however it is closely regulated by governmental agencies.

Interactive CW Animation
Heme group

Heme stability

Freshly butchered meat has a deep purplish-red colour due to the presence of **myoglobin**. Myoglobin is a water-soluble protein that stores oxygen for aerobic metabolism in the muscle. It consists of a protein portion and a nonprotein porphyrin ring with a central iron atom. The iron atom is an important player in meat colour. The factors that primarily determine the colour of meat are the oxidation state of the iron and the compounds that are attached to the iron portion of the molecule. Different types of muscle also have different amounts of myoglobin and therefore result in slightly different colours.

As oxygen from the air comes into contact with the exposed meat surfaces it is absorbed and binds to the iron. This added oxygen causes the meat to turn to a bright red colour as the new pigment **oxymyoglobin** is formed. This is the usual colour of meat that we see in our local butcher shop or supermarket. In both myoglobin and oxymyoglobin, the iron is in the Fe^{2+} state. An electron can be lost from this ion; that is, the iron undergoes oxidation and **metmyoglobin** is formed. This pigment has a brown colour and results in meat that looks like it has gone bad. The conversion between these three forms is readily done.

myoglobin (Mb)	$\rightleftharpoons$	oxymyoglobin (MbO_2)	$\rightleftharpoons$	metmyoglobin (MMb)
purple-red, Fe^{2+}		red, Fe^{2+}		brown, Fe^{3+}

Figure 6.4.7 The bright red colour of this meat is due to the oxymyoglobin formed when oxygen from the air comes into contact with the exposed meat surface.

AS F.4.6
Compare the two processes of non-enzymatic browning (Maillard reaction) and caramelization that cause the browning of food. © IBO 2007

The pigments myoglobin, oxymyoglobin and metmyoglobin are freely interchangeable, depending on the conditions at which the meat is stored. Storing the meat in conditions free of oxygen can be achieved by using packaging films with low gas permeabilities. The meat can also be packed in an oxygen-reduced atmosphere. An oxygen-reduced atmosphere is one in which oxygen is replaced with another gas such as carbon dioxide, which is then used as the storage gas.

Non-enzymatic browning of food

Browning of food refers to the process by which food turns brown. This process can occur via either enzymatic or non-enzymatic means. We will focus on the non-enzymatic browning of foods in this section. The most common type of non-enzymatic browning occurs in the **Maillard reaction**. In the early

twentieth century, Louis-Camille Maillard happened upon what came to be known as the Maillard reaction when he was trying to find out how amino acids linked up to form proteins. He discovered that when he heated sugars and amino acids together, the mixture slowly turned brown. In this type of reaction the actual chemical composition of the food is changed. A condensation reaction occurs between the amino acids of a protein and either glucose or lactose (both **reducing sugars**). During this condensation reaction a molecule of water is removed. The optimum range for the Maillard reaction is pH 4–7, which is the most common pH of foods.

WORKSHEET 6.2
Colour in foods

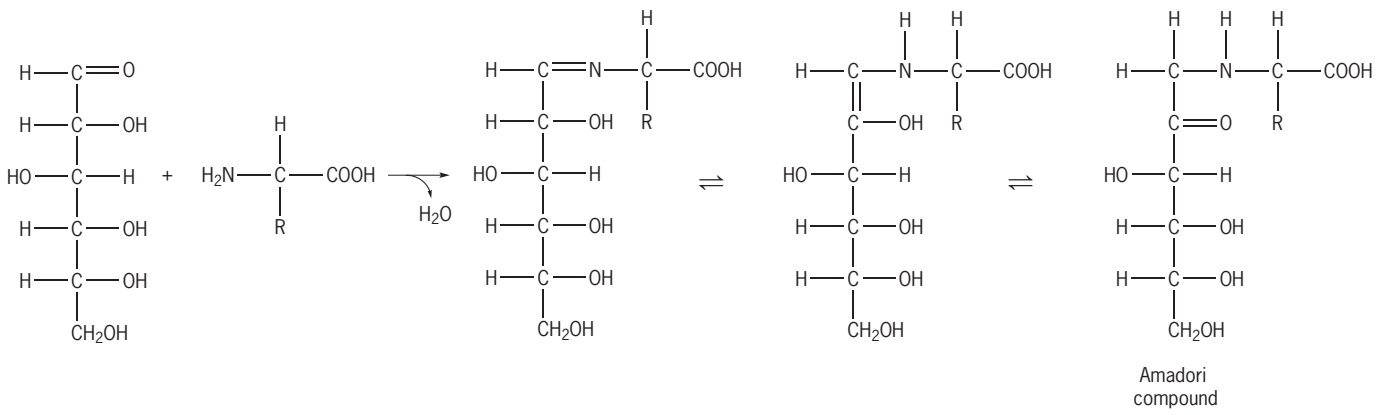


Figure 6.4.8 The first step in the Maillard reaction between glucose and an amino acid.

The Maillard reaction occurs readily when you are frying meat and has an optimum temperature of 149–260°C, although it can occur at temperatures as low as 49°C. The browning of meat results in the production of the hundreds of different flavour compounds that makes browned meat so tasty. The presence of the amino acid lysine results in the most browning colour and cysteine the least colour.

Maillard reactions are important in baking, frying or otherwise heating of nearly all foods. Maillard reactions are (partly) responsible for the flavour of bread, cookies, cakes, meat, beer, chocolate, popcorn, cooked rice, toffees, caramels and fudges.

The caramel colour and flavour may be undesirable for some types of foods. For example, in the Second World War soldiers complained that their powdered eggs turned brown and developed unappealling flavours. Laboratory studies showed that the unappetizing tastes were coming from the browning reaction. The eggs were stored at room temperature, but the concentration of amino acids and sugars in the dehydrated mix was high enough to produce a reaction.

A second way to cause browning in foods is **caramelization**, which leads to desirable colour and flavour in bakery goods, coffee, beverages, beer and peanuts. Undesirable effects of caramelization include a burned sugar smell and blackening. Caramelization occurs during dry heating and roasting of foods with high sugar content and without-nitrogen containing compounds (for example proteins). Conditions for caramelization are optimal when pH levels are either below 3 or above 9 and the temperature is above 120°C. Generally caramelization can be thought of as the removal of water from a sugar, although in reality it involves a series of reactions to reach the desired outcome. As the process occurs, volatile chemicals are released, producing the characteristic caramel flavour.

Figure 6.4.9 The brown crust formed on freshly baked bread is due to the Maillard reaction between amino acids and reducing sugars.

AS F.4.5
Discuss the safety issues associated with the use of synthetic colourants in food. © IBO 2007

Food colouring regulation

The colourants permitted for use varies greatly among countries and can be traced as far back as a 1396 edict in France that prohibited the use of colouring agents in butter. Many of the synthetic food colourings have been found to have adverse effects; they may be carcinogenic, or cause hyperactivity and allergic reactions. Most countries now require that all ingredients of food must be placed on nutritional labels. This legislation is not mandatory in all countries, which may put the safety of some of your favourite foods into question. For example, that can of fruit cocktail sitting in your pantry may contain maraschino cherries, naturally white, that have been dyed with FD&C Red No. 3—Erythrosine (E127). It is an approved food colourant in both Australia and the US. It was slated to be banned for use in the US in 1983, but lobbying from industry and manufacturing prevented the government from banning it. So this controversial chemical is still widely used in some parts of the world, but it has been banned from some grocery market shelves in the UK.

International trade is becoming increasingly important, therefore colour legislation is now of international concern. As a way to reduce the potential controversy with cross-border shipping of food products, many retailers and consumer are now focused on buying more organic and natural foods. However, the term 'natural', as it pertains to colours, has never been legally defined and has no universally accepted definition. Therefore natural foods may be no safer than foods that have synthetic colourings added to them.

Section 6.4 Exercises

1 Explain the importance that food colouring plays in the food industry.

2 **a** Explain the occurrence of colour in naturally occurring pigments.
b Explain what is meant by a complementary colour.
c If violet wavelength of light is absorbed, what is the colour of the food?
d If orange wavelength of light is absorbed, what is the colour of the food?

3 Distinguish between a dye and a pigment in food chemistry.

4 Explain how the astaxanthin found in blue-green coloured live lobsters turns into the more commonly associated red-coloured carotenoid in cooked lobsters.

5 By examining the structures of the four common pigments—anthocyanins, carotenoids, chlorophyll and heme—suggest the functional groups necessary for a substance to be coloured.

6 Explain the dependence of anthocyanin colour on pH.

7 Explain how cooking foods that are high in anthocyanins in aluminium or iron pans may influence the final colour of the foods.

8 Food is sometimes canned to preserve it. Explain why canned food that has been stored for some time may become discoloured.

9 Explain the difference in colour you see between fresh and canned peas, commenting specifically on the chemistry involved. How may one eliminate the differences between the peas?

10 Comment on the connection between oxymyoglobin, myoglobin and metmyoglobin.

11 Compare the two processes of browning in terms of chemical composition of the food involved, factors that increase the rate of the browning, and products and examples.

12 Compare the governmental regulations regarding food colouring in your country with those of another country. Discuss your findings in light of a global marketplace.

6.5 GENETICALLY MODIFIED (GM) FOODS

AS F.5.1 Define a genetically modified food. © IBO 2007

Genetics has come a long way since Gregor Mendel, an Austrian monk, first experimented with pea plant hybrids back in the 1800s. Scientists are now able to insert foreign **genes** into an organism and switch other genes off. All **DNA** (deoxyribonucleic acid) manipulation techniques fall under the umbrella of **genetic engineering**, an area of genetics that has brought **genetically modified (GM) foods** into the public arena amid great controversy.

What are GM foods?

Genetically modified foods are foods created for human or animal consumption that have been produced from genetically altered plants or animals. Often only one gene has been altered out of the tens of thousands of different types of genes carried by the organism. Generally, a new gene is inserted to give the organism a particular characteristic. **Recombinant DNA technology** is used to combine genes from different organisms. If an existing gene has been switched off, the technique is known as **gene silencing**.

Crops are the most common type of genetically modified food. The inserted gene may produce large or small changes; it can make a crop pest-resistant, faster growing, hardier or more nutritious. The organism containing the foreign gene is said to be **transgenic**. The first transgenic plants were created in 1983—antibiotic-resistant tobacco was created by inserting bacterial genes into the plants. An organism that has had its DNA manipulated is also called a **genetically modified organism (GMO)**.

The first GM food offered to the public was the Calgene company's Flavr Savr tomato in 1994, engineered to prevent rotting. Today, millions of farmers worldwide plant transgenic crops. Most of these are designed to be **herbicide** and/or pest resistant. Herbicides are used to kill weeds. The GM food market is growing, especially in developed countries. However, certain regions—notably Europe—are resisting GM foods due to the public backlash. A 1999 study (since widely criticized by the scientific community) found certain strains of GM potatoes to be toxic to rats. This led to an anti-GM food campaign in Europe and the negative feeling still lingers today.

Figure 6.5.2 Anti-GM sentiments are common, especially in Europe.

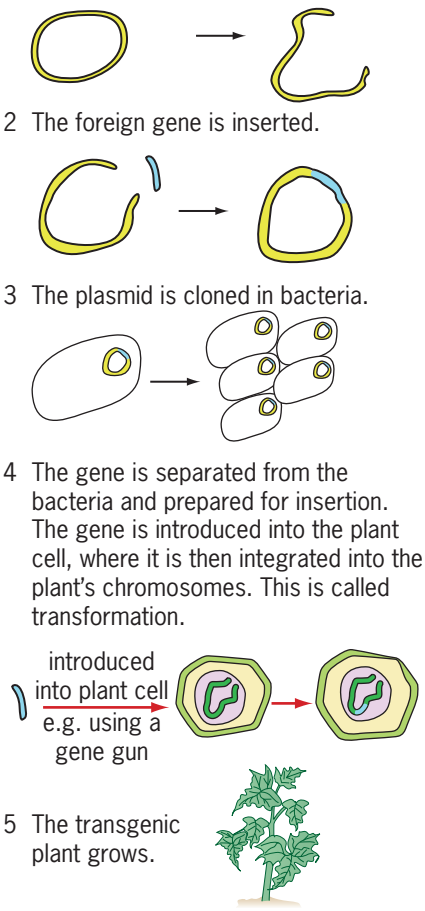


Figure 6.5.1 How a new gene is inserted into a plant. Note that several different methods can be used.

THEORY OF KNOWLEDGE

Since there is no concept of absolute certainty in science, all predictions made about the cause and effects of genetic modification are tentative until they can be falsified at some later stage. So, in the absence of certain knowledge how do you know if GM foods produced from genetically modified organisms (GMO) are safe?

The Cartagena Protocol on Biosafety is an international agreement that aims to provide a good level of protection against the risks posed by GMOs. Part of the protocol is the adoption of the 'Precautionary Principle', which states that when there is scientific uncertainty about the cause and effect of a GMO then it should not be used in food until there is convincing evidence that the risks have been identified and that the associated uncertainties are very small. The purpose of this is to protect humans and the environment *before* a food product is introduced rather than finding out the harmful effects later, when the damage has already been done.

The implication of the Precautionary Principle for scientists is that it requires them to prove the risk of adverse effects before implementation. This is different from the traditional approach where the responsibility for proving a food is safe lies not only with the creator but also with other scientists and experts who try to prove it is not.

Critics claim that since scientists will never be able to prove that GMOs are absolutely safe, this principle will stop innovative advances in food technology. If 'sound science' is used then the risks should only be accepted if and when they occur. However, supporters say the intention is not to insist on absolute certainty but rather to make scientists more morally responsible for the effects of their discoveries by placing the burden of doubt on them rather than on society, other researchers and the potential victims if things go wrong.

- Can you think of examples of any products once claimed to be safe but later found to be harmful?
- What support is there for a counter claim that the Precautionary Principle actually encourages advances in food technology? Consider the groups, individuals, research or types of products and technologies that might benefit from it.
- What do you consider to be the characteristics of 'sound science'?

The promises of GM foods

AS F.5.2
Discuss the benefits and concerns of using GM foods.

Proponents of GM foods claim that they are the answer to the world's food shortage problems, offering increased quality and quantity of crops. With Earth's population expected to top 9 billion people by 2050, maintaining an adequate and nutritious food supply will be challenging, particularly in poor and drought-stricken areas. The benefits of GM crops are outlined in table 6.5.1.

TABLE 6.5.1 BENEFITS OFFERED BY GM CROPS

Advantageous characteristic	Why is it advantageous?	Examples
Pest resistance	Pestilence can cause the loss of vast quantities of crops, resulting in financial ruin for farmers and starvation in the least economically developed countries. In addition, the use of chemical pesticides can be a health hazard, and run-off can contaminate water supplies.	*Bacillus thuringiensis* (Bt) is a soil bacterium that produces proteins that are toxic to many pests. Inserting the gene that codes for this protein into the genome of plants confers pest protection. Currently Bt cotton, corn and potato crops are grown.
Herbicide resistance	Farmers need to get rid of weeds so that their crops can grow effectively and this necessitates the use of herbicides. Care must be taken so that the crops aren't damaged.	By creating crops resistant to one strong herbicide, the number of herbicide applications can be reduced. The Monsanto company has created soybeans resistant to their herbicide Roundup. Generally, only one Roundup application is needed to kill the surrounding weeds.
Enhanced taste and quality	Slow ripening allows fruit and vegetables to stay on the tree or vine longer, which enhances flavour. Longevity of crops also maximizes profit and increases storage time (shelf life).	The previously mentioned Flavr Savr tomato had delayed ripening to allow it to stay on the vine longer and become tastier. Many other slow-ripening fruit varieties have now been created. A new flavour-enhanced GM tomato produces geraniol, a fragrant substance found in fruit and flowers.
Faster maturing	If plants that normally take years to produce fruit could be made to start producing fruit earlier, it would be of great benefit to people needing food quickly, and could also result in significant profit for farmers.	Fruit and nut trees that yield much earlier than usual are in the development stage, and could soon become commercially available.
Increased nutrients	Malnutrition is of serious concern in the least developed countries, leading to poor health and low life expectancies. Kwashiorkor (protein deficiency) is common, as is blindness from vitamin A deficiency.	'Golden rice' is a yellow transgenic rice that contains high amounts of beta-carotene (vitamin A) due to the insertion of daffodil and bacterium genes. The project has been widely criticized because of the time and money expended.
Resistance to extremes of climate	Frosts, droughts and high levels of salinity can devastate crops. Creating plants that can resist climate extremes increases the total area of land that can be farmed.	Drought-resistant tobacco plants needing only 30% as much water as usual have been created, and research continues on other types of plants.
Disease resistance	Many microbes can affect the health of plants. These microbes are hard to fight once they have taken hold of a crop.	Inserting genes to offer resistance to specific microbes is highly advantageous. A transgenic papaya resistant to the papaya ringspot virus is common in Hawaii.
Increased yield	Low yields translate into higher prices for consumers. If good crop yields can be maintained, fruit and vegetables will remain affordable.	Conflicting reports exist as to whether or not GM crops offer increased yields. Some farmers claim the yields for GM and non-GM are similar. An extensive 2003 study on Bt cotton grown in India showed yields were increased by 80% compared to non-Bt cotton.
Allergy reduction or elimination	The incidence of life-threatening allergies is on the rise. If allergens could be eliminated from foods, less vigilance would be needed. Lives could be saved and public health costs could be reduced.	Gene technology can be used to 'switch off' genes that produce allergy-causing proteins in foods such as nuts and wheat. In 2002, researchers created the first hypoallergenic soybeans by silencing one gene.

Figure 6.5.3 Golden rice was engineered to combat vitamin A deficiency in least developed countries.

While it is mainly crops that are being genetically engineered for commercial purposes, extensive animal research is underway. The aims are to create animals with:

- faster growth
- greater muscle mass
- disease resistance
- good feed efficiency
- specific protein composition of milk
- better overall health
- increased yield of wool, milk or eggs
- less waste.

Feed efficiency refers to the conversion of feed weight into meat. Feed efficiency depends on an animal's metabolism, appetite, size and growth rate. It is used to measure productivity, which is usually defined as quantity of milk or meat produced per feed unit.

Mad cow disease (spongiform encephalopathy) resistant cows have been produced by researchers (through gene silencing), as have low-fat bacon-producing pigs which carry a spinach gene. It is expected that the first genetically modified animal cleared for human consumption could hit supermarket shelves within a year—a salmon with growth rates up to ten times that of unmodified salmon.

The environmental benefits of genetically engineered foods are considerable. Pest- and herbicide-resistant crops reduce the dependence of farmers on chemical means of insect and weed control. More efficient growth patterns and hardier plants conserve resources such as water, soil and energy. Specially designed GM feed has been shown to decrease nitrogen and phosphorus excretion from pigs and poultry. This natural waste management reduces contamination of waterways through run-off.

Reasons for concern

The potential benefits of GM foods are indisputable, but the voices of the detractors are too loud to be ignored. Many religious groups view genetic engineering as humans 'playing God'. Some evolutionists believe the results may be unpredictable since combining genes from unrelated organisms, such as pigs and spinach, scrambles genetic lines created by millions of years of natural selection. Clearly, genetic modification in any form should not be approached lightly. Just because we *can* do it, it doesn't mean that we *should* do it.

One argument for GM foods is that they could end world hunger. People who disagree with this view say there is more than enough food in the world; the problem is that it is unevenly distributed. It seems everyone has an opinion on GM foods, even the Vatican, which recently endorsed them to end starvation in the least developed countries. Interestingly, some African nations, including Zambia and Zimbabwe, are opposed to GM crops and have refused international food aid containing GM grain. Some of the specific concerns related to GM foods are listed in table 6.5.2.

TABLE 6.5.2 CONCERNS ABOUT GM FOODS

Concern	Description	Reason for this concern
Pesticide-resistant insects	If the natural pesticide produced by the plants kills susceptible insects, the pests remaining will have natural resistance and will soon be present in large numbers.	Bt-resistant bollworms have been found in a number of Biotech(Bt)-cotton crops in the past five years.
Transfer of inserted gene to another organism (out-crossing)	If cross-pollination occurs, for example, between a GM crop and a weed, the inserted gene may become part of a non-target organism. This could also occur if pollen from a GM plant is blown to a farm where there are non-GM crops.	Monsanto has brought lawsuits against farmers with GM crops they claim are the result of cross-pollination. Researchers are working to create plants with pollen that doesn't contain the inserted gene. Currently there are claims that GM out-crossing has produced herbicide-resistant superweeds.
Harm to other organisms	The inserted gene may have side-effects that are perilous to organisms that happen to come into contact with the GMO.	A 1999 study found that GM corn kills monarch butterfly caterpillars. There is concern that beneficial insects like bees may be killed along with the harmful ones.
Increased incidence of allergies	It is possible that combining genes from two organisms could create a new, potent allergen that could affect people involved in the processing of the plant. Also, someone may eat a GM food thinking it is safe, not knowing that it has had a gene inserted which codes for a protein they are allergic to.	Serious allergies are on the increase in developed nations, and fear of anaphylaxis is common. There have been no reported allergies to GM foods thus far. Caution is needed though, because some of the proteins appearing are completely new to the human diet.
Unknown human health impacts	The risk of changing the composition of a balanced diet by altering the natural nutritional quality of foods has effects that are unpredictable.	As yet, there doesn't seem to be a problem. This is very much a case of fear of the unknown, a fear which may yet prove to be justified.
Irresponsible application of the technology	Some of the research being conducted seems to have no real application, and there are still questions about the wisdom of combining genes from totally unrelated organisms.	The creation of fish and rabbits that glow in UV light sparked heated public debate about the responsible use of gene manipulation technologies.
Animal ethics	The animals don't get to choose what is being done to them, and some may suffer in the process.	Most transgenic animals are mice, bred for research into disease. Animal rights groups say they are 'bred to suffer'. Other transgenics have had problems such as short lives. It is complicated in that the changes generally benefit humans, not the animal.
Antibiotic resistance	Early on, GM researchers used antibiotic resistance to select cells which had taken up a new gene.	The fear is that eating antibiotic-resistant GM plants will create similar resistance in humans. Researchers are working on ways to remove the antibiotic resistant gene from GM plants.
Access and choice	It is the wealthier countries that control the research and the majority of GM crops.	Less-developed nations may become even more dependent on developed nations. Nations totally dependent on international food aid are angry that they don't really have a choice about GM foods; they either eat them or starve.

Figure 6.5.4 A label from an Australian product containing GM soy.

Not all countries have policies on GM foods. Labelling laws vary widely. In Europe, where there is much negativity towards GM foods, the market is rigorously regulated and all GM foods are clearly labelled. It is rare to find whole GM products. Mostly, GM ingredients appear in processed food. In Australia, if the GM ingredient has identical make-up to the natural product, no labelling is required. If new DNA or new protein is present, the GM ingredient is indicated.

THEORY OF KNOWLEDGE

The general public and experts perceive the risk associated with genetically modified (GM) foods differently. Scientist using expert opinion and scientific reasoning are concerned with the effects of genetically modified organisms (GMOs), and the amount of exposure needed to cause harm. The public, however, is more concerned with intuition, value judgments, the opinions of others and questions of morality and safety. To make things more complicated these perceptions vary according to age, gender, income level, education, culture, country and the type of biotechnology product.

Scientists also judge the acceptability of risk differently. For example, although the risks associated with driving a car are greater than eating GM food, generally speaking the public perceives food products made from GMOs as riskier. A scientist using reasoning to evaluate the probability of harm would claim that if you have a 1 in 70 chance of dying in a car accident and a 1 in 10 million chance of developing an extreme allergic reaction to a GMO, then driving a car carries a greater risk.

Considering the difficulty in managing public perception risk analysis is used to help make important societal decisions about the safety of emerging biotechnologies. The value of risk analysis is that it uses the weight of the evidence and reasoning provided by scientists as well as public perception to inform policy making and implementation. Risk analysis is powerful because, if done well, it can foster consensus and decision making on issues of great importance to society.

- Suggest some possible reasons for these differences in public perception.
  - People in the US are more accepting of GM foods than people in the European Union and Japan.
  - People in the European Union and Japan think that the risks associated with GM foods is more than outweighed by their cheaper cost.
  - Consumers in China are prepared to pay more for GM soybeans and rice than non-GM alternatives.
  - Medicines produced from biotechnology products are more likely to be supported than genetically modified foods.
  - People are generally more concerned with the risks associated with specific products produced from biotechnology than with the process of producing these products.
- What are some of the issues with relying too heavily on surveys to determine the public's perception of GM foods?
- Suggest some possible questions a scientist and a member of the public might ask about the safety of GM foods, and comment on the nature of the differences.
- In 2002 the Zambian government said it would not accept genetically modified food aid from the World Health Organization (WHO) for the 2.5 million Zambians facing severe food shortages and famine because in the face of scientific uncertainty the GM corn being offered might adversely affect human heath. In response Jorgen Schlundt, of the WHO food safety program, said 'We accept the right of any government to reject GM food, but we believe the GM foods on the market do not represent a health risk. They have gone through risk assessment and have actually been eaten for many years.' Whose view, that of the Zambian government or the World Health Organization, is morally right in this case and how do you know? What arguments and counter arguments did you use in coming to your decision?

Section 6.5 Exercises

1 Define the following terms:
   - **a** genetically modified food
   - **b** transgenic
   - **c** recombinant DNA technology
   - **d** gene silencing

2 Describe three examples of GM crops, and explain how the inserted gene has been advantageous.

3 Explain how GM foods could potentially benefit:
   - **a** farmers
   - **b** the environment
   - **c** least developed nations.

4 Describe the difference between gene insertion and gene silencing, and predict whether opponents to gene insertion would be just as opposed to gene silencing.

5 Explain how GM foods could negatively impact:
   - **a** other animals
   - **b** humans
   - **c** the environment.

6 Explain how economics, religion, society and history each play a role in the genetically modified foods debate.

6.6 TEXTURE

Food texture is related primarily to how a food feels in the mouth. Words often used to describe texture include soft, smooth, thick, crunchy, crumbly, creamy, juicy, watery and chewy. Texture is an important factor in the enjoyment of food; perhaps more people would like oysters if they were less slimy. The texture of a food is determined by its ingredients and structure. Foods can change texture after periods of storage, or if the temperature changes. How water is dispersed in the food is significant, as is the dispersal of other components. For example, as cakes lose water, they become hard and stale; as crackers absorb water, they lose their crispness.

Food texture can be measured by a sensory panel of people, or by instrumental methods. The instruments measure properties such as ripeness, compressibility, elasticity, tensile strength and 'crunch'.

Figure 6.6.1 People prefer 'low energy' textures for dessert, such as creamy and spongy.

Dispersed systems

AS F.6.1 Describe a dispersed system in food. © IBO 2007

AS F.6.2 Distinguish between the following types of dispersed systems: suspensions, emulsions and foams in food. © IBO 2007

Although most foods look homogeneous, many are actually **dispersed systems**—kinetically stable mixtures where one phase is dispersed throughout another phase, with the two phases being largely immiscible. 'Kinetically stable' means that the dispersed system can last for a long time, despite the fact that most are thermodynamically unstable. The phase that is dispersed is (unsurprisingly) called the **dispersed phase**, or discontinuous phase. The other phase is called the **continuous phase**.

Various types of dispersed systems exist in foods. Margarine is an **emulsion**. The froth that forms the head on the top of a glass of beer is a **foam**. A mixture of cornstarch and water is a **suspension**. The different types of dispersed systems can sometimes be hard to tell apart—the differences are often quite subtle. Many are colloidal. In a **colloid**, the dispersed particles are so small that they cannot be seen with a light microscope; they are smaller than 1 μm.

Dispersed systems can be classified according to the nature of the dispersed and continuous phases, as outlined in table 6.6.1. In a suspension, solid particles are dispersed throughout a liquid. Traditionally, a mixture was only referred to as a suspension if the particles were non-colloidal, but this distinction has disappeared in recent times. In an emulsion, droplets of one liquid are dispersed throughout another liquid which is generally of quite different polarity, for example oil spread throughout water. In a foam, gas bubbles are spread through a liquid, with the total volume of the gas being greater than that of the liquid.

TABLE 6.6.1 DIFFERENT TYPES OF DISPERSED SYSTEMS FOUND IN FOODS

Name of dispersed system	Dispersed phase	Continuous phase	Example
Suspension *or* sol	Solid	Liquid	Hot chocolate
Emulsion	Liquid	Liquid	Mayonnaise
Foam	Gas	Liquid	Whipped cream

The stability of a dispersed system depends on factors such as the size of the dispersed particles, the viscosity of the continuous phase and temperature. Emulsions and foams may be stabilized by the use of additives.

Figure 6.6.2 Three types of dispersed systems.

Emulsions

Food emulsions are either water droplets dispersed in oil (water-in-oil; w/o) or oil droplets dispersed in water (oil-in-water; o/w). Ready-to-eat emulsions include milk (o/w), mayonnaise (o/w) and margarine (w/o). Emulsions look cloudy because the dispersed droplets scatter light.

AS **F.6.3 Describe the action of emulsifiers.© IBO 2007**

If a surfactant is added to an emulsion, the surface energy between oil and water can be lowered and the emulsion can be stabilized. Surfactants can be classed as anionic (negatively charged), cationic (positively charged), non-ionic (uncharged) or zwitterionic (both positively and negatively charged), depending on the nature of the hydrophilic head. Surfactants that stabilize emulsions and foams are called **emulsifiers**. Like all surfactants, these molecules have hydrophilic heads and hydrophobic tails. They are soluble in both lipids and water because they are capable of bonding well with both. Emulsifiers collect at the surface of dispersed droplets and act as an interface, helping the droplets stay suspended and preventing their coalescence.

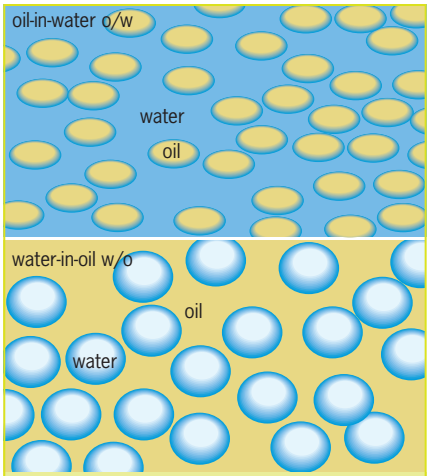


Figure 6.6.3 The two emulsion types.

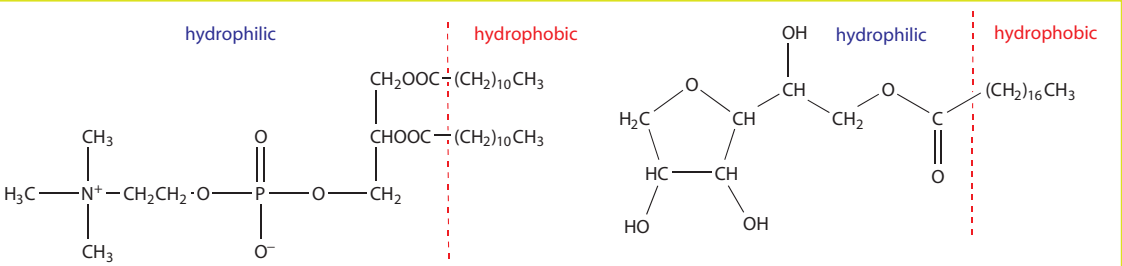


Figure 6.6.4 Lecithin (left) is a natural emulsifier found in eggs. Sorbitan monostearate (right) is a synthetic emulsifier used in foods.

Pr **PRAC 6.6 Investigating emulsions**

A shaken mixture of oil and water will naturally settle out into two layers if left, because the two liquids are more stable if separated. Thermodynamically unstable emulsions such as mayonnaise do the same thing if left for a long time; the suspended droplets coalesce (join together) to reduce their surface area and hence lower their total surface energy. **Stabilizers** are food additives designed to keep two normally immiscible ingredients in a homogeneous state after mixing. They help to contribute to a smooth texture. Gelatin is a common example. They are often used in conjunction with surfactants to stabilize emulsions.

If you try to make an emulsion at home by shaking a mixture, large droplets are produced and the emulsion is unstable. Using commercial processors, it is possible to produce colloidal dispersed droplets which, with the help of additives, can remain dispersed for a long time and are hence kinetically stable. When an emulsifier is present, it adsorbs at the oil–water interfaces. The hydrophilic head will bond with the water while the hydrophobic tail will be oriented to bond with the oil. Figure 6.6.5 shows the orientation of an emulsifier in a w/o emulsion.

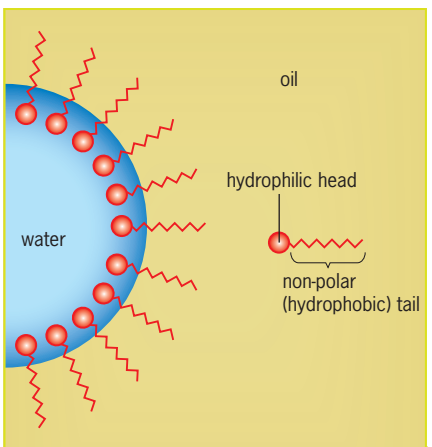


Figure 6.6.5 The emulsifiers orient themselves to maximize bonding strength.

In o/w emulsions, if the emulsifier has an ionically charged hydrophilic head, the dispersed droplets will repel each other when they come into contact.

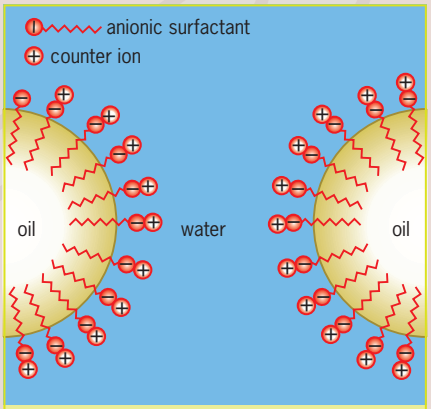


Figure 6.6.6 The electrostatic double layer formed by the ionic surfactant and its counter ion stabilizes the emulsion. This diagram shows an anionic surfactant.

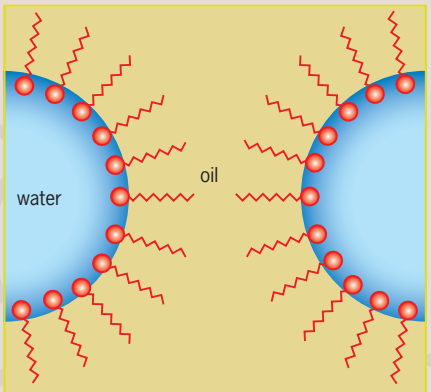


Figure 6.6.7 A water-in-oil emulsion is stabilized by the barrier formed at the surface of the dispersed phase.

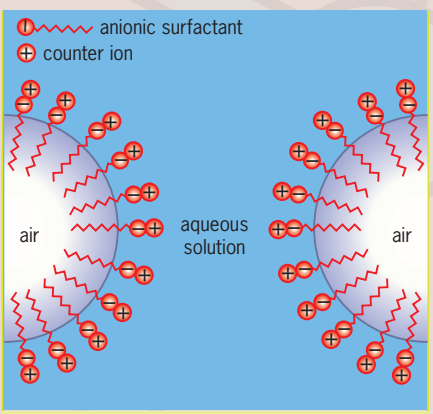


Figure 6.6.8 Stabilization of a foam by emulsifiers.

Counter ions come into effect by forming an electrostatic double layer that prevents the droplets coming together (figure 6.6.6).

In o/w emulsions, if the hydrophilic head is not charged, a hydrate shell forms around the polar groups to prevent coalescence. The tails of the emulsifiers protruding from the dispersed water droplets form a bilayer barrier that is difficult for the water molecules to break.

Making emulsions

Making emulsions requires mechanical energy because an energy input is needed to reduce the size of the droplets in the dispersed phase. Not all emulsions need to be well stabilized. Margarine actually contains relatively large water droplets, but this is not important because the droplets become trapped when the surrounding oil is chilled and becomes semi-solid. Emulsions are generally made in two steps—mixing and stabilization. One method of stabilizing emulsions, besides using emulsifiers, is to add substances that increase the viscosity of the continuous phase. Unstable emulsions may invert—an o/w emulsion can become a w/o emulsion and vice versa. This is what happens when milk is creamed until it turns into butter.

Emulsions are not equilibrium systems; droplets are not constantly coalescing and reforming. To make new droplets, energy must always be put into the system. While most emulsions are not thermodynamically stable, **microemulsions**, in which the droplet size is less than 100 nm, are transparent and thermodynamically stable. Microemulsions require the presence of an emulsifier and co-surfactant. Table 6.6.2 shows three examples of emulsions and the stabilizing substances present.

TABLE 6.6.2 EMULSION EXAMPLES

Emulsion	Type	Dispersed phase	Continuous phase	Stabilising compounds
Milk	o/w	Butterfat triglycerides, other lipids (4%) 1–10 μm size	Aqueous solution containing minerals, carbohydrates and proteins	Phospholipids, casein, lipoproteins
Salad dressing	o/w	Vegetable oil (30%) 1–5 μm size	Aqueous solution containing, for example, sugar, salt and flavours	Synthetic surfactants, hydrocolloids
Butter	w/o	Buttermilk	Various lipids	Phospholipids in buttermilk

Foam stabilization

All foams are thermodynamically unstable. Additives are needed to stabilize them. Surfactants can stabilize gas bubbles in a similar manner to the stabilization of the dispersed phase in an oil-in-water emulsion. If the bubble contains air, it is non-polar, so the hydrophobic tail of the surfactant will be oriented inward, as shown in figure 6.6.8.

Section 6.6 Exercises

1 Food-texture analysis is the process of measuring the properties related to how a food feels in the mouth. Discuss the use of human subjects and instrumental analysis as a means to comment on food texture.

2 Many food ingredients are completely immiscible and so will form separate phases within a food. However the sizes of these phases can be very small, so to the naked eye the food will appear homogeneous.

a Describe a dispersed system in food.

b Copy and complete the following table with respect to dispersed systems.

		Continuous phase		
		Solid	Liquid	Gas
Dispersed phase	Solid	Frozen food	Molten chocolate	
	Liquid			
	Gas			

3 Define and give two examples of each of the following types of dispersed systems in foods.

a Suspensions

b Emulsions

c Foams

4 Oil and water are normally immiscible. List what is needed in order for an emulsion to be made between oil and water.

5 Describe the three different roles that emulsifiers play in the food industry.

6 Eggs can be used as emulsifiers, coagulants and foams. Indicate what roles eggs could be playing in each of the following situations.

a A permanent emulsion is formed

b Too much beating causes low volume and the egg breaks down into curds

c The yolk is the part of the egg specifically used for this function

d Irregular-shaped air cells form

e A thin film of egg white protein surrounds each cell

f Air is incorporated into the egg white

g Aids in the mixing of oil and water

7 Stabilizers are used in such food products as peanut butter and ice cream. Describe the role that stabilizers play in the development of a more appealing texture in food.

CHEM COMPLEMENT

The perfect soufflé

Many amateur chefs are frustrated in their attempts to make that most delicate of dishes, the soufflé. A soufflé is an egg-white foam with something added, such as a béchamel sauce for a savoury soufflé. Eggs are used for three purposes in cooking: emulsifying, foaming and coagulation on heating. Egg-white proteins include conalbumin, ovomucin and phosvitin. During whipping, proteins denature and form aggregates. In particular, ovomucin forms a surface-active film between the air bubble and the liquid, which stabilizes the bubble. Globulins contribute by making the continuous phase very viscous.

The more the whites are whipped, the smaller the bubbles and the more stable the foam. Egg whites are soufflé-ready when they can support the weight of an egg in its shell. Tiny bits of yolk in the whites can destroy their whippability. The fats in the yolk bond to egg-white proteins and stop them forming a stable bubble–liquid interface. Yolks are added only after the whites are well whipped. Overwhipping can also be a problem. Egg whites that have been whipped too much will 'weep' as the water separates from the proteins.

Figure 6.6.9 Creating the perfect soufflé can be a challenging task!

6.7 OXIDATIVE RANCIDITY

HL

Rancidity was first discussed in section 6.3, page 341. Oxidation of fats results in the replacement of an oxygen atom for a hydrogen atom in a fatty acid molecule. This substitution destabilizes the molecule and its unsaturated components are broken down into more volatile molecules that are characterized by an unpleasant change in the flavour and odour of a food. As previously mentioned, unsaturated fats are more susceptible to oxidation than saturated fats. In fact, any degree of unsaturation in fats will mean that oxidative rancidity is unavoidable.

AS F.7.1 Describe the steps in the free-radical chain mechanism occurring during oxidative rancidity. © IBO 2007

Step 1: Initiation Homolytic fission of the R–H bond, where RH is a fatty acid that can lose a hydrogen to form free radicals

The C–H bonds are strong with an average bond enthalpy of 412 kJ mol^{-1}, so the initial formation of the **free radical** requires considerable energy to overcome this high activation energy.

$$RH \rightarrow R\cdot + H\cdot$$

Step 2: Propagation Addition of oxygen through propagation

Propagation is a free-radical chain reaction in which the free radicals are the propagators for the addition of oxygen to the fatty acid, RH, to form lipid hydroperoxides, ROOH.

$$R\cdot + O_2 \rightarrow ROO\cdot$$
$$ROO\cdot + RH \rightarrow ROOH + R\cdot$$

Alternatively, it is also possible that oxygen can react directly with the fatty acid to produce the hydroperoxide without going through the free radical mechanism.

$$RH + O_2 \rightarrow ROOH$$

Step 3: Peroxide decomposition

Decomposition of hydroperoxides occurs quickly due to the relative weak nature of the O–O bond in the hydroperoxide. This results in the formation of a series of aldehydes, alcohols, alkanes, ketones and esters.

To illustrate these reactions, the hydroperoxide is shown as ROOH or HR_2COOH, where 'R' and 'HR_2C' represent straight and branched hydrocarbon chains.

$$HR_2COOH \rightarrow HR_2CO\cdot + OH\cdot$$

Formation of aldehyde: $HR_2CO \rightarrow RCHO + R\cdot$

Formation of alcohol: $OH\cdot + R'H \rightarrow HR'_2COH + R'\cdot$

Step 4: Termination

The formation of some of these products results in the termination of the oxidative rancidity process.

Formation of alkane: $R\cdot + R\cdot \rightarrow RR$

Formation of ketone: $HR_2CO\cdot + R'\cdot \rightarrow R_2CO + R'H$

Formation of ester: $R\cdot + ROO\cdot \rightarrow ROOR$

or $ROO\cdot + ROO\cdot \rightarrow ROOR + O_2$

Section 6.7 Exercises

1 Explain the importance of free radicals in oxidative rancidity.

2 Describe the mechanism for the addition of oxygen through propagation.

3 Explain how the termination of oxidative rancidity is dependent on free radicals.

4 State the equations that occur during the formation of aldehydes, ketones, alcohols, alkanes and esters during oxidative rancidity.

6.8 ANTIOXIDANTS

HL

AS **F.6.8 Explain the differences between the three main types of antioxidants. © IBO 2007**

In section 6.3 we learned that an antioxidant is a substance that delays the onset or slows the rate of oxidation of a food. They are used to extend the shelf life of food. Antioxidants can be subdivided into three main groups. The first type of antioxidants, the radical scavengers (quenchers, AH), work by inhibiting the formation of free radicals in the first step of the auto-oxidation process or they interrupt the propagation of the free-radical chain. Examples of radical scavengers include BHA (butylated hydroxyanisole), BHT (butylated hydroxytoluene), TBHQ (tertiary butylhydroquinone) and tocopherols (vitamin E). What do these structures have in common?

Free radical scavengers (quenchers) form stable and less-reactive free radicals. If the number of free radicals can be kept low enough, oxidation will not occur.

$$ROO\cdot + AH \rightarrow ROOH + A\cdot$$

The second type of antioxidants are chelating agents that work to decrease the number of free metal ions present so that there is less chance of the oxidation reaction being catalysed. A **chelating agent** is a chemical that forms a complex with divalent (2+) metal ions so that they cannot deteriorate the food. Examples of such chelating agents include salts of EDTA (ethylenediaminetetracetic acid), and natural chelating agents including rosemary, tea and ground mustard. EDTA has a coordination number of six and forms octahedral chelates (complexes). It is commonly found in such food products as beer and mayonnaise.

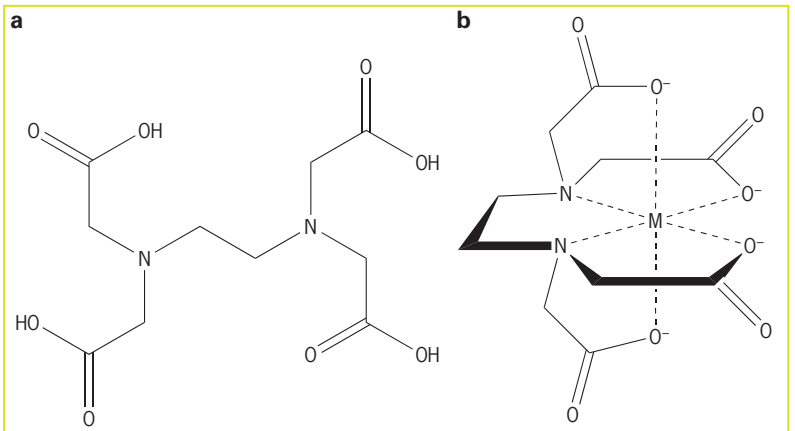


Figure 6.8.1 Structures of (a) EDTA and (b) an EDTA–metal chelate.

The third group of antioxidants is the reducing agents (electron donors) and other such agents that remove or reduce concentrations of oxygen. Specific examples of such reducing agents include vitamin C and various carotenoids (refer to section 6.4). As an antioxidant, vitamin C's primary role is to neutralize free radicals. Since ascorbic acid is water-soluble, it can work both inside and outside the cells to combat free radical damage. Vitamin C is an excellent source of electrons; therefore, it can donate electrons to free radicals so that they are no longer available to be used in the oxidation process. When vitamin C is oxidized, it forms dehydroascorbic acid, as seen in the reaction in figure 6.8.2.

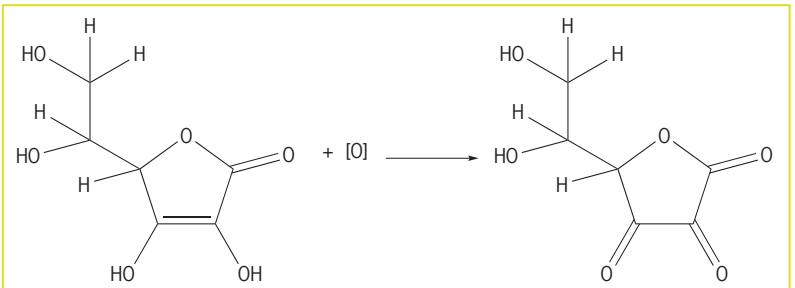


Figure 6.8.2 Oxidation of vitamin C to dehydroascorbic acid occurs readily, thus making it a good antioxidant.

Section 6.8 Exercises

1. Compare the three different types of antioxidants in terms of their reactions, their sources and their structures.
2. Discuss the chemistry and bonding that occurs when a chelate of EDTA–metal forms.
3. List the functional groups that are common in the three different types of antioxidants.
4. Complete the reaction that occurs when vitamin C is oxidized to dehydroascorbic acid.

6.9 STEREOCHEMISTRY IN FOOD

HL

AS F.9.1 Explain the three different conventions used for naming the different enantiomeric forms. © IBO 2007

Recall from our earlier work (see *Chemistry: For use with the IB Diploma Programme Higher Level*, pages 324–331) that molecules that are not identical to their mirror image are called **enantiomers** (from the Greek *enantio* meaning 'opposite'). Enantiomers cannot be superimposed on each other. Enantiomers are said to be **chiral** (from the Greek *cheir* meaning hand). The most commonly encountered cause of chirality is that the molecule contains a carbon atom bonded to four different substituents, CWXYZ. Such a carbon is called an asymmetric centre, and is also referred to as a chiral carbon. Molecules containing asymmetric carbon atoms are chiral molecules. Enantiomers (**optical isomers**) have identical physical and chemical properties, with two exceptions:

- their direction of rotation of plane-polarized light
- their interaction with other optically active compounds.

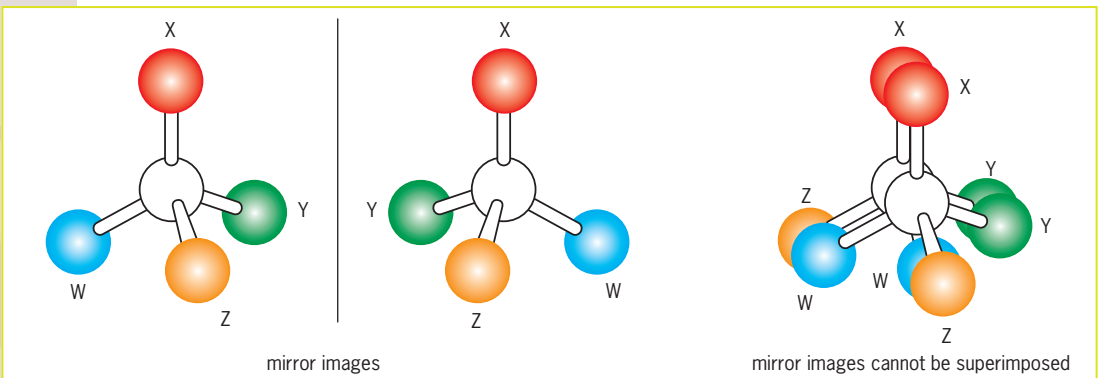


Figure 6.9.1 Enantiomers are non-superimposable mirror images.

Naming enantiomers

Three different conventions are used for naming the different enantiomeric forms. One we have considered is based on the direction of rotation of plane-polarized light. Thus enantiomers that rotate light clockwise, to the right, are said to be **dextrorotatory** (from the Latin *dextro* meaning 'right'). Those that rotate light counter clockwise, to the left, are said to be **levorotatory** (from the Latin *levo* meaning 'left'). By convention, rotation to the right is given a plus sign (+), while rotation to the left is given a minus sign (–). Enantiomers rotate plane-polarized light to the same extent, but in opposite directions. One is therefore the –(l) form, the other the +(d) form. The +(d) and –(l) notation thus labels enantiomers according to the direction in which they rotate the plane of polarized light.

The second convention, commonly used for carbohydrates and amino acids, is the D and L notation, based on the arrangement in space of parts of the molecule. For D and L isomers, glyceraldehyde is a reference molecule, as shown in figure 6.9.2. Note the use of **Fischer projections**. In Fischer projections horizontal groups are understood to project above the plane of the page, while vertical groups project below the plane of the page.

Carbohydrates are assigned a D or L by reference to the (+) glyceraldehyde molecule, which is assigned the letter D. Carbohydrates in nature are formed from D-saccharides. For amino acids the 'CORN' rule is used. With the exception of Gly, the α-carbon atom of all amino acids is asymmetric. The CORN rule arranges the substituents, **CO**OH, **R**, **N**H_2 and H, around the chiral α-carbon, with the hydrogen pointing away. If the substituents are arranged clockwise it is the D isomer. If they are arranged anticlockwise it is the L isomer. All naturally occurring amino acids have the L configuration. The D- and L- notation is based on the glyceraldehyde configurations, and *not* on the actual direction of rotation of plane-polarized light. The d(+)/l(–) property does not correspond to the D/L configurations; that is, the L configuration does not refer to levorotation (although L-glyceraldehyde is levorotatory). In fact, nine of the L-amino acids commonly found in proteins are dextrorotatory.

The third convention, the R and S notation, is a more general approach to specifying the configuration of groups around the chiral carbon. It is used for many stereoisomers, including triglycerides. This notation labels the chiral carbon according to a system by which the substituents are assigned a priority. The priorities are determined according to a system proposed by Cahn, Ingold and Prelog. The highest priority is given to the substituent in which the first atom has the highest atomic number. If there are two or more substituents with the same first atom, the priority is based on the second atom or combined second atoms. Figure 6.9.4 provides examples of the use of these priorities. The molecule is orientated so that the substituent of lowest priority is directed away from the viewer, and the remaining groups are directed toward the viewer in a tripodal fashion. If the direction of decrease in order of priority is clockwise, the configuration is R (Latin, *rectus* meaning right); if anticlockwise, it is S (Latin, *sinister* meaning left). The R configuration is not necessarily dextrorotatory or levorotatory; that is, the direction of rotation of the R configuration varies with the particular example.

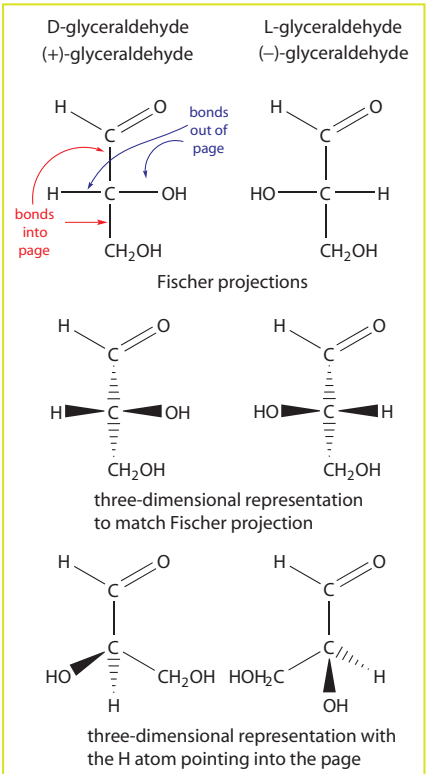


Figure 6.9.2 Glyceraldehyde is used as a reference molecule for naming enantiomers.

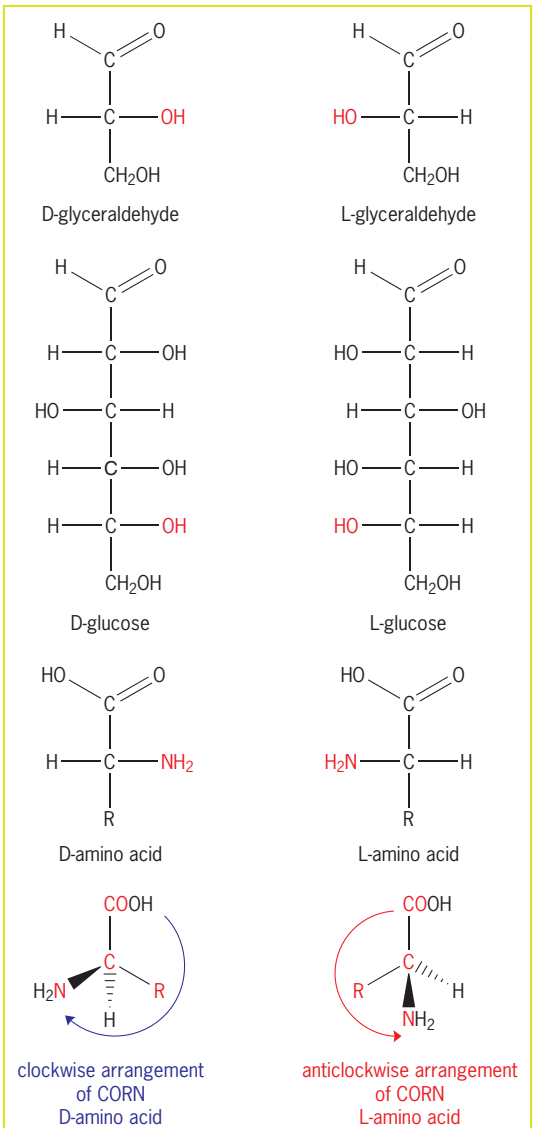


Figure 6.9.3 Monosaccharides and amino acids are named using the D and L notation.

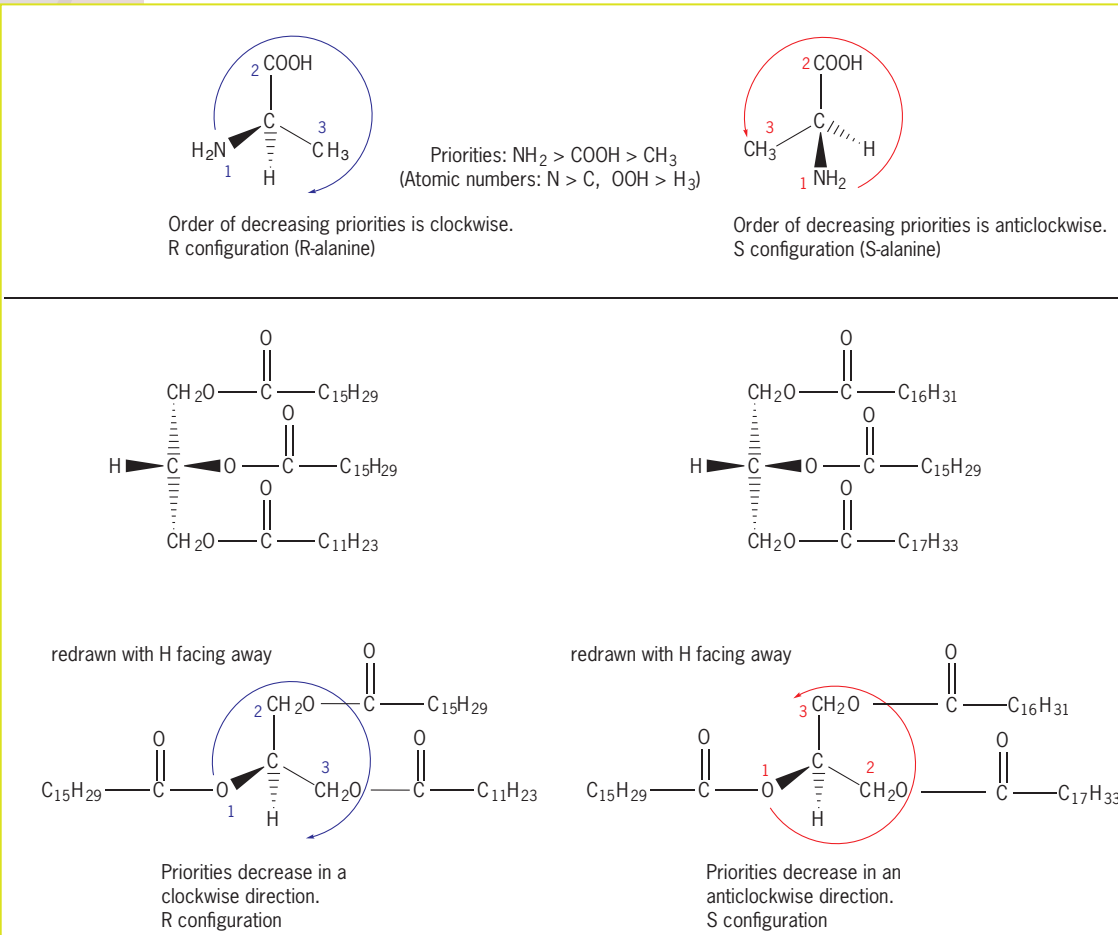


Figure 6.9.4 The R and S notation for enantiomers uses a system of priorities for groups attached to the chiral carbon.

AS F.9.2 Distinguish between the properties of the different enantiomeric forms of stereoisomers found in food. ©IBO 2007

Enantiomers in food

The only source of optically active substances is living organisms, which usually produce mostly one enantiomer of each compound. Thus most amino acids occur in the L-form, while most monosaccharides occur in the D-form. The occurrence of the other enantiomers (D-amino acids and L-saccharides) is very rare. It seems that biological processes produce optically pure compounds, whereas chemical synthesis generates a **racemic mixture** (a mixture of both enantiomers). This is perhaps not surprising, since enzymes, the important biological catalysts, are themselves optically active. We have seen evidence of the differing biological activities of enantiomers in the case of drug molecules (see *Chemistry: For use with the IB Diploma Programme Higher Level*, p. 329, and section 4.8, p. 242 of this book).

The differing biological activities of enantiomers are also important in the tastes of foods:

- D-amino acids taste sweet, while their L configuration is tasteless.
- The L form of the amino acid glutamic acid tastes meaty or brothy, but the D form and the racemate do not have this taste.
- L-aspartame tastes sweet, D-aspartame tastes bitter.

- D-glucose is sweet and nutritious and an important component of our diet, whereas L-glucose is tasteless and cannot be metabolized by the body.
- S-(+)-carvone tastes of caraway seeds and dill, while R-(–)-carvone tastes of spearmint.

These substances have different tastes because taste is the result of interactions between these chiral molecules and chiral receptors in the taste buds of your mouth.

Chiral structure is found to be important in perception of aroma, as well as taste. Enantiomers may differ in aroma, or in the intensity of their aroma. For example:

- R-(+)-limonene smells of oranges and S-(–)-limonene smells of lemons.
- R-(–)-1-octene-3-ol exhibits the aroma of fresh mushrooms, but S-(+)-1-octene-3-ol has a musty, grassy aroma.

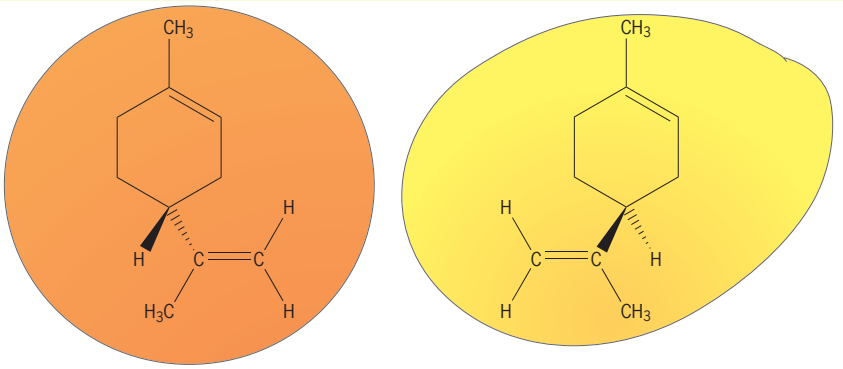


Figure 6.9.5 Enantiomers of the same compound produce the different smells of oranges and lemons.

- S-(+)-carvone has a caraway aroma, and R-(–)-carvone has a mint aroma.

These substances have different aromas because the perception of smell at the molecular level is the result of interactions between the flavour molecules and olfactory receptors. As these receptors are proteins, and are therefore chiral molecules, interactions show enantiomeric selectivity. Isomers of flavour may also differ in their biological activity. For example, S-(+)-carvone inhibits the growth of potato sprouts. There is an attempt to use this phenomenon to increase the shelf-life of potatoes.

Because the different enantiomeric forms vary in their tastes, odours and toxicity, they can be used to determine the authenticity of food and the extent of processing. For example:

- Natural raspberry flavour is due to R-alpha-ionone; synthetic raspberry flavourings contain both the R and S isomers. Other synthetically made foods often contain a racemic mixture of each enantiomer.
- Natural mint contains pure R configuration monoterpenoids. The presence of S enantiomers in commercial oil would indicate contamination by other naturally occurring compounds, or the production of the oil by synthetic means.

We have seen that many natural products are purely of one isomeric form. However, it has been found that in some food ingredients, particularly aromatic substances, there is a predominance of one isomer, but a small amount of the other isomer is present in a characteristic proportion. In recent years the determination of these proportions has been of particular interest in the analysis of aromatic food compounds.

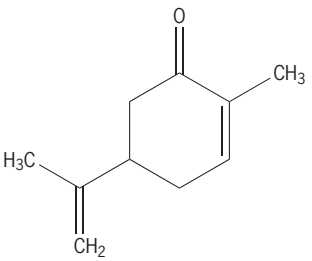
Monitoring of the enantiomers present in foods is being used increasingly. Development of sensitive and specific chromatographic procedures for differentiating enantiomers has enabled this type of monitoring to occur. The use of specific biosensors to selectively analyse for particular enantiomers may in the future make this type of monitoring and authentication a much quicker and more routine process. Already such sensors are used for the assay of D-alanine and L-lactic acid.

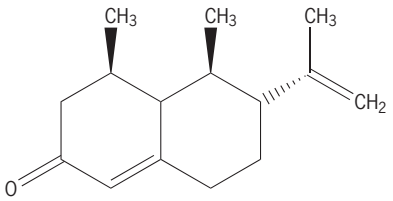
Figure 6.9.6 Is this made from real raspberries or is it a synthetic mixture? Analysis of enantiomers might provide the answer.

Section 6.9 Exercises

1 Define the term *enantiomer*.

2 Identify the chiral centres in each of the molecules shown.





3 Explain the +(d) and –(l) system of naming enantiomers.

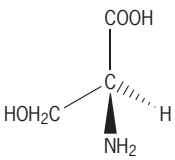
4 Explain the D and L system of naming enantiomers commonly used for carbohydrates and amino acids.

5 **a** Explain how enantiomers of a compound are named using the R and S system.

b State one advantage of the R and S notation compared to the D and L notation.

6 In the R and S system of naming enantiomers does D-glyceraldehyde become R-glyceraldehyde or S-glyceraldehyde? Explain your choice.

7 A representation of one enantiomer of serine is shown. This enantiomer rotates plane-polarized light to the right. Name this molecule using each of the d(+) and l(–), D and L, and R and S systems of naming enantiomers.



8 Discuss three examples to illustrate why stereochemistry is important to the food industry.

9 Explain how the two enantiomeric forms of carvone can be used to determine the authenticity of food products.

6.10 CHEMICAL STRUCTURE AND COLOUR

HL

AS F.10.2 Explain why anthocyanins, carotenoids, chlorophyll and heme form coloured compounds while many other organic compounds are colourless. © IBO 2007

One obvious observation of various foods is that some are coloured while others are not. It is the specific chemical composition of foods that either allow them to be coloured or colourless. Earlier in section 6.4 (p. 351) we saw that when white light passes through or is reflected by a coloured substance, a characteristic portion of the mixed wavelengths is absorbed. The remaining light will then assume the complementary colour to the wavelengths absorbed. There is a direct link between the wavelength of visible light absorbed and the complementary colour that is seen as shown in the colour wheel.

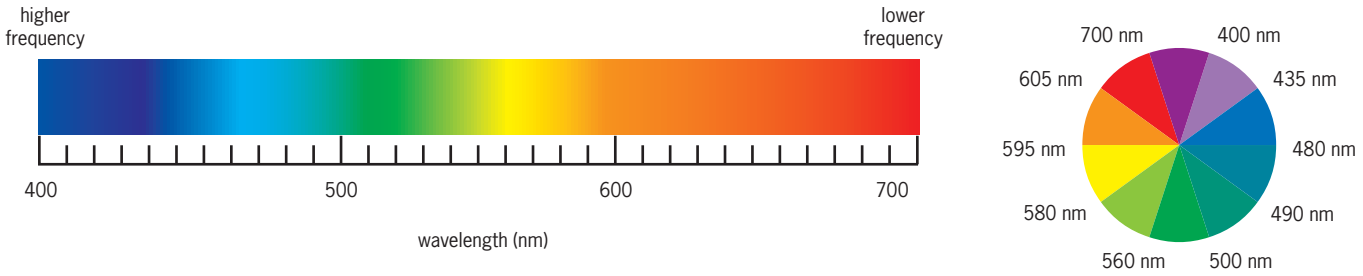


Figure 6.10.1 The link between the visible spectrum of light and the colour wheel. Food absorbs one wavelength of light and we see the complementary colour.

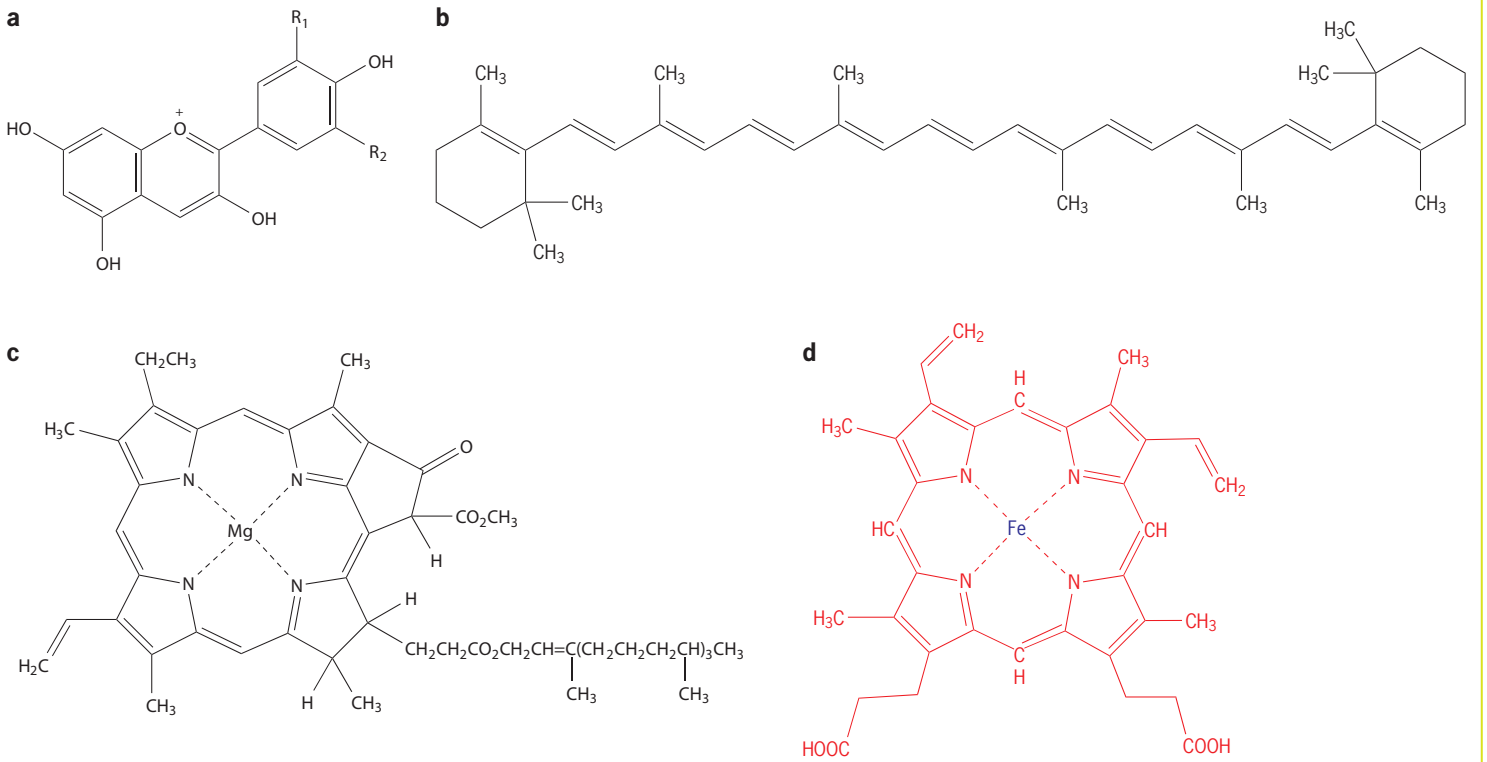


Figure 6.10.2 Notice the extensive conjugation found in these natural pigments: (a) anthocyanin (b) β-carotene (c) chlorophyll *a* and (d) heme.

The specific part of the molecule that absorbs the visible light is called a **chromophore**. A chromophore is a region in a molecule where the energy difference between two different molecular orbitals falls within the range of the visible spectrum. Visible light that hits the chromophore is absorbed when electrons jump from their ground state to their excited state. The energy levels the electrons jump between are the extended pi orbitals created by a series of alternating single and double bonds, which are commonly found in aromatic systems (benzene rings). These alternating single and double bonds lead to a series of **conjugated** pi electrons. This arrangement is seen in the structures of some common food colourings (figure 6.10.2). The more extensive the conjugation, the longer the wavelength of the light that is absorbed.

It is the existence of these chromophores that allow these molecules to be coloured. If an organic molecule does not have conjugated pi electrons then it is colourless.

AS F.10.1
Compare the similarities and differences in the structures of the natural pigments: anthocyanins, carotenoids, chlorophyll and heme. © IBO 2007

AS F.10.3
Deduce whether anthocyanins and carotenoids are water- or fat-soluble from their structures. © IBO 2007

Structure of some natural pigments

As well as having conjugated pi electrons, there are several other properties of natural pigments that give them their unique colouration.

Anthocyanins

Over 500 anthocyanins have been isolated from plant parts, but all are based on a single basic core structure, the flavonoid skeleton, $C_6C_3C_6$, with conjugated double bonds.

Seven different substituents are possible on a flavonoid skeleton. These substituents are usually hydrogen atoms, hydroxyl groups or methoxy groups. The different combinations of these substituents result in different **anthocyanidins** being formed, each able to absorb light of different wavelengths.

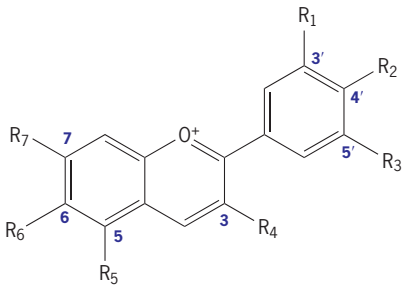


Figure 6.10.3 The flavonoid skeleton on which all of the 500 isolated anthocyanins are based.

TABLE 6.10.1 COMMON ANTHOCYANIDINS AND THEIR MAXIMUM ABSORBANCE

Anthocyanidin	R_1	R_2	λ_{max} (nm)
Pelargonidin	H	H	520
Cyanidin	OH	H	535
Peonidin	OCH_3	H	532
Delphinidin	OH	OH	546
Putinidin	OCH_3	OH	543
Malvidin	OCH_3	OCH_3	542

When anthocyanidins are combined with sugars such as glucose and galactose, colourful anthocyanins are produced. A common anthocyanin is cyanidin-3-glucoside is found in strawberries and results in the red colour that is common in ripe strawberries.

Anthocyanins are water-soluble due the addition of phenolic groups (–OH attached to a benzene ring). The addition of these hydroxyl groups makes these chemicals polar, which allow them to dissolve in polar water.

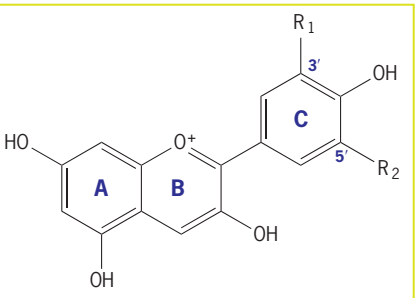


Figure 6.10.4 Basic structure of an anthocyanidin. When the R_1 and R_2 groups are changed, the resulting anthocyanidin absorbs visible light of a different wavelength.

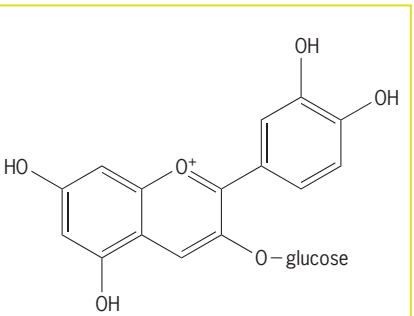


Figure 6.10.5 Cyanidin-3-glucoside, the anthocyanin responsible for the red colour of strawberries.

Carotenoids

Conjugation of pi electrons is also an important factor in the production of colour in carotenoids. Common examples of carotenoids are α- and β-carotene which can be found in carrots, sweet potatoes, kale, spinach, cantaloupe and apricots. Both forms of carotene are precursors for the formation of vitamin A (vitamin A_1, retinol) which is used in the body to maintain healthy vision.

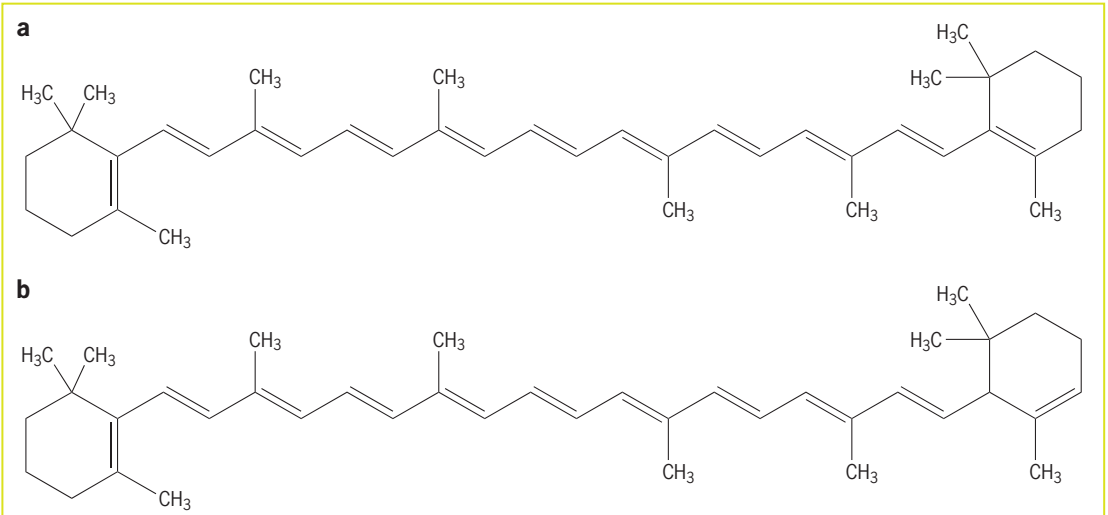


Figure 6.10.6 Two common carotenoids are (a) α-carotene and (b) β-carotene. Notice that they only differ in structure by the position of one double bond. Can you find that difference?

α-carotene and β-carotene can be split into simpler molecules to produce retinal, which has an aldehyde group at the right-hand end of the molecule instead of the hydroxyl group found in vitamin A. Retinal can be reversibly reduced to produce retinol, vitamin A.

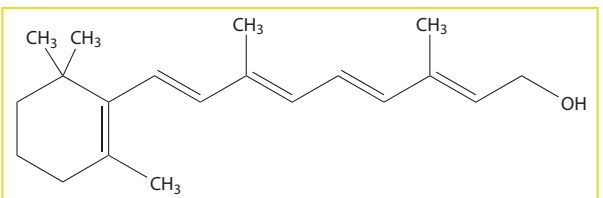


Figure 6.10.7 Vitamin A, a common carotenoid. Can you find the similarities between its structure and the structures of α-carotene and β-carotene?

The majority of carotenoids are derived from a 40-carbon polyene chain, which can have cyclic end groups. When oxygen-containing functional groups are added to the polyene chain, the carotenoids are called **xanthophylls**. Xanthophylls are yellow pigments that have molecular structures based on carotenes. Unlike carotenes, which have long alkene chains, xanthophylls have some hydrogen atoms that are substituted by either hydroxyl groups or oxygen atoms. The yellow colour of chicken egg yolk comes from hens ingesting the xanthophyll lutein.

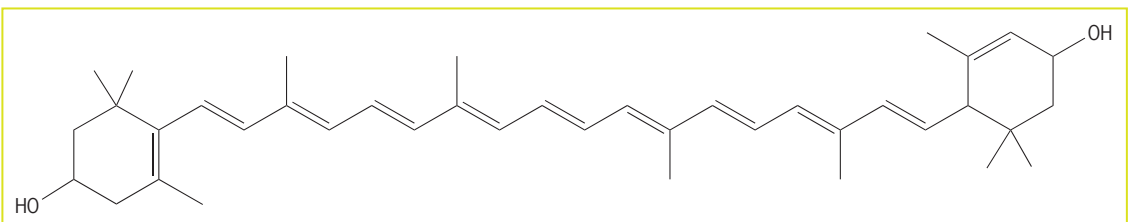


Figure 6.10.8 The xanthophyll lutein, which is found in chick egg yolks.

Carotenoids are fat-soluble due to their long non-polar side chains, which cancel out the polar hydroxyl side chains found in vitamin A or xanthophylls.

Molecules containing porphins

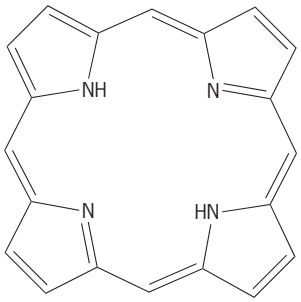


Figure 6.10.9 The general structure of a porphin molecule.

Both chlorophyll and heme contain a planar **heterocyclic** unit called a **porphin**. Around the perimeter of the macrocycle ring, there is a cyclic chain of sp^2 hybridized carbon atoms, all of which are part of a conjugated double bond system, giving the molecule its aromatic character. The aromatic character of porphin derives from both its conjugation as well as its planar geometry, meaning that all the atoms lie in a single plane. Porphin-containing molecules are coloured due to the conjugation of the macrocycle ring. Chlorophyll is green in colour and heme is red in colour.

When the central nitrogen atoms lose their hydrogens, they can act as ligands and bond to a metal ion in the centre of the macrocycle ring. When porphins have substituents in position 1 to 8 of the macrocycle ring they are called **porphyrins**. An example of a porphyrin is the common pigment called chlorophyll. Chlorophyll occurs naturally in five different forms, the two most common forms are chlorophyll *a* and chlorophyll *b*. Chlorophyll has a magnesium ion in the central position. The two most common forms of chlorophyll differ only in one substituent: CH_3 in chlorophyll *a* and CHO in chlorophyll *b*.

A **heme** molecule (figure 6.10.2d) also has a porphyrin macrocycle ring in which an iron(II) ion is held rigidly by the nitrogen atoms in the centre of the macrocycle. Heme molecules are found in both myoglobin and hemoglobin. Myoglobin is the primary pigment in muscle tissue and is a water-soluble protein that stores oxygen for aerobic metabolism in the muscle. It is a complex that consists of a protein portion (globin) and a porphyrin ring with a central iron atom (heme). The main function of hemoglobin is to transport oxygen from the lungs to the tissues and then transport CO_2 back from the tissues to the lungs. One hemoglobin molecule has the ability to transport up to four oxygen molecules.

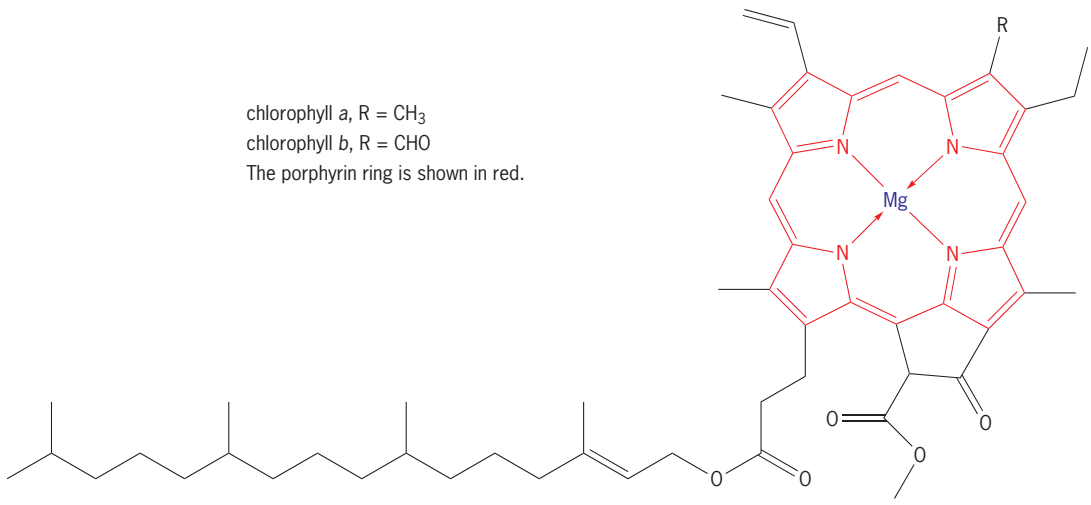


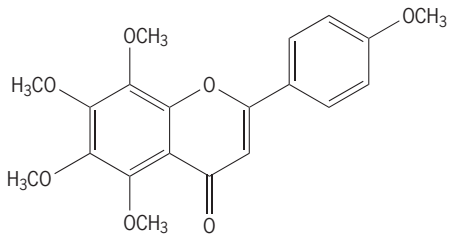
Figure 6.10.10 Chemical structures of chlorophyll *a* and chlorophyll *b*. Notice the porphyrin macrocycle ring and the magnesium ion at the centre of the structure.

Section 6.10 Exercises

1 Explain in terms of the chemistry involved why some chemicals such as pigments are coloured and why others are colourless.

2 Explain how the wavelength of energy absorbed by a molecule relates to the colour the food appears.

3 Tangeritin, shown below, is found in tangerine and other citrus peels. It strengthens the walls of cells and protects the fruit from invasion by pests. Tangeritin can be used as a cholesterol-lowering agent and to combat Parkinson's disease. It has been shown to be an anti-cancer agent.



a Identify the class of pigments to which tangeritin belongs and explain your choice.

b Explain whether or not you would expect tangeritin to be coloured.

4 Explain why different anthocyanins can have different colours.

5 **a** Explain the difference between carotenes and xanthophylls.

b Give an example of a common carotene and a common xanthophyll.

c Explain whether you would expect carotene and xanthophylls to be fat-soluble or water-soluble.

6 Explain how carotene is a precursor to vitamin A.

7 Define the terms *porphin*, *macrocycle* and *porphyrin*.

8 Explain the importance of metal ions in the porphin type molecules of chlorophyll and heme.

Chapter review questions and tests are available on student CD.

Chapter 6 Summary

Terms and definitions

2-Amino acids A carboxyl group and an amino group are bonded to the same carbon atom.

Anthocyanidin Benzopyrilium salt; a natural antioxidant.

Anthocyanin Sugar molecule bound to an anthocyanidin; a natural antioxidant.

Antioxidant Substance which slows or delays the oxidation of a food.

Astaxanthin Carotenoid pigment; antioxidant.

Auto-oxidation Reaction of fats and oils occurring by addition of oxygen across the carbon–carbon double bond.

Browning Process of food turning brown.

Caramelization Browning sugar in the presence of heat.

Carbohydrate Polymer of monosaccharides (contain C, H and O only).

Carotenoids Yellow-coloured pigment; natural antioxidant.

Cellulose Polymer of β-glucose.

Chelating agent Substance which forms coordinate covalent bonds with a metal ion.

Chiral Relating to a molecule that cannot be superimposed on its mirror image.

Chlorophyll Pigment in green plants which initiates photosynthesis.

Chromophore A group such as a carbon–carbon double bond which is able to absorb UV or visible radiation.

Colloid Dispersed system with particles smaller than 1 μm.

Condensation Joining of smaller molecules to make a larger molecule and water.

Conjugated Every second bond is a double bond.

Continuous phase Phase in a dispersed system through which another phase is dispersed.

Curing Meat preservation by adding sugar, salt, nitrates and/or nitrites.

Dextrorotatory Rotates to the right.

Dipeptide Two amino acids joined by a peptide bond.

Disaccharide Two monosaccharides joined by an ether link.

Dispersed phase In a dispersed system, this phase is spread throughout another phase.

Dispersed system System where one phase is spread throughout another.

DNA Deoxyribonucleic acid.

Dye Synthetic coloured species.

Emulsifier Surfactant which stabilizes emulsions.

Emulsion Dispersed system; one liquid dispersed throughout another of different polarity.

Enantiomer One of a pair of mirror-image molecules.

Endogenous naturally present.

Fatty acid Long chain carboxylic acid.

Fischer projections Method for representing stereochemical relationships. Horizontal bonds project towards the viewer, vertical bonds are into the page.

Foam Dispersed system; gas dispersed in a liquid.

Food Substance intended for human consumption.

Free radical Species with an unpaired electron.

Gene Segment of DNA which codes for a particular characteristic.

Gene silencing Technique which deactivates a gene in an organism.

Genetic engineering Field of science involved with manipulating DNA.

Genetically modified (GM) food Food produced from genetically altered organism.

Genetically modified organism (GMO) Organism with gene inserted or silenced.

Glycerol 1,2,3-propanetriol.

Glycogen Branched polymer of α-glucose.

Glycosidic bond (ether link) –O– link between two carbons.

Heme Porphyrin ring containing iron; constituent of hemoglobin and myoglobin.

Herbicide Weed-killer.

Hermetic sealing Airtight sealing.

Heterocyclic compound Cyclic compound in which one or more atoms in the ring are elements other than carbon.

High density lipoprotein (HDL) Lipoproteins with larger protein proportion.

Humectant Substance which absorbs moisture.

Hydrogenation Addition of hydrogen.

Hydrolysis Splitting up of large molecules using water.

Hydrolytic rancidity Triglycerides breaking into glycerol and fatty acids.

Hydroperoxide Functional group made up of two oxygen atoms bonded by single bonds to each other and a hydrogen atom (ROOH).

Levorotatory Rotates to the left.

Lipid Any triglyceride (fat or oil).

Lipoprotein Emulsion of proteins and cholesterol.

Low density lipoprotein (LDL) Lipoprotein with larger cholesterol proportion.

Macronutrient Present in large quantities in the diet.

Maillard reaction Non-enzymatic browning reaction.

Metmyoglobin Myoglobin containing iron(III).

Microemulsion Emulsion with droplet size less than 100 nm.

Micronutrient Present in small quantities in the diet.

Monosaccharide Simple sugar with empirical formula CH_2O.

Mono-unsaturated fatty acid Fatty acid containing one carbon–carbon double bond.

Myoglobin Purplish red iron-containing meat pigment.

Nutrient Substance which provides energy, or regulates growth, maintenance and repair of the body's tissues.

Optical isomers Isomers with different spatial orientation.

Oxidation Loss of electrons or gain of oxygen.

Oxidative rancidity Involves oxidation across the double bond of an unsaturated lipid.

Oxymyoglobin Myoglobin with oxygen bound to the iron(II) ion.

Pasteurization Heating milk to a certain temperature for a fraction of a second.

Peptide link –CONH– bond.

Phenolic A compound with a hydroxyl group bonded to an aromatic ring.

Photo-oxidation Oxidation that is catalysed by light.

Pigment Natural coloured species.

Polypeptide Polymer of amino acids.

Polysaccharide Carbohydrate polymer.

Polyunsaturated fatty acid Fatty acid containing more than one carbon–carbon double bond.

Porphin Planar heterocyclic molecule.

Porphyrin Large molecule with four lobes.

Protein Polymer of amino acids having multiple structures (primary, secondary, tertiary).

Racemic mixture Mixture containing equal quantities of enantiomers.

Rancidity Reactions that produce offensive odours and flavours in fats and oils.

Recombinant DNA technology Technology that allows foreign genes to be inserted.

Reducing sugar Sugar that forms aldehyde or ketone in basic solution.

Refractive index A measure for how much the speed of light is reduced inside the substance.

Saturated fatty acid Fatty acid containing only carbon–carbon single bonds.

Shelf life Time during which a food maintains its quality.

Stabilizer Food additive designed to maintain homogeneous texture.

Starch Mixture of amylase and amylopectin.

Suspension Dispersed system; solid dispersed throughout a liquid.

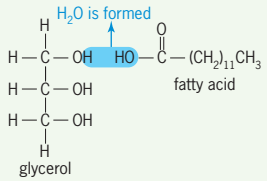
Transgenic Containing a gene from another organism.

Water activity (a_w) The amount of water in the food not bound to food molecules.

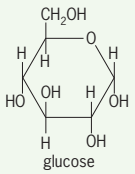
Xanthophyll Yellow, oxidized carotenoid pigment.

Concepts

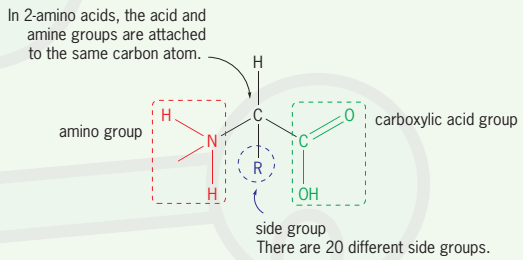
- A food is a substance intended for human consumption, whereas a nutrient is a substance that provides energy, or regulates growth, maintenance and repair of the body's tissues
- Lipids (fats and oils) are triesters (triglycerides) formed from three long-chain fatty acid molecules and one glycerol molecule.



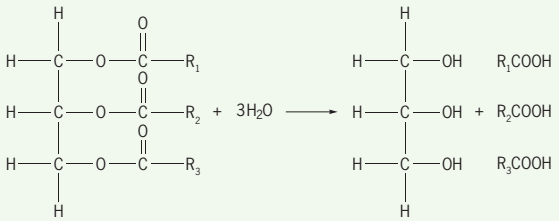
- The simplest carbohydrates are monosaccharides with the empirical formula CH_2O.



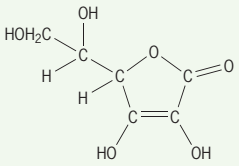
- Proteins are made up of 2-amino acids.

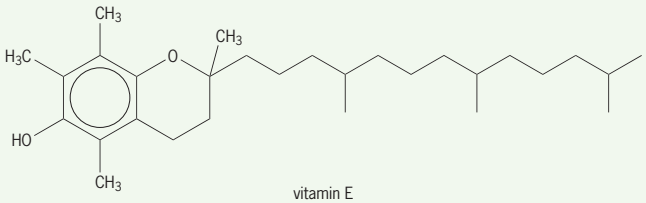


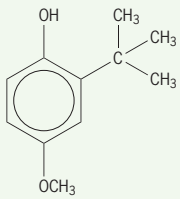
- Saturated fatty acids have no carbon–carbon double bonds; unsaturated fatty acids have at least one, and often many, carbon–carbon double bonds, which create a 'kink' in the carbon chain.
- The melting points of saturated fats are usually higher than those of unsaturated fats due to the close packing and greater van der Waals' forces possible between the hydrocarbon chains of saturated fats. It is possible to predict the stability and melting points of fats and oils from their structure.
- Hydrogenation of unsaturated fats is a process in which hydrogen is added across the double bonds, making the degree of saturation greater.
- Advantages of hydrogenation of fats and oils include increased melting point and decreased rate of oxidation.
- Disadvantages of hydrogenation of fats and oils includes formation of *trans* fatty acids and the healthier nature of unsaturated fats which are changed by hydrogenation.
- Shelf life is how long a food can be stored before the quality diminishes. It is affected by pH, water activity, additives, reactive species, microbes, temperature, light, humidity and exposure to air.
- Rancidity results in the production of small, volatile organic compounds with 'off' flavours and smells.
- Hydrolytic rancidity, which can be accelerated by lipases, produces glycerol and fatty acids, as shown.

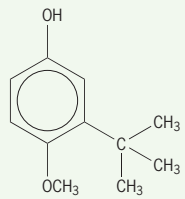


- Oxidative rancidity occurs via a free radical mechanism and produces peroxides, alcohols, hydrocarbons, aldehydes and ketones.
- Food preservation methods include salting, pickling, fermentation, heat, additives, refrigeration, freezing, hydrogenation, drying, modified atmospheres and acidifying.
- Antioxidants slow or delay oxidation.
- Natural antioxidants such as vitamins C and E are found in foods such as nuts and citrus fruits.
- Synthetic antioxidants, including BHA and BHT, are phenolic compounds. They are cheaper and more effective than their natural counterparts, but natural antioxidants are preferred by consumers.

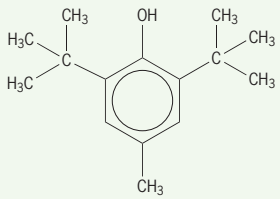

vitamin C


vitamin E


(2-BHA)


(3-BHA)

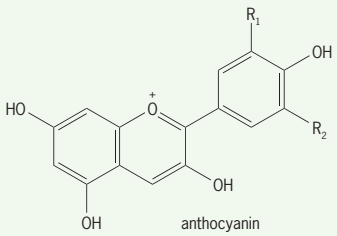
BHA


BHT

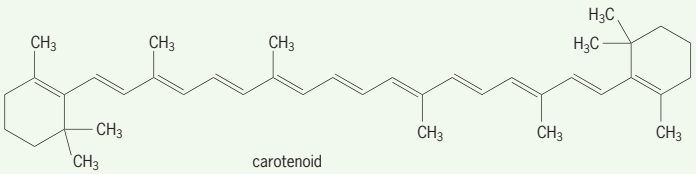
- Foods with antioxidant properties that have been used for centuries include green tea, oregano and chocolate.
- The different colours of foods are the result of some wavelengths of visible light being absorbed by the chemical bonds and substituents of the substances in food and other wavelengths being reflected.
- Pigments are insoluble natural colours that are found as suspensions, in the cells of plants and animals.
- Dyes are water-soluble substances that can be added to foods to change or enhance their natural colours.

Pigment	Colours	Some sources
Anthocyanins	Orange-red to red-blue	Blackcurrant, raspberry,
Carotenoids	Yellow to orange to red	Oranges, carrots, tomatoes,
Chlorophyll	Green to olive-green	Green leafy plants.
Heme and myoglobin	Red to red-purple	Meat, poultry and fish

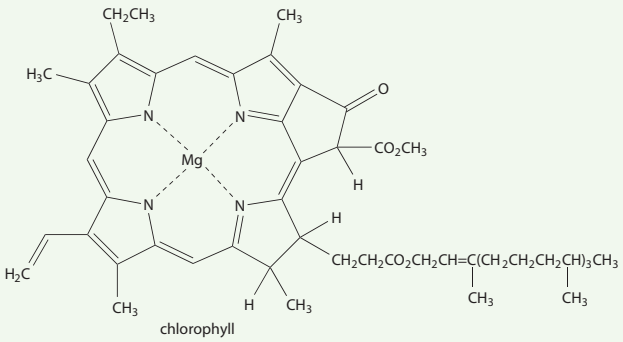
- Anthocyanin stability is affected by changes in pH, temperature, oxygen, UV light and different co-factors such as metal ions.


anthocyanin

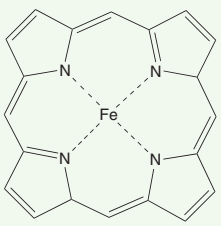
- Carotenoids are susceptible to oxidation catalysed by light, metals and hydroperoxides.


carotenoid

- Chlorophyll stability is affected by changes in oxygen levels, light, heat and pH.


chlorophyll

- Myoglobin, the molecule responsible for the colour of meat, consists of a protein portion and a non-protein porphyrin ring with a central iron atom. The iron atom is an important player in meat colour. The factors that primarily determine the colour of meat are the oxidation state of the iron and which compounds are attached to the iron portion of the molecule.

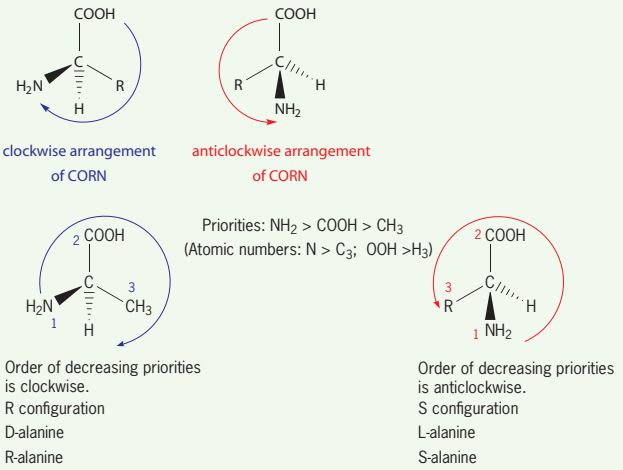

heme

- The two processes of non-enzymatic browning of food are the Maillard reaction and caramelization.
- At present there are no international standards with regards to synthetic food colourants.
- Genetically modified (GM) foods are produced by genetically modified organisms (GMOs).
- Most GM foods are crops.
- GM crops offer pest resistance, herbicide resistance, disease resistance, increased quality, faster maturing and increased yield.
- Disadvantages of GM crops include consumer negativity, rise of pesticide-resistant pests and herbicide-resistant weeds, out-crossing, increased allergies and unknown human impacts.
- Texture relates to how a food feels in the mouth.
- Dispersed systems include foams, emulsions and suspensions.
- Emulsifiers stabilise emulsions by lowering surface energy.

- The free-radical mechanism that occurs during oxidative rancidity involves three steps: initiation, propagation and termination.

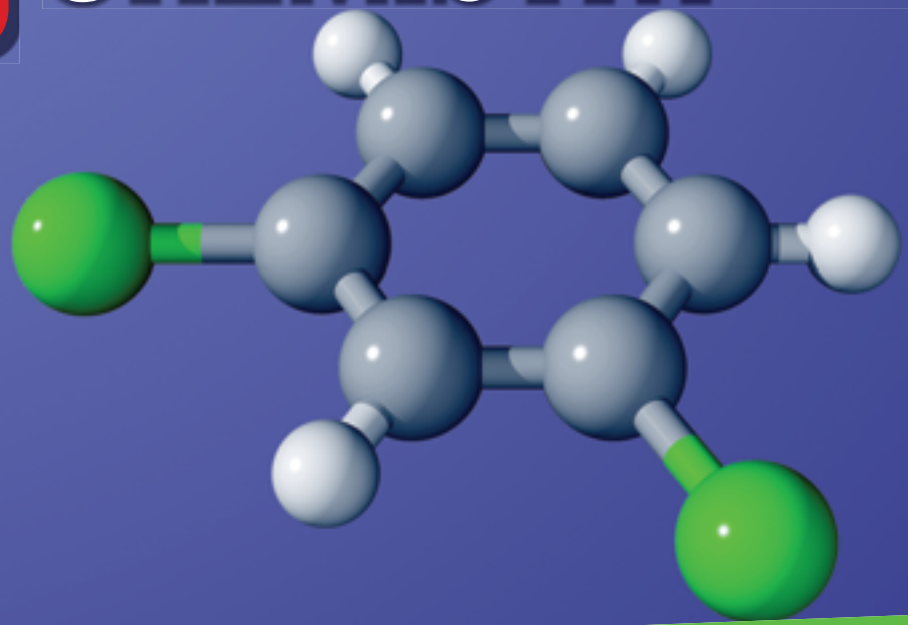
HL

- There are three main groups of antioxidants: radical scavengers, chelating agents and reducing agents.
- Enantiomers are named according to three different conventions that are based on the chiral carbon:
  - (+) and (–) depending on the direction in which they rotate plane-polarized light
  - D and L depending upon their resemblance to the (+) glyceraldehyde molecule and the 'CORN' rule for amino acids
  - R and S depending upon the priorities placed on the four substituents based on their molecular masses and then whether this priority is clockwise (R) or anticlockwise (S)



- Enantiomeric forms of stereoisomers found in food may cause the food to have a different taste, or smell or to have different degrees of toxicity.
- Anthocyanins are coloured due to the presence of a $C_6C_3C_6$ flavonoid skeleton with conjugated double bonds.
- Heme and chlorophyll are coloured as they contain porphins whose structure contains a cyclic system of conjugated double bonds.
- Due to their structures, anthocyanins are water-soluble and carotenoids are fat-soluble.

7 FURTHER ORGANIC CHEMISTRY



Chapter overview

This chapter covers the IB Chemistry syllabus Option G: Further Organic Chemistry.

By the end of this chapter, you should be able to:

- describe and explain the electrophilic addition mechanisms of the reactions of alkenes with halogens and hydrogen halides
- describe and explain the mechanism for the addition of hydrogen cyanide to aldehydes and ketones
- use equations to describe the hydrolysis of cyanohydrins to form carboxylic acids
- describe and explain the mechanism for the elimination of water from alcohols to form alkenes
- use equations to describe the reactions of 2,4-dinitrophenylhydrazine with aldehydes and ketones
- describe and explain the structure of benzene using physical and chemical evidence
- describe and explain the relative rates of hydrolysis of benzene compounds halogenated in the ring and in the side-chain
- use equations to describe the formation and reactions of Grignard reagents
- deduce reaction pathways for reactions involving alkenes, alcohols, aldehydes and ketones
- describe and explain the acidic properties of substituted phenols and carbboxylic acids in terms of bonding
- compare and explain the relative basicities of ammonia and amines

- use equations to describe the reactions of acid anhydrides to form carboxylic acids, esters, amides and substituted amides HL
- explain the reactions of acyl chlorides with nucleophiles in terms of an addition–elimination mechanism
- describe and explain the mechanisms for the nitration, chlorination, alkylation and acylation of benzene
- describe and explain the directing effects and relative rates of reaction of different substituents on a benzene ring
- deduce reaction pathways for reactions involving acidic anhydrides, acyl chlorides, benzene and their derivatives.

IBO Assessment Statements G1.1 to G10.4

We have seen in the Standard Level and Higher Level courses that the study of the millions of known organic compounds is simplified by studying the structures and reactions of homologous series of compounds containing particular functional groups. We have also looked at how certain organic reactions occur and at the mechanism of these reactions, and have begun to see how an understanding of reaction mechanisms allows us to choose conditions to favour the synthesis of desired organic compounds. This chapter extends the study of reactions of compounds containing particular functional groups, and of the mechanisms of these reactions. The chemistry and reactions of the benzene containing compounds will also be introduced.

7.1 ELECTROPHILIC ADDITION REACTIONS

AS G.1.1
Describe and explain the electrophilic addition mechanisms of the reactions of alkenes with halogens and hydrogen halides. © IBO 2007

We know from earlier studies that the chemistry of the alkenes is the chemistry of the carbon–carbon double bond. The double bond consists of a strong sigma (σ) bond and a weak pi (π) bond. The σ bond electrons are sheltered between the carbon nuclei, and are therefore not easily attacked. The pair of electrons in the π bond is less firmly held to the bonding carbon nuclei. The π electron clouds above and below the plane of the atoms serve as a source of electrons. This electron richness and accessibility leaves the alkene double bond open to reaction with an **electrophile** (an electron-loving reactant), resulting in a range of **electrophilic addition** reactions.

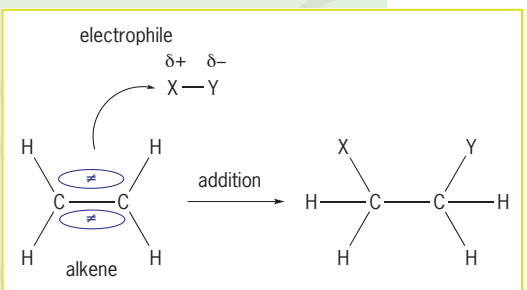


Figure 7.1.1 The alkene double bond is open to reaction with electrophiles.

Electrophiles that react with the electrons of the carbon–carbon double bond include positively charged reactants such as H^+, and neutral reactants such as HBr and Br_2. Electrophilic addition reactions of alkenes are a general reaction type that allows the synthesis of a wide variety of compounds.

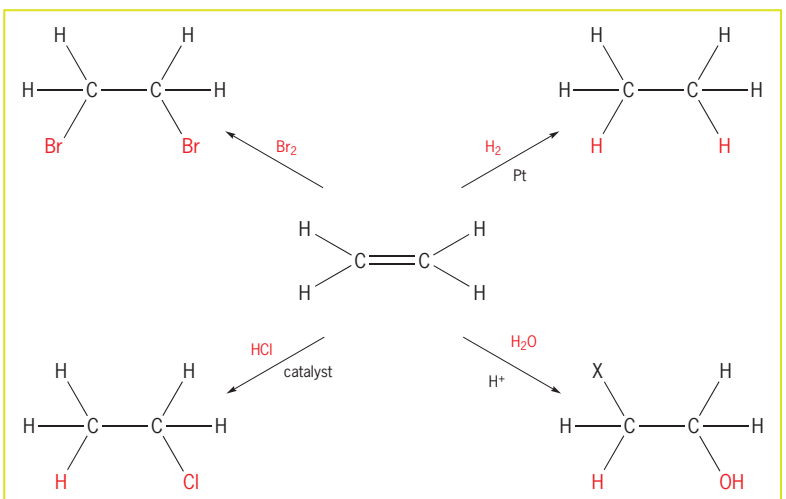


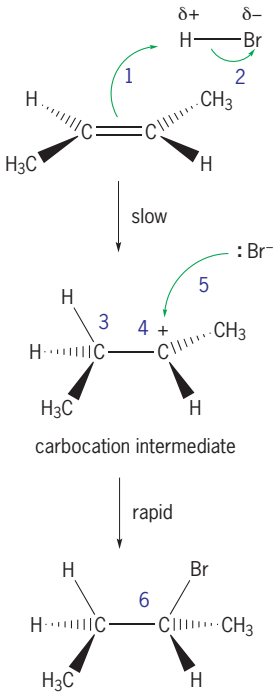
Figure 7.1.2 Alkenes undergo a range of electrophilic addition reactions.

The electrophilic addition mechanism is illustrated by the reaction of an alkene with hydrogen bromide, HBr (shown in figure 7.1.3). Reaction begins with the attack of the π electrons of the double bond on the electrophile (the δ+ hydrogen end of the HBr molecule). Two electrons from the double bond form a new σ bond with the hydrogen, and a carbon–hydrogen bond forms. This leaves the other carbon atom of the double bond with a positive charge and a vacant p orbital. A **carbocation** intermediate is thus formed. The bond of the H–Br breaks heterolytically, leaving the negatively charged bromide ion, Br^-. The carbocation intermediate accepts an electron pair from this bromide ion to form a new C–Br bond. The initial carbocation formation step is slow. Once formed the carbocation intermediate reacts rapidly with the bromide ion to produce the final bromoalkane product.

Other asymmetrical reactants such as ICl add similarly across a carbon–carbon double bond. The electrophilic part (δ+) of the reactant is attacked by the electron-rich double bond. A carbocation is formed, and the reaction is then completed by the nucleophilic part (δ-) of the reactant.

The electrophilic addition mechanism can also be illustrated by the addition of a symmetrical reactant such as a bromine molecule across a carbon–carbon double bond. When a bromine molecule approaches the π cloud of the double bond, the outer electrons of the nearest bromine atom are repelled, and so the bromine molecule becomes polarized. The bromine closer to the π cloud is now positively polarized, and is attracted to the electron cloud. Two electrons from the π cloud form a new C–Br, and the Br–Br bond breaks heterolytically. As a result, a positively charged carbocation intermediate is formed, and the other bromine becomes a bromide ion. This bromide ion then attacks the carbocation, resulting in the dibromo product.

This two-step mechanism raises the possibility that in the rapid second step the addition process may be completed by any nucleophile present. For example, if addition of bromine was conducted in aqueous solution, another nucleophile present, H_2O, might react with the carbocation to form the 1-bromo-2-hydroxy product, rather than the dibromo product. This may occur as the H_2O nucleophile is far more abundant than the Br^- nucleophile (see figure 7.1.6).



1 The electophile HBr is attacked by the π electrons of the double bond.

2 The H–Br bond breaks heterolytically.

3 A new C–H bond forms.

4 The other carbon atom now has a positive charge.

5 Br^- donates an electron pair to the positively charged carbocation.

6 A new C–Br bond forms.

Figure 7.1.3 The electrophilic addition mechanism illustrated by the reaction of an alkene with hydrogen bromide.

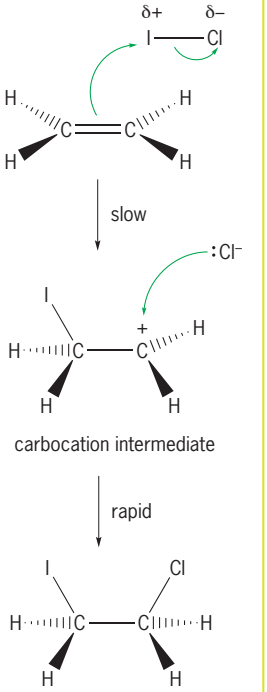


Figure 7.1.4 The electrophilic addition mechanism illustrated by the reaction of an alkene with iodochloride.

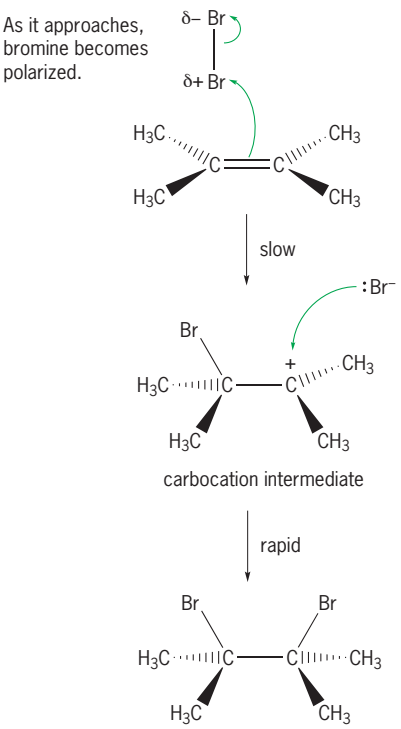


Figure 7.1.5 The electrophilic addition mechanism is illustrated by the reaction of an alkene with bromine.

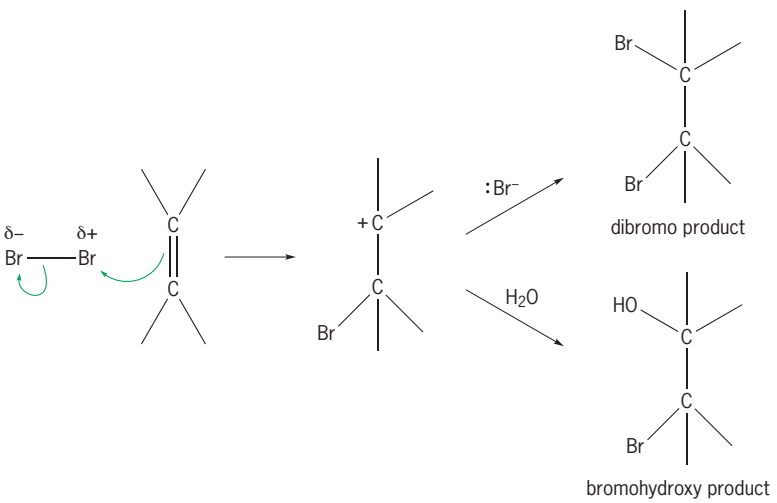


Figure 7.1.6 Electrophilic addition may be completed by any nucleophile present. Reaction of an alkene with bromine may produce a dibromo product or a bromohydroxy product.

Orientation of addition: Markovnikov's rule

AS G.1.2 Predict and explain the formation of the major product in terms of the relative stabilities of carbocations. © IBO 2007

In the examples considered above, the alkene reacting was symmetrical; that is, the alkyl groups on either side of the double bond were the same. When an asymmetrical electrophile is added across the double bond of an asymmetrical alkene, such as 2-methylpropene, a problem of orientation of addition arises. For example, hydrochlorination of 2-methyl propene involves competing additions, as shown in figure 7.1.7. It is found that only the 2-chloro-2-methylpropane product is formed. How can this be explained?

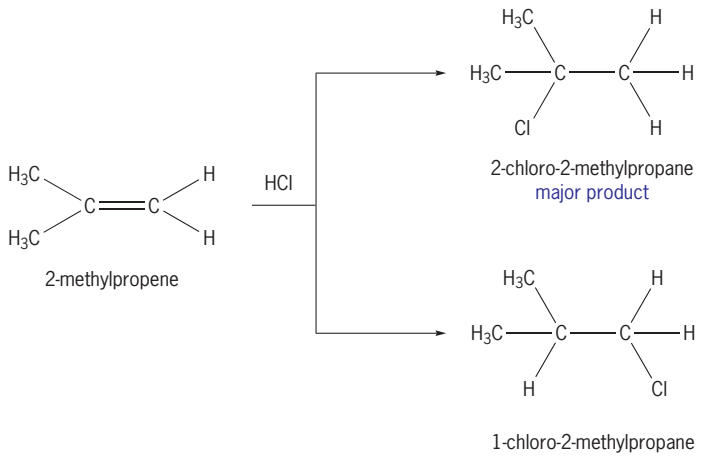


Figure 7.1.7 Addition of an asymmetrical electrophile to an asymmetrical alkene produces two products, but not in equal quantities.

Explanation lies in the reaction mechanism, and can be understood in terms of the stability of the intermediate carbocation. Addition to form the 2-chloro-2-methylpropane product takes preference due to the higher stability of the tertiary carbocation relative to the primary carbocation.

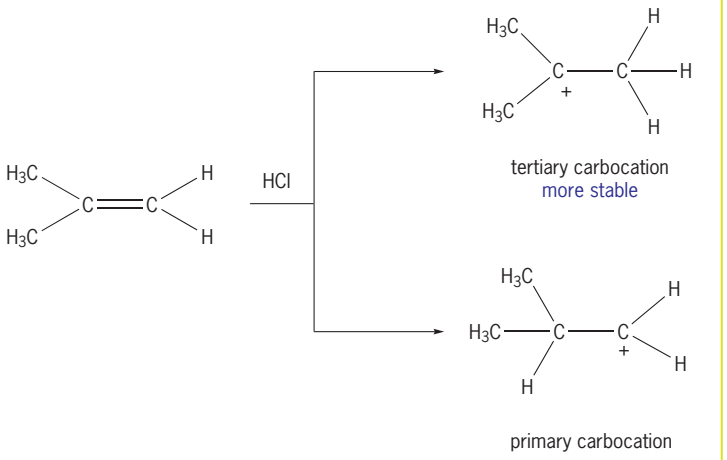


Figure 7.1.8 The tertiary carbocation intermediate is more stable, so the reaction to produce this cation is favoured.

Stability of carbocations was considered in *Chemistry: For use with the IB Diploma Programme Higher Level*, chapter 9, when looking at the **nucleophilic substitution** reactions (p. 299). The idea of stabilizing the charge on the carbocation is also an important consideration here in the electrophilic addition reaction. Alkyl groups have a tendency to 'push' the R–C bonding electron pair towards the positively polarized carbon atom. This effect is known as the **positive inductive effect**. Positive inductive effects of alkyl groups in a carbocation stabilize the ion. The effect of the electron-'pushing' alkyl groups is to allow the positive charge of the carbocation to be dispersed, and so to be more stabilized. This stabilization of the positive charge increases the likelihood of formation of the carbocation. The order of stability of carbocations is therefore related to the number of alkyl groups attached to the central, positively charged carbon atom. The three alkyl groups in a tertiary carbocation exert a greater positive inductive effect than the two groups in a secondary carbocation. These two groups in turn exert a greater effect than the one group in a primary carbocation, which in turn has a greater effect than the hydrogen atom of a methyl carbocation.

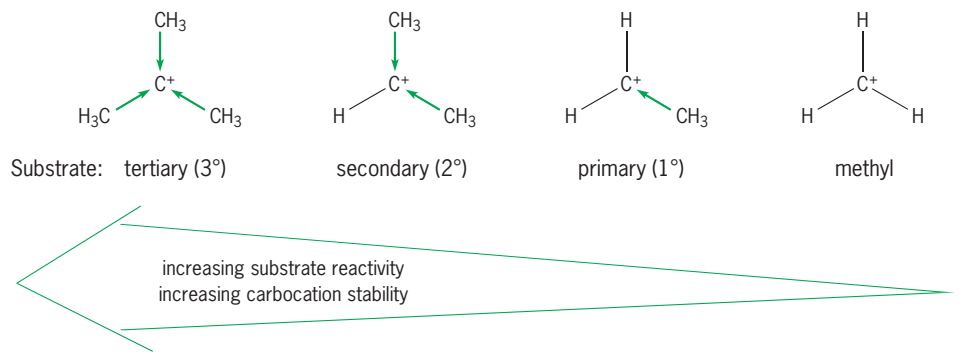


Figure 7.1.9 Stability of carbocations affects the reactivity of the substrate molecules.

WORKSHEET 7.1
Electrophile addition reactions

Thus when HBr is added to an asymmetrical alkene, the H attaches to the carbon of the double bond that has the least number of alkyl groups attached. This, in turn, forms the most highly substituted, and hence the most stable, carbocation. The Br^- ion then subsequently attaches to the carbon with the most alkyl groups in the rapid second step. When both ends of the double bond have the same degree of substitution with alkyl groups, a mixture of products results.

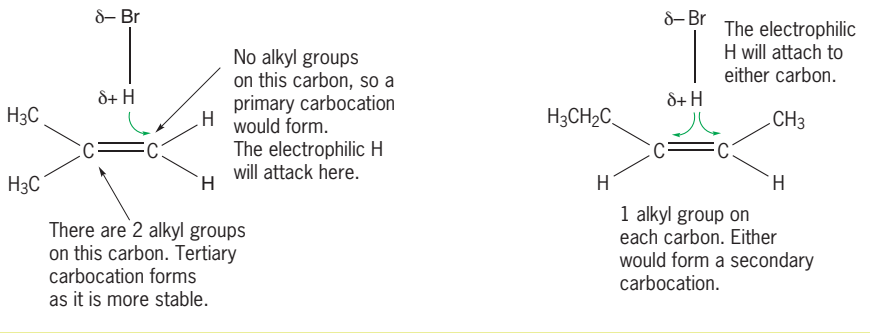


Figure 7.1.10 Stability of carbocations explains the orientation of addition of HBr to an asymmetrical alkene.

This orientation of addition of asymmetrical reactants to asymmetrical alkenes was proposed by Russian chemist Vladimir Markovnikov in 1869, and became known as **Markovnikov's rule:**

> **In the addition of HX to an alkene, the H attaches to the carbon with fewer alkyl substituents, and the X attaches to the carbon with more alkyl substituents.**

It is also sometimes stated in the form that:

> **Addition will take place so as to give the most stable intermediate carbocation.**

Section 7.1 Exercises

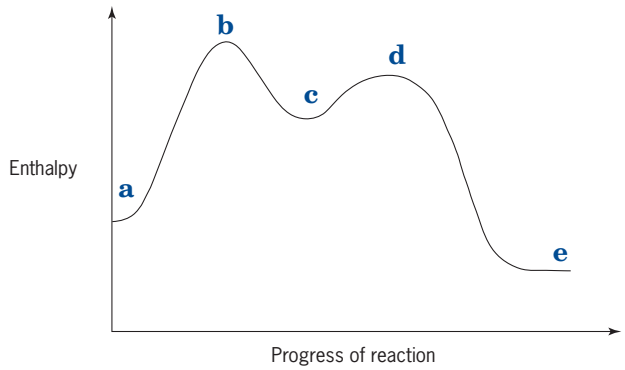
1 With the aid of examples and diagrams, explain the meaning of each of the following terms.

- **a** Nucleophile
- **b** Electrophile
- **c** Carbocation
- **d** Addition reaction
- **e** Positive inductive effect of alkyl groups

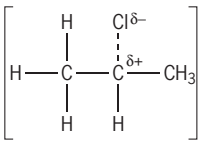
2 Consider the reaction between but-2-ene and bromine.

- **a** If the bromine is added drop-wise to the but-2-ene, what would you expect to observe?
- **b** Name the type of reaction occurring between but-2-ene and bromine.
- **c** Draw a mechanism for the reaction.

3 The graph below shows the energy profile for an electrophilic addition reaction. The structures of five molecules, transition states or intermediates are also shown, labelled V to Z. Use the letters V to Z to identify the structures present at each of the locations labelled **a** to **e** on the energy profile.



V

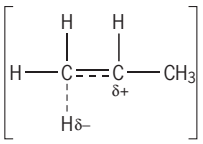


W CH_2CHCH_3

X $CH_3C^+HCH_3$

Y $CH_3CHClCH_3$

Z



4 **a** When ethene is reacted with bromine in a non-polar solvent, 1,2-dibromoethane forms.

Draw a reaction mechanism for this reaction.

b Predict the product formed when ethene is reacted with bromine in an aqueous solvent.

c Account for the different products formed in the reactions in parts **a** and **b**.

5 When HBr is added to 2-methylpropene two isomeric products may form.

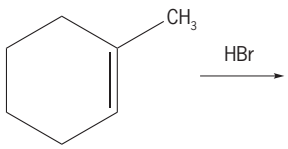
a Give the structural formulas and names of the products.

b Draw the mechanism for the reaction.

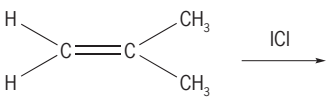
c In terms of the mechanism explain why one product is favoured over the other.

6 Predict the major product in each of the following reactions. Explain your choices.

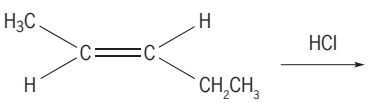
a



b



c



7.2 ELIMINATION REACTIONS

G.3.1
Describe, using equations, the dehydration reactions of alcohols with phosphoric acid to form alkenes. © IBO 2007

G.3.2
Describe and explain the mechanism for the elimination of water from alcohols. © IBO 2007

In the addition reactions considered in section 7.1, the double bond of an alkene molecule was removed as atoms were added to the carbon atoms on either side of the double bond. The reverse of this addition process is the removal of atoms from a molecule in such a way as to generate a double bond. The dehydration (removal of water) of alcohols by an **elimination reaction** to form an alkene is an example of such a reaction.

In the Standard Level course we looked at the reactions of alcohols to form aldehydes, ketones and carboxylic acids. Alcohol reactions typically may occur at either the O–H bond or the C–O bond. One of the most valuable reactions of alcohols involving the C–O bond is dehydration to give alkenes. The C–O bond and a neighbouring C–H are broken and an alkene is formed. Dehydration is often carried out by treatment of the alcohol with acid. While sulfuric acid is a good dehydrating agent, it is also an oxidizing agent, so that, in practice, dehydration is often performed using phosphoric acid, at temperatures from 100°C to 200°C. The acid functions as a catalyst to protonate the oxygen atom in the alcohol.

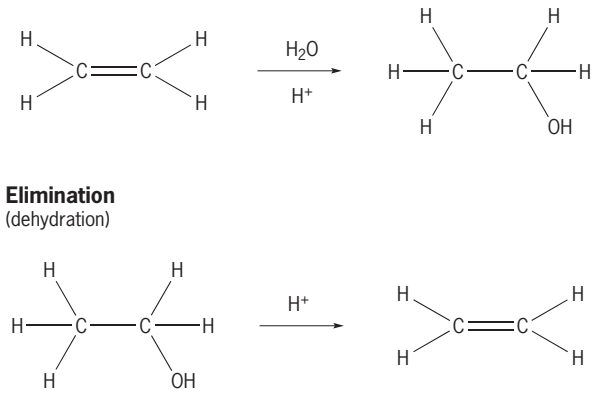


Figure 7.2.1 Addition of water to ethene produces ethanol. Elimination of water from ethanol produces ethene.

These acid-catalysed dehydrations proceed by an elimination reaction mechanism involving three steps, as shown in figure 7.2.2. The mechanism involves protonation of the alcohol oxygen, in which two electrons from the oxygen atom bond to H^+ to give a protonated intermediate. As the C–O bond breaks, the two electrons stay with the oxygen, leaving a carbocation intermediate. Finally, loss of a proton (H^+) from the neighbouring carbon atom occurs, with the two electrons forming a π bond between the carbon nuclei.

This type of dehydration of alcohols does not proceed equally well for all alcohols. Only tertiary alcohols are normally dehydrated with phosphoric acid. Secondary alcohols can be made to react, but severe conditions are required (75% H_2SO_4, 100°C). Primary alcohols are even less reactive than secondary ones, and very harsh conditions are necessary to cause dehydration (95% H_2SO_4, 150°C). The reactivity order for acid-catalysed dehydrations is therefore primary less than secondary, and secondary less than tertiary. This order of reactivity can be understood by considering the mechanism of the reaction, and largely depends on the stability of the intermediate carbocation. Tertiary alcohols always react fastest in these reactions because they lead to highly stabilized, tertiary carbocation intermediates.

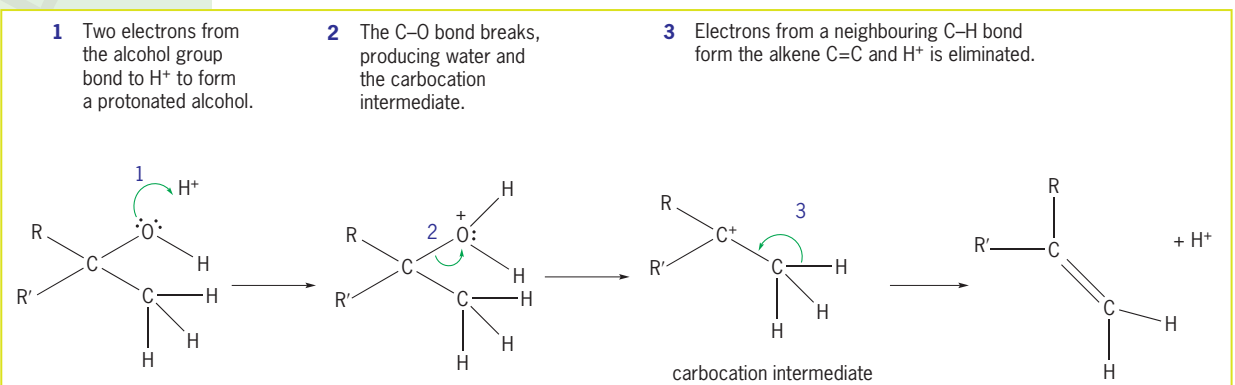


Figure 7.2.2 Dehydration of an alcohol by an elimination reaction mechanism.

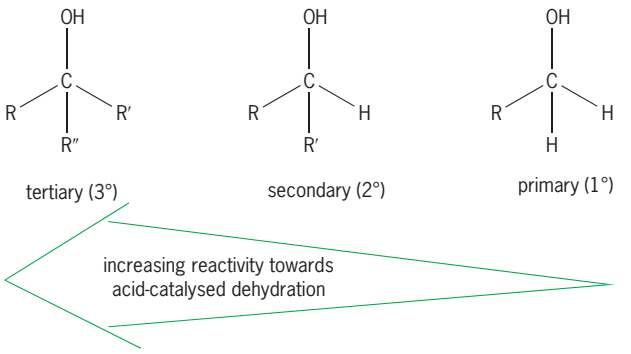


Figure 7.2.3 Reactivity of alcohols towards acid-catalysed dehydration matches the order of stability of carbocations shown in figure 7.1.9.

Section 7.2 Exercises

1 In what way can the dehydration of alcohols reaction be thought of as the reverse of an addition reaction to an alkene?

2 Explain why it is preferable to dehydrate alcohols with phosphoric rather than sulfuric acid.

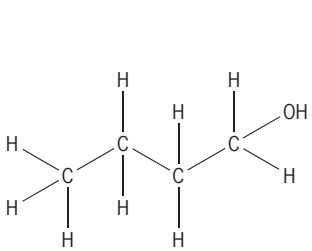
3 Butan-2-ol can be converted to but-2-ene.

a Name the reagent used for this conversion.

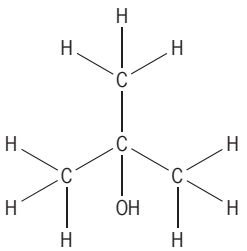
b Identify the type of reaction.

c Draw the mechanism for the reaction.

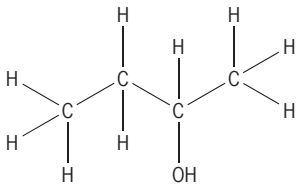
4 The structures of four alcohols are shown below.



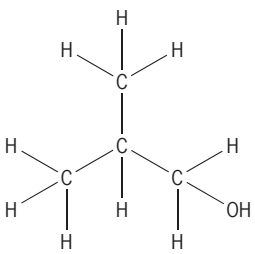
I



II



III



IV

a Draw a possible product of the acid-catalysed dehydration of each alcohol.

b Identify the alcohol that would be expected to show the greatest reactivity towards acid-catalysed dehydration. Explain your choice.

7.3 NUCLEOPHILIC ADDITION REACTIONS

We saw in the Standard Level course that aldehydes were readily oxidized to form carboxylic acids, but ketones were generally inert towards oxidation. We now examine the most general reaction type of both aldehydes and ketones—**nucleophilic addition** reactions.

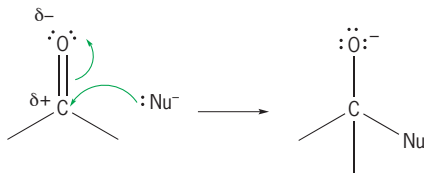


Figure 7.3.1 Polarity of the carbonyl bond leads to attack by nucleophiles.

Owing to the high electronegativity of oxygen, the carbonyl group is strongly polarized, with the electrons shifted towards the oxygen atom. The carbonyl carbon is consequently electron-deficient, while the oxygen is electron-rich. The partially charged carbonyl carbon is therefore readily attacked by a nucleophile. Attack by nucleophilic reactants at the carbonyl carbon is the prominent type of reaction of aldehydes and ketones.

The attacking nucleophile can be either negatively charged ($Nu:^-$) or neutral (:Nu–H), but it must have at least one non-bonding pair of electrons for coordinating with the carbonyl carbon. If it is neutral, the nucleophile usually carries a hydrogen atom that can subsequently be eliminated. Examples of negatively charged nucleophiles are HO^-, H^-, RO^- and NC^-. Neutral nucleophiles include H_2O, ROH, NH_3 and RNH_2.

The mechanism for nucleophilic addition is shown in figure 7.3.2 and involves the initial attack of the nucleophile on the carbonyl carbon, forming a new Nu–C bond. The carbonyl oxygen gains an unshared electron pair, forming an **alkoxide ion** RO^- as an intermediate. The electron-rich oxygen of the alkoxide ion then transfers its electron pair to a proton, thus completing the overall addition of Nu–H to the carbonyl group.

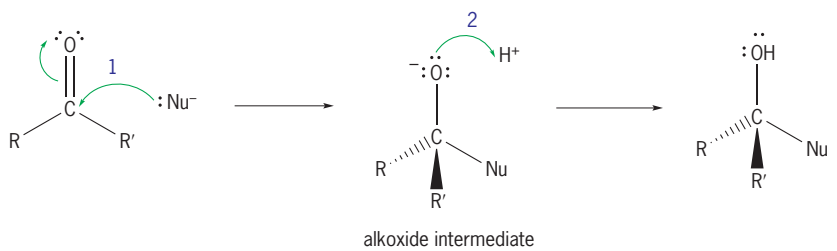


Figure 7.3.2 The nucleophilic addition reaction mechanism.

AS **G.2.1 Describe, using equations, the addition of hydrogen cyanide to aldehydes and ketones. © IBO 2007**

The addition of hydrogen cyanide (HCN) to aldehydes and ketones is an example of the nucleophilic addition reaction. Hydrogen cyanide adds to the carbonyl groups of aldehydes and most ketones to form **cyanohydrins** (hydroxynitriles), RCH(OH)CN or RR′C(OH)CN.

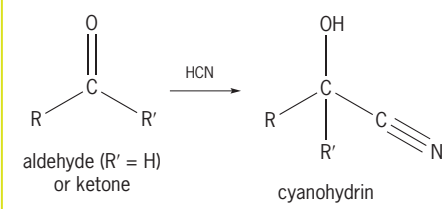


Figure 7.3.3 Reaction of an aldehyde or ketone with hydrogen cyanide produces a cyanohydrin.

The neutral HCN molecule is a poor nucleophile, so reaction occurs very slowly with pure HCN. However, if a small amount of base is added the cyanide ion, CN^-, forms. The CN^- is a strong nucleophile. The mechanism of cyanohydrin formation, shown in figure 7.3.4, begins with a nucleophilic attack by the cyanide ion on the carbonyl carbon. The alkoxide intermediate is formed. The alkoxide ion then becomes protonated to yield the cyanohydrin.

G.2.2
Describe and explain the mechanism for the addition of hydrogen cyanide to aldehydes and ketones.

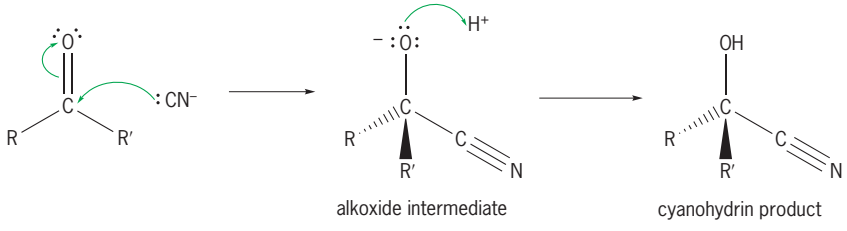


Figure 7.3.4 The nucleophilic addition reaction mechanism for the reaction of a ketone with hydrogen cyanide.

WORKSHEET 7.2
Nucleophilic addition reactions

Cyanohydrins are useful intermediates in organic syntheses. The CN^- group is readily hydrolysed by hot, aqueous acid to yield a carboxylic acid. Cyanohydrins may also undergo reduction with $LiAlH_4$ to yield primary amines. Cyanohydrins therefore provide a useful method of transforming an aldehyde or ketone into a different functional group.

G.2.3
Describe, using equations, the hydrolysis of cyanohydrins to form carboxylic acids.

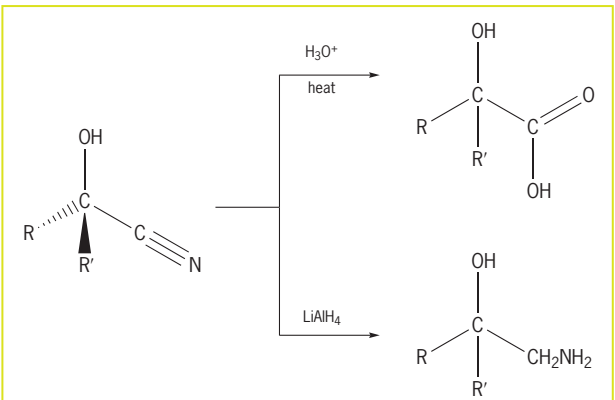


Figure 7.3.5 Cyanohydrins react to produce carboxylic acids and amines.

Addition–elimination reactions

Aldehydes and ketones also react with amines in nucleophilic addition reactions. Recall that the amine functional group is made up of a nitrogen atom bonded to two hydrogen atoms. In a similar way to the hydrogen cyanide reaction, the nucleophile, in this case NH_3 or an amine, attacks the positively polarized carbonyl carbon atom, producing the alkoxide intermediate. In the next step, however, the alkoxide ion does not become protonated to form the alcohol group. Elimination of a water molecule occurs to yield a product with a C=N bond.

G.4.1
Describe, using equations, the reactions of 2,4-dinitrophenylhydrazine with aldehydes and ketones.

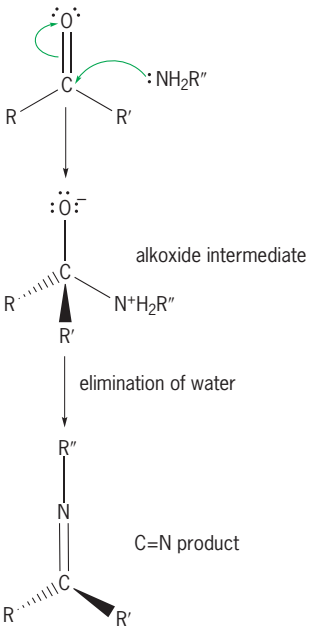


Figure 7.3.6 The addition–elimination mechanism for the reaction of ketones with amines.

Although the mechanism involves addition and elimination, the reaction may be thought of more simply as a **condensation reaction** in which two molecules combine with the elimination of a simple molecule (water).

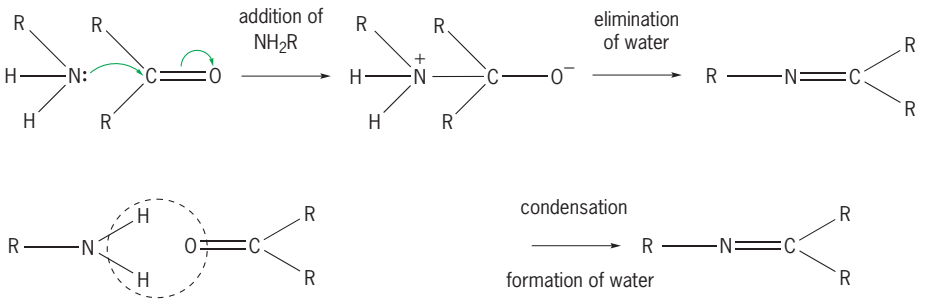


Figure 7.3.7 Addition–elimination may be thought of simply as a condensation reaction.

An important example of this kind of reaction involves the reaction of aldehydes or ketones with a number of derivatives of ammonia, in particular, with **hydrazine**, phenylhydrazine and 2,4-dinitrophenylhydrazine (the structures of which are shown in figure 7.3.8). Aldehydes and ketones react with these ammonia derivatives to form hydrazone, phenylhydrazone and 2,4-dinitrophenylhydrazone products respectively.

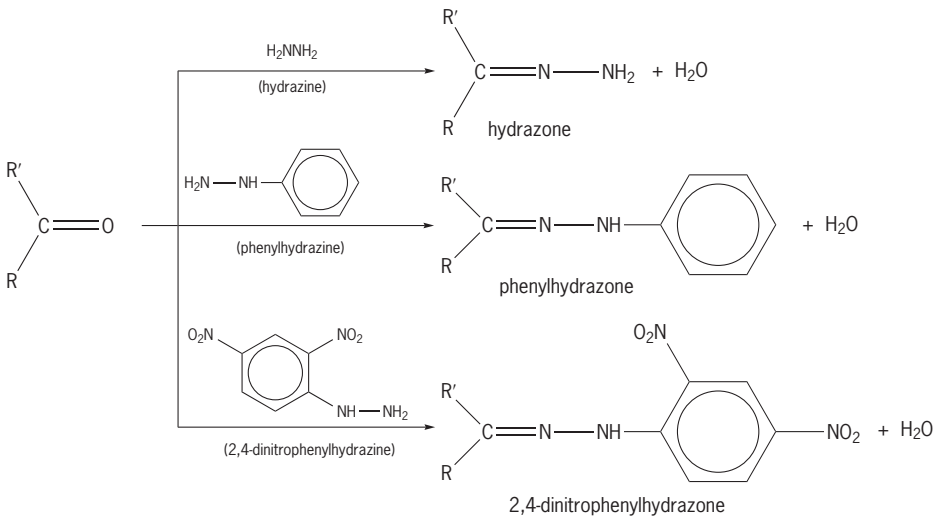


Figure 7.3.8 Reaction of aldehydes and ketones with hydrazines produces hydrazone products.

These hydrazone condensation products are crystalline solids, often red or orange, which can be readily purified by recrystallization. They have sharp characteristic melting points and so they can be used to identify the original aldehyde or ketone.

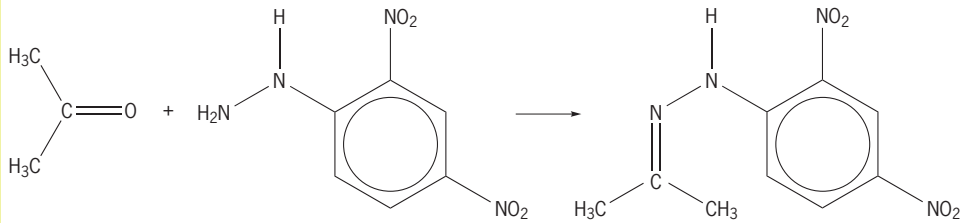


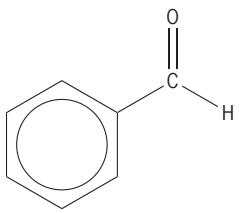
Figure 7.3.9 2,4-dinitrohydrazones have distinctive characteristics that may be used analytically.

Section 7.3 Exercises

1 Draw diagrams to illustrate:

- **a** the carbonyl functional group
- **b** a cyanide ion
- **c** a cyanohydrin
- **d** an alkoxide ion
- **e** hydrazine
- **f** a hydrazone.

2 This question refers to the molecule shown at right.



- **a** Circle the carbonyl functional group in the molecule.
- **b** Indicate the polarity of the carbonyl group in the molecule.
- **c** Draw the product formed when the molecule reacts with hydrogen cyanide.
- **d** Name the type of reaction occurring when the molecule reacts with hydrogen cyanide.

3 Write equations for the following reactions.

- **a** Butanal with hydrogen cyanide
- **b** Butanone with hydrogen cyanide

4 **a** Draw the mechanism for the reaction of hydrogen cyanide with:

- **i** propanal
- **ii** propanone.

b Write an equation for the acid hydrolysis of the product of the reaction in part **a ii**.

5 Show reaction sequences to illustrate the use of nucleophilic addition reactions involving hydrogen cyanide in the conversion of:

- **a** ethanal to 2-hydroxypropanoic acid
- **b** butanone to 2-hydroxy-2-methylbutanoic acid.

6 **a** Draw a structural diagram of the organic product of the reaction of propanone with 2,4-dinitrophenylhydrazine.

b Explain how the product in part **a** may be used to identify the original reacting ketone.

7.4 ORGANOMETALLIC CHEMISTRY

AS **G.6.1**
Outline the formation of Grignard reagents. © IBO 2007

Nucleophilic addition reactions are also used to produce alcohols from aldehydes and ketones. This involves R^- serving as the nucleophile to form the intermediate alkoxide ion, and the subsequent protonation of the alkoxide ion to yield the alcohol.

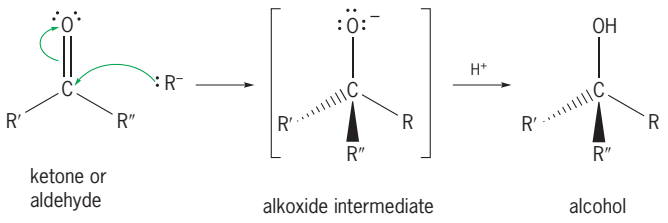


Figure 7.4.1 Nucleophilic addition of aldehydes and ketones with the R^- electrophile produces alcohols.

Generation of the R^- nucleophile involves us in a branch of organic chemistry known as organometallic chemistry. Compounds containing a carbon atom covalently bonded to a metal atom are known as **organometallic compounds**. One important group of these compounds is the **Grignard reagents**, named after their discoverer Victor Grignard, who shared the 1912 Nobel Prize in Chemistry for their discovery. Grignard reagents can be prepared by reacting halogenoalkanes, RX, with magnesium metal in ether solvent to yield alkylmagnesium halides, RMgX.

AS **G.6.2**
Describe, using equations, the reactions of Grignard reagents with water, carbon dioxide, aldehydes and ketones. © IBO 2007

Grignard reagents are highly reactive. They react with a wide variety of inorganic compounds, including water and carbon dioxide, and with most types of organic compounds. Reactivity of Grignard reagents is due to a highly polar covalent C–Mg bond. The carbon atom is negatively polarized with respect to the metal. This is not the typical polarity of the carbon atom, and this makes these Grignard reagents unique in their chemistry. The polarized carbon–magnesium bond makes the carbon atom electron-rich, and so it is nucleophilic. Because of this nucleophilic character, Grignard reagents react with electron-deficient centres. We will consider three reactions as examples.

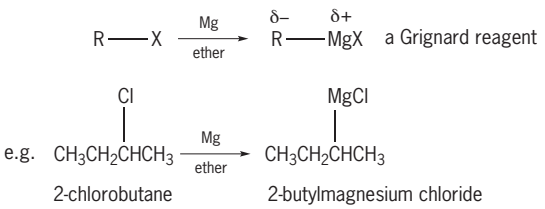


Figure 7.4.2 Grignard reagents are prepared by reacting halogenoalkanes with magnesium metal in ether solvent.

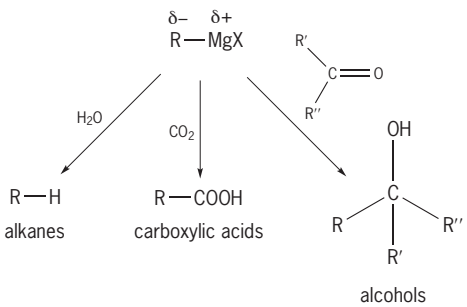


Figure 7.4.3 Grignard reagents are highly reactive. They react with water, carbon dioxide and carbonyl compounds.

Grignard reagents react with water to yield hydrocarbons.

$$CH_3CH_2MgBr \xrightarrow{H_2O} CH_3CH_3 + Mg(OH)Br$$

During their preparation and in reactions, Grignard reagents must therefore be kept in completely dry conditions or they will react readily with any water present to form an alkane. The sequence of Grignard formation followed by hydrolysis forms a useful method for converting a halogenoalkane into an alkane.

$$CH_3CH_2Br \xrightarrow{Mg} CH_3CH_2MgBr \xrightarrow{H_2O} CH_3CH_3 + Mg(OH)Br$$

Grignard reagents also react readily with carbon dioxide to produce carboxylic acids. The reaction is carried out either by pouring the Grignard reagent over dry ice (solid CO_2) or by bubbling a stream of dry CO_2 through the Grignard reagent solution.

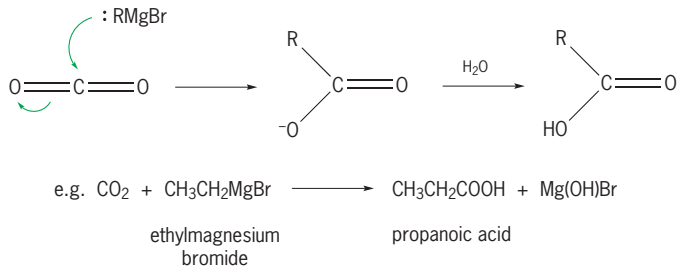


Figure 7.4.4 Grignard reagents react with carbon dioxide to produce carboxylic acids.

As seen earlier, Grignard reagents also react with aldehydes and ketones to yield alcohols in a nucleophilic addition reaction. The class of alcohol obtained depends on the type of carbonyl compound used; methanal yields primary alcohols, other aldehydes yield secondary alcohols and ketones yield tertiary alcohols.

Note that the reaction of Grignard reagents with aldehydes and ketones (and other reactants) enables the carbon chain of these organic molecules to be increased by one or several carbon atoms (depending on the alkyl used in the Grignard reagent). This lengthening of the carbon chain is a useful synthetic feature. Similarly, recall that the carbon chain of an aldehyde or ketone can be increased by one carbon using nucleophilic addition with hydrogen cyanide, and, from the Higher Level course, that nucleophilic substitution of a halogenoalkane with hydrogen cyanide also increases the carbon chain length by one carbon atom.

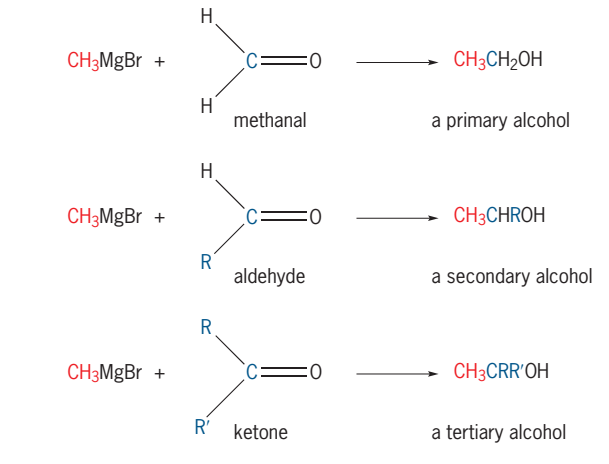


Figure 7.4.5 Grignard reagents react with aldehydes and ketones to produce alcohols.

CHEM COMPLEMENT

The Nobel Prize

The Nobel Prize is the most prestigious prize that can be awarded to a scientist, writer, economist or peace advocate. Alfred Nobel (1833–1896) was the Swedish-born inventor of dynamite. Nobel opened a nitroglycerin factory in 1865 that soon blew up! Nobel's discovery in 1867 that adsorbing nitroglycerin onto cellulose would make it more stable made him a very wealthy man. Despite the fact that he had given the military a powerful new weapon, Nobel was a strong supporter of peace movements. Nobel also derived great pleasure from literature. He wrote poetry and drama, and may have ended up a full-time writer had not his scientific interests consumed him. Nobel never married, which allowed him to leave his entire estate to a fund, the interest of which, '…shall be annually distributed in the form of prizes to those who, during the preceding year, shall have conferred the greatest benefit on mankind.' One must wonder to what degree Nobel was tormented by the fact that his most famous invention, dynamite, was henceforth used to blow up countless people and places.

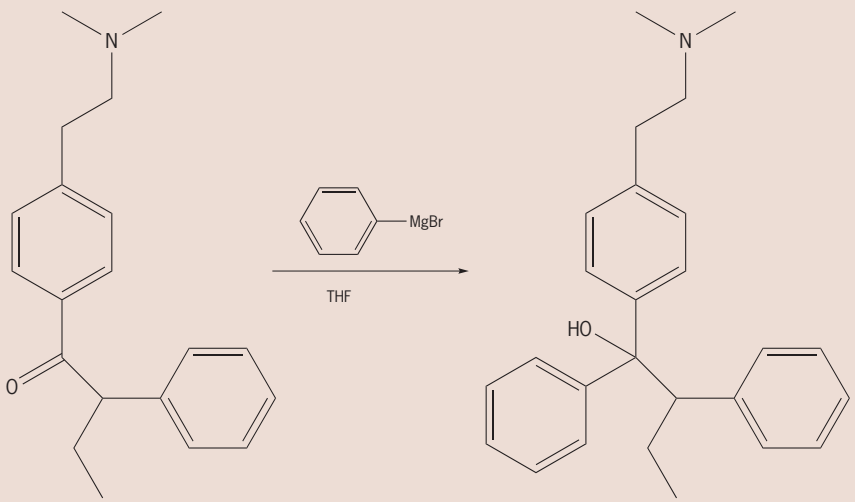
THEORY OF KNOWLEDGE

The conversion of ammonium cyanate into urea by Friedrich Wöhler in 1828 is considered to be the birth of organic synthesis. However, it was to be nearly 80 years before a shift in thinking would occur. In the early 1900s Victor Grignard discovered an organometallic reaction involving alkyl-magnesium halides that could be used to form new carbon–carbon bonds, allowing larger organic molecules to be synthesized much more simply, quickly, cheaply and with a higher yield. These revolutionary reagents were soon being used for a wide range of important applications, such as the synthesis of flavour enhancers, perfumes, pain killers and the production of tamoxifen, a drug used in the treatment of breast cancer, shown below.

The Grignard reaction was superior to any other known at the time, and greatly extended the limits of knowledge at the time, creating new opportunities for research and innovation. Victor Grignard shared the 1912 Nobel Prize in Chemistry for his discovery. By the 1970s French Chemist Yves Chauvin had made a further breakthrough by identifying the catalysts and step-by-step mechanism for a new reaction called metathesis. By the 1990s, effective catalysts for metathesis had been found by Americans Richard Schrock and Robert Grubbs.

Like the discovery of Grignard reagents, metathesis has been heralded as one of organic chemistry's most important new mechanisms, because it has further advanced our knowledge of how to assemble complex structures simply and quickly, creating opportunities for producing even more new molecules for use in medicine, agriculture, and the perfume, pigment and polymer industries. Chauvin, Schrock and Grubbs were awarded the Nobel Prize in 2005 for their work.

- Since the first synthesis of urea in 1928 the field of organic chemistry has grown considerably. What is the meaning of the phrase 'to grow'? Does the word growth predispose us to draw certain conclusions about how knowledge is acquired?
- What might be some of the different reasons for the growth of the field of organic synthesis in the last 40 years?
- What determines the validity of a reaction mechanism? What requirements are needed for the formation of the correct product?
- The short history of organic synthesis has yielded two Nobel prizes in Chemistry for discoveries of great benefit to society. What does this imply about the nature of this branch of chemistry?
- Comment on the claim that: 'Imagination will soon be the only limit to what molecules can be built!'

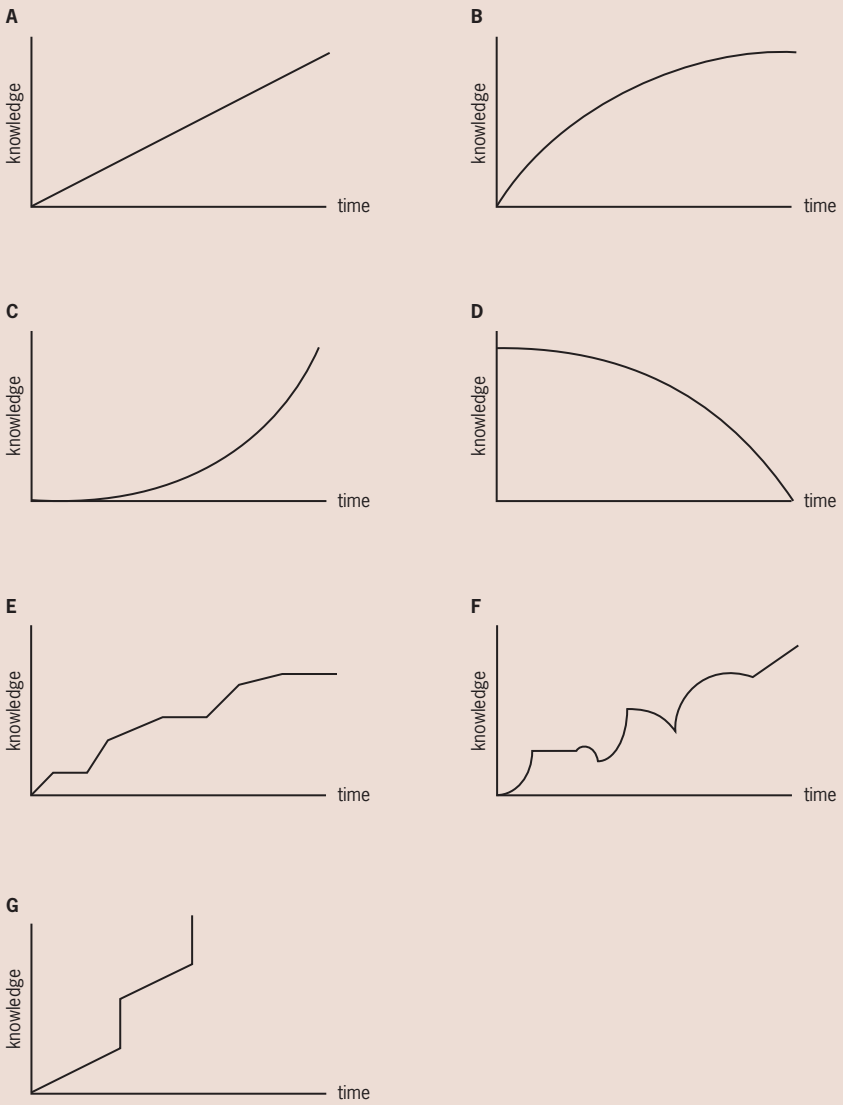


- In the 2005 Nobel Prize presentation speech, Professor Per Ahlberg of the Royal Swedish Academy of Sciences used a dance to help members of the audience visualize the abstract nature of the metathesis mechanism.

  The catalyst, in which the metal holds its carbon partner with both hands, 'dances' among carbon–carbon pairs, which are molecules with carbon–carbon double bonds. When the 'catalyst pair' meets a carbon–carbon pair, the two dancing pairs unite in a round dance. After a while they let go of each others hands, leave their old dance partners and dance along with their new ones. The new 'catalyst pair' is now ready to catch another dancing carbon–carbon pair for a new round dance.

  Does the use of symbols like a 'dance' increase our clarity of understanding in chemistry or do symbols just create ambiguity? Are symbols more suited to other areas of knowledge such as English or the Arts?

- Knowing what you know now about the growth of scientific knowledge, consider the sketch graphs below and decide which one most accurately reflects the way in which our knowledge of chemistry has accumulated with time. Be able to defend your choice and offer justification for why you rejected the others.



Section 7.4 Exercises

1 Write an equation for the preparation of a Grignard reagent using:

a chloroethane and magnesium

b 1-bromobutane and magnesium.

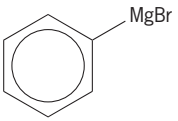
2 Why must water be excluded during the preparation of a Grignard reagent?

3 Draw the major product expected for each of the following reactions.

a CH_3CH_2MgBr + ethanal $\xrightarrow{H_2O}$

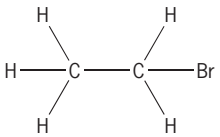
b $CH_3(CH_2)_3MgBr$ + propanone $\xrightarrow{H_2O}$

c $CH_3(CH_2)_2CH_2Br \xrightarrow[\text{ether}]{Mg} \xrightarrow[H_2O]{CO_2}$

d 

+ 3-methylbutanal $\xrightarrow{H_2O}$

4 The equations below show conversions of one organic compound to another in two-step processes. Complete the equations by adding any missing structural formulas and names of compounds.

a  $\xrightarrow[\text{ether}]{Mg}$ **ii** $\xrightarrow{\text{ethanal}}$ **iii**

i ethylmagnesium bromide **iv**

b **i** $\xrightarrow[\text{ether}]{Mg}$ **ii** $\xrightarrow{\textbf{iv}}$ **v**

1-chloropropane **iii** butanoic acid

5 Describe one useful synthetic feature of the use of Grignard reagents.

AS **G.7.1 Deduce reaction pathways given the starting materials and the product.**

7.5 REACTION PATHWAYS

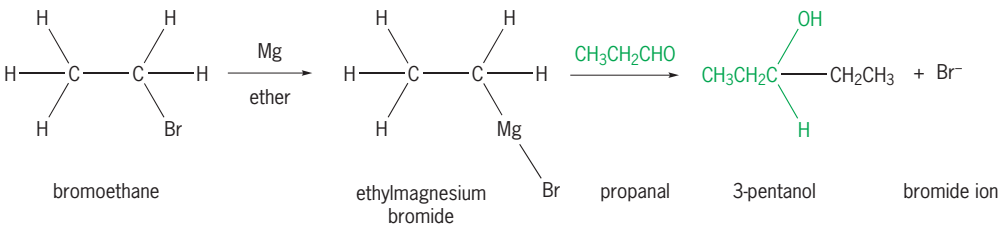
To summarize the reactions discussed so far in this chapter we may construct a 'web' of reaction pathways, as shown in the figure 7.5.1. You should be able to apply these reactions to hydrocarbons with up to six carbon atoms in their chain.

Worked example

Show a reaction pathway to produce pentan-3-ol using propanal and bromoethane.

Solution

Notice that the conversion involves an increase in chain length by two carbon atoms (from prop = 3 to pent = 5). Recall that increasing the chain length can be done using a Grignard reagent. A suitable pathway is shown below.



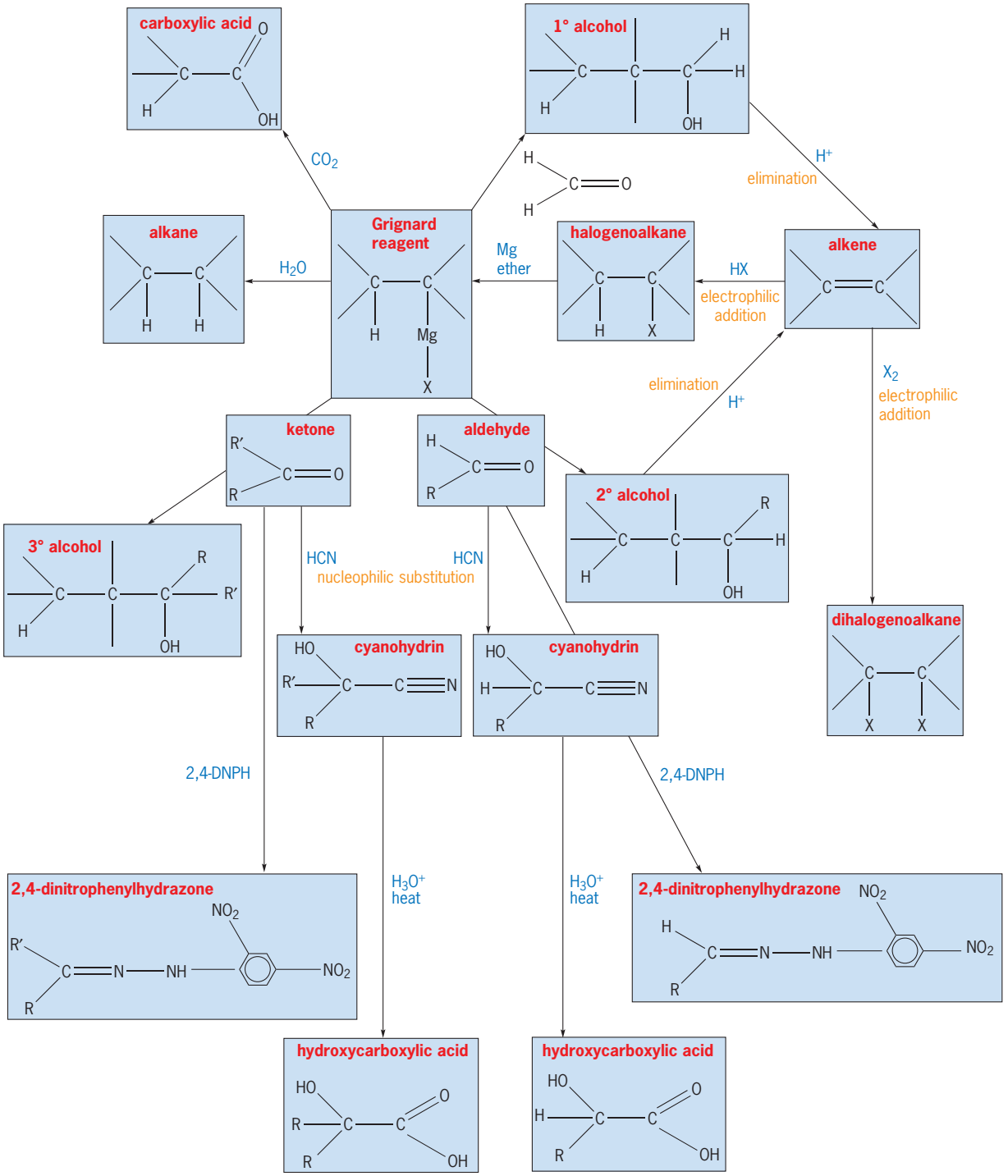


Figure 7.5.1 A 'web' of reaction pathways

In practice, many organic syntheses do not proceed by a simple, short pathway. Often many steps are involved. In developing a multi-step organic synthesis, reaction mechanisms are a powerful tool. They aid in understanding how known reactions proceed, and provide a guide to the conditions that will favour the formation of a desired product.

Section 7.5 Exercises

Draw possible two-step reaction pathways for the synthesis of the following compounds from the given starting compounds. Include structural formulas, and any inorganic reagents and conditions needed for the pathway.

a 2, 2-Dimethylpropanoic acid from 2-bromo-2-methylpropane

b Propanoic acid from bromoethane

c 2-Hydroxy-2-methylpentanoic acid from pentan-2-one

d Butanone from but-1-ene

e 3-Ethyl-3-hydroxyhexane from pentan-3-one and 1-bromopropane

f Ethane from bromoethane

7.6 ARENES

AS **G.5.1**
Describe and explain structure of benzene using physical and chemical evidence. © IBO 2007

Compounds containing the benzene ring are known as **arenes**, with benzene itself being the simplest arene. We were introduced to the benzene ring structure in the Standard Level course, and here we will look more closely at this unique structure and the types of physical and chemical evidence on which our description of the structure is based.

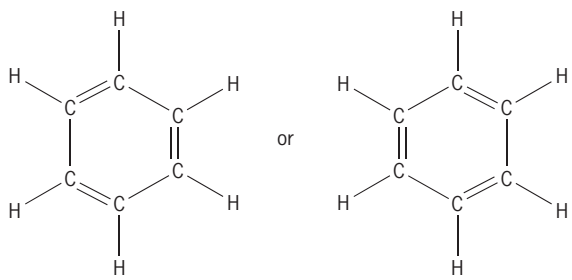


Figure 7.6.1 An early proposed benzene structure-1,3,5-cyclohexatriene.

Benzene (C_6H_6) has been known since the early 1800s, but its structure was not satisfactorily explained until nearly 100 years later. The difficulty was not in the complexity of the structure, but in the limitations of the bonding models. Proposed structures had several double or triple bonds. The problem with these structures was that benzene did not behave chemically as expected of a substance with multiple bonds in its molecules. In 1865, August Kekulé proposed that the molecule might have a ring structure, 1,3,5-cyclohexatriene.

Confusion arose when the products of the reaction of benzene with bromine were examined. Benzene reacts slowly with bromine in the presence of iron to give the substitution product C_6H_5Br, not the expected addition product, $C_6H_6Br_2$. When C_6H_5Br is reacted with more bromine, three isomeric $C_6H_4Br_2$ compounds are formed, not four as would be expected for reaction of 1,3,5-cyclohexatriene.

Kekulé suggested that the failure to isolate the fourth isomer was accounted for by the rapid interconversion of the two 1,2-dibromocyclohexatriene molecules. Kekulé's proposed structure was widely criticized at the time. It failed to account for the lack of reactivity of benzene compared to other alkenes, and the occurrence of a substitution product rather than an addition product.

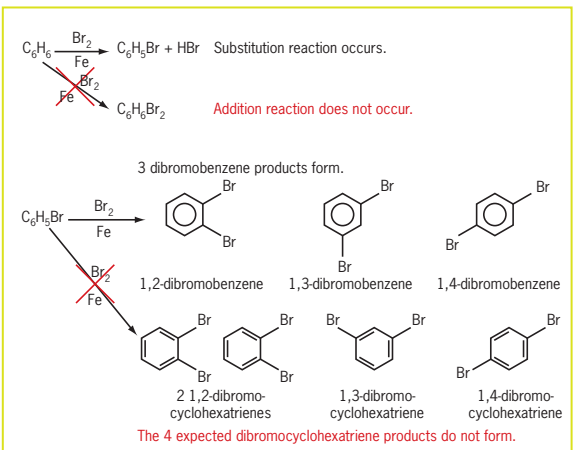


Figure 7.6.2 The reactions of benzene do not match those expected of 1,3,5-cyclohexatriene.

CHEM COMPLEMENT

Dreaming of molecules

In 1858, August Kekulé (of the University of Bonn) had proposed that the key to the diversity of carbon compounds was that carbon could form chains. This answer was said to have come to him while dozing during a London bus ride.

> I fell into a reverie and lo! the atoms were gambolling before my eyes … I saw how, frequently, two smaller atoms united to form a pair, how a larger one embraced the two smaller ones; how still larger ones kept hold of three or even four of the smaller ones; whilst the whole kept whirling in a giddy dance. I saw how the larger ones formed a chain … I spent part of the night putting on paper at least sketches of these dream forms. (A. Kekulé)

In 1865, Kekulé offered an answer to the question of the structure of benzene. He proposed that the benzene molecule might have a ring structure. This answer was said to have come to him while dozing by the fire.

> I was sitting at my textbook, but the work did not progress; my thoughts were elsewhere. I turned my chair to the fire, and dozed. Again the atoms were gambolling before my eyes. My mental eye, rendered more acute by repeated visions of this kind, could now distinguish larger structures of manifold conformations; long rows, sometimes more closely fitted together; twisting and turning in snake-like motion. But look! What was that? One of the snakes seized hold of its own tail, and the form whirled mockingly before my eyes. As if by a flash of lightning I woke; … I spent the rest of the night working out the consequences of my hypothesis. Let us learn to dream gentlemen [*and ladies*] and then perhaps we shall learn the truth. (A. Kekulé)

The unexpected lack of reactivity of benzene was evident in a comparison of the reactions of cyclohexane, cyclohexene and benzene (table 7.6.1). A further measure of the stability of benzene was seen in a comparison of heats of hydrogenation (table 7.6.2). If the structure of benzene is that of 1,3,5-cyclohexatriene, the expected heat of hydrogenation would be approximately three times that of cyclohexene. The actual value is considerably less than this, indicating that benzene has considerable 'extra' stability.

TABLE 7.6.1 SOME REACTIONS OF CYCLOHEXANE, CYCLOHEXENE AND BENZENE

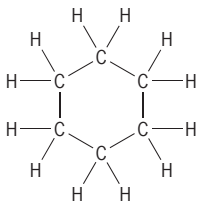
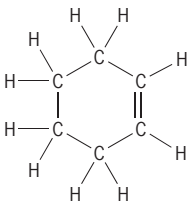
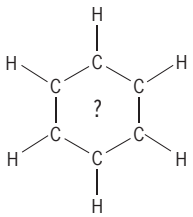
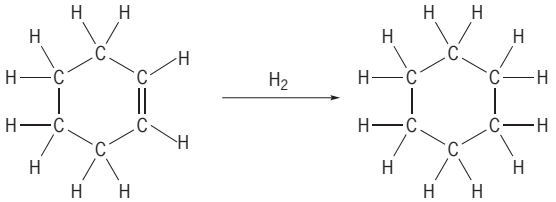
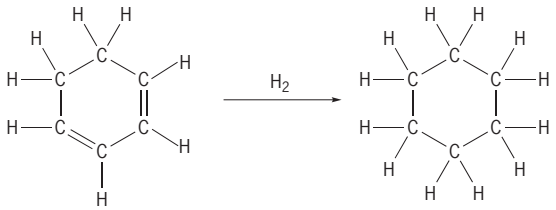
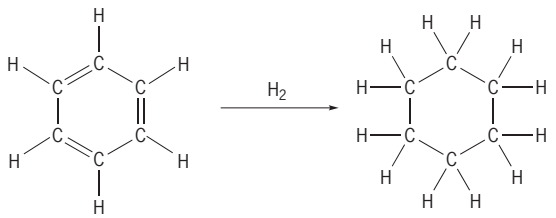
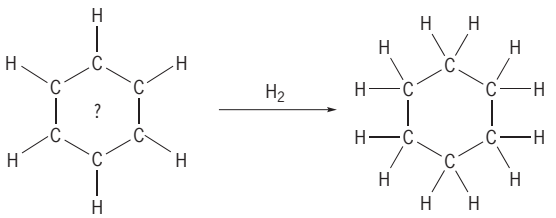
Reagent	Cyclohexane	Cyclohexene	Benzene
			
Br_2/CCl_4 (in dark)	No reaction.	Br_2 is decolourized readily.	No reaction with Br_2 alone. In the presence of Fe filings, Br_2 is decolourized slowly and HBr is evolved.
$KMnO_4/H^+$	No reaction.	MnO_4^- is decolourized readily.	No reaction.
H_2/Ni	No reaction.	One mole absorbs one mole of H_2 at room temperature.	One mole absorbs 3 moles of H_2 slowly at 150°C.

TABLE 7.6.2 HEATS OF HYDROGENATION OF CYCLIC ALKENES AND BENZENE

Reaction	$\Delta H_{hydrogenation}$ (kJ mol^{-1})
cyclohexene → cyclohexane 	–118
1,3-cyclohexadiene → cyclohexane 	–230
cyclohexatriene → cyclohexane 	Approximate expected value $-118 \times 3 = -336$
benzene → cyclohexane 	–206

It was later found that all the carbon–carbon bonds in benzene are the same length, and intermediate (0.139 nm) between that of a single (0.154 nm) and a double (0.134 nm) carbon–carbon bond. To account for these unusual properties, and to best represent benzene's structure, we use the concept of **resonance structures**.

The benzene molecule is seen as a hybrid of two equivalent structures in which the carbon–carbon bonds lie between single and double bonds. It is important to note here that benzene has a single, unchanging structure that combines the characteristics of both resonance forms. This resonance adds stability to the structure. We cannot easily represent the structure with our usual conventions so it is sometimes drawn with a ring or dotted line to represent the equivalence of the carbon–carbon bonds.

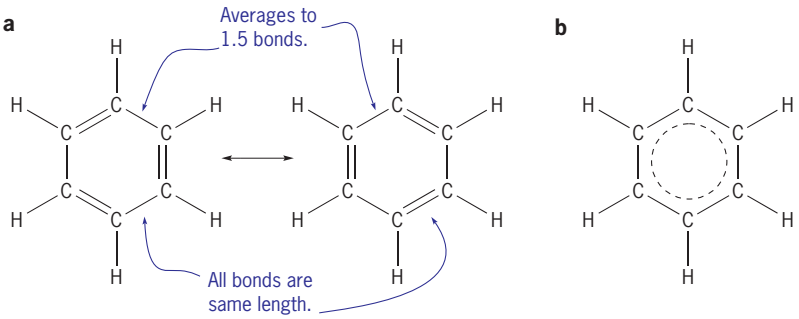


Figure 7.6.3 (a) The structure of benzene is a hybrid of two resonance forms. (b) An alternative representation of the structure of benzene.

The molecular orbital description of benzene helps to see the equivalence of the bonds. Benzene is a planar molecule with a hexagonal shape. All bond angles are 120°. The six carbon atoms are sp^2 hybridized, and each carbon has a p orbital perpendicular to the plane of the ring structure. These p orbitals overlap with each other so that the six π electrons are completely delocalized around the ring. There are therefore two donut-shaped clouds of electrons, one above and one below the ring.

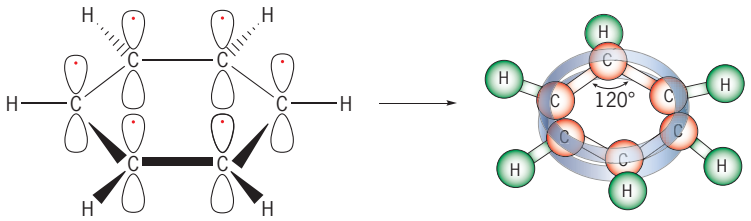


Figure 7.6.4 The molecular orbital representation of the structure of benzene.

This delocalization of electrons above and below the plane of the benzene ring confers considerable stabilization, so that it generally involves a considerable amount of energy to activate the ring before any reaction can occur. As a result of this stabilization, the benzene ring requires higher temperatures and usually catalysis before it can react. As we have seen, even if it can react under such conditions, its reaction rate is still much slower than that of alkenes. Unlike the alkenes, benzene undergoes **electrophilic substitution** rather than electrophilic addition. The substitution reactions (to be considered in section 7.9) allow the six delocalized electrons of benzene to be regenerated after attack, so that the resonance stabilization of the ring is preserved.

To highlight this difference in reactivity due to the delocalization of electrons around the ring, we will compare the relative rates of reaction of halogenated arenes, in which the halogen is attached directly to the ring, with those in which it is present in a side chain. First consider the hydrolysis reaction of bromomethylbenzene shown in figure 7.6.5.

AS G.5.2
Describe and explain the relative rates of hydrolysis of benzene compounds halogenated in the ring and in the side-chain. © IBO 2007

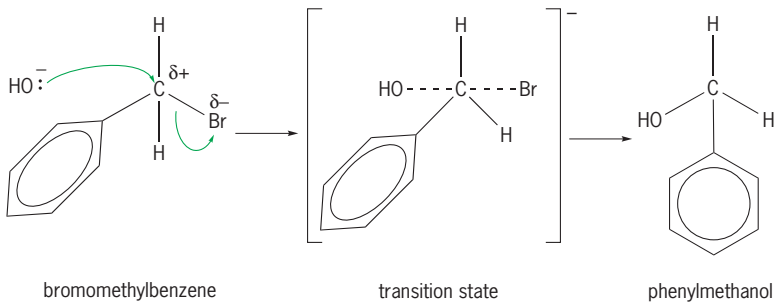


Figure 7.6.5 Hydrolysis reaction of bromomethylbenzene by S_N2 reaction.

Bromomethylbenzene undergoes a typical nucleophilic substitution reaction, S_N2 (as considered previously in chapter 11 of *Chemistry: For use with the IB Diploma Programme Standard Level* and chapter 9 of *Higher Level*). An aqueous hydroxide ion acts as the nucleophile, the bromide ion is the leaving group, and the product is phenylmethanol.

Now consider the hydrolysis reaction of bromobenzene shown in figure 7.6.6. Here the bromine is directly bonded to the ring. Nucleophilic substitution is severely hindered because the nucleophile (OH^-) is repelled by the electron cloud surrounding the benzene ring. In addition, the C–Br bond is strengthened by the interaction of the non-bonding electrons on the bromine with the delocalized electrons of the benzene ring. This is evident in a shortening and strengthening of the C–Br bond (compared with that in bromomethylbenzene). The bromide ion is therefore a poorer leaving group in bromobenzene, and nucleophilic substitution will occur very slowly, if at all. In fact, as we will see later (section 7.9), bromobenzene typically undergoes electrophilic substitution reactions, rather than nucleophilic substitution.

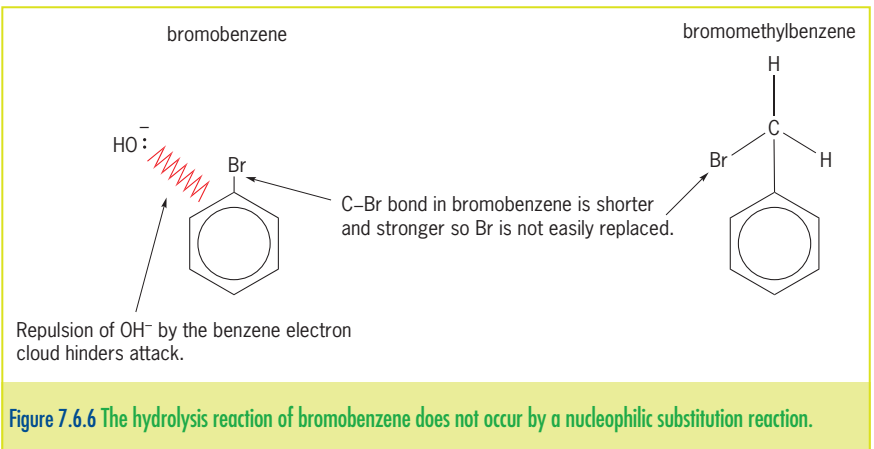


Figure 7.6.6 The hydrolysis reaction of bromobenzene does not occur by a nucleophilic substitution reaction.

Section 7.6 Exercises

1 In terms of structure and bonding, explain why the benzene molecule is planar, whereas cyclohexane is non-planar.

2 Outline two types of physical evidence that support the model of benzene as a hybrid of two resonance structures.

3 Outline two types of chemical evidence that support the model of benzene as a hybrid of two resonance structures.

4 **a** Write semi-structural formulas for bromobenzene and bromomethylbenzene.

b Compare the C–Br bond in bromobenzene and bromomethylbenzene. Which bond is:

i longer?

ii stronger?

c Account for the differences in the bond lengths and strengths stated in part **b**.

5 $C_6H_5CH_2Cl$ and C_6H_5Cl are warmed separately with aqueous sodium hydroxide. Which compound would react more slowly? Explain your choice.

7.7 ACID–BASE REACTIONS

The strength of an acid, HA, is a measure of its extent of hydrolysis; the greater the strength, the greater the extent of the hydrolysis. Acid strengths are compared quantitatively by considering the following equilibrium:

$$H_2O + HA \rightleftharpoons H_3O^+ + A^-$$

$$K_a = \frac{[H_3O^+][A^-]}{[HA]}$$ where K_a is the acidity constant.

For weak acids, the values of K_a are very small, and so pK_a is used, where

$$pK_a = -\log_{10} K_a.$$

Note that the smaller the value of pK_a, the greater is the strength of that acid. Factors that may affect the acidity of an organic acid, HA, include the electronegativity of A, the strength of the H–A bond, and any features that stabilize the conjugate base A^- with respect to HA.

AS G.8.1
Describe and explain the acidic properties of phenol and substituted phenols in terms of bonding. © IBO 2007

AS G.8.2
Describe and explain the acidic properties of substituted carboxylic acids in terms of bonding. © IBO 2007

To illustrate these factors we will compare the acid strengths of methanol, CH_3OH ($pK_a = 15.5$) and methane, CH_4 ($pK_a = 50$). Methanol, though a very weak acid, is still considerably more acidic than methane. The oxygen atom is more electronegative than the carbon atom, so the O–H bond in methanol breaks more readily than the C–H bond in methane. In addition, the conjugate base, CH_3O^- is more stable than CH_3^-.

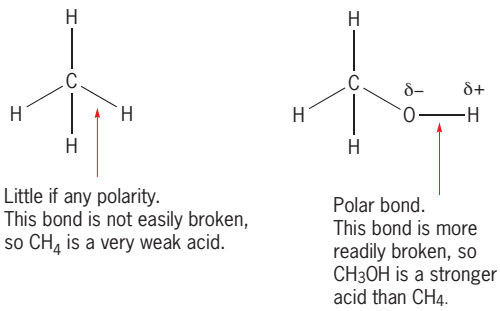


Figure 7.7.1 Methanol is more acidic than methane.

Now we compare the acid strengths of methanol, CH_3OH ($pK_a = 15.5$), and methanoic acid, HCOOH ($pK_a = 3.75$). Methanoic acid is far more acidic than methanol. This is in part due to the electron-withdrawing carbonyl group. This electron withdrawal weakens the O–H bond, so the O–H bond in methanoic acid breaks more readily than the O–H bond in methanol. In addition, the conjugate base, $HCOO^-$, is stabilized by resonance structures. In the case of methanol there is no such resonance stabilizing the conjugate base, the alkoxide anion, CH_3O^-. Alcohols in general are therefore very much less acidic than carboxylic acids.

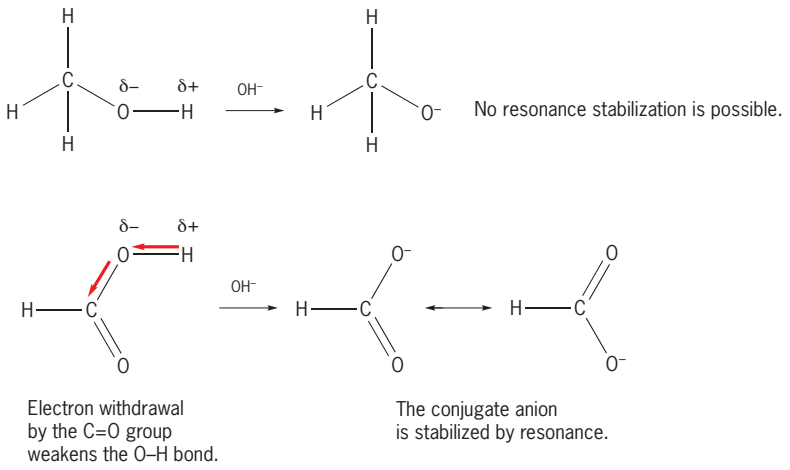


Figure 7.7.2 Methanoic acid is more acidic than methanol.

CHEM COMPLEMENT

Phenols in the headlines

> 'Nalgene to stop making bottles with BPA. Hard-plastic Nalgene water bottles made with bisphenol A (BPA) will be pulled from stores over the next few months because of growing customer concern over whether the chemical poses a health risk' (*Sydney Morning Herald*, April 2008)

What is BPA, and why the customer concerns? Bisphenol A (BPA) is an organic compound containing two phenol functional groups. It is used as a starting material for the production of polycarbonate plastics and synthetic resins.

Polycarbonate plastic is used to make a variety of common products, including containers that carry foods such as drink bottles, baby bottles and the lining of tins used for tinned food. While products containing BPA have been used for some 50 years, the use of BPA has become controversial.

BPA belongs to a group of compounds that can act in a similar way to some hormones, and so may disrupt hormone functioning. Some studies in laboratory animals suggest that low levels of consumed BPA may have an effect on the reproductive system. Given that in some circumstances chemicals in food containers may move into the foods they contain, there are concerns over the use of BPA. Some have responded with bans on the use of BPA in products for infants.

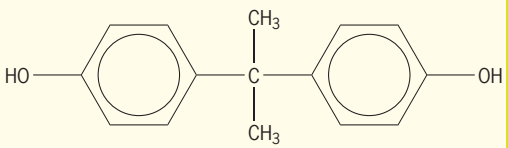


Figure 7.7.5 Bisphenol A (BPA). Its use in food containers has become controversial.

The substituted arene, phenol, is a benzene ring with an attached –OH group. Phenols are mildly acidic. When comparing the acid strength of phenols and alcohols, the stability of the conjugate base is again critical. The conjugate base of phenol, the **phenoxide ion**, $C_6H_5O^-$, is stabilized by the delocalization of the negative charge by interaction with the π orbitals of the benzene ring. As a result, phenols ($pK_a = 0.42–10.00$) are stronger acids than alcohols ($pK_a > 15.5$).

Phenols are, however, weaker acids than carboxylic acids. This can also be explained in terms of the stabilization of the conjugate anion. Delocalization of the negative charge in the carboxylate anion, COO^-, involves resonance structures of identical energy content, and involves two highly electronegative oxygen atoms (see figure 7.7.2). In the phenoxide anion, however, the resonance structures are not all of equal energy. Those that involve a negative charge on a carbon atom are of higher energy, and are therefore less stable than the form that has the negative charge on the oxygen atom (see figure 7.7.3).

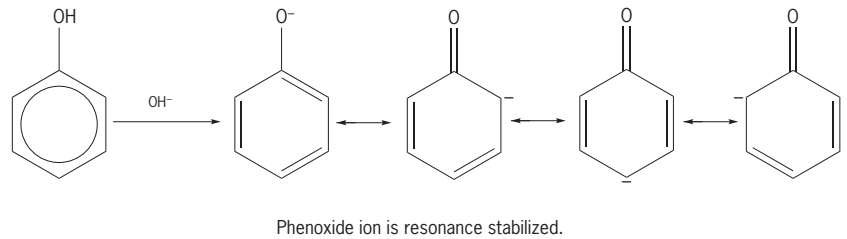


Figure 7.7.3 Resonance stabilization of the conjugate base leads to greater acidity for phenols than alcohols.

As a result of the three factors affecting the acidity of an organic acid, HA—electronegativity of A, the strength of the H–A bond and the stabilization of the conjugate anion A^-—the general order of acidity is alcohols less than phenols, which in turn are less than carboxylic acids.

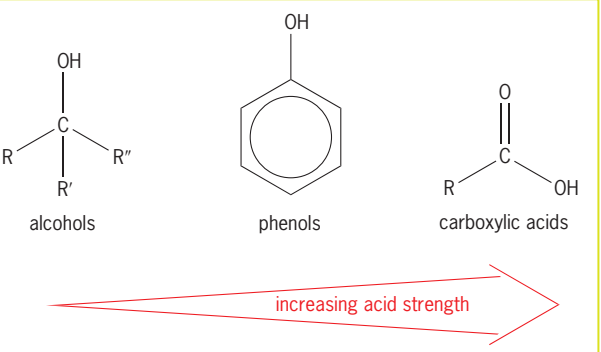


Figure 7.7.4 Relative acid strengths of alcohols, phenols and carboxylic acids.

The effects of substituents on acidity

The ease of breaking of the acidic H–A bond, and the stability of the conjugate anion, are both affected by substituent groups. Any substituent groups that can stabilize the resulting anion will increase the acidity of HA. Conversely, any groups that destabilize the anion will decrease the acidity of HA. Generally, electron-withdrawing substituents increase the acidity, whereas electron-donating substituents decrease the acidity of the molecule.

We have seen that alcohols are very weak acids. The conjugate base of an alcohol is an alkoxide ion. The more stable the alkoxide ion, the greater will be the acidity of HA. Electron-withdrawing groups will increase acidity as they help to disperse the negative charge on the alkoxide ion. The more electron-withdrawing groups present, the greater its acidity. Electron-donating groups tend to intensify the negative charge on the alkoxide ion, making it more unstable. Recall that alkyl groups tend to be electron donating, so that the more alkyl groups that are attached to the acid molecule, the weaker its acidity will be. The acid strength of primary alcohols is therefore greater than that of secondary alcohols, which in turn is greater than that of tertiary alcohols.

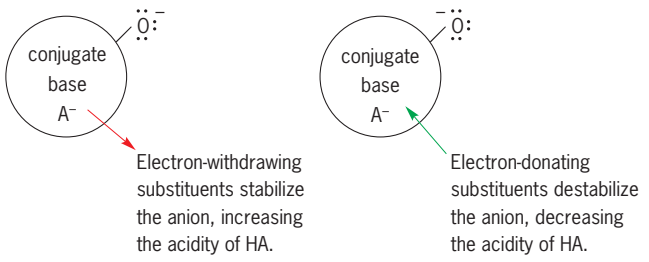


Figure 7.7.6 The acidity of the acid HA is influenced by its substituents.

As for alcohols, the presence of electron-withdrawing groups (such as chlorine and the nitro group) on the ring structure of phenol will help to disperse the negative charge and thus stabilize the conjugate anion, phenoxide. As a result, acidity will be higher. Electron-donating groups (such as alkyl groups) on the ring structure will intensify the negative charge and destabilize the phenoxide anion, thus lowering the acidity of the substituted phenol. The more substituents there are, the greater the effect on the acidity will be.

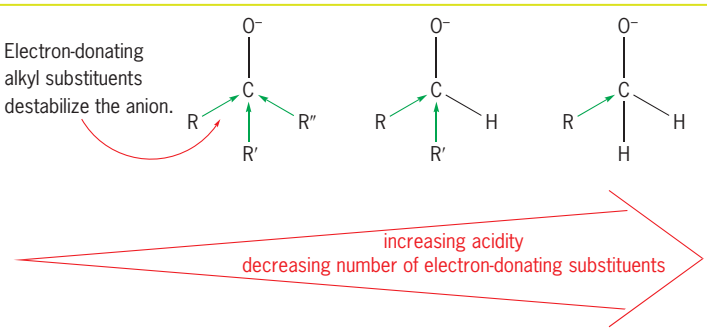


Figure 7.7.7 Primary alcohols are more acidic than tertiary alcohols.

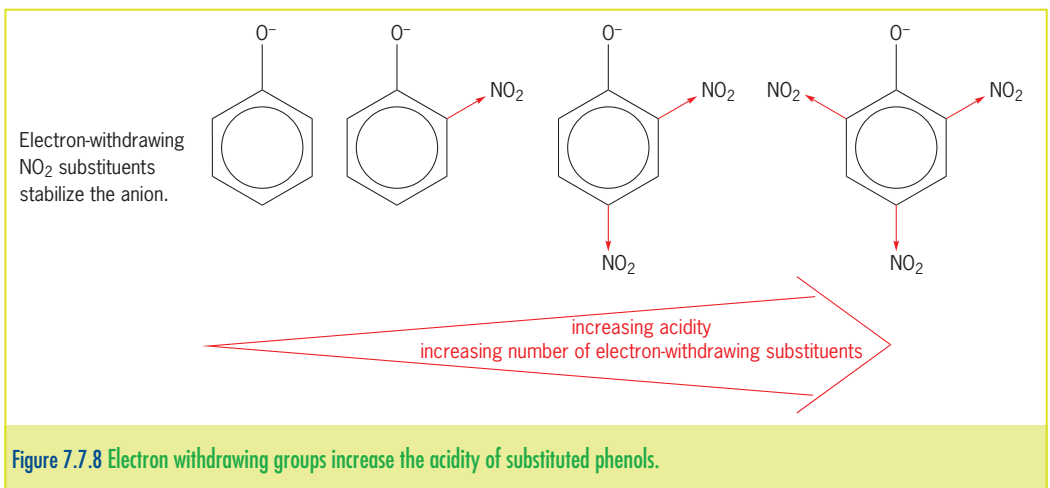


Figure 7.7.8 Electron withdrawing groups increase the acidity of substituted phenols.

Electron-donating substituents in a carboxylic acid will also decrease acid strength. The electron-donating group pushes electrons towards the carbonyl carbon in the acid molecule, thus reducing its charge. This in turn reduces the polarity of the O–H bond, strengthening the bond and consequently it will not break easily. In addition, electron-donating substituents will push electrons towards the COO^- group of the conjugate anion, further intensifying the negative charge and so further destabilizing the anion. These effects may be seen in the differing pK_a values for acids with electron-donating alkyl groups. Thus methanoic acid ($pK_a = 3.75$), with only a hydrogen atom attached to the COOH group is more acidic than ethanoic acid ($pK_a = 4.76$), with an electron-donating CH_3 group attached.

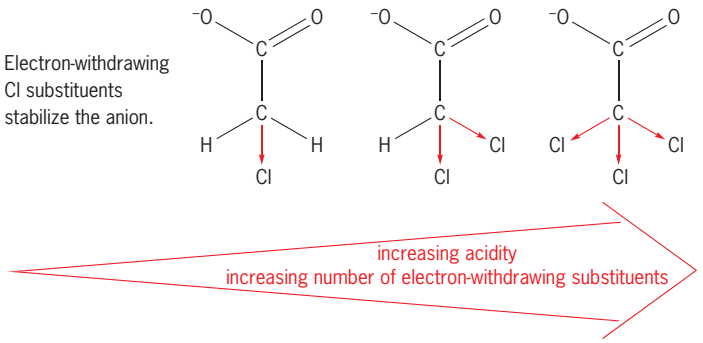
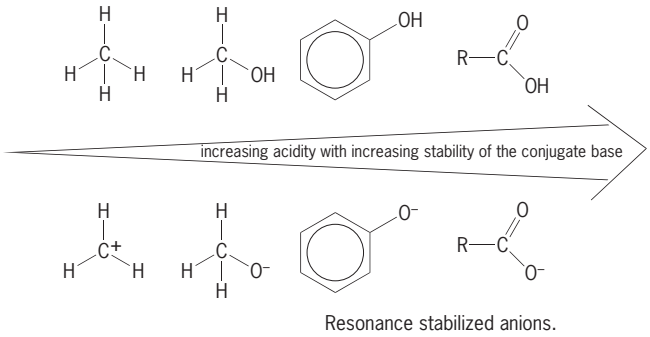
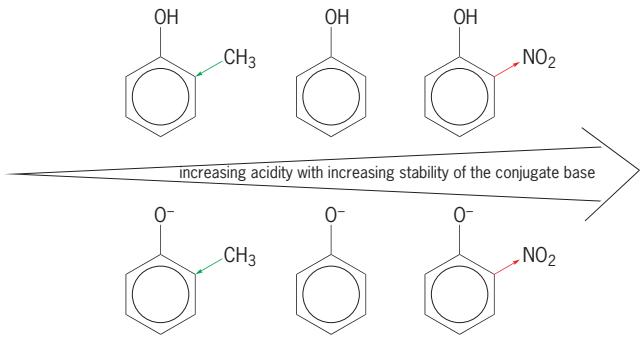
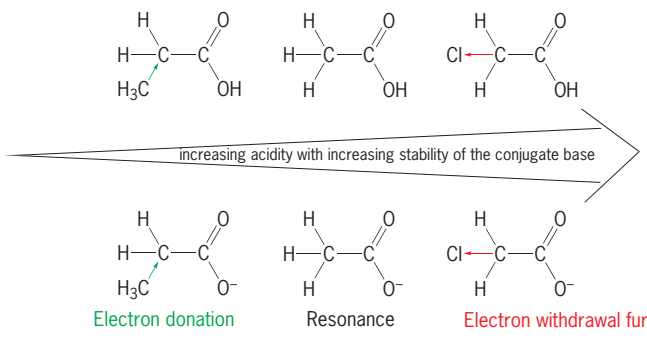


Figure 7.7.9 Electron-withdrawing groups increase the acidity of substituted carboxylic acids.

Electron-withdrawing groups, such as the chlorine atom, increase the acidity, as they help to weaken the O–H bond and stabilize the resulting COO^- anion by dispersing its negative charge. Inductive effects are additive; the more electron-withdrawing groups attached, the stronger the acid. Thus chloroethanoic acid is a stronger acid than ethanoic acid, and trichloroethanoic acid is much stronger again. The inductive effect of substituents on acidity decreases rapidly when the substituents are placed further away from the carbonyl group. For example, $CH_3CH_2CHClCOOH$ ($pK_a = 2.9$) is a stronger acid than $CH_3CHClCH_2COOH$ ($pK_a = 4.1$).

TABLE 7.7.1 A SUMMARY OF THE COMPARISON OF ACIDIC PROPERTIES OF ALCOHOLS, PHENOLS, CARBOXYLIC ACIDS AND THEIR SUBSTITUTED DERIVATIVES.

Type of compound	Comparison of acidic properties
Alkanes, alcohols, phenols and carboxylic acids	 Resonance stabilized anions.
Substituted phenols	 Electron donation destabilizes the anion. Resonance stabilized anion. Electron withdrawal further stabilizes the anion.
Substituted carboxylic acids	 Electron donation destabilizes the anion. Resonance stabilized anion. Electron withdrawal further stabilizes the anion.

Base strength

AS **G.8.3**
Compare and explain the relative basicities of ammonia and amines. © IBO 2007

The strength of a base is a measure of its extent of hydrolysis; the stronger the base, the greater the extent of hydrolysis. Base strengths are compared quantitatively by considering the following equilibrium:

$$H_2O + B: \rightleftharpoons BH^+ + OH^-$$

$$K_b = \frac{[BH^+][OH^-]}{[B:]}$$ where K_b is the basicity constant.

For weak bases, the values of K_b are very small, and so pK_b is used, where

$$pK_b = -\log_{10} K_b.$$

Note that the smaller the value of pK_b, the greater is the strength of the base.

The base strength of amines in water depends on the availability of the electron pair for donation to a proton. This availability will be increased by any electron-donating substituents, since these substituents push electrons towards the nitrogen and make it denser in negative charge. Recalling that alkyl groups are electron donating, it would be expected that the order of basicity for amines would therefore be NH_3 less than primary amine, less than secondary amine, less than tertiary amine. If we consider the four bases and their pK_b values shown in figure 7.7.10, we see that the predicted order is not entirely correct. The tertiary amine is a weaker base than expected.

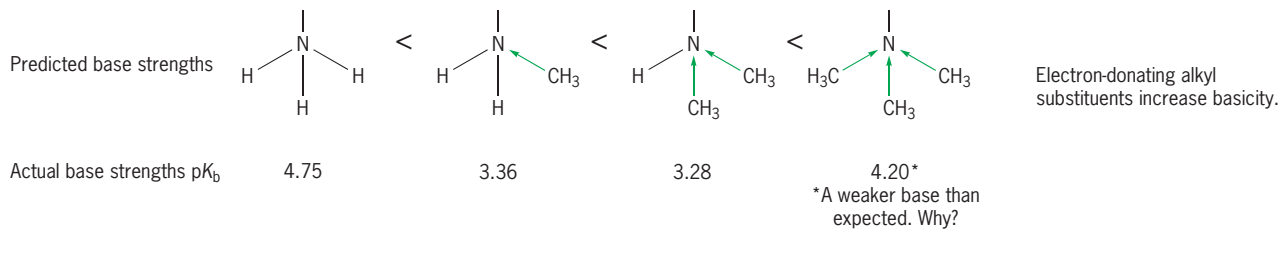


Figure 7.7.10 Predicted order of base strengths, and pK_b values. Tertiary bases are weaker than expected.

This unexpected order can be explained by taking another factor into account: the ease with which the protonated amine can be solvated by water molecules. Solvation with water molecules by the formation of intermolecular hydrogen bonds reduces the charge density on the nitrogen atom of the protonated amine, thus stabilizing the cation. Solvation with water molecules depends on the number of hydrogen atoms on the nitrogen that can form intermolecular hydrogen bonds with water molecules. The protonated secondary amine is more stable than the protonated tertiary amine, as it can form more intermolecular hydrogen bonds with water.

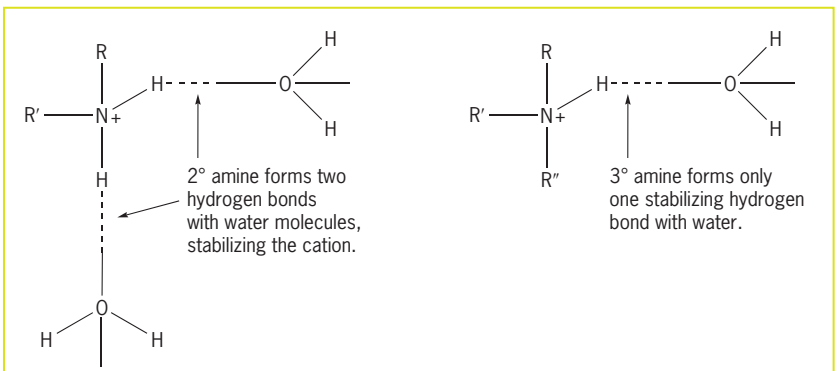


Figure 7.7.11 The conjugate acid of a base is stabilized by solvation with water molecules.

WORKSHEET 7.3
Organic acid–base reactions

The overall effect of both of the factors, the electron-releasing groups producing a greater charge density on the nitrogen atom and the stabilizing effect of solvation of the protonated amine product, accounts for the observed order of basicity: NH_3 is less basic than a tertiary amine, which is less basic than a primary amine, which in turn is less basic than secondary amine.

When the amine group is attached to a benzene ring, the non-bonding pair of electrons on the nitrogen is delocalized by interaction with the π orbitals of the ring. The non-bonding pair on nitrogen is thus less readily available for coordination with protons, and so aniline ($C_6H_5NH_2$) would be expected to be less basic than other amines. The high pK_b value of 9.38 confirms this.

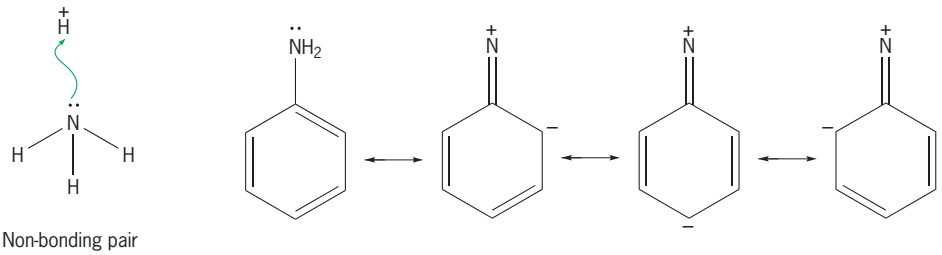


Figure 7.7.12 Interaction of the non-bonding pair on the nitrogen atom with the ring electrons reduces the base strength of $C_6H_5NH_2$.

Note that when amines act as bases they form ammonium salts. These salts will react with warm aqueous sodium hydroxide to release the amines. For example, methylamine is released when methylammonium chloride is reacted with sodium hydroxide.

$$CH_3NH_3Cl(s) + NaOH(aq) \rightarrow CH_3NH_2(aq) + NaCl(aq)$$

Section 7.7 Exercises

1 **a** State how the strength of an acid is related to the K_a of the acid, and to its pK_a value.

b State how the strength of a base is related to the K_b of the base, and to its pK_b value.

2 By referring to their structures, explain the difference in the acid strengths of:

a ethanol and phenol

b ethanol and ethanoic acid

c ethanoic acid and propanoic acid

d chloroethanoic acid and dichloroethanoic acid.

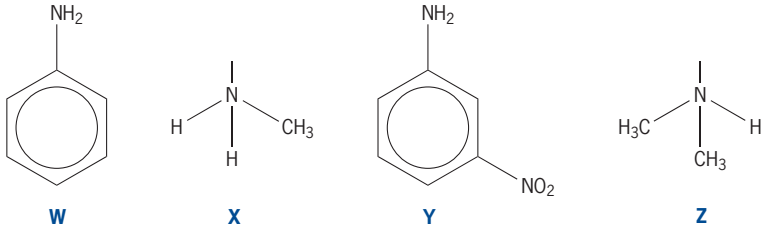
3 By referring to their structures, explain why:

a aminoethane is more basic than ammonia

b 2-aminopropane is more basic than 1-aminopropane

c 2,4,6-trinitrophenol is more acidic than phenol.

4 Consider the following molecules.

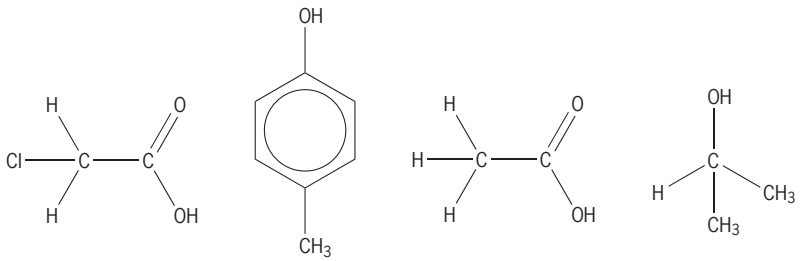


W X Y Z

a List molecules W to Z in order of increasing basicity, clearly stating which is the most basic.

b Explain your choices in part **a**.

5 Consider the following molecules.



P Q R S

a List molecules P to S in order of increasing acidity, clearly stating which is the most acidic.

b Explain your choices in part **a**.

7.8 ADDITION–ELIMINATION REACTIONS

HL

Acyl chlorides

Acyl chlorides, also known as acid chlorides, are derived from the COOH group of carboxylic acids by replacing the –OH group with the –Cl group. Examples include ethanoyl chloride, CH_3COCl, and butanoyl chloride, $CH_3CH_2CH_2COCl$. Acyl chlorides are among the most reactive of carboxylic acid derivatives, and can readily be converted to many other kinds of compounds. The chlorine atom readily undergoes substitution by other nucleophiles such as HO^-, RO^- and NH_2^- to produce acids, esters and amides respectively.

The reaction mechanisms are similar to the addition–elimination reaction of aldehydes and ketones discussed in section 7.3. Conversion to acids occurs by hydrolysis. Reaction involves the attack on the positively polarized carbonyl carbon by the nucleophile water. The tetrahedral

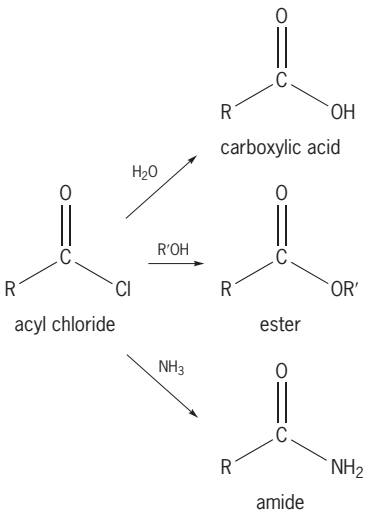


Figure 7.8.1 Acyl chlorides are readily converted to carboxylic acids, esters and amides.

AS **G.9.2**
Describe, using equations, the reactions of acyl chlorides with nucleophiles to form carboxylic acids, esters, amides and substituted amides. © IBO 2007

AS G.9.3
Explain the reactions of acyl chlorides with nucleophiles in terms of an addition–elimination mechanism.

intermediate then undergoes elimination of Cl^- and H^+ to give the final carboxylic acid product, and hydrogen chloride.

Recall that, as for the **addition–elimination** reactions of aldehydes and ketones, this reaction can be thought of as a condensation reaction in which two molecules combine with the elimination of a small molecule (in this case HCl). The hydrolysis reaction is faster in alkaline solution, where the stronger nucleophile OH^- reacts in place of water. The OH^- ion also serves to 'scavenge' the HCl produced and so prevent it from causing side reactions.

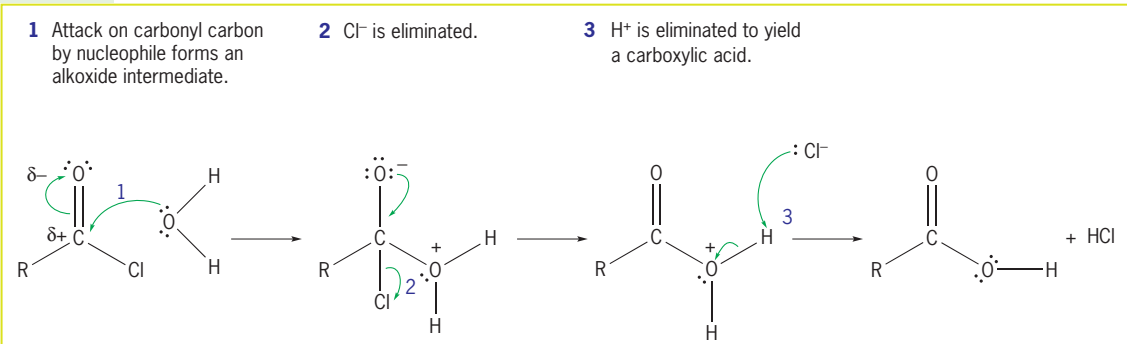


Figure 7.8.2 Reaction of an acyl chloride with water by the addition–elimination mechanism.

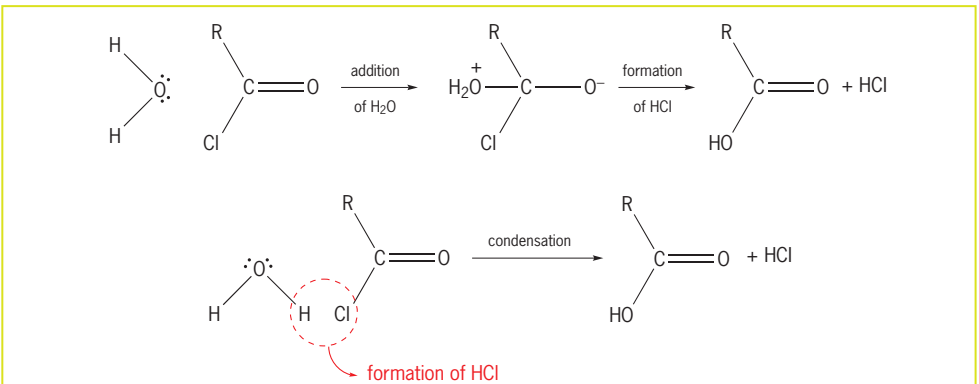


Figure 7.8.3 Addition–elimination reactions may be thought of more simply as condensation reactions.

Conversion of acyl chlorides to esters occurs by reaction with alcohols. Reaction is analogous to that with water, and involves the attack on the carbonyl carbon by the nucleophile ROH. The tetrahedral intermediate again undergoes elimination of Cl^- and H^+ to give the final ester product, and hydrogen chloride (figure 7.8.4)

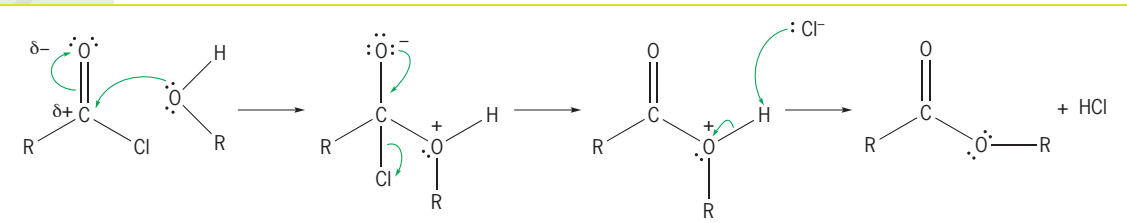


Figure 7.8.4 Reaction of an acyl chloride with an alcohol by the addition–elimination mechanism.

Conversion to amides occurs by reaction with ammonia or an amine. Reaction is analogous to that with water, and involves the attack on the carbonyl carbon by the nucleophile NH_3 or an amine. The tetrahedral intermediate again undergoes elimination of Cl^- and H^+ to give the final amide product. Reaction with a primary amine yields a monosubstituted amide, whereas secondary amines react to produce disubstituted amides. Since HCl is eliminated in these reactions, twice as many mole of ammonia or amine compared to the amount in mole of acyl chloride must be used. One mole reacts with the acyl chloride, and one mole reacts with the HCl by-product to form an ammonium chloride salt (see figure 7.8.5).

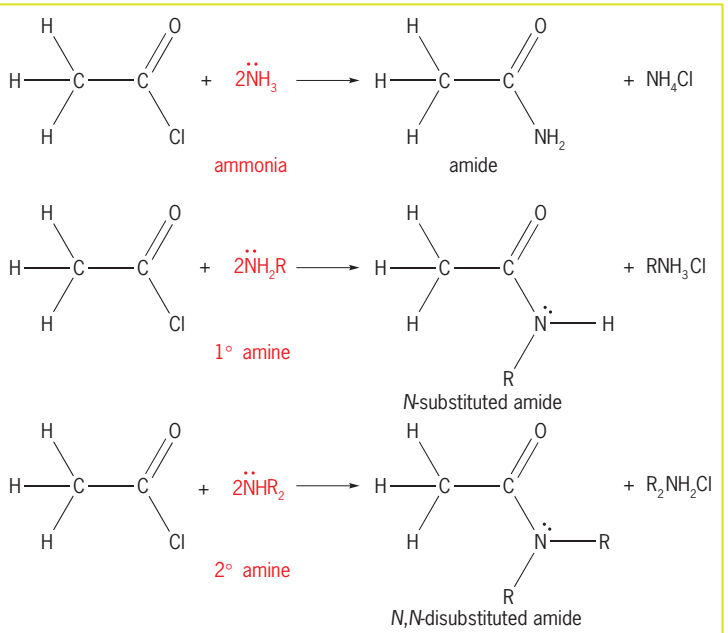


Figure 7.8.5 Reaction of acyl chlorides with ammonia or an amine produces amides by the addition–elimination mechanism.

Acid anhydrides

Acid **anhydrides** are derived from carboxylic acids by replacing the OH– group by a RCOO– group. Examples include ethanoic anhydride, $(CH_3CO)_2O$, and propanoic anhydride, $(CH_3CH_2CO)_2O$. The reactions of acid anhydrides are similar to those of acid chlorides, although anhydrides react more slowly. Thus acid anhydrides react with water to produce carboxylic acids, with alcohols to produce esters and with amines to produce amides. Examples of each of these reactions are shown in figure 7.8.6.

Notice that in each reaction only 'half' of the acid anhydride molecule is used. The other 'half' acts as a leaving group in the reaction mechanism, producing an acid by-product in each reaction (in the case of reaction with ammonia, the acid by-product forms an ammonium salt due to the basic conditions). Anhydrides are therefore inefficient to use, and acyl chlorides are more often used in the synthesis of acids, esters and amides in addition–elimination reactions.

AS **G.9.1**
Describe, using equations, the reactions of acid anhydrides with nucleophile to form carboxylic acids, esters, amides and substituted amides.

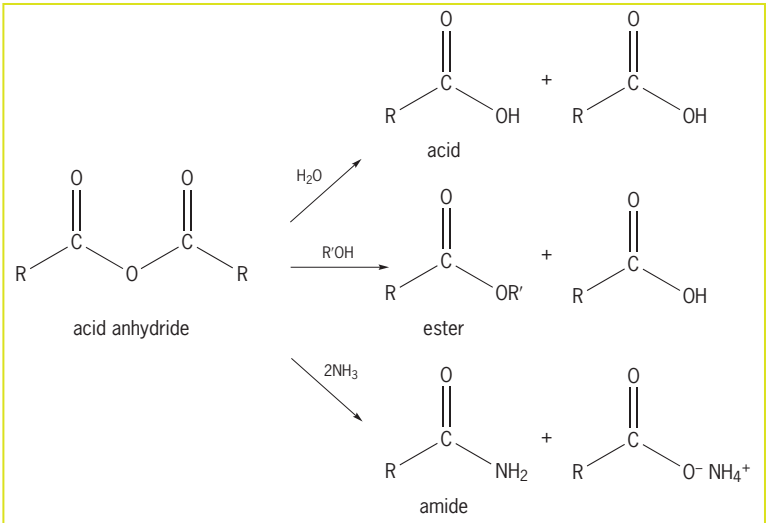


Figure 7.8.6 Acid–anhydrides react to produce carboxylic acids, esters and amides.

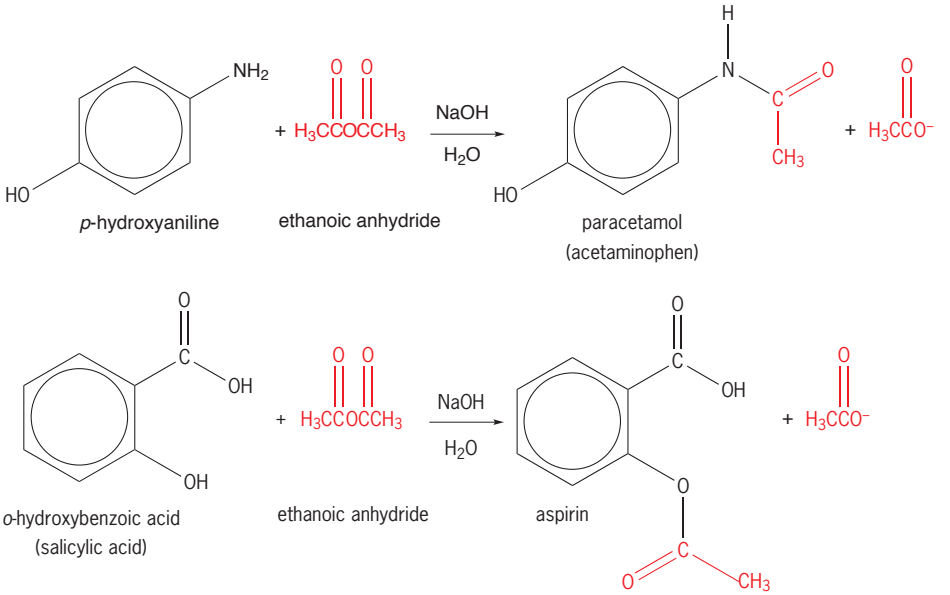


Figure 7.8.7 Ethanoic anhydride is used in the production of pain-killers and analgesics such as paracetamol (also called acetaminophen) and aspirin.

Section 7.8 Exercises

1 Draw and name an example of each of the following molecules.

a An acyl chloride containing five carbon atoms

b An acid anhydride containing four carbon atoms

c An ester containing five carbon atoms

d An amide containing six carbon atoms

e A substituted amide containing four carbon atoms

2 Draw a reaction mechanism for each of the following reactions.

a $CH_3C(=O)Cl + CH_3OH \longrightarrow$

b $CH_3CH_2C(=O)Cl + CH_3NH_2 \longrightarrow$

c $CH_3CH_2CH_2C(=O)Cl + H_2O \longrightarrow$

3 Explain, with the aid of a diagram, how the addition–elimination reaction of an acyl chloride to form an ester can be thought of as a condensation reaction.

4 Write balanced equations, using structural formulas, for:

a salicylic acid + $(CH_3CO)_2O$ (ethanoic anhydride) $\longrightarrow$

b $CH_3CH_2NH_2$ (ethanamine) + $(CH_3CH_2CO)_2O$ (propanoic anhydride) $\longrightarrow$

c $HN(CH_3)_2$ (*N*,*N*-dimethylamine) + $(CH_3CO)_2O$ (ethanoic anhydride) $\longrightarrow$

5 Explain why using reactions of acyl chlorides to produce esters and amides is preferred to reactions using acid anhydrides.

7.9 ELECTROPHILIC SUBSTITUTION REACTIONS

HL

We saw earlier that the delocalization of electrons above and below the plane of the benzene ring confers considerable stabilization, so that it generally involves a considerable amount of energy to activate the ring before any reaction can occur. As a result of this resonance stabilization, benzene undergoes electrophilic substitution rather than electrophilic addition reactions. These substitution reactions allow the six delocalized electrons of benzene to be regenerated after attack, so that the resonance stabilization of the ring is preserved. Electrophilic substitution reactions of benzene produce a vast range of compounds with a wide variety of uses.

AS **G.10.1**
Describe, using equations, the nitration, chlorination, alkylation and acylation of benzene. © IBO 2007

AS **G.10.2**
Describe and explain the mechanisms for the nitration, chlorination, alkylation and acylation of benzene. © IBO 2007

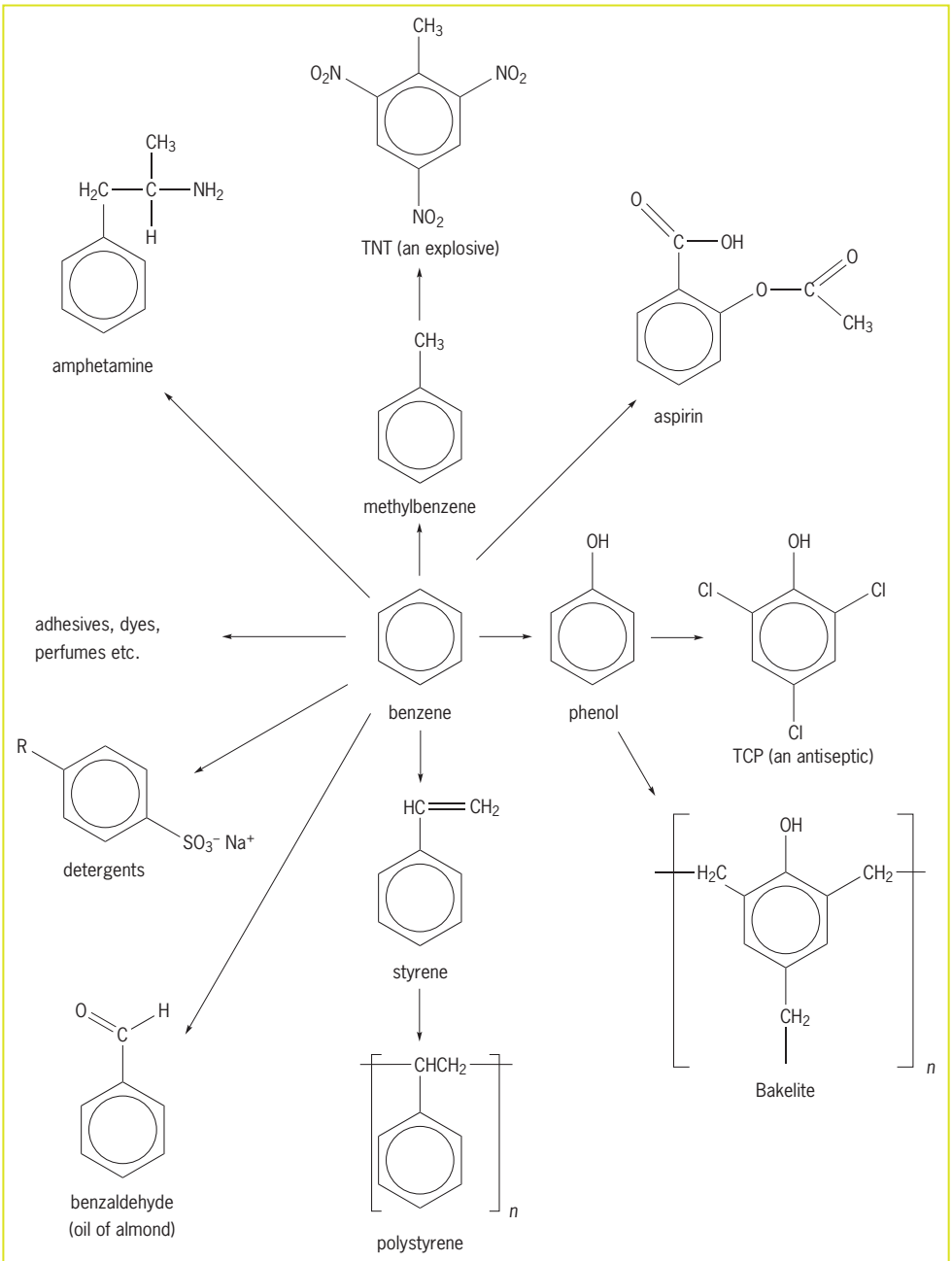


Figure 7.9.1 Benzene and its derivatives have a wide variety of uses.

The general reaction of electrophilic substitution, shown in figure 7.9.2, can be represented using an electrophile, E^+. Reaction involves attack by E^+ on the electron cloud of the benzene ring. The electrophile E^+ takes two electrons to form a bond to a ring carbon and to generate a positive charge in the ring. The resulting intermediate carbocation no longer has the cyclic system of electrons, because in the formation of this ion the electron from one ring carbon is no longer available for integration into the π system. The four remaining electrons of this intermediate are delocalized over the remaining five carbons. Resonance stabilization of the positive charge partially compensates for the loss of the delocalized π system. Because of the disruption of the ring stabilization in this step, the formation of this carbocation intermediate is endothermic and rather slow, and this becomes the rate-limiting step for the reaction. Formation of the final product occurs by loss of the ring proton from the position of electrophilic attack. Electrons from this C–H bond are used in reforming the π system so that the benzene ring structure is restored. The overall result is substitution of the electrophile E for hydrogen. Examples of electrophilic substitution reactions of benzene are shown in figure 7.9.3.

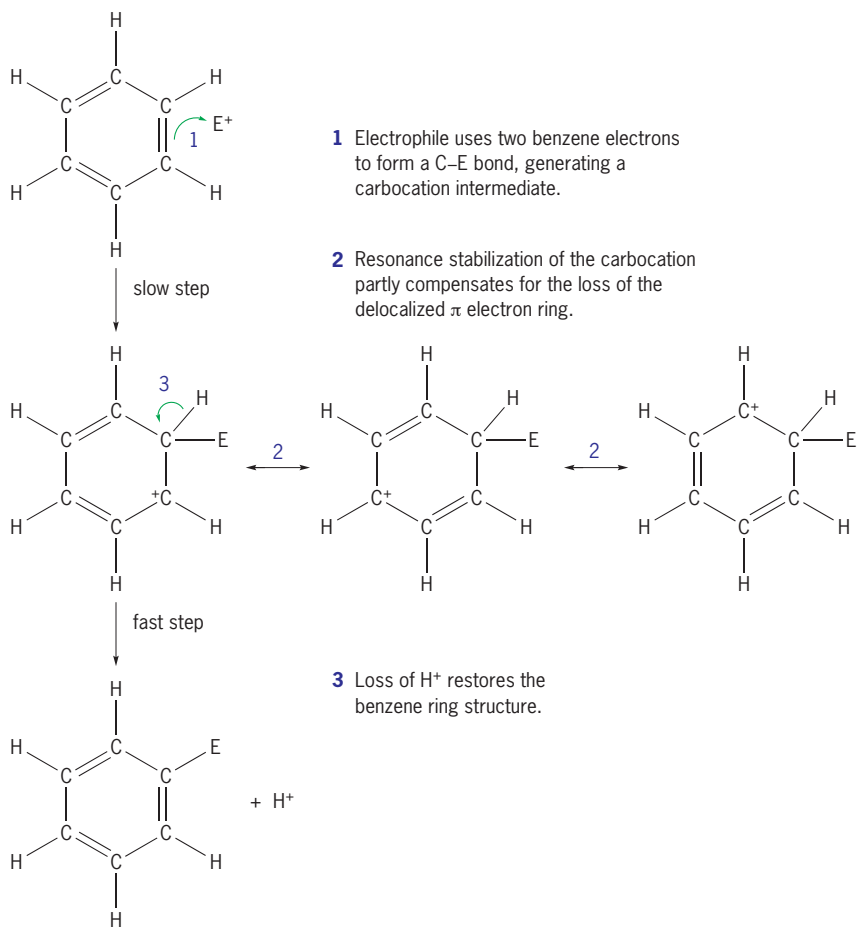


Figure 7.9.2 The electrophilic substitution reaction mechanism.

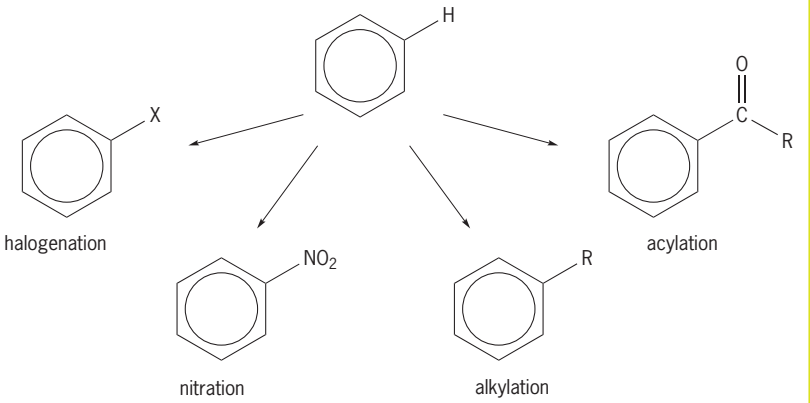


Figure 7.9.3 Some of the electrophilic substitution reactions of benzene.

Chlorine reacts readily with alkenes, but does not react readily with benzene. In order to chlorinate benzene in a substitution reaction, a Lewis acid (such as $AlCl_3$ or $FeCl_3$) is added as a catalyst. This catalyst produces a more reactive electrophile, and in the presence of this Lewis acid catalyst benzene reacts readily with chlorine to produce chlorobenzene.

The mechanism of *chlorination* involves a vacant orbital on the Lewis acid. This orbital can form a complex with chlorine, withdrawing electrons from the chlorine molecule. This complex acts as the electrophile to transfer a positive chlorine ion Cl^+ to the benzene ring. The function of the Lewis acid is therefore to break the chlorine molecule heterolytically, and so to encourage the formation of the C–Cl bond. The intermediate carbocation is stabilized by resonance. In the final stage, loss of the ring proton at the position of attack restores the ring structure, HCl is produced, and the Lewis acid catalyst is regenerated.

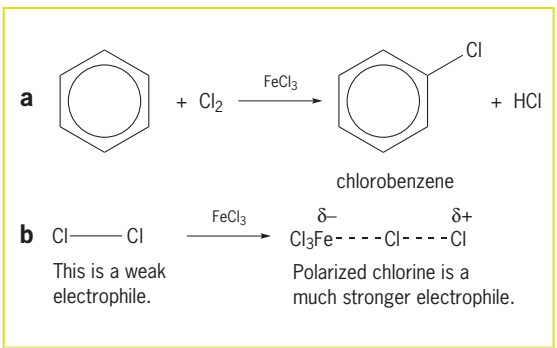


Figure 7.9.4 (a) Chlorination of benzene requires a catalyst. (b) The catalyst creates a stronger electrophile.

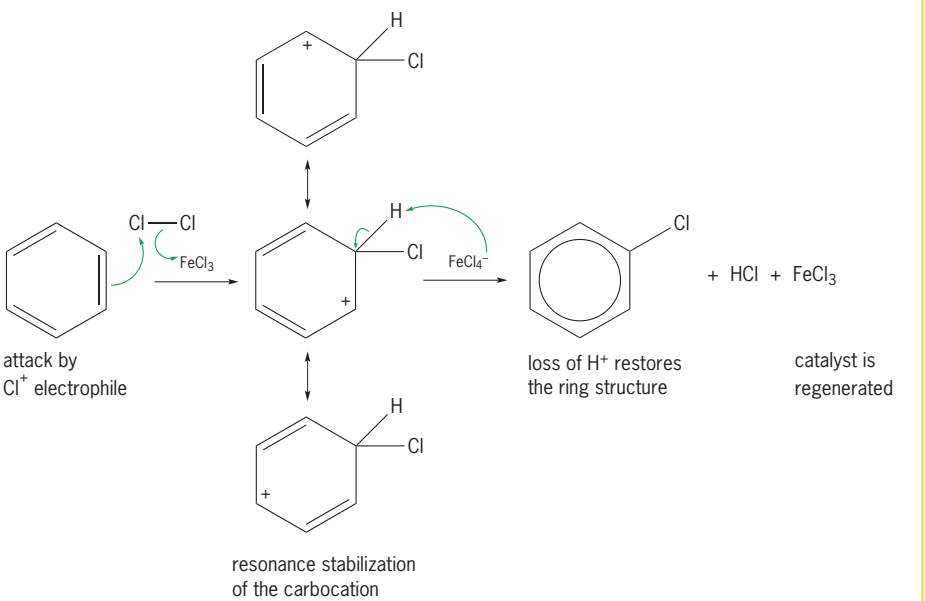


Figure 7.9.5 Chlorination of benzene by the electrophilic substitution mechanism.

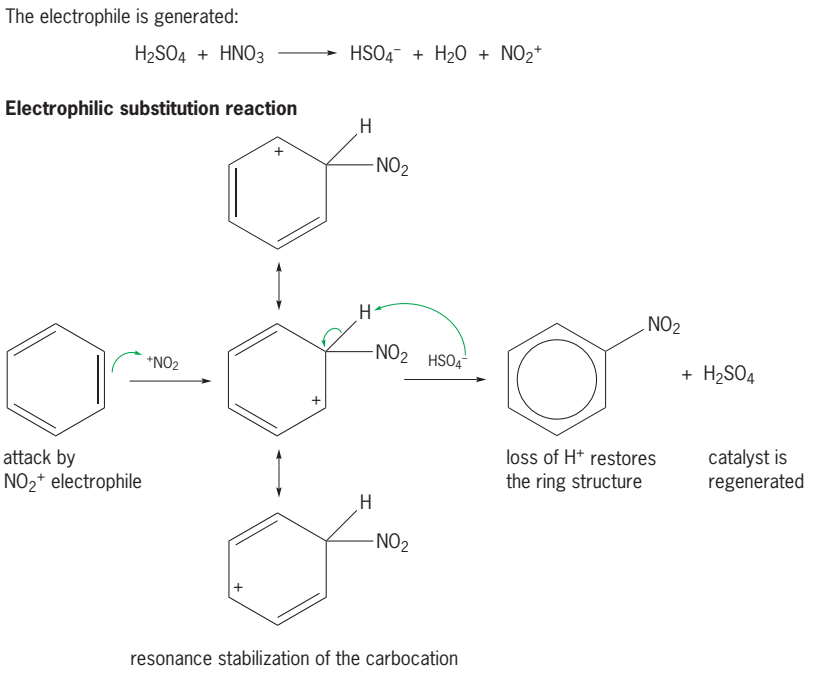


Figure 7.9.6 Nitration of benzene by the electrophilic substitution mechanism.

Nitration of benzene involves the use of a heated mixture of concentrated nitric acid and concentrated sulfuric acid. The benzene reacts by substituting a ring hydrogen with a nitro $–NO_2$ group. The electrophile is generated when nitric acid acts as a base to accept a proton from the stronger sulfuric acid. The protonated nitric acid then dissociates to generate the reactive attacking electrophile – the nitronium ion NO_2^+. This ion attacks the ring, forming a carbocation intermediate. Reaction proceeds with the loss of the ring hydrogen at the site of electrophile attack to restore the ring structure (see figure 7.9.6).

Alkylation, the addition of an alkyl group to the benzene ring involves reaction with a halogenoalkane in the presence of an $AlCl_3$ catalyst. The reaction occurs in a similar way to chlorination. The catalyst withdraws the pair of electrons from the R–X bond into its vacant orbital, leaving the R^+ electrophile. The benzene ring provides two electrons to form the new C–C bond, and the carbocation intermediate is generated. Loss of a proton then yields the alkylated benzene product.

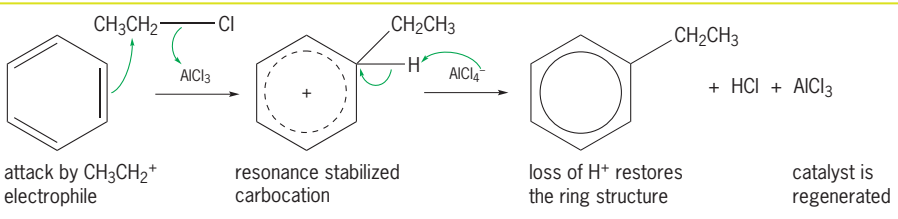


Figure 7.9.7 Alkylation of benzene by the electrophilic substitution mechanism.

Acylation, the addition of an RCO group to the benzene ring involves a similar reaction; this time between benzene and an acyl chloride (ROCl) in the presence of $AlCl_3$ catalyst. The function of the catalyst is the same as in alkylation: to produce a stronger electrophile, RCO^+.

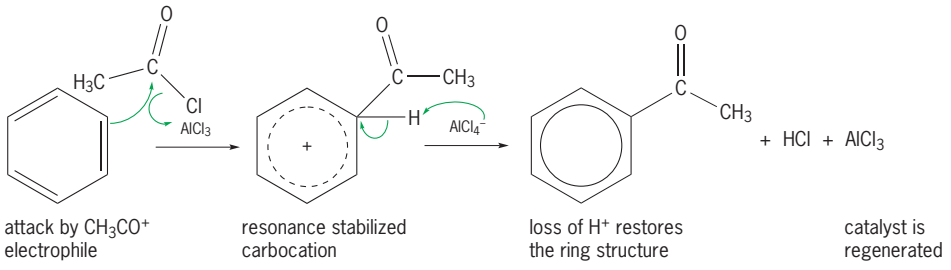


Figure 7.9.8 Acylation of benzene by the electrophilic substitution mechanism.

The $AlCl_3$ catalyst is sometimes called a halogen carrier. The French chemist Charles Friedel and the American chemist James Craft realized the significance of these halogen carriers in the production of positively charged electrophiles to be used in substitution reactions of benzene. These types of alkylation and acylation reactions using $AlCl_3$ catalysts are therefore sometimes known as **Friedel–Crafts reactions**.

Effects of substituents on electrophilic substitution reactions

AS G.10.3
Describe, using equations, the nitration, chlorination, alkylation and acylation of methylbenzene. © IBO 2007

AS G.10.4
Describe and explain the directing effects and relative rates of reaction of different substituents on a benzene ring. © IBO 2007

When substituted arenes undergo electrophilic attack, the groups that are already present on the benzene ring will affect both the rate of reaction (reactivity) and the site of attack (orientation). We have seen that electrophilic substitution of the benzene ring and its derivatives depends largely on two factors: a high electron density in the benzene ring to attract the attacking electrophiles, and stabilization of the positive intermediate carbocation. The effect of substituents on the reactivity of substituted benzenes can be understood in terms of inductive and resonance effects on these two factors.

If a substituent S can donate electrons to the ring, then the ring will have higher electron density and it will be more susceptible to electrophilic attack than benzene. This electron-releasing effect can also stabilize the positive intermediate carbocation formed during electrophilic attack, so that these electron-releasing substituents activate the attached ring towards further electrophilic substitutions. For example, consider the difference in reactivity of benzene and methylbenzene with the electrophile E^+. The CH_3 group, like all alkyl groups, has a positive inductive effect, that is, an electron-donating effect towards the attached benzene ring. This enhances the electron density on the benzene ring and stabilizes the intermediate carbocation, so that the CH_3 activates the benzene ring towards electrophilic substitutions. Methylbenzene will therefore be more reactive to electrophilic substitution reactions than benzene. Another electron-donating, **activating group** is the hydroxyl group. Phenol will therefore be more reactive to electrophilic substitution reactions than benzene.

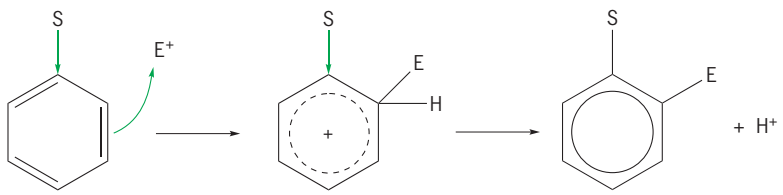


Figure 7.9.9 Electron-donating substituents activate the benzene ring towards electrophilic substitution.

Conversely, an electron-withdrawing substituent will reduce the electron density on the benzene ring to which it is attached, thus destabilizing the positive intermediate carbocation. The effect of electron-withdrawing substituent is to deactivate the substituted benzene towards further

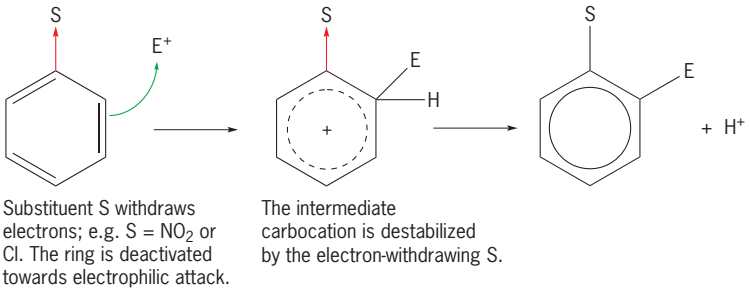


Figure 7.9.10 Electron-withdrawing substituents deactivate the benzene ring towards electrophilic substitution.

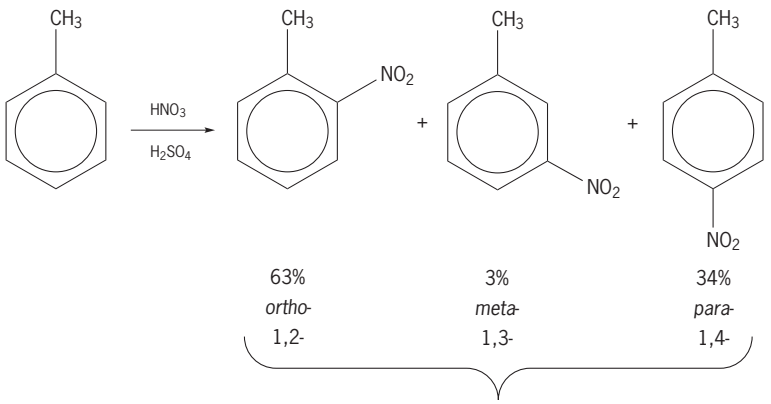


Figure 7.9.11 Three products are possible when methyl benzene undergoes mononitration — but not all are produced to the same extent.

electrophilic substitution. For example, in nitrobenzene the $-NO_2$ substituent is electron withdrawing. It therefore deactivates the attached benzene ring towards electrophilic substitution, and so reaction of nitrobenzene is slower than reaction of benzene. Another electron-withdrawing, **deactivating group** is the halogen group.

In addition to altering the rate of reaction, substituents on the benzene ring can affect the *orientation of attack by electrophiles*. Reactions of monosubstituted benzene compounds with electrophilic reagents can produce three isomeric disubstituted products: the 1,2–(known as *ortho, o*), 1,3–(*meta, m*) and 1,4–(*para, p*) disubstituted derivatives. For example, three products are possible when methylbenzene undergoes mononitration.

Although three products may be possible, only two occur to any extent. Why? It is found that certain substituents on the benzene ring direct further electrophilic substitution to give predominantly *ortho* (1,2- and *para* (1,4-) substituted products, while other substituents tend to direct substitution predominantly to the *meta* (1,3-) position. The explanation for these directing effects again lies in the stability of the intermediate carbocation. We will consider four cases to illustrate these directing effects.

Consider the reaction of methylbenzene with the electrophile E^+. The CH_3 group has a positive inductive effect, so reaction will be enhanced. Three intermediate carbocations are possible. These will be resonance stabilized, as shown in figure 7.9.12. All three intermediate carbocations are stabilized, but the *ortho* and *para* forms are the most stable because, in these forms, the positive charge is present on the carbon attached to the CH_3 group. The positive inductive (electron-donating) effect of the CH_3 group will stabilize these ions. The methyl group, and all other electron-donating groups, are therefore *ortho–para* directors.

Consider the reaction of phenol with the electrophile E^+. The OH group activates the ring, so reaction will be enhanced. Three intermediate carbocations are possible. These will be resonance stabilized, as shown in figure 7.9.13. All three intermediate carbocations are stabilized, but the *ortho* and *para* forms are the most stable. Four resonance structures can be drawn for these carbocations, including the relatively stable form in which the non-bonding electron pair from the oxygen atom is donated to form an extra bond to the ring carbon. The OH group is therefore an activating *ortho–para* director. Phenol is, in fact, so reactive that it can react with chlorine without the aid of a halogen carrier such as $AlCl_3$. Reaction proceeds rapidly to produce 2,4,6-trichlorophenol, known as TCP. TCP has been used in a variety of applications including fungicides, herbicides, insecticides and antiseptics.

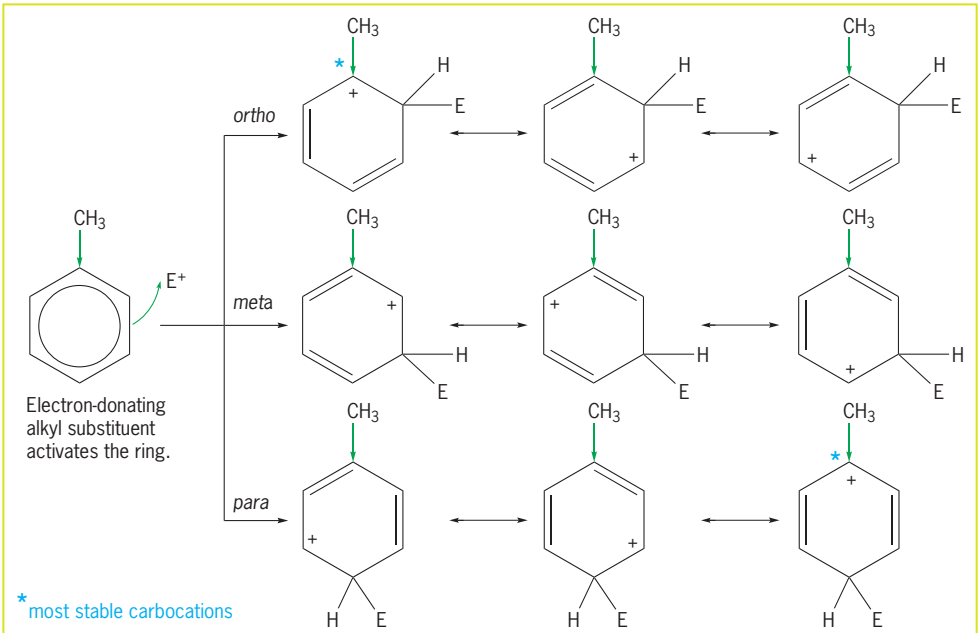


Figure 7.9.12 Carbocation stability explains the *ortho–para* directing effects of electron-donating substituents in electrophilic substitution reactions.

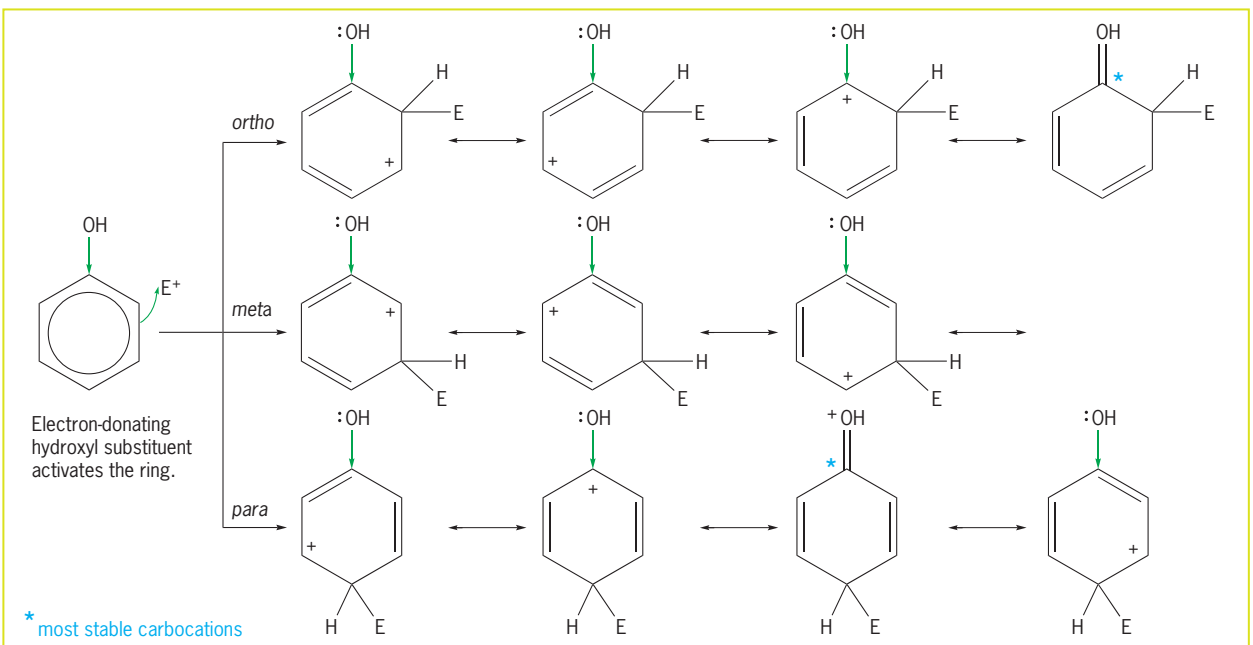


Figure 7.9.13 Carbocation stability explains the *ortho–para* directing effect of phenols in electrophilic substitution reactions.

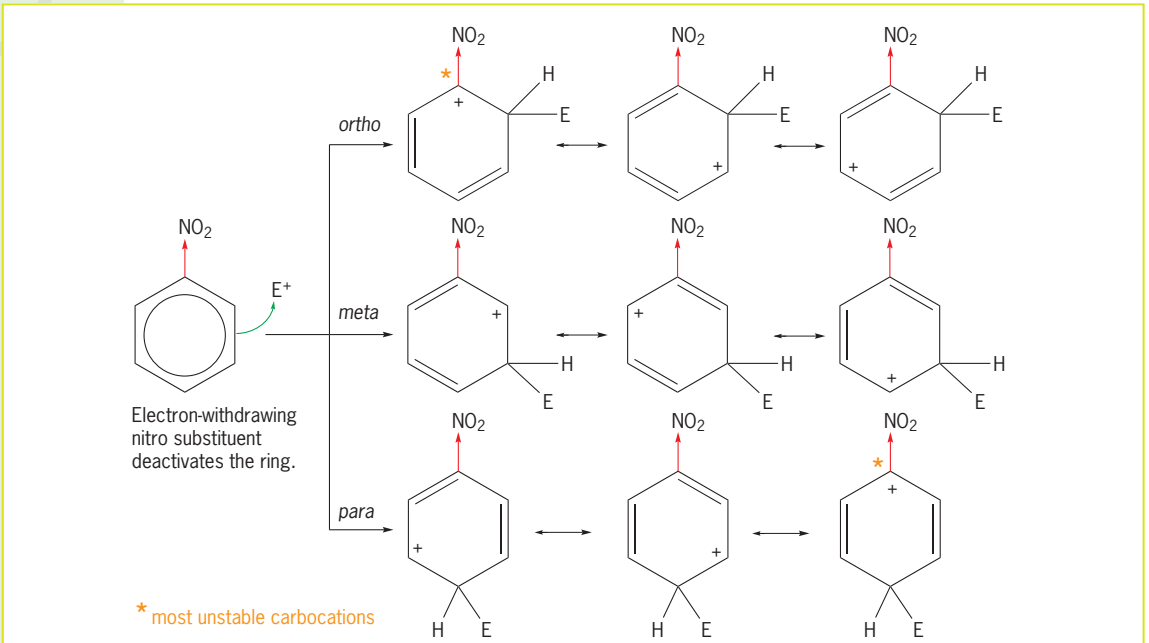


Figure 7.9.14 Carbocation stability explains the *meta*-directing effects of most electron-withdrawing substituents in electrophilic substitution reactions.

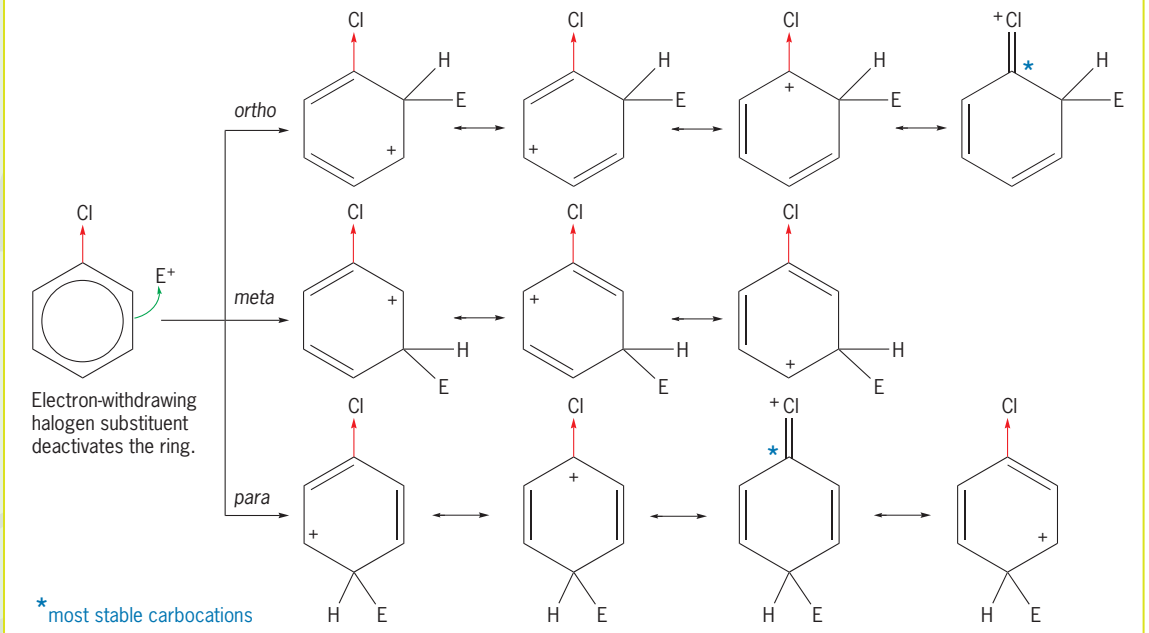


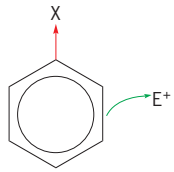
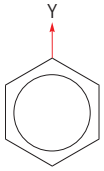
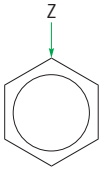
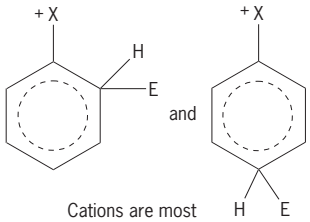
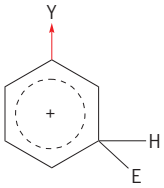
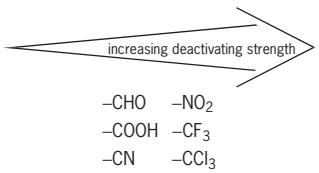
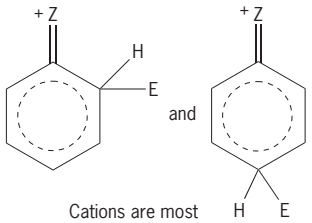
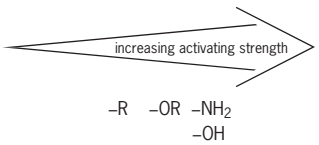
Figure 7.9.15 Carbocation stability explains the *ortho–para* directing effects of halogen substituents in electrophilic substitution reactions.

Consider the reaction of nitrobenzene with the electrophile E^+. The NO_2 group has a negative inductive effect, so reaction will be slowed, compared with that of benzene. Three intermediate carbocations are possible. These will be resonance stabilized, as shown in figure 7.9.14. Here the *ortho* and *para* carbocations include a highly unstable form in which the positive charge is present on the carbon attached to the NO_2 group. The electron-withdrawing effect of the NO_2 group destabilizes these cations. The NO_2 group, and other electron-withdrawing groups are therefore *meta* directors.

Consider the reaction of chlorobenzene with the electrophile E^+. The Cl atom has a negative inductive effect, so reaction will be slowed, compared with benzene. Three intermediate carbocations are possible. These will be resonance stabilized, as shown in figure 7.9.15. Here the *ortho* and *para* carbocations include the relatively stable form in which the positive charge is present on the electron-rich chlorine atom. The halogens therefore display the unusual property of being deactivating, but *ortho–para* directing. The effects of various substituents on reactivity and orientation of electrophilic attack on arenes are summarized in table 7.9.1.

WORKSHEET 7.4
Reactions of benzene

TABLE 7.9.1 EFFECT OF SUBSTITUENTS ON ELECTROPHILIC SUBSTITUTION

Deactivating *ortho–para* (2-, 4-) directors (halogens)	Deactivating *meta* (3-) directors	Activating *ortho–para* (2-, 4-) directors
 X = halogen Electron withdrawal deactivates the ring and destabilizes the carbocation.	 Y = e.g. NO_2 Electron withdrawal deactivates the ring and destabilizes the carbocation.	 Z = e.g. OH Electron donation activates the ring and stabilizes the carbocation.
 Cations are most stable when X carries the positive charge.	 Most stable cation.  –CHO –NO_2 –COOH –CF_3 –CN –CCl_3	 Cations are most stable when Z carries the positive charge.  –R –OR –NH_2 –OH

We have seen that the presence of substituents on a benzene ring will have effects on both the *rate* and *orientation* of further electrophilic substitution reactions. The chlorination of methylbenzene, shown in figure 7.9.16 illustrates these effects. The chlorination reaction also serves to illustrate another important aspect of organic synthesis reactions, that of the choice of reaction conditions. Depending on the conditions chosen, methylbenzene reacts with chlorine to produce a variety of chlorinated products, either by electrophilic substitution or nucleophilic substitution.

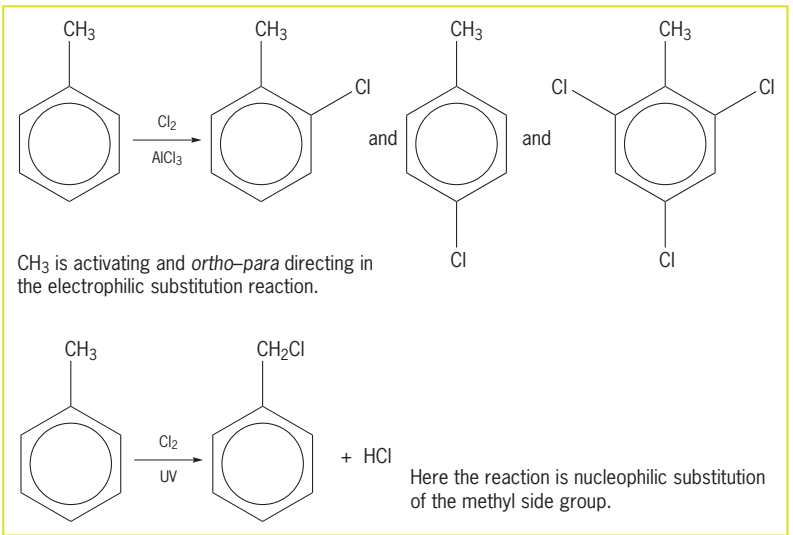
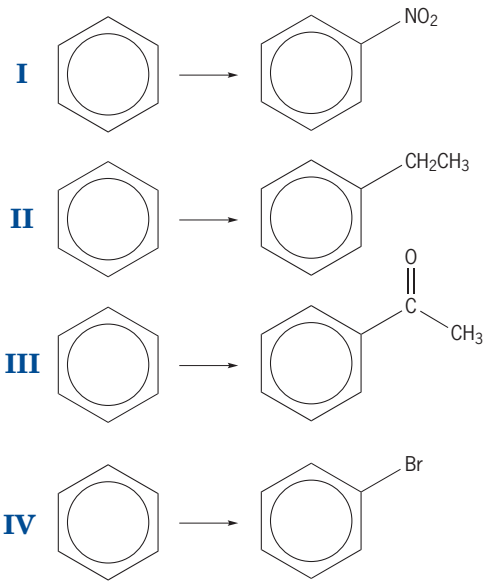


Figure 7.9.16 Reaction of methylbenzene with chlorine produces different products under different reaction conditions.

Section 7.9 Exercises

1 This question concerns the conversions numbered I to IV below.



a For each conversion (I to IV) listed above:

i identify the reagent and catalyst used

ii show how the electrophile is formed from the reagents.

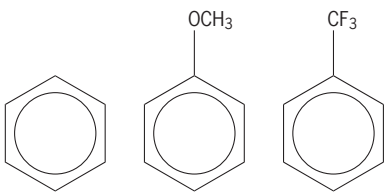
b Draw a mechanism for the reaction in conversion II.

2 Explain why the nitration of:

a methylbenzene is faster than the nitration of benzene

b chlorobenzene is slower than the nitration of benzene.

3 List these compounds in order of increasing reactivity towards electrophilic substitution reaction. Explain your choice in terms of the mechanism of the reaction and the structure of intermediates in the reaction.



4 Draw and name the predicted products of the mononitration of:

a benzoic acid (C_6H_5COOH)

b aniline ($C_6H_5NH_2$)

5 Methylbenzene reacts with chlorine under different conditions to produce three isomeric products (C_7H_7Cl).

a Draw structures for the three products.

b Two of the isomers form in the same type of reaction.

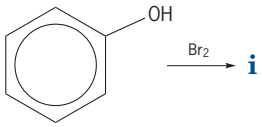
i Give the conditions necessary for this reaction.

ii Draw a mechanism for this reaction.

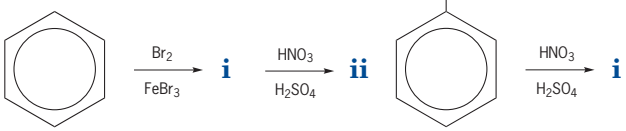
c Name the conditions and the type of reaction necessary to produce the third isomer.

6 Draw the structures for the products expected to be obtained from each of the following reactions. Explain your choice in each reaction.

a



b



c

7.10 REACTION PATHWAYS

HL

To summarize the reactions discussed in sections 7.8 and 7.9, we may construct a 'web' of reaction pathways, as shown in the figure 7.10.1.

AS **G.11.1 Deduce reaction pathways given the starting materials and the product. © IBO 2007**

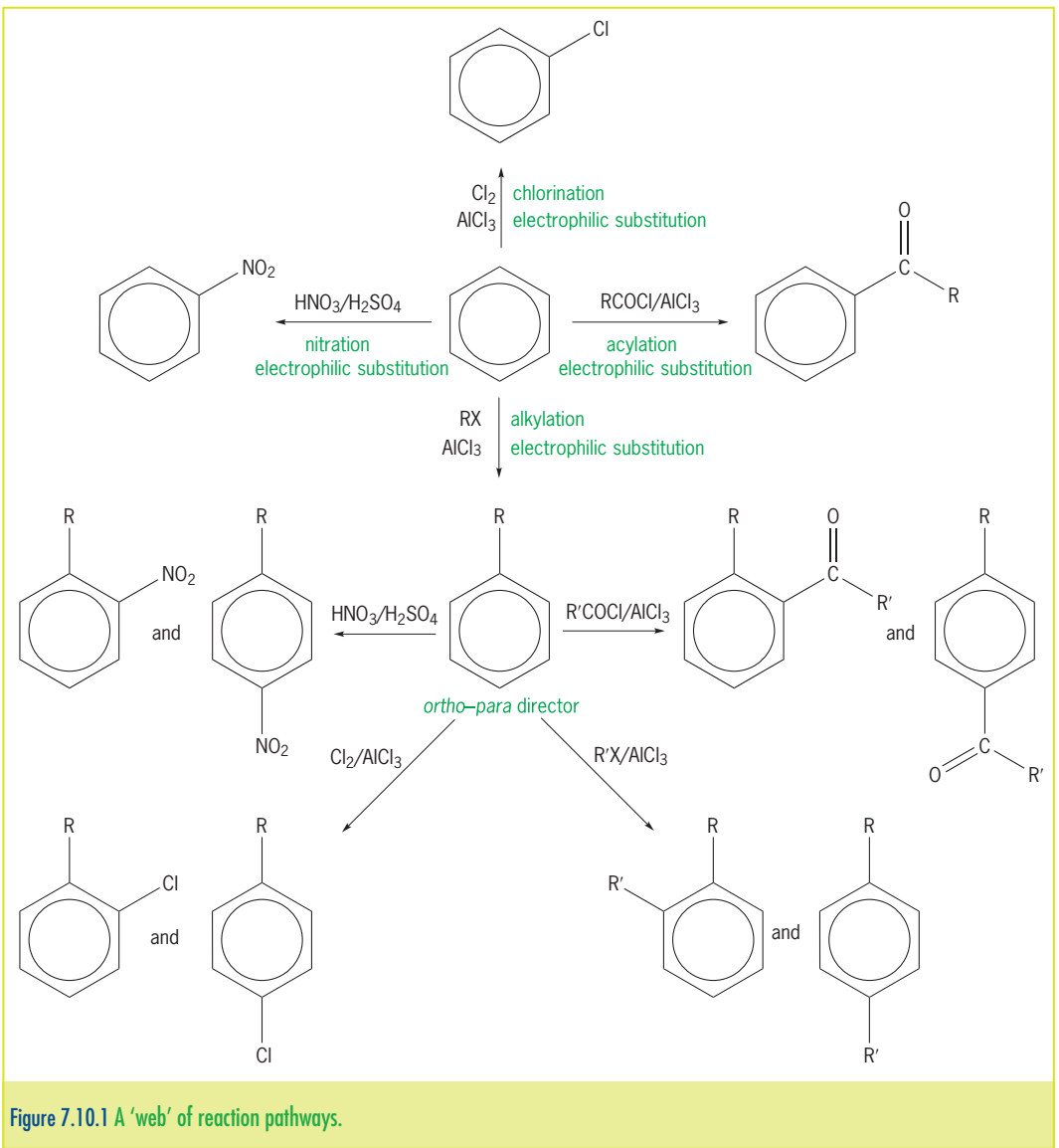


Figure 7.10.1 A 'web' of reaction pathways.

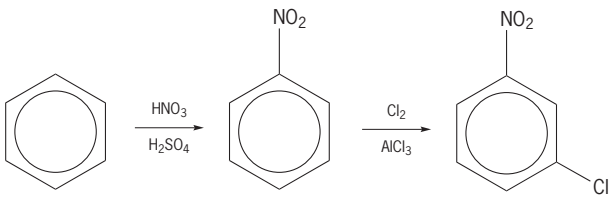
Worked example

Show a reaction pathway to the two-step conversion of benzene to 3-nitrochlorobenzene.

Solution

Note that the substituents on the ring are in the *meta* position. The nitro group is a *meta* director. The chlorine atom is an *ortho–para* director. This means that the nitration of the ring must occur first.

A suitable pathway is shown below.



Section 7.10 Exercises

Draw possible reaction pathways of no more than two steps for the synthesis of the following compounds from the given starting compounds. Include structural formulas, and any inorganic reagents and conditions needed for the pathway.

a Ethanoic acid from ethanoyl chloride

b 3-Nitromethylbenzene from benzene

c Propanamide from propanoic anhydride

d Phenylmethanol from methylbenzene

Chapter review questions and tests are available on student CD.

Chapter 7 Summary

Terms and definitions

Activating group An electron-donating group, such as hydroxyl, attached to the benzene ring that increases the reactivity of the ring towards electrophilic substitution reaction. All activating groups are *ortho–para* directors.

Acyl chloride The carboxylic acid derivative of general formula RCOCl. Also known as an acid chloride.

Acylation In electrophilic substitution, the substitution of an acyl group for a hydrogen atom on the benzene ring.

Addition–elimination A reaction in which two reactants join, then a small molecule is eliminated from the product of the addition. Examples include the reaction of aldehydes and ketones with hydrazines, and the reactions of acyl chlorides with nucleophiles such as water, alcohols and amines. Addition–elimination reactions may be thought of more simply as condensation reactions.

Alkoxide ion The anion, RO^-, formed when an alcohol loses a proton.

Alkylation In electrophilic substitution, the substitution of an alkyl group for a hydrogen atom on the benzene ring.

Anhydride The carboxylic acid derivative of general formula $(RCO)_2O$.

Arene A compound containing a benzene ring.

Carbocation A hydrocarbon cation, R_3C^+. This planar, positively charged species is formed as an intermediate during a number of organic reactions.

Condensation reaction A reaction in which two molecules containing functional groups react and join together with the loss of a small molecule (often water).

Cyanohydrins Compounds with the general formula RCH(OH)CN or RR′C(OH)CN. Also known as a hydroxynitriles. Formed by the nucleophilic addition of hydrogen cyanide to an aldehyde or ketone.

Deactivating group An electron-withdrawing group, such as nitro, attached to the benzene ring that decreases the reactivity of the ring towards electrophilic substitution reaction. Most deactivating groups are *meta* directors (but halogens are deactivating *ortho–para* directors).

Electrophile An electron-loving species. A reactant with a full or partial positive charge that is attracted to a centre of negative charge.

Electrophilic addition The addition of an electrophile to an alkene to produce a saturated product.

Electrophilic substitution A reaction in which an electrophile, (E^+), reacts with a benzene ring and substitutes for one of the hydrogens on the ring.

Elimination reaction A reaction in which a single reactant splits into two products. An example is the dehydration of alcohols to form water and an alkene.

Friedel–Crafts reaction Electrophilic substitution in which an arene reacts with acyl chloride or a carboxylic acid anhydride in the presence of aluminium chloride. An acyl group becomes bonded to the ring.

Grignard reagent A reactive organomagnesium halide, RMgX, prepared by reaction of a halogenoalkane, RX, with magnesium in ether.

Hydrazines The compound H_2NNH_2 and derivatives of this compound. For example, phenylhydrazine, $H_2NNH(C_6H_5)$, and 2,4-dinitrophenylhydrazine, $H_2NNH(C_6H_3(NO_2)_2)$. Derivatives produced by reaction with aldehydes and ketones (hydrazones) have distinctive melting points and may be used to identify the reacting compound.

Markovnikov's rule In the addition of HX to an alkene, the H attaches to the carbon with fewer alkyl substituents, and the X attaches to the carbon with more alkyl substituents.

Nitration In electrophilic substitution, the substitution of a nitro (NO_2) group for a hydrogen atom on the benzene ring.

Nucleophilic addition A reaction in which a nucleophile adds to the electrophilic carbonyl group of a ketone or aldehyde.

Nucleophilic substitution A reaction in which one nucleophile replaces another nucleophile attached to a carbon atom in a saturated compound.

Organometallic compounds Compounds containing a carbon atom covalently bonded directly to a metal atom, e.g. Grignard reactions.

Phenol A compound with an OH group bonded directly to a benzene ring, ArOH.

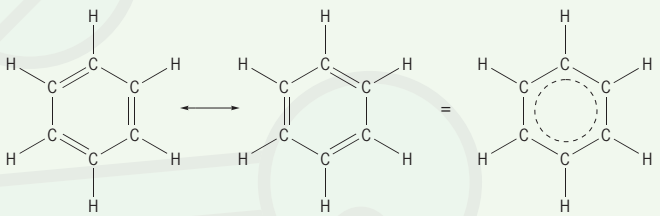
Phenoxide ion The anion, ArO^-, formed when phenol loses a proton.

Positive inductive effect The tendency of alkyl groups to 'push' the bonding electron pair towards the positively charged carbon of a carbocation. One effect of electron-'pushing' alkyl groups is to allow the positive charge of the central carbon in the carbocation to be dispersed, and so to be stabilized.

Resonance structures Representations of the structure of a molecule which differ in the positions of their electrons only. The structure of the molecule cannot be adequately represented by a single representation, and so is seen to be a hybrid or average of the various resonance structures.

Concepts

- Reaction types considered in this chapter are summarized in the table on pages 435–438. Summaries are also shown in figures 7.5.1 (p. 405) and 7.10.1 (p. 431).
- The benzene molecule is seen as a hybrid of two equivalent resonance forms in which the carbon–carbon bonds lie between single and double bonds. This resonance adds stability to the structure.



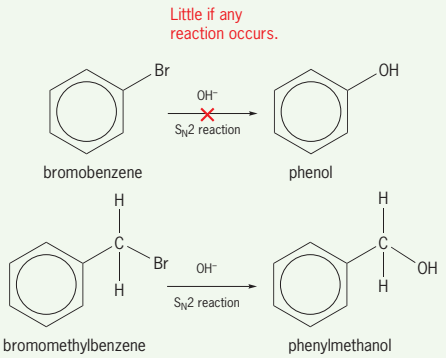
Evidence for this structure is both physical:

  - Bond lengths are intermediate between those of single and double bonds.
  - Only three structural isomers of $C_6H_4X_2$ are found.

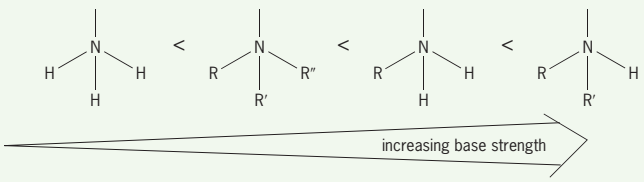
  and chemical:

  - The enthalpy of hydrogenation of benzene is less than that expected for 1,3,5-cyclohexatriene.
  - Benzene undergoes substitution reactions rather than addition reactions.

- Hydrolysis of benzene compounds halogenated in the ring proceeds very slowly when compared with the rate of hydrolysis of compounds halogenated in the side chain. These two hydrolysis reactions proceed by different mechanisms.

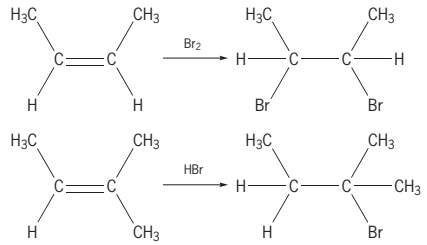
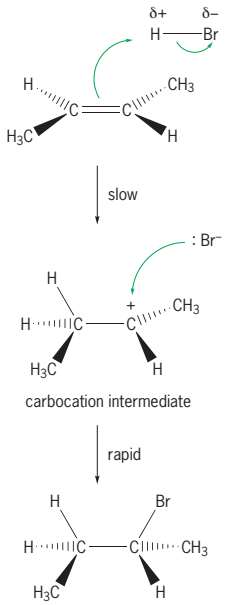
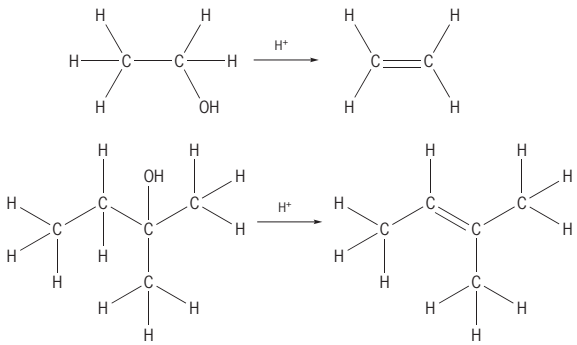
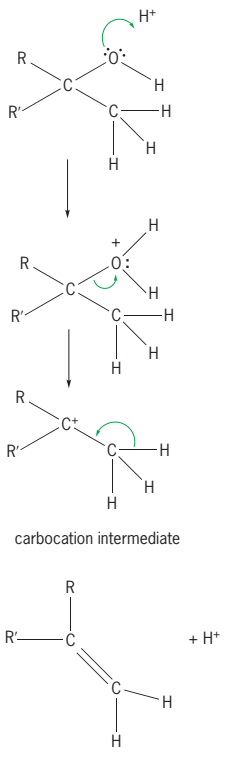


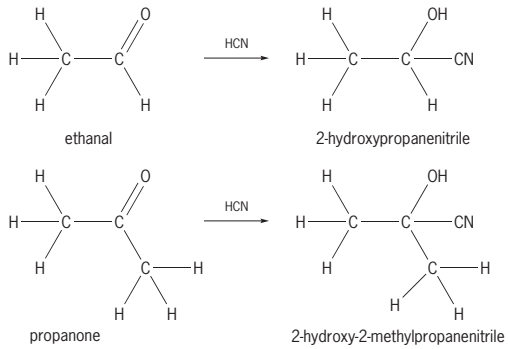
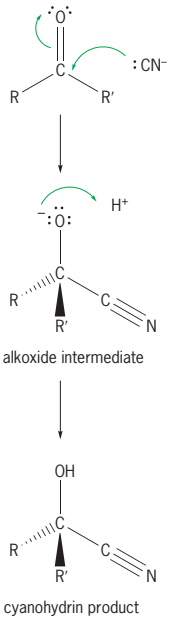
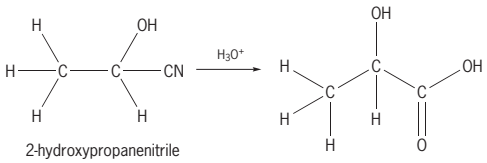
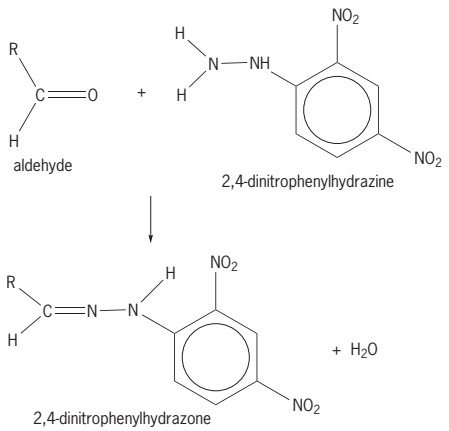
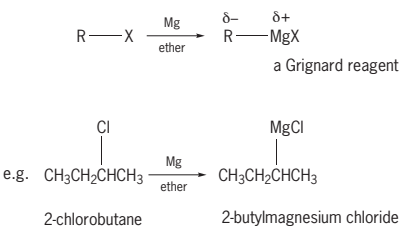
- The relative acid strengths of alcohols, carboxylic acids, phenols and their substituted derivatives may be explained in terms of the strength of the H–A bond, the electronegativity of A, and features that stabilize the conjugate base A^-. These effects are summarized in table 7.7.1 (p. 414).
- The relative base strengths of ammonia and amines may be explained in terms of the stabilizing, positive inductive effect of alkyl groups attached to the basic nitrogen atom, and the stabilization of the positively charged conjugate acid by solvation with water molecules. The order of basicity resulting from these effects is shown below.

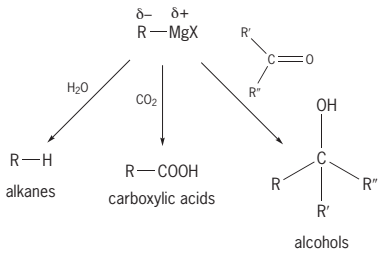
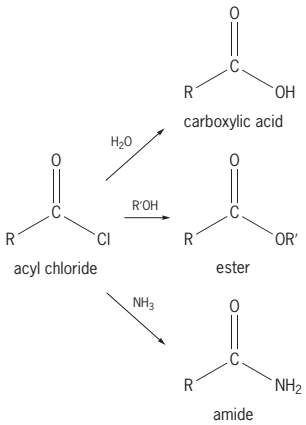
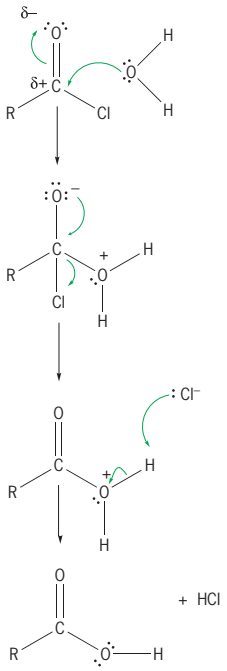
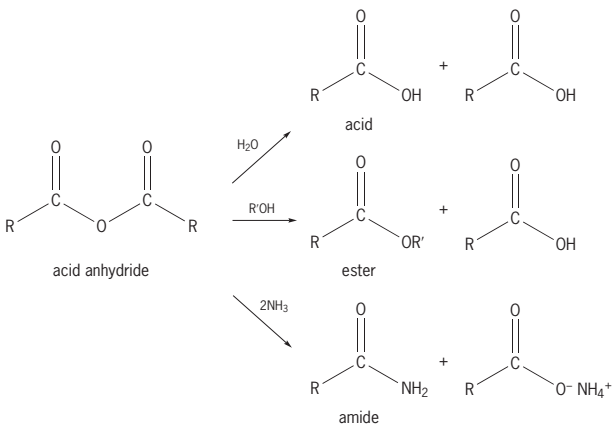


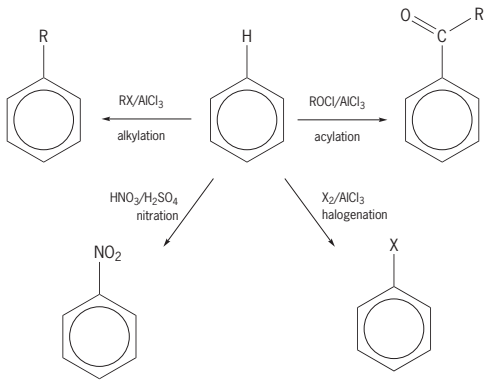
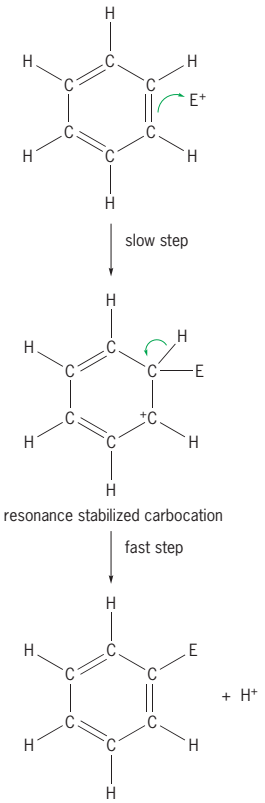
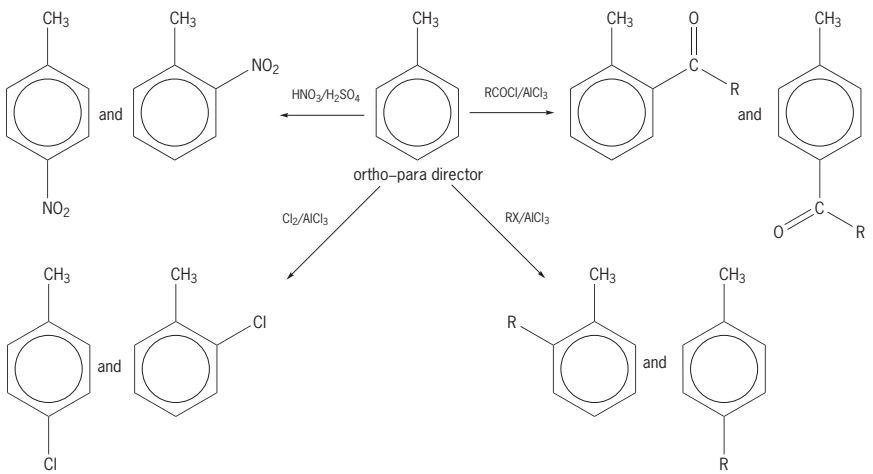
- Benzene and its derivatives undergo electrophilic substitution reactions. Substituents on the benzene ring affect both the rate of reaction and orientation of the substitution in these reactions. The effects are summarized in table 7.9.1 (p. 429).

HL

Reaction type and example	Reaction mechanism	Special features and factors affecting the reaction
Electrophilic addition  	 	Markovnikov's rule In the addition of HX to an alkene, the H attaches to the carbon with fewer alkyl substituents, and the X attaches to the carbon with more alkyl substituents. Explanation for this rule lies in the stability of the intermediate carbocation. The order for carbocation stability is 1° < 2° < 3°. Stability relates to the positive inductive (electron-donating) effect of alkyl groups.
Elimination (of water from an alcohol)  	 	Reaction is faster for 3° alcohols due to increased carbocation stability.

Reaction type and example	Reaction mechanism	Special features and factors affecting the reaction
Nucleophilic addition  	 	Nucleophilic addition of HCN to an aldehyde or a ketone produces a cyanohydrin.
Hydrolysis of cyanohydrins  	Not required	Reaction of an aldehyde or a ketone with HCN, and subsequent hydrolysis to produce a carboxylic acid adds one carbon to the carbon chain of the carbonyl compound.
Addition–elimination (for ketones and aldehydes)  	Not required, but may be thought of as a condensation reaction	Production of characteristic 2,4-dinitrophenylhydrazones may be used to identify the reacting aldehyde or ketone.
Formation of Grignard reagents  	Not required	Grignard reagents are organometallic compounds. The carbon atom is negatively polarized.

Reaction type and example	Reaction mechanism	Special features and factors affecting the reaction
Grignard reactions  	Not required	Grignard reagents are highly reactive. They may be used to increase the carbon chain length of organic molecules.
Higher level only HL		
Addition–elimination (for acyl chlorides)   		Acyl chlorides are reactive. The chlorine atom is readily substituted by other nucleophiles to produce acids, esters and amides.
Reactions of acid anhydrides  	Not required	Acid anhydrides react more slowly than acyl chlorides, and are inefficient (as they produce acid by-products).

Reaction type and example	Reaction mechanism	Special features and factors affecting the reaction
Electrophilic substitution (of benzene)  	 	Electrophilic substitution allows the resonance stabilization of the benzene ring to be maintained. A halogen carrier catalyst such as $AlCl_3$ is required for several of the substitution reactions. The carrier enables the production of a more reactive electrophile.
Reactions of methyl benzene  	Reaction mechansim not required	The methyl group is an electron-donating group and therefore activates the ring towards substitution. It is an *ortho–para* director.

SOLUTIONS

These solutions have been simplified. Fully explained answers, including fully worked solutions to calculations and answers requiring a diagram, may be found on the *Chemistry: For use with the IB Diploma Programme Options* student CD packaged with this book.

Chapter 1 Modern analytical chemistry

Section 1.1

1 Identify, confirm the structure of substances, analyse composition of substances, determine purity.

2 As E increases, λ decreases. Long λ = radio waves, IR; short λ = UV, X-rays

3 wavenumber = $\frac{1}{\lambda}$

4 violet

6 **a** A, B, E **b** B **c** E **d** C, D

8 molecules vibrate

Section 1.2

1 Energy is absorbed and bending or stretching vibrations occur.

3 **a** 500 cm^{-1} **b** 1250 cm^{-1} **c** 4000 cm^{-1}

5 Glass and plastic absorb IR radiation.

6 3200–3600 cm^{-1}: O–H
1050–1410 cm^{-1}: C–O
2850–3100 cm^{-1}: C–H

7 2850–3100 cm^{-1}

8 **a** **i** C–O **ii** C=O
b **i** 900–1400 cm^{-1}

9 B

10 Peak A: C–H
Peak B: C=O
Peak C: C–O

Section 1.3

1 Sample is bombarded by high energy electrons.

3 Molecular ion, CH_3OH^+: 32
CH_3^+: 15
OH^+: 17

4 **a** $CH_3CH_2COCH_2CH_2CH_2CH_3^+$
b $COCH_2CH_2CH_2CH_3^+$ or $CH_3CH_2COCH_2CH_2^+$
c $CH_3CH_2CO^+$
d $CH_3CH_2^+$

7 Compound C could be ethanol, CH_3CH_2OH.

8 Compound D is pentane, C_5H_{12}.

Section 1.4

1 strong magnetic field and radio waves

2 Spin: spinning nuclei create a magnetic field.

4 Number of peaks: the number of different ^1H environments
Relative size of the peaks: relative numbers of H atoms with the same environment

6 possible

7 **a** 2 **b** 1:6 **c** Y

8 **a** 2 **b** 3:1

Section 1.5

1 Electrons are excited to higher energy levels.

2 to produce a hot flame

3 **a** gaseous
b monochromator
c The lamp used as A contains the same metal as the metal being analysed.

6 **b** **i** 0.47 mg/100 cm^3 **ii** 2.0×10^{-4} mol dm^{-3}
c 2.8 mg

7 **a** 0.0125 g

8 **a** 0.026 mg

Section 1.6

2 It has a greater rate of adsorption and desorption.

3 $R_f = \frac{\text{distance moved by component}}{\text{distance moved by solvent}}$

4 **a** **i** water coated onto the fibres of the paper
ii a solvent
b **i** layer of fine powder such as aluminium oxide coated onto the chromatography plate
ii a solvent
c **i** aluminium oxide or silicon dioxide packed into the column
ii a solvent

7 **a** 4 in total, 2 in each dye
b yellow 0.17; blue 0.37; orange 0.25; red 0.32

8 **a** Examples: R_f is used to identify the components; both are simple techniques
b Examples: TLC is faster, more sensitive, capable of greater resolution, uses wider range of solvents

9 **a** Components adsorb to the stationary phase and desorb into the mobile phase to different degrees, depending on their size and polarity. These differing adsorptions and desorptions result in different rates of movement, and hence separation is achieved.
b B and E

Section 1.7

1 UV: 200–400 nm; visible: 400–800 nm

2 UV light has higher energy.

3 Chlorophyll *a*: 430 nm and 660 nm; chlorophyll *b*: 470 nm and 625 nm

5 nature, charge and size of the transition metal ion; stereochemistry of the complex

7 **a** 17 000 μg cm^{-3}

8 **a** II, III, I
b Molecule I might be coloured.

9 **a** III
b The greater the number of double bonds, the greater the maximum wavelength.

Section 1.8

1 **a** tetramethylsilane

2 **a** It is the position of a peak compared to the position of the TMS peak.

4 $H_3C–O–CH_3$ would have only one peak, $H_3C–CH_2OH$ would have three peaks.

5 The NMR spectrum of each would have only one peak.

7 **a** 3

b 3 ^1H environments, but 4 Cs, ∴ one of the carbons has no hydrogens attached to it.

c neighbouring C has 3 hydrogens

d neighbouring C is C=O

8 **b** It increases the chemical shift.

c peaks labelled as Z

d I: ethyl methanoate; II: propanoic acid

Section 1.9

1 Greater surface area results in more effective separation.

2 **a** **i** an inert gas

ii viscous, high boiling point liquid coated onto fine granules of an inert solid

b **i** a solvent under high pressure

ii fine particles of a solid such as Al_2O_3

3 separation of temperature-sensitive materials

5 **a** can be vaporized without decomposing or is already in the gaseous form

b non-volatile molecular substance with a large relative molar mass (≥ 300 g mol^{-1})

6 **a** retention time **b** peak areas **c** retention time

7 A: butane; B: pentane; C: octane; D: decane

10 **a** 15 cm^3 **b** redox titration or HPLC

Chapter 2 Human biochemistry

Section 2.1

1 **a** almonds

2 1850 kJ

3 water and small quantities of minerals and vitamins

4 2.78 kJ g^{-1}

5 2.65 kJ g^{-1}

Section 2.2

2 **a** It is the pH at which the amino acid carries no net electrical charge.

3 **a** **i** alanine

ii glutamic acid

iii asparagine

4 **a** amide **c** water

5 **b** **ii** six

6 **a** four **b** three

Section 2.3

1 **b** Structural isomers have the same molecular formula, but different structural formulas.

d 18 molecular mass units less

4 **a** condensation reaction

b H_2O

5 **a** amylose, amylopectin, cellulose and glycogen

6 **a** polysaccharide **b** polysaccharide

c disaccharide **d** monosaccharide

e monosaccharide **f** disaccharide

g polysaccharide **h** polysaccharide

Section 2.4

2 **a** more cholesterol in LDL than HDL

b causes cardiovascular disease

3 **c** ester functional group

4 **a** Unsaturated fats have at least one carbon–carbon double bond.

b The melting point of an unsaturated fat will be lower.

c The unsaturated molecules do not pack as closely together as saturated molecules.

5 **a** They are essential in the diet of all humans. They cannot be manufactured in the body.

6 The mass of iodine that reacts with 100 g of the lipid.

7 **a** 5 **b** 420

8 **a** 1 **b** 74.9

10 **a** pH above 7 and bile **b** adipose cells

Section 2.5

1 Four hydroxyl groups can form hydrogen bonds with water.

2 They make up more than 0.005% of body weight.

3 co-factors of enzymes

5 **a** water-soluble **b** water-soluble

c fat-soluble **d** fat-soluble

6 **a** lack of vitamin C **b** lack of vitamin D

c lack of thiamine **d** lack of iron

e lack of proteins

Section 2.6

1 **a** to function as chemical messengers

b endocrine

2 **a** testosterone or estradiol

b progesterone or testosterone

5 **a** Stimulates bone growth and appetite, and induces puberty in males.

b Used in sports and body building to increase muscle mass.

Section 2.7

1 **a** Enzymes are biological catalysts.

b Enzymes have faster reaction rates and operate under milder conditions.

2 **a** $HCO_3^-(aq) + H^+(aq) \rightarrow H_2CO_3(aq)$

c The enzyme is denatured.

d It is not altered.

4 **a** K_m = substrate concentration at which $V = \frac{1}{2}V_{max}$

b 7 mmol dm^{-3} **c** enzyme A **d** enzyme B

5 **a** **i** K_m will increase. **ii** K_m will stay the same.

6 pH, temperature, heavy-metal ions such as Ag^+, Hg^{2+}, Pb^{2+}

Section 2.8

1 **b** the identity of the purine or pyrimidine base and the sugar

3 **b** **i** GCATGC **ii** GCAUGC

4 **a** Base is adenine, sugar is deoxyribose, so found in DNA.

b Neither RNA or DNA.

5 hydrogen bonding between pairs of purine and pyrimidine bases

7 **a** DNA and mRNA

b mRNA and tRNA

8 paternity testing, authenticating foods, matching organ donors with recipients

10 **a** PCR: polymerases chain reaction

b It is used to amplify the sample by separating the stands and replicating them.

Section 2.9

1 Aerobic respiration of glucose releases more energy.

2 Glucose is oxidized, oxygen is reduced.

3 Cu^{2+} and Fe^{2+}

4 coordinate (dative) bonding

Chapter 3 Chemistry in industry and technology

Section 3.1

1 Iron is less reactive than aluminium.

2 **a** haematite, magnetite

b 0.96 tonne

4 **a** $C(s) + O_2(g) \rightarrow CO_2(g)$ then $CO_2(g) + C(s) \rightarrow 2CO(g)$

b $3Fe_2O_3(s) + CO(g) \rightarrow 2Fe_3O_4(s) + CO_2(g)$

$Fe_3O_4(s) + CO(g) \rightarrow 3FeO(s) + CO_2(g)$

$FeO(s) + CO(g) \rightarrow Fe(s) + CO_2(g)$

c $CaO(l) + Al_2O_3(l) \rightarrow CaAl_2O_4(l)$

5 Pig iron is hard and brittle.

7 **a** harder, stronger, lower ductility and malleability

b harder (slightly), tougher, stronger, more corrosion and shock resistant

c corrosion and abrasion resistant

d harder, stronger, improves heat resistance, prevents cracking and distortion

8 **a** annealing and tempering

b quenching

11 **a** Anode(+): $C(s) + 2O^{2-}(l) \rightarrow CO_2(g) + 4e^-$

Cathode(–): $Al^{3+}(l) + 3e^- \rightarrow Al(l)$

12 Pure aluminium is soft and lacks tensile strength.

Section 3.2

1 coal, crude oil, natural gas

2 **a** Chemical derived from petroleum.

b Liquid or gaseous hydrocarbons used to make intermediate or primary chemicals.

4 For example: $C_2H_6(g) \rightleftharpoons C_2H_4(g) + H_2(g)$

6 **a** Increased branching reduces melting point and tensile strength.

b Cross-linking increases hardness and rigidity.

c Melting point and strength increase as the chain length increases.

9 Advantages:

- durable
- chemically inert
- affordable

Disadvantages:

- Many are not recyclable and end up in landfill.
- Many produce toxic gases when burned.
- Most are not biodegradable.

Section 3.3

1 **a** heterogeneous **b** homogeneous

c heterogeneous

2 **a** increases surface area **b** heterogeneous

c maximizes yield

3 The iodide ions are not consumed (participate in the reaction, but are re-formed in the last step of the reaction) and increase reaction rate.

5 selectivity, efficiency, durability, cost, availability

7 They are rechargeable.

Section 3.4

1 A liquid crystal substance flows like a liquid, but retains some semblance of crystalline order.

2 calculators, computers, watches, bulletproof clothing, thermometers

3 **b** **i** a biphenyl nitrile **ii** soap

4 They are aligned along a major axis and in layers that can slide past each other.

7 They have chemical stability, fast response to voltage changes and are stable over a wide range of operating temperatures.

Section 3.5

1 1 nm = 10^{-9} m. Nanotechnology entails research and technology in the 1–100 nm range.

2 It allows atoms to be manipulated one by one.

4 **a** A structure like graphite that has been rolled into a cylinder.

Section 3.6

1 Monomers used in addition polymerization must be unsaturated. Monomers used in condensation polymerization must contain at least two functional groups per molecule.

2 **a** hydroxyl and carboxyl

b hydrogen (on phenol) and aldehyde

c hydroxyl and isocyanate

5 It increases the conductivity of polyethyne.

8 A process called 'chain transfer to polymer' occurs in LDPE production, which produces many branches. This doesn't occur in HDPE production due to the structure of the Ziegler–Natta catalysts.

Section 3.7

1 Doping is the manipulation of a semiconductor's conductivity by adding impurities.

3 Intrinsic semiconductors are pure crystals. Extrinsic semiconductors have properties that are largely determined by the doping agent.

4 Light hits n–p junction → electrons move to n side → holes move to p side → load connected → electrons move externally through load from n side to p side.

5 Yes; it is essentially linear and has significant polarity.

6 The molecules are arranged in layers, each slightly twisted with respect to the layers above and below. Within each layer, the molecules are aligned along a major axis.

Section 3.8

2 **a** Membrane cell does not use toxic asbestos, produces very pure caustic soda and uses less energy.

b Mercury cell produces concentrated caustic soda that needs no further processing.

c Diaphragm cell does not use mercury (a neurotoxin) and uses less energy.

3 **a** Anode: $2Cl^-(aq) \rightarrow Cl_2(g) + 2e^-$

Cathode: $Na^+(aq) + e^- \rightarrow Na(l)$

b Anode: $2Cl^-(aq) \rightarrow Cl_2(g) + 2e^-$

Cathode: $2H_2O(l) + 2e^- \rightarrow 2OH^-(aq) + H_2(g)$

c Anode: $2Cl^-(aq) \rightarrow Cl_2(g) + 2e^-$ (anode)

Cathode: $2H_2O(l) + 2e^- \rightarrow 2OH^-(aq) + H_2(g)$

Chapter 4 Medicines and drugs

Section 4.1

1 A drug affects how the body works; a medicine has the function of improving health.

3 The rate of absorption in the stomach is small.

4 rectal administration

5 intravenous, intramuscular, subcutaneous

6 ED_{50}: the amount that would produce a therapeutic response in 50% of the population

LD50: the size of the dose that would be enough to kill 50% of the population

7 The therapeutic window is the ratio of LD_{50} to ED_{50}.

8 It is an effect that is not the intended effect of the drug.

9 **a** tolerance

Section 4.2

1 Examples: magnesium oxide, magnesium hydroxide, calcium carbonate

2 **a** $NaHCO_3(aq) + HCl(aq) \rightarrow NaCl(aq) + H_2O(l) + CO_2(g)$

b $Mg(OH)_2(aq) + 2HCl(aq) \rightarrow MgCl_2(aq) + 2H_2O(l)$

c $CaCO_3(s) + 2HCl(aq) \rightarrow CaCl_2(aq) + H_2O(l) + CO_2(g)$

4 dimethicone

5 Magnesium hydroxide is more effective.

6 Sodium hydroxide is corrosive.

Section 4.3

1 **a** It is an anti-inflammatory drug.

b They are antipyretic drugs, i.e. reduce fever.

c Can lead to ulceration and bleeding of the stomach.

2 **a** morphine, codeine and diamorphine (heroin)

b They all contain ether, alkene and tertiary amine functional groups at same location and a benzene ring.

3 **a** pain relief **b** constipation

4 **a** 50 000 **b** morphine

6 **b** tertiary amine

7 **a** hydroxyl **b** $-OCOCH_3$

c **i** 14.03 **ii** hydroxyl group

8 **b** a hydroxyl group and a carboxyl group

c ester group

Section 4.4

1 **a** calm and relax the central nervous system

b induce sleep

2 **a** acidified potassium dichromate

b orange-yellow to green

c redox reaction

5 The performance of the other drug is often significantly enhanced; e.g. alcohol taken with aspirin increases risk of stomach bleeding.

6 Depressants relieve (diminish the feelings of) depression.

Section 4.5

2 **a** Both have a benzene ring and an amine group.

b Adrenaline has two hydroxyl groups bonded to the benzene ring, amphetamine has none.

4 nicotine

7 sleeplessness, anxiety and irritability

8 Both have tertiary amine functional groups.

Section 4.6

1 **a** bacteria

b They interfere with the cell's ability to synthesize cell walls.

2 Benzylpenicillin was found to be deactivated by stomach acid, so it had to be injected.

3 Bacteria developed resistance to penicillinase.

4 It leads to development of penicillin-resistant populations.

6 antibiotics in animal feedstock

Section 4.7

2 They have no ribosomes or reproduction enzymes.

3 **a** The virus uses the reproductive mechanism of the cell.

b Bacteria multiply (reproduce) asexually.

5 Acyclovir inhibits the viral DNA polymerase, terminating the nucleotide chain.

6 HIV binds to surface proteins of cells of the immune system.

Section 4.8

4 Chirality refers to the ability of a compound to exist as optical isomers.

6 Enantiomers are mirror images of each other.

7 **a** Four: 3 C, 1 N

b The reactivity of the β-lactam ring makes penicillin able to bond to the enzyme involved in the synthesis of bacterial cell walls.

8 Polar hydroxyl groups in morphine are replaced by non-polar ester groups, facilitating transport into the non-polar environment of the central nervous system.

Section 4.9

1 Molecules in a compound library can be screened for a particular use, such as biological activity in a required field.

2 A lead compound is a drug that is known to show biological activity.

3 parallel synthesis, solid phase synthesis (or mix and split)

4 Saves time and synthesis allows more molecules to be investigated.

5 Acidic groups can lose hydrogen ions to form negative ions and then form ionic salts.

6 amine group

7 **a** A chiral auxiliary determines the formation of just one enantiomer.

b It prevents the preparation of a racemic mixture or a second potentially harmful enantiomer.

8 **a** It is a mixture of two different enantiomers.

b An undesirable or harmful biological effect could occur.

Section 4.10

1 Imitate the way in which some hormones (e.g. adrenaline) affect the body.

3 Both mimic the action of serotonin in the brain; both have an indole ring structure.

4 Serotonin is a neurotransmitter. It modulates a range of conditions such as aggression, body temperature, mood and gastrointestinal operation.

6 THC (tetrahydrocannabinol)

Chapter 5 Environmental chemistry

Section 5.1

2 Incomplete combustion of octane:
$2C_8H_{18}(l) + 17O_2(g) \rightarrow 16CO(g) + 18H_2O(l)$

Complete combustion of octane:
$2C_8H_{18}(l) + 25O_2(g) \rightarrow 16CO_2(g) + 18H_2O(l)$

Incomplete combustion produces CO.

5 Catalytic converters work to convert the toxic by-product of combustion CO to less toxic CO_2: $2CO(g) + O_2(g) \rightarrow 2CO_2(g)$

6 In alkaline scrubbers, a basic mixture of limestone and lime is used to convert SO_2 (a source of acid rain) into less harmful slag that can either be placed in landfill or used to make gypsum.

10 Generally air pollution is changing the natural cycle of the planet.

Section 5.2

1 Acid deposition refers to the process by which acidic particles, gases (dry deposition) and precipitation (wet deposition) leave the atmosphere.

2 Normal deposition is subject to reaction with carbon dioxide to produce carbonic acid: $H_2O(l) + CO_2(g) \rightarrow H_2CO_3(l)$

5 $CaCO_3(s) + H_2SO_4(aq) \rightarrow CaSO_4(aq) + H_2O(l) + CO_2(g)$

8 Scrubbing converts the harmful pollutant SO_2 (a source of acid rain) to a non-toxic sludge.

9 Limestone in soils and rocks creates a natural buffer zone and counteracts the effects of acid deposition.

10 Several of the same pollutants are involved.

11 Deterioration of plants, wildlife and water supplies will ultimately lead to a decrease in economies dependent on the environment.

Section 5.3

3 The main natural contributor to greenhouse effect is water vapour. The main anthropogenic source is carbon dioxide.

5 Agricultural by-products are responsible for the biggest percentage of both methane and nitrous oxide. The biggest source of carbon dioxide is the combustion of fossil fuels for power production.

8 Because they are cheaper than their alternatives, CFCs are smuggled across borders from countries where they are still manufactured to countries where they have been banned.

Section 5.4

2 The natural formation of ozone in the atmosphere:
$O_2 + \text{light energy} \rightarrow 2O\cdot$
$O\cdot + O_2 \rightarrow O_3$

The natural depletion of ozone in the atmosphere:
$O_3 + \text{light energy} \rightarrow O_2 + O\cdot$
$O_3 + O\cdot \rightarrow 2O_2$

3 In the lower atmosphere, ozone is an air pollutant and one of the main constituents of smog.

5 CFCs are used as coolants in old refrigerators and air conditioners, as solvents in cleaners, as foaming agents in fire extinguishers and as propellants in aerosols.

6 NO_x form as the result of high temperature combustion within internal combustion engines.

8 CFCs contain chlorine that is weakly bonded to the remainder of the molecule. As a result, it is readily removed from the CFC molecule in the upper atmosphere, where it is free to catalyse the destruction of ozone molecules.

Section 5.5

1 Aquatic plants and animals rely on oxygen gas that has been dissolved in water for respiration.

2 Solubility of gases decreases with increased temperature; therefore, as the temperature increases, dissolved oxygen levels decrease.

3 The amount of dissolved oxygen needed by aerobic decomposers to break down the organic materials in a given volume of water at a certain temperature over a specified time period is called biochemical oxygen demand (BOD).

6 Eutrophication is the process by which lakes, estuaries and other still bodies of water receive higher than normal levels of nutrients, resulting in an excessive growth of plants—an algal bloom.

7 It is an unnatural increase or decrease in water temperature due to anthropogenic effects.

Section 5.6

1 As there is such a small percentage of fresh water that is ready for human consumption, it is a very important commodity.

9 If water is to be used for drinking water, all three stages are important.

Section 5.7

1 Soil is important in our environment as a large portion of the world's population relies on agriculture as a means of income.

2 Soil degradation occurs when actions result in the soil being unhealthy or infertile.

3 These practices can be hard to do on a large scale and maintain profits.

4 The organic constituents in the soil, including undecayed plant and animal tissues, their partial decomposition products and the soil biomass.

5 SOM retains water and water-soluble plant nutrients so that they can be taken up by plants roots.

6 The dark colour of humus helps to absorb heat from the Sun and therefore helps to warm cold soil, creating a better environment for plant growth.

Section 5.8

1 As a more consumerist society we have also become a 'throw away' society.

3 Recycling is the salvaging of usable metals, glass, plastic and paper products from municipal solid waste and selling them to manufacturing industries.

4 Examples: To reduce you could use rechargeable batteries. To reuse you could organize a clothes or sports equipment swap in your school. To recycle you could organize and use a waste management program at school or in your home.

Section 5.9

2 The bond in O_2 is stronger than the bond in O_3, so a shorter wavelength of light (i.e. higher energy) is needed to break it.

4 World weather and wind patterns concentrate the harmful CFCs and NO_x there.

5 The ice crystals found in the polar regions provide a surface on which the reaction takes place.

Section 5.10

1 Photochemical smog forms when a series of primary pollutants and secondary pollutants interact under the influence of sunlight.

Section 5.11

1 sulfur dioxide, particulate matter and nitrogen oxides

2 **a** Aging coal-fired power plants burn sulfur-contaminated coal, producing SO_2.

b SO_2 is also produced in the smelting of metal-sulfate ores in the production of iron, nickel and steel.

c NO_x forms as the result of high-temperature combustion within internal combustion engines.

4 In the atmosphere, ammonia neutralizes the acids formed to a large extent, to form ammonium salts.

Section 5.12

1 SOM enhances the ability of the soil to buffer changes in pH, it acts like a trap binding organic and inorganic materials and pollutants.

2 CEC is defined as the capacity of a soil to exchange cations with the soil solution.

3 Acidic: H^+, Fe^{2+}, Mn^{2+} and Al^{3+}; Basic: Ca^{2+}, Mg^{2+} and K^+

4 The optimal pH level for most groups is in the range of 6.0 to 8.0 or only slightly acidic or slightly basic.

5 $K_{sp} = [Ba^{2+}][SO_4{}^{2-}]$

$K_{sp} = [Fe^{3+}][OH^-]^3$

$K_{sp} = [Ca^{2+}]^3[PO_4{}^{3-}]^2$

6 1.1 g

7 $K_{sp} = 3.0 \times 10^{-15}$

8 0.0080 mol dm^{-3}

Chapter 6 Food chemistry

Section 6.1

1 **a** A food is any substance intended for human consumption whereas a nutrient is a substance obtained from food.

b Carbohydrates, lipids, proteins, minerals, vitamins and water.

c Provision of energy.

2 Lipids supply energy and essential fatty acids, act as vitamin carriers and increase the palatability of food.

4 **a** carbohydrate (monosaccharide)

b protein (tripeptide)

c lipid (triglyceride)

5 Amino acids and monosaccharide both contain polar groups (NH_2 and OH) which are able to hydrogen bond with water molecules.

6 **a** The straight-chain form contains an aldehyde group at carbon 1.

b The α and β isomers differ only in the orientation of the OH groups on carbons 1 and 4.

7 **a** Refer to figure 6.1.10, p. 327.

b $C_{12}H_{22}O_{11}$ and H_2O

8 **a** Refer to figure 6.1.4, p. 325.

b **i** glutamic acid **ii** asparagine

9 **a** peptide (or amide) **c** water

10 **b** ester

Section 6.2

1 **a** fatty acid and glycerol

b A fatty acid is a long chain molecule of general formula C_xH_yCOOH. A fat forms by reaction of three fatty acids and one molecule of glycerol ($C_3H_8O_3$).

4 Olive oil (of plant origin) contains high levels of unsaturated fatty acids; lard contains saturated fats.

5 **a** Reaction with oxygen to form a variety of compounds including hydroperoxides, aldehydes and ketones.

b Reaction with oxygen, in the presence of light, to produce a variety of compounds.

c Reaction with water to form products containing hydroxyl groups.

6 Hydrogenation is the addition of hydrogen to the double bonds in fatty acid chains.

7 Refer to table 6.2.2, p. 338.

Section 6.3

1 Shelf life refers to the period of time during which a food may be stored under recommended conditions before its quality diminishes.

4 Rancidity is a noticeable deterioration in sensory quality due to the decomposition of triglycerides in the food.

7 Examples: refrigeration, freezing, vacuum packaging

8 **a** a substance that slows the rate of oxidation of another substance by being preferentially oxidized itself

b green leafy vegetables

c selenium, β-carotene, vitamin E

9 Antioxidants generally are phenolics with one or more reactive hydroxyl groups.

Section 6.4

1 Colour adds to the aesthetics of the food we eat, which makes many foods more marketable and more enjoyable to eat.

2 **c** yellow-green

d blue

3 Pigments are natural and dyes are synthetic.

5 All these pigments contain conjugated double bonds.

8 Fruits and vegetables that are manufactured in cans containing ferric or aluminium ions can become discoloured when the ions leach from the can into the food.

Section 6.5

1 **a** A food created for human or animal consumption which has been produced from genetically altered plants or animals.

b An organism that contains a foreign gene.

c A genetic engineering technique used to combine genes from different organisms.

d A genetic engineering technique that 'switches off' a gene, i.e. stops it from being expressed.

4 Gene insertion requires a gene from one organism to be inserted into the chromosomes of another organism. Gene silencing involves an existing gene being prevented from expressing itself, which is achieved using one of a number of techniques.

Section 6.6

2 **a** Dispersed systems are kinetically stable mixtures in which one phase is dispersed throughout another phase, with the two phases being largely immiscible.

3 **a** Solid particles are dispersed throughout a liquid, e.g. cornstarch in water, hot chocolate.

b Droplets of one liquid are dispersed throughout another liquid of different polarity, e.g. mayonnaise, cold cream.

c Gas bubbles are dispersed throughout a liquid, e.g. whipped cream, beer head.

4 An emulsifier will be needed, as will a means of creating very small droplets of the dispersed phase. A stabilizer might also be added.

6 **a** emulsifier **b** coagulant **c** emulsifier

d foam **e** foam **f** foam

g emulsifier

7 Stabilizers are food additives designed to keep two normally immiscible ingredients in a homogeneous state after mixing.

Section 6.7

1 Free radicals are involved in all steps of oxidative rancidity.

2 Propagation is a free-radical chain reaction in which free radicals are propagators for the addition of oxygen to a fatty acid, RH, to form lipid hydroperoxides, ROOH.

$R\cdot + O_2 \rightarrow ROO\cdot$

$ROO\cdot + RH \rightarrow ROOH + R\cdot$

Section 6.8

2 EDTA is a chelating agent that forms a complex with divalent metal ions. EDTA binds to these metal ions through four carboxyl and two amine groups, via a complex that occurs due to the formation of coordinate bonds.

3 Radical scavengers (quenchers): phenol (hydroxyl and a benzene ring)

Antioxidants: amine group in EDTA

Reducing agent: alcohol, ketone and ether linkage in vitamin C

Section 6.9

1 A molecule that is not identical to its mirror image. Enantiomers are non-superimposable mirror images of one another.

3 Enantiomers that rotate plane-polarized light clockwise are dextrorotatory, +(d). Those that rotate it to the left are levorotatory, –(l).

4 This system is based on the arrangement in space of parts of the molecule, with glyceraldehyde as the reference molecule.

5 **a** This system labels the chiral carbon according to the priority of the substituents around it.

6 R-glyceraldehyde

7 Rotates light to the right, hence d(+) serine.

Anticlockwise CORN arrangement, hence L-serine.

Anticlockwise order of decreasing priorities, hence S-serine.

8 Different enantiomeric forms have different tastes, aromas and biological activity

9 The presence of both enantiomers in a food product containing carvone may indicate a synthetic origin and/or contamination by another product.

Section 6.10

2 The more extensive the conjugation is in a food, the less energy or the longer the wavelength of light that is absorbed.

3 **a** It is an anthocyanin due to its flavonoid skeleton.

b Due to its extensive conjugation, tangeritin would be expected to be coloured.

4 When the substituent changes on the flavonoid skeleton, different anthocyanidins are formed, each able to absorb different wavelengths of light.

5 **a** Carotene has a long non-polar group; xanthophylls have several polar hydroxyl groups.

b Carotene: β-carotene found in carrots

c Both are fat-soluble.

7 Porphin: a planar heterocyclic unit which has around its perimeter a cyclic chain of sp^2 hybridized carbon atoms.

Macrocycle: a ring that contains a large number of atoms, usually classified as more than nine atoms.

Porphyrin: occur when porphins have substituents in position 1 to 8 of the macro cycle ring.

8 Chlorophyll has a magnesium ion in the central position of the macrocycle ring. Heme has an iron ion in the central position of the macrocycle ring. Without these metal ions, the properties of these biological molecules would not be the same.

Chapter 7 Further organic chemistry

Section 7.1

1 **a** an electron-rich, nucleus loving species

b an electron-deficient, electron loving species

c a hydrocarbon cation, R_3C^+

d a reaction in which a double bond is replaced by a single bond when new atoms add to the carbon atoms on either end of a double bond

e the tendency of alkyl groups to 'push' the bonding

2 **a** The bromine will be decolourized.

b electrophilic addition

3 **a** W **b** Z **c** X **d** V **e** Y

4 **b** 1-bromo-2-hydroxyethane

6 **a** 1-bromo-1-methylcyclohexane

b 2-chloro-1-iodo-2-methylpropane

c 2-chloropentane and 3-chloropentane

Section 7.2

2 Sulfuric acid is an oxidizing agent, in addition to being a dehydrating agent.

3 **a** 75% H_2SO_4 at 100°C

b elimination

4 **b** II. Tertiary alcohols are the most reactive to dehydration as they produce the most stable carbocation intermediate.

Section 7.3

2 **d** nucleophilic addition

3 **a** $CH_3CH_2CH_2CHO + HCN \rightarrow CH_3CH_2CH_2CH(OH)CN$

b $CH_3CH_2COCH_3 + HCN \rightarrow CH_3CH_2C(OH)(CH_3)CN$

6 **b** The hydrazone product has a characteristic melting point which may be used to identify the original ketone.

Section 7.4

1 **a** $CH_3CH_2Cl + Mg \rightarrow CH_3CH_2MgCl$

b $CH_3CH_2CH_2CH_2Br + Mg \rightarrow CH_3CH_2CH_2CH_2MgBr$

2 Grignard reagents react readily with water to yield hydrocarbons.

4 **a** **i** bromoethane **iv** 2-butanol

b **iii** propylmagnesium chloride **iv** CO_2

5 Use of Grignard reagents allows the carbon chain length of organic compounds to be increased by one or several carbon atoms.

Section 7.6

1 Three atoms bond to each C atom in C_6H_6 forming a triangular planar structure. Four atoms bond to each C in C_6H_{12} to form a tetrahedral structure around each carbon.

2 Examples:

- C–C bond lengths in C_6H_6 are all equivalent and mid-way between those of single and double bonds.
- Only 3 isomeric form of $C_6H_4Br_2$ (not 4 as would be expected for 1,3,5-cyclohexatriene)

3 Benzene undergoes substitution reactions rather than addition reactions; heats of hydrogenation of benzene are lower than that expected for a cyclohexatriene structure.

4 **a** C_6H_5Br and $C_6H_5CH_2Br$

b **i** C–Br in bromomethyl benzene

ii C–Br in bromobenzene

5 Substitution reaction of C_6H_5Cl would be severely hindered due to repulsion of the OH^- nucleophilic by the delocalized electron cloud of the benzene ring. Reaction would be expected to be slow. Also, C–Cl bond in C_6H_5Cl is stronger.

Section 7.7

1 **a** The larger the K_a value and the smaller the pK_a, the stronger the acid.

b The larger the K_b value and the smaller the pK_b, the stronger the base.

4 **a** Y < W < X < Z

5 **a** S < Q < R < P

Section 7.8

3 Two molecules combine with the elimination of a small molecule (HCl).

5 Reactions of acid anhydride molecules are inefficient. Reactions of acyl chlorides are more 'atom' efficient.

Section 7.9

1 **a** **i** I Nitration, HNO_3/H_2SO_4

II Alkylation, $CH_3CH_2Cl/AlCl_3$

III Acylation, $CH_3COCl/AlCl_3$

IV Bromination, $Br_2/FeBr_3$

b I $H_2SO_4 + HNO_3 \rightarrow HSO_4^- + H_2O + NO_2^+$

II $CH_3CH_2Cl + AlCl_3 \rightarrow AlCl_4^- + CH_3CH_2^+$

III $CH_3COCl + AlCl_3 \rightarrow AlCl_4^- + CH_3CO^+$

IV $Br_2 + FeBr_3 \rightarrow FeBr_4^- + Br^+$

2 **a** The methyl group has an electron-donating inductive effect.

b The chloro group has an electron-withdrawing inductive effect.

3 Y < X < Z

4 **a** *meta*-nitro

b *ortho*-nitroaniline and *para*-nitroaniline

5 **b** **i** electrophilic substitution reaction; $Cl_2/FeCl_3$

c free radical reaction; Cl_2/UV

6 **a** OH is an activating *ortho–para* director.

b Br substituted in the first step is a deactivating *ortho–para* director.

c CH_3 is an activating *ortho–para* director.

INDEX

Page numbers in **bold** refer to key terms in **bold** type in the text. These terms are also defined in the 'terms and definitions' section of each chapter.