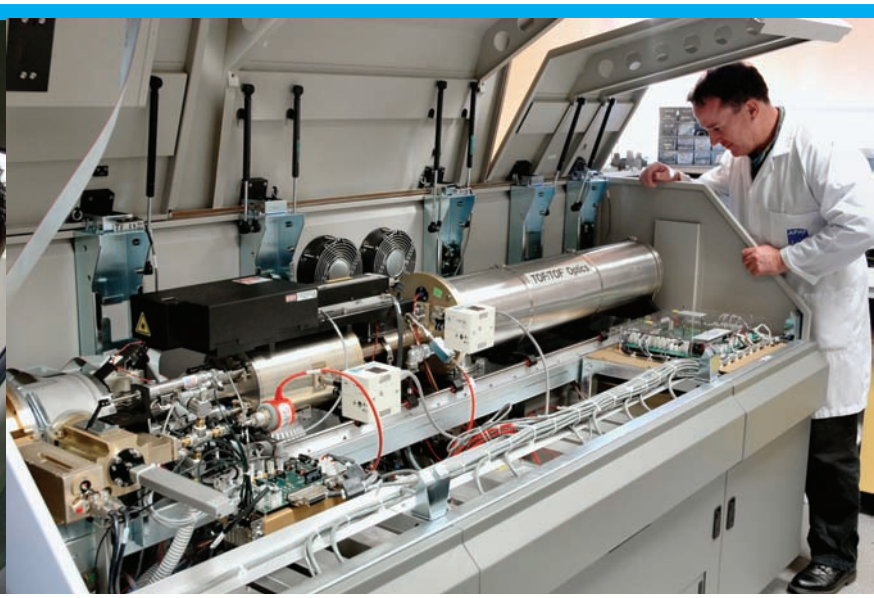
OPTION **3**

Forensic chemistry

Forensic chemistry is the branch of chemistry that provides information for use in courts of law or in public discussion and debate. For example forensic chemistry can provide information showing that:

- a sample of wine contained a substance known not to occur in the natural product (and so prove that the wine had been adulterated)
- the paint mark on one car came from another car (and so help establish that the first car was involved in a collision with the second one)
- that some hairs found at a crime scene came from a particular person (and so help identify the guilty party)
- that a sample of an athlete's urine contains prohibited substances (and so provide evidence of illegal use of performance-enhancing drugs)
- that certain compounds found in a waterway came from a particular factory (and so provide information about the polluting activities of that factory).

Forensic chemistry is a special branch of analytical chemistry—the branch of chemistry concerned with identifying the substances present in samples and how much of each there is. The first step in identifying a substance often involves



working out the broad class of compound the substance belongs to. That is where we shall start—first deciding whether a sample is organic or inorganic, and if organic performing tests to determine which class the substance belongs to. This will be done in the first part of Chapter 13.

Many forensic samples are biological in origin, so we will survey some major classes of biological (or biochemical) compounds such as carbohydrates (or sugars) and proteins. Chapter 13 will concentrate on carbohydrates while proteins will be treated in Chapter 14. Today there is considerable emphasis on using DNA to identify the person responsible for a biological sample found at a crime scene—samples such as blood, semen, saliva or hair follicle. Therefore in Chapter 14 we shall also discuss the structure of DNA and how it can be analysed for forensic purposes.

Samples for forensic analysis are often very small and frequently need to be analysed without being destroyed. A range of instrumental methods capable of handling very small samples without destroying much of them will be described in Chapter 15—mass spectrometry, atomic emission spectroscopy, gas–liquid chromatography and high performance liquid chromatography.

FORENSIC CHEMISTRY—METICULOUS HAND COLLECTION OF SAMPLES AND SOPHISTICATED INSTRUMENTAL ANALYSIS

A first step in a forensic investigation is to collect evidence (samples) for analysis from the crime scene in a thorough and systematic way without contamination (far left and above left). Then very sophisticated instruments such as the mass spectrometer in the photo (second left) are used to identify and trace the origins of such samples. Today DNA analysis is widely used not only to identify culprits but also to identify victims of disasters such as the December 2004 tsunami in S. E. Asia (above right): a simple swab of saliva from inside the cheek is all that is needed to collect enough DNA for accurate analysis

CHAPTER 13

Classifying and identifying substances (including sugars)

IN THIS CHAPTER

The job of a forensic chemist is to identify materials and to trace their origins. Often samples for identification are quite small—only a few milligrams. With all samples for analysis, but particularly with very small ones, an important consideration is avoiding contamination.

13.1 AVOIDING CONTAMINATION OF SAMPLES

There are three main sources of contamination of samples for analysis:

- *inadvertent addition of extraneous material* in the collection, handling and transporting of the sample before it reaches the laboratory; this is a reason for careful sealing of athletes' urine or blood samples immediately after collection, of the meticulous packaging of samples (in suitable bags, tins or specimen jars) from a crime scene, and the use of scrupulously clean bottles for storage of environmental water samples.
- *an unclean laboratory*: an old laboratory with flaking paint, corroding fittings and instruments, and dust-releasing floor or walls, free flows of dirty air through the laboratory and small amounts of stray chemicals from dirty benches or poorly maintained instruments can all lead to contamination of samples and so invalidate the analyses done.
- *a careless analyst*: not wearing suitable protective clothing such as laboratory coat, gloves, face mask and hair cover can result in the analyst 'shedding' contaminants such as dust, dandruff, sweat and associated dirt, and breath-born impurities into samples so producing erroneous results.

Careful handling of samples and meticulous attention to detail in performing analyses are essential for avoiding all of these three modes of contamination. This is true for all forms of analytical work, not just forensic work, especially when very small samples are involved.

However for forensic work there is an additional requirement and that is the validation of the *chain of custody* of the sample.

Chain of custody

For forensic purposes there must be witnesses to the collection of the sample and the sealing of the container into which it was placed (for example in collecting urine samples from athletes for testing for performance-enhancing drugs), and detailed documentation of the passage of the sample from the collection point to the laboratory where the analysis was to be performed, and then witnesses to the breaking of the seal of the sample and the actual performance of the analysis. This is to ensure that there was no tampering with the sample between collection and analysis or substitution of a different sample for the original one.

The importance of this aspect of sample handling was demonstrated in a highly publicised trial in California, USA, of a nationally famous sports star, O.J. Simpson, in 1997 for the murder of his estranged wife. The DNA evidence against him was very convincing, but his lawyers argued that there was sufficient laxness in the chain of custody of the sample for there to be some doubt about whether the sample analysed had not been contaminated or substituted for; they raised sufficient doubt in the minds of jury members that Simpson was acquitted.

One safeguard for ensuring the legitimacy of the sample being analysed is to provide, at the time of collection of the sample, an identical sealed sample to the person being investigated and so allow that person to arrange their own analysis of the sample. This happens in the routine testing of athletes for prohibited substances and in the blood testing of persons suspected of driving a car under the influence of alcohol or drugs.

13.2 THE NEED FOR ACCURACY

Accuracy in forensic chemistry is particularly important because faulty forensic evidence could:

- *convict an innocent person* and force them to spend many years in jail
- result in a *person or company having to pay massive fines* for offences such as polluting the environment (by apparently exceeding allowed emissions) or selling allegedly contaminated products, all on the basis of inaccurate analyses.

In addition, even if the overall conclusions of the forensic investigations are correct, sloppy or inconsistent work can allow a defence counsel to cast doubt over the validity of the whole forensic evidence because of one small fault in part of it.

The classic example of faulty forensic work in Australia in recent decades was the Azaria Chamberlain case of the 1980s. Baby Azaria's mother was convicted of murdering her daughter, partly on the forensic evidence that there was foetal blood in the family motor car. After the mother had spent several years in jail, the baby's jacket was found in circumstances that supported the mother's claim that a dingo had taken her baby. In the enquiry that followed new forensic examination

showed that the so-called foetal blood was actually adult blood contaminated with rust (consistent with the parents' original story of giving a ride to an injured adult). Lindy Chamberlain was eventually declared innocent and released.

Another aspect of accuracy is *completeness or thoroughness*. Sometimes there is nothing wrong with the forensic results as such, except that not all desirable forensic examinations have been done: early evidence points towards one person, a few forensic tests support those suspicions so the investigation is ended, whereas further forensic tests might have revealed evidence that would clear the original suspect or implicate another person. Forensic chemists need to be aware of the dangers of stopping an investigation before it is complete.

While accuracy is important in forensic chemistry, it is just as important in all branches of chemistry. Careless results by a research worker can send colleagues on wild goose chases following up the work or disproving it. Inaccurate analyses by quality control chemists can cause a chemical factory to discard batches of product or modify production procedures unnecessarily and at great cost. Poor experimental technique in synthesis can result in synthesised chemicals being contaminated or yields being very low. Faulty monitoring of discharges to air and water can lead to a factory inadvertently exceeding allowable limits and harming people in surrounding areas.

All chemists need to ensure that their work is accurate, thorough and reliable.

Ethical issues in a forensic investigation

Apart from avoiding contamination of samples and doing accurate analyses, forensic chemists must be aware of several ethical issues associated with their work. They must:

- *ensure that the analyses they do are accurate, complete and objective*, and that they do not just find a piece of information that will help build a case against a prime suspect and go no further
- *ensure that samples from all suspects are analysed* and not stop work when one sample appears to incriminate one particular suspect
- *report their scientific findings objectively* without exaggerating their accuracy or reliability and not obscure uncertainties in the results with large amounts of interpretation and opinion
- *not cover up findings* that cast doubt upon the guilt of the accused (such as by not reporting that certain analyses with unfavourable results were performed)
- *present their findings in clear and simple language* that non-chemists can understand and in such a way that juries can assess them fairly.

Overall, forensic chemists must act as unbiased scientists reporting their objective findings and not appear to be part of the prosecution team building a case against a particular suspect. This will sometimes involve resisting pressure from law enforcement officers to produce a particular result (when the analysis is inconclusive) or to confine their analyses to selected samples only.

13.3 ORGANIC OR INORGANIC?

One of the first tasks of an analyst is to determine whether a sample is an organic or inorganic substance, because the methods of analysis differ quite significantly for these two classes.

Originally the term organic compound was used for substances that were present in living matter. It had been thought that some 'life force' was required to

make them. However early in the nineteenth century it was shown that organic compounds could be made in a chemical laboratory. Today

Organic compounds are compounds of carbon (with the exception of CO, CO_2, and carbonates, hydrogen carbonates and cyanides of metals).

Some organic compounds that we have already considered are the alkanes and alkenes of Sections 9.9 and 9.10 of *CCPC*, ethanol of Section 1.16, other alkanols, alkanoic acids and esters of Sections 5.16 to 5.19, and the haloalkanes of Section 7.10. Most of the other compounds dealt with in this book so far have been inorganic ones.

Organic compounds almost always contain hydrogen as well as carbon (exceptions being the CFCs of Section 7.10). Other elements commonly present in organic compounds are oxygen, nitrogen, halogens (F, Cl, Br, I) and sulfur.

Distinguishing between organic and inorganic compounds

The simplest method of distinguishing between organic and inorganic compounds is to *heat a sample of the compound in air*. If the compound burns or reacts with air or decomposes to leave no residue, the compound is almost certainly organic. A quantitative analysis for carbon and hydrogen would confirm it.

13.4 IDENTIFYING SOME CLASSES OF INORGANIC COMPOUNDS

If the substance to be identified turns out to be inorganic (by the test from Section 13.3), a useful starting point for identifying the substance is to determine whether it is acidic, basic or neutral.

Determining acidity or basicity of a sample

If the sample exists as a solution, this can be done by using a suitable pH paper or a pH meter (Sections 4.2, 4.18 and 4.19). If using pH paper, the forensic chemist would put a drop of the solution onto the paper rather than dipping a piece of paper into the sample (so as not to contaminate the sample). Similarly drops of indicator would not be added to the sample. Dipping the glass electrode of a pH meter into the solution would not contaminate it. The pH tells whether the substance is acidic (pH < 7), neutral (pH = 7) or basic (pH > 7).

If the sample was a solid, the chemist would try to dissolve a small portion of it in water: this in itself provides useful information (is the compound soluble or insoluble?). If the substance is soluble, the pH of the solution would be determined as just described.

If the sample is insoluble, its acidity or basicity would be determined first by adding a drop of sodium carbonate solution to a small portion of the solid and heating it and observing whether any bubbles of gas form. If so, the substance is acidic (reacts with carbonate to form carbon dioxide).

If there was no reaction with sodium carbonate, drops of dilute nitric acid would be added to a fresh portion of the sample and the mixture warmed. If the sample dissolved, then it must have reacted with nitric acid (since it had already been established that it was insoluble in water) and so it must be a base (such as an insoluble oxide, hydroxide or carbonate).

One source of inorganic compounds that often has to be examined by forensic chemists is soil.

13.5 ANALYSIS OF SOILS

Soil is a complex mixture of inorganic materials (such as clay, silt, sand and gravel), decaying organic matter (called *humus*), water, air and living organisms.

Soil is formed by the combination of weathering of rocks, sedimentation, and decay of living matter. The solids in soils are typically 5% organic matter and 95% inorganic material. The amount of water present in soils depends upon their porosity and texture.

The order of size of particles in soils is clay < silt < sand < gravel.

Soils differ from one another in the proportions of these inorganic materials present. In addition the chemical composition of these materials in soils varies from place to place. Consequently an analysis of soil found at a crime scene can provide forensic chemists with useful evidence, such as the source of the sample.

Forensic chemists often need to identify the source of a sample of soil found at a crime scene. If the sample is greater than a couple of grams, the following tests can be performed:

- colour (soils differ in colour from red to brown to black)
- texture (depends upon the relative amounts of sand, clay, humus, etc and varies from place to place)
- appearance when shaken with water (the different components of soil settle out at different rates depending on their density—in the order, gravel, sand, silt—with clays usually staying in suspension and organic matter often floating on top)
- pH (universal indicator or an indicator paper is dipped into the solution of the previous test—after it has settled).

If these tests do not identify the source of the soil, or if the sample available is well less than a gram, then forensic analysis of soils is generally done by detecting the presence of (and determining the amounts of) the *less common* elements in soil. This is usually done by either atomic emission spectroscopy (Sections 6.17 and 15.12) or atomic absorption spectroscopy (Section 6.18). If a data base of soil compositions is available the source of a particular sample can be identified.

Analysis of soil from a muddy footprint left in the room in which a crime was committed may show that the perpetrator had recently been in (or came from) a particular locality; this may help narrow down the list of suspects. Alternatively analysis of the soil in the grooves of a suspect's shoes may show that she or he was recently at the crime scene.

Other inorganic material

Other inorganic materials that can be analysed to provide forensic evidence include metal and alloy fragments, residues from certain building materials such as plasterboard and concrete, pigments from paints and various types of glass. Determination of concentrations of metal ions, particularly unusual ones, can provide information about the origin of such samples. Forensic chemists would generally use atomic emission spectroscopy (Sections 6.17 and 15.12) to perform such analyses.

Inorganic analysis, particularly for uncommon metal ions, often provides useful information in forensic investigations.

13.6 IDENTIFYING CLASSES OF ORGANIC COMPOUNDS

Having determined that a compound is organic (i.e. a carbon compound) from the test in Section 13.3, the analyst's next job is to work out which class of carbon compound the unidentified substance belongs to.

So far in this course we have met several different classes of carbon compound as summarised at the start of Section 13.3. There are tests that we can perform to determine the class of compound an unidentified sample belongs to. Tests relevant to the main classes we have examined so far are given in Table 13.1. The classes of compound are listed in a convenient sequence for testing.

TABLE 13.1 Tests for distinguishing between some classes of carbon compounds

Class of compound[a]	Test and the result if the compound belongs to the class
alkanoic acid (or any carboxylic acid)	Add drops of sodium carbonate solution to a sample; bubbles of colourless gas form (CO_2). This is a normal reaction of an acid with a carbonate (Section 6.12).
alkanol (or any alcohol)	Dry a sample carefully with granules of calcium chloride, then add a small piece of sodium; bubbles of colourless gas form (H_2). (Alkanoic acids react in the same way but would have been identified by the previous test.)
alkene	In the absence of strong u.v. light, add drops of bromine in an organic solvent such as chloroform or hexachloroethane; the bromine solution loses its colour. Aqueous bromine can also be used: it is decolorised by alkenes (Section 1.4). (Many alkanols also react with aqueous bromine so if using this reagent it is necessary to perform the sodium test first to eliminate the possibility of the compound being an alkanol.)
alkane or ester	If all the above tests are negative, the compound is an alkane or ester. Add drops of a solution of bromine in hexachoroethane and expose the mixture to u.v. light; the colour slowly fades (Section 1.3)
ester	If the compound has a pleasant smell, it is probably an ester. We really need to measure its infrared spectrum (see below) to confirm that it is an ester

a *The only classes of compounds considered here are those that are part of the NSW HSC core syllabus.*

Alkanol plus sodium reaction

Sodium metal reacts with an alkanol to form hydrogen gas and what is called an **alkoxide ion**:

for example for ethanol:

$$2CH_3{-}CH_2{-}OH(l) + 2Na(s) \longrightarrow H_2(g) + 2CH_3{-}CH_2{-}O^- + 2Na^+$$
(solution in ethanol)

This particular alkoxide is called *ethoxide*; it forms by removal of an H^+ from ethanol.

An example will illustrate the use of these tests.

Example 1

Describe how you would distinguish between members of the following groups of substances. All the compounds listed are colourless liquids. Write equations for all reactions which occur.

a hexane and 2-hexene

b ethanol, propanoic acid and ethyl ethanoate (acetate)

a Add drops of reddish-brown bromine solution to samples of each of the liquids. With one liquid the whole mixture takes on a pale brown colour (that is, no reaction), whereas with the other liquid the colour disappears. The latter liquid is 2-hexene, because alkenes decolorise bromine solutions whereas alkanes do not.

$$CH_3-CH_2-CH_2-CH=CH-CH_3 + Br_2 \longrightarrow CH_3-CH_2-CH_2-\underset{\displaystyle Br}{\underset{|}{CH}}-\underset{\displaystyle Br}{\underset{|}{CH}}-CH_3$$

b We have here an alkanol, an alkanoic acid and an ester. Hence add drops of sodium carbonate solution to samples of each liquid. With two of the liquids there is no reaction, while with the third liquid bubbles of colourless gas form. This third liquid is propanoic acid. To samples of the other two liquids add granules of calcium chloride to dry them then add a small piece of sodium to each dried sample. Bubbles of colourless gas (H_2) form with one liquid but not with the other. The one that formed bubbles is ethanol, so the other must be ethyl ethanoate.

Equations for the reactions involved are:

$$2CH_3-CH_2-COOH + Na_2CO_3(aq) \longrightarrow 2CH_3-CH_2-COO^- + 2Na^+ + CO_2(g) + H_2O(l)$$

$$2CH_3-CH_2-OH + 2Na(s) \longrightarrow 2CH_3-CH_2-O^- + 2Na^+ + H_2(g)$$

School laboratory versus forensic chemist's laboratory

The tests described in Table 13.1 can be used in a school laboratory to identify the class a particular compound belongs to (providing a sample of at least several milligrams is available). However they are not sensitive enough for forensic chemists who would normally use **infrared (i.r.) spectroscopy** for this purpose.

Infrared (i.r.) spectroscopy

This is a technique in which a sample to be analysed is irradiated with i.r. radiation of changing wavelength, with the amount of radiation passing through the sample being measured. The percentage of the incident radiation that is absorbed by the sample is recorded as a function of wavelength of the incident radiation (either on a pen recorder in older-style instruments or on a computer screen in newer types). Such a graph is called an *infrared spectrum*. Several such i.r. spectra are shown in Figure 13.1.

Different functional groups absorb i.r. radiation at different wavelengths, and so an i.r. spectrum tells the analyst which functional groups are in the sample. In

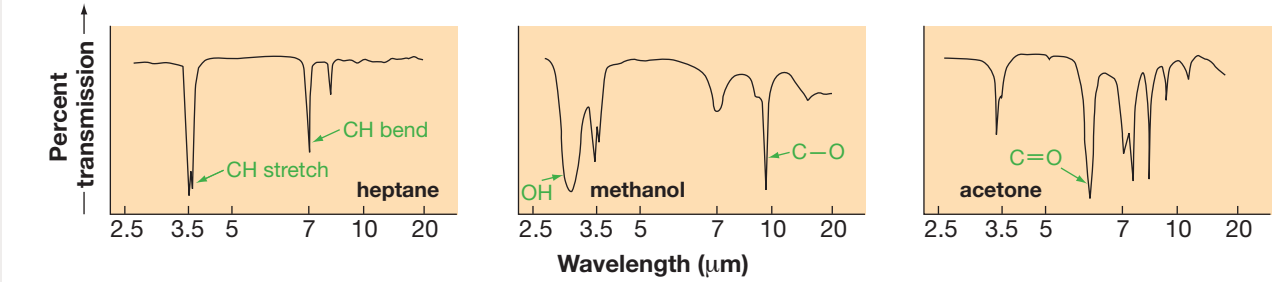


FIGURE 13.1
Infrared spectra of heptane, methanol and acetone. The functional groups or structural features that lead to major absorption peaks are shown beside the peaks

many cases, comparison of the measured spectrum of an unidentified compound with standard spectra of known compounds provides clear identification of the compound. Infrared spectroscopy is a quick and sensitive method of analysis and is widely used by forensic chemists.

Identifying metal ions in samples

A similar situation exists for metal ions. In Chapters 6 and 8 we explored a range of test tube reactions that could be used in a school laboratory to identify the cation or cations present in a sample. In Chapters 5 and 6 we looked at some methods of quantitative analysis (volumetric and gravimetric) that can be performed in school laboratories. These qualitative and quantitative methods require samples of at least 1 to 10 mg. Forensic chemists often have samples that contain only microgram amounts of key cations. Hence for analysis of cations forensic chemists generally use instrumental techniques such as atomic emission spectroscopy (Sections 6.17 and 15.12) and atomic absorption spectroscopy (Section 6.18). These techniques are much more sensitive than school laboratory methods and provide more accurate results more rapidly (though at much greater cost).

Exercises

1 Three colourless liquids (at room temperature), A, B, and C, were tested to determine whether they were organic or inorganic. When gently heated in a stainless steel spoon-type spatula over a small Bunsen burner flame, C burnt with a clearly visible flame while A and B just vaporised. When analysed quantitatively for C, H and Cl, they gave the following results:

A: 13% C; 0% H, 57% Cl

B: 0% C; 0% H; 83% Cl

C: 68% C, 14% H; 0% Cl

Which of these compounds is(are) organic and which inorganic? Justify your choices.

2 Write chemical equations for the reactions involved in the tests described in Table 13.1. Use the following compounds: acetic acid, methanol, 2-hexene, pentane and ethyl acetate.

3 In the laboratory, how would you distinguish between the members of the following pairs of substances? Describe fully what you would do, and what you would observe with each substance. Write equations for all reactions that occur.

a hexane and 1-hexene
b pentane and 2-butanol
c 2-propanol and propanoic acid
d glucose and sodium sulfate
***e** 1-pentene and methanoic acid
***f** 1-pentanol and 2-hexene

4 Describe tests you would use to identify the members of the following sets of compounds. Write equations for all reactions involved.

- **a** 2-pentene, 1-hexanol, propanoic acid
- **b** heptane, 3-heptene, ethyl propanoate
- ***c** 1-heptene, 2-propanol, hexane

5 Two compounds, P and Q, had molecular weights between 65 and 75. After careful drying, P reacted with sodium to form bubbles of colourless gas while Q did not. Q rapidly decolorised drops of a solution of bromine in hexachloroethane while P decolorised this solution only very slowly. Identify the two compounds and explain your reasoning.

Often the sample a forensic chemist is required to identify will not be a simple organic compound such as one of those in Table 13.1, but rather will be one of the many complex compounds that occur in living matter. Let us consider three major classes of such compounds, namely

- carbohydrates
- proteins
- DNA (deoxyribonucleic acids).

We shall look at the structures of these compounds and see what information a forensic chemist can get from them.

13.7 SIMPLE CARBOHYDRATES

Carbohydrates are compounds of C, H and O, having the general formula $C_x(H_2O)_y$ where x and y may be the same or different.†

These compounds are also called **sugars**. In everyday language the word 'sugar' is used for one particular compound, sucrose; however in science 'sugar' is a general name used for many carbohydrates—for the so-called mono- and disaccharides (explained below). Glucose and sucrose are carbohydrates: glucose has the formula $C_6H_{12}O_6 = C_6(H_2O)_6$, while sucrose is $C_{12}H_{22}O_{11} = C_{12}(H_2O)_{11}$.

The simplest carbohydrates have three to six carbon atoms per molecule. The most common ones are those with five or six carbon atoms; they are called *pentoses* and *hexoses* respectively. Three important hexoses are glucose, fructose and galactose; all are isomers of $C_6H_{12}O_6$. The structural formulae of these three sugars are shown in Figure 13.1 in two different forms. In the left-hand structures all the atoms are shown. This gives rise to a rather cluttered structure. Hence we often omit the C atoms as in the right-hand structures (and as was done in Section 1.13 and Figure 1.7 for glucose). In such diagrams *there is a C atom at every junction of bonds* (or in different words, at every apex of the structure). These 'unlabelled carbon' structures are widely used in chemistry, particularly when cyclic (ring) compounds are involved.

The five- and six-membered rings are not strictly planar—they are buckled as was shown in Section 1.13 and Figure 1.7—but for simplicity we often draw

† This is the definition given in the NSW Board of Studies Chemistry Stage 6 Syllabus 2002. It is not strictly correct because there are many substances considered to be carbohydrates which do not conform to this formula, notable examples being deoxyribose (involved in DNA) and Vitamin C (ascorbic acid). A more accurate definition is that carbohydrates are polyhydroxylated aldehydes and ketones or compounds that hydrolyse to these. However aldehydes and ketones are not part of this syllabus, so we probably need to stick with the less accurate definition.

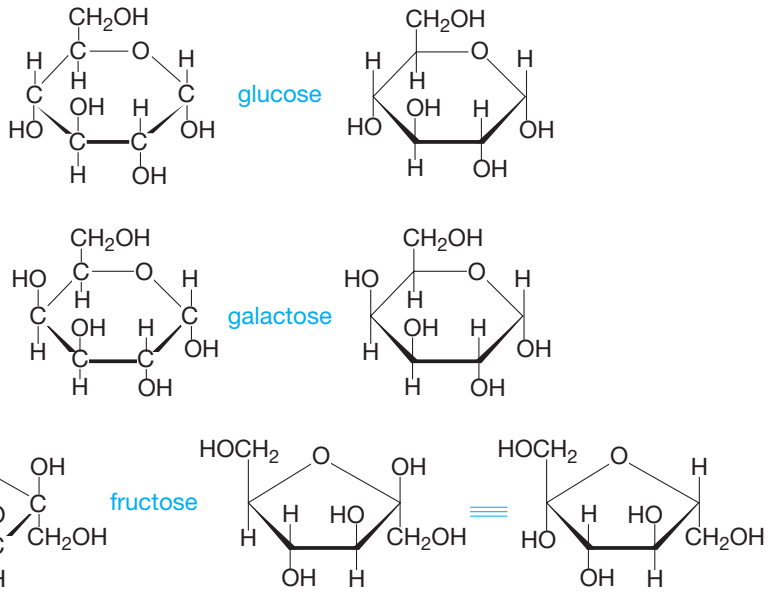


FIGURE 13.2
Structures of three common monosaccharides. The left-hand structures show all the atoms in the molecules; in the right-hand ones the carbon atoms are not labelled. The right-hand structure for fructose is just the middle structure flipped through 180°

them planar as in Figure 13.2. We use thick and wedge-shaped bonds in these rings to indicate that the plane of the ring is perpendicular to the paper. The H atoms and OH and CH_2OH groups attached to the ring are above and below that plane (though not actually perpendicular to the plane). The positioning of these groups above or below the plane is important, because it determines which compound the structure represents. Glucose and galactose differ only in the location of one OH group (Fig. 13.2). Despite this they are distinctly different compounds with noticeably different properties.

It is important that you carefully study a molecular model of glucose (or at least look closely at the photographs in Figure 1.7) to ensure that you fully understand the general structure of the six-membered ring hexoses. Similar models of fructose are shown in Figure 13.3. Fructose is a hexose that contains a five-membered ring.

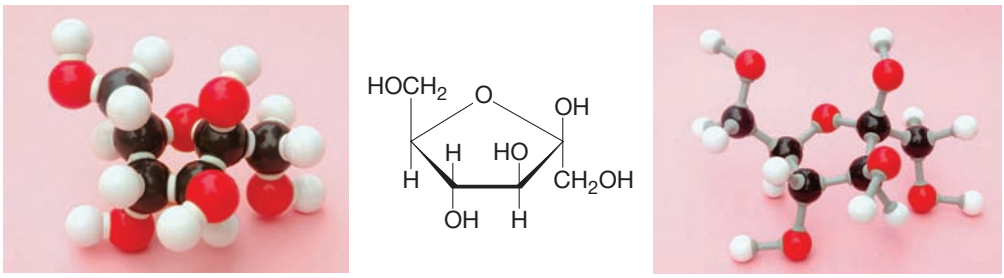
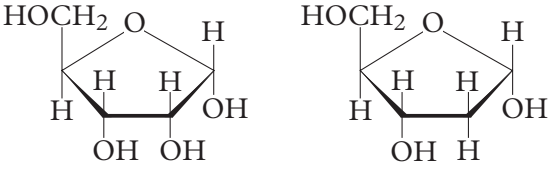


FIGURE 13.3
Space-filling (left) and ball-and-stick (right) molecular models of fructose along with the structure drawn in a similar orientation. Black is C, white is H and red, O

An important pentose is ribose, Structure (1). It is a basic unit in ribonucleic acid, RNA. Deoxyribose, Structure (2), is a reduced form of ribose. 'Deoxy' means without oxygen; the molecule has one less oxygen than ribose. Deoxyribose is one of the structural units in deoxyribonucleic acid, DNA, which will be discussed starting at Section 14.12.



(1) (2)

The simple C_3 to C_6 sugars are called *monosaccharides*. The word means 'one sugar'. We shall see in Section 13.9 that two simple carbohydrate molecules can join together to form what is called a *disaccharide* (two sugars). Many sugar molecules can join together to form what is called a *polysaccharide* (many sugars). These di- and polysaccharides react with water, with a suitable catalyst, to form single sugars or monosaccharides.

Recall from Section 5.5 that reaction with water is called *hydrolysis*.

Hence we have the definition:

Monosaccharides are carbohydrates that cannot be hydrolysed into two or more simpler sugars.

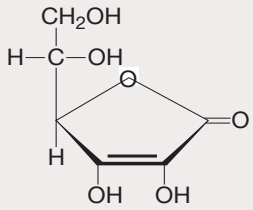
Exercises

6 Upon analysis it was found that a particular compound contained 39.8% carbon and 6.7% hydrogen with the balance being oxygen. Calculate its empirical formula. Could this compound be a carbohydrate? Give your reason.

7 What is the **(a)** molecular formula and **(b)** empirical formula of:

i ribose (Structure 1 in Section 13.7) and **ii** deoxyribose (Structure 2)?

8 Ascorbic acid (vitamin C) has the structure:



a What is its **(i)** molecular formula **(ii)** empirical formula?
b What functional groups are present?
c By the definition in Section 13.7 is ascorbic acid a carbohydrate? Explain.

9 Covalent molecular compounds, and hence most organic compounds, are generally either liquids or soft solids with low melting points and are usually insoluble in water. On the other hand, carbohydrates are hard solids that can be heated to quite high temperatures and then they decompose rather than melt, and they are readily soluble in water. Account for these differences in properties.

13.8 DISACCHARIDES

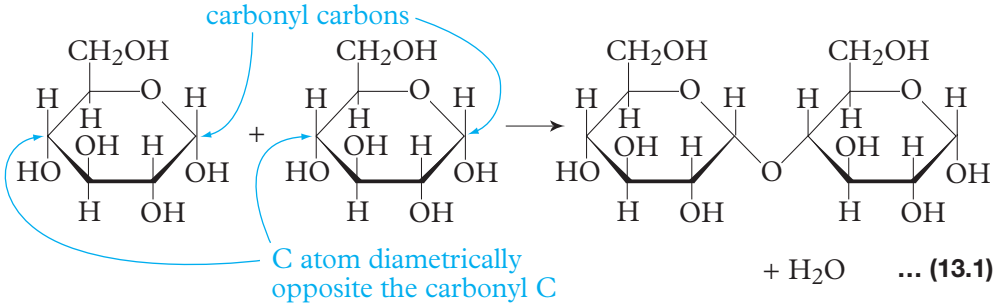
Disaccharides are carbohydrates that can be hydrolysed into two simple sugars.

Common disaccharides are *sucrose* from cane sugar, *maltose* from malt and *lactose* from milk.

Maltose forms when two glucose molecules join together. The union is between the OH on what is called the *carbonyl carbon* of one molecule (the C that has an O and an OH attached to it†) and the OH of the diametrically opposite C of the other molecule as in Equation 13.1. A molecule of water is eliminated.

† more about this in Section 13.10

glucose + glucose ⟶ maltose + water

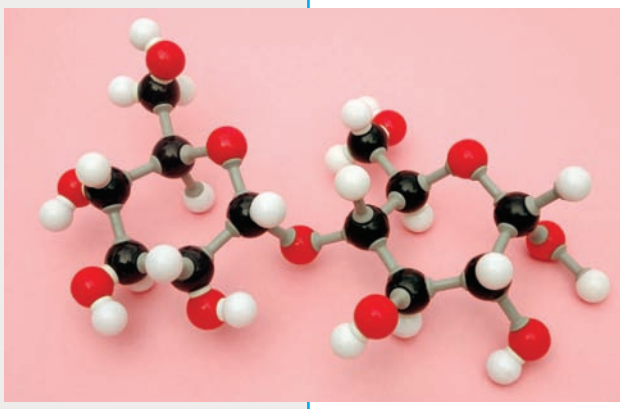


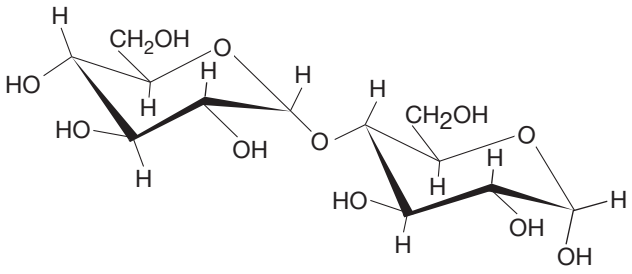
+ H_2O **... (13.1)**

As will be explained in Section 13.10, there are two isomers of glucose, called α- and β-glucose. The glucose in Figure 13.2 and Equation 13.1 is α-glucose; it has the OH on the carbonyl carbon on the opposite (above/below) side of the ring from the CH_2OH group (OH below the ring, CH_2OH above the ring). We say that maltose has an α *linkage*; the linkage is always from the carbonyl OH of one glucose molecule to the particular alcohol OH of the other as shown above. Although it would be possible to join the left-hand glucose to one of the other alcohol groups on the right-hand molecule, that would not be maltose.

Figure 13.4 shows a ball-and-stick molecular model of maltose, along with the structural formulae in the more accurate buckled-ring form oriented in a similar way to the model.

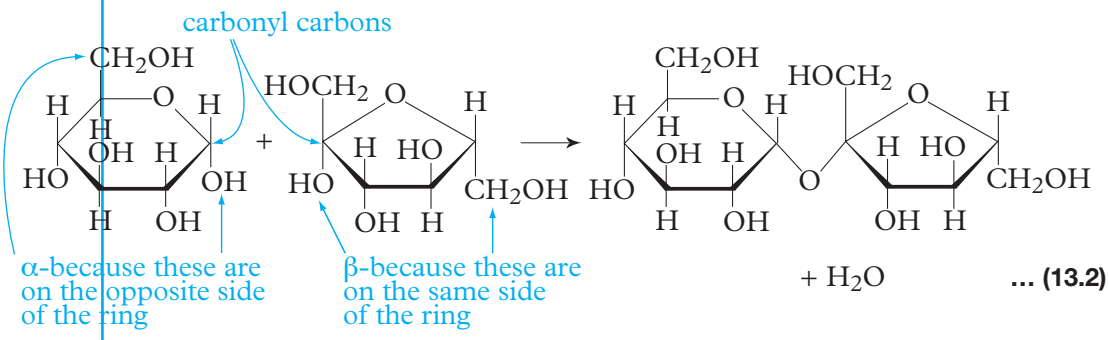
FIGURE 13.4
A ball-and-stick model of maltose, showing the α-linkage and the way the actual structure relates to the structural formula we draw





Sucrose is a disaccharide formed by linking the carbonyl carbons of α-glucose and β-fructose:

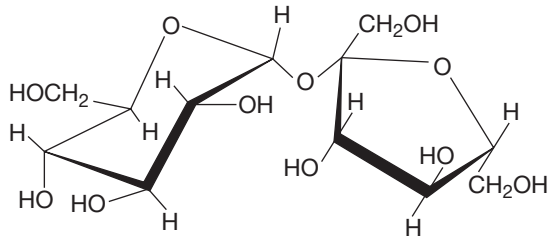
α-glucose + β-fructose ⟶ sucrose + water



+ H_2O **... (13.2)**

The fructose structure used here is the right-hand structure from Figure 13.1—the one with the carbonyl carbon on the left.

Figure 13.5 shows a ball-and-stick model of sucrose along with its molecular structure oriented in approximately the same way as in the model.



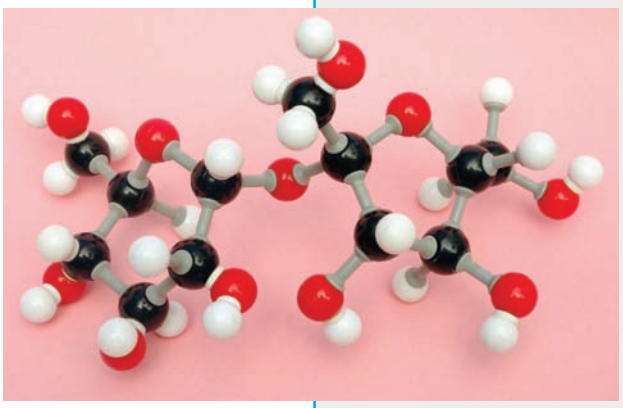


FIGURE 13.5
A ball-and-stick model of sucrose with its structure for comparison

Condensation reactions

A reaction in which two molecules join together with the elimination of a small molecule, such as water, is called a **condensation reaction**.

Formation of a disaccharide from two monosaccharides is a condensation reaction. Similarly a reaction in which many molecules join together end-to-end with the elimination of a small molecule between each pair of joining molecules is also called condensation. This happens when polysaccharides are formed (next section). It is also the way that condensation polymers such as nylon and polyester were formed in Section 1.12.

Disaccharides can be *hydrolysed* into two monosaccharides. In the laboratory we hydrolyse disaccharides by gently heating them in aqueous acid solution.

With aqueous H^+:

$$\left.\begin{array}{l}\text{maltose} + \text{water} \longrightarrow \text{glucose} + \text{glucose} \\ \text{sucrose} + \text{water} \longrightarrow \text{glucose} + \text{fructose}\end{array}\right\} \quad \ldots (13.3)$$

In living matter this is brought about by enzymes (biological catalysts, p. 298 *CCPC*).

Exercises

10 Write an equation using molecular formulae for a condensation reaction that can occur between glucose and fructose.

11 Lactose is a disaccharide formed by making a link from the carbonyl OH of the β isomer of galactose to the same OH on glucose as is used for maltose. Draw the structure of lactose. α-galactose is shown in Figure 13.2.

12 **a** Draw structures of two disaccharides that could be formed between ribose (Structure (1) in Section 13.7) and glucose.

b Write an equation (using molecular formulae) for the reactions forming these disaccharides. Why does one such equation describe both reactions?

13 Write equations using molecular formulae for the two reactions given in Equation 13.3.

13.9 POLYSACCHARIDES

Polysaccharides are carbohydrates that consist of a large number of monosaccharide molecules joined to one another in a long string.

The linkage is the same as in disaccharides. The main polysaccharides are:

- *starch* and *cellulose* from plants and
- glycogen from animals.

All three are made from glucose units, but these units are joined together in slightly different ways.

Polymerisation of glucose to cellulose was briefly mentioned in Section 1.10. The terms polymerisation, polymer and monomer were defined in Sections 1.12 and 1.13. Starch, cellulose and glycogen are polymers. The monomer from which all three are made is glucose.

Many of our staple foods have a high starch content

The difference between cellulose on the one hand, and starch and glycogen on the other, is that they use different isomers of glucose.

Cellulose consists of long chains of β-glucose units joined together. This is shown in Figure 13.6. In (a) the glucose units have been drawn using the planar rings we have mainly been drawing. However, as already mentioned, six-membered rings are not strictly planar. (b) shows cellulose using the accurate geometry of these rings. For simplicity the H and OH groups have been omitted. The CH_2OH group is shown; remember that in β-glucose this CH_2OH is on the same side of the ring as the OH on the carbonyl carbon. Because of this β-arrangement, it is necessary to invert every second glucose unit in the cellulose structure.

The structures in Figure 13.7 may help you convert from planar rings to 'bent' rings. (a) is the planar structure we have been using, except that the H atoms directly attached to C atoms are omitted. This is another convention that chemists often use (in addition to not labelling C atoms).

At every apex of a structure there is a C atom and attached to that C atom is the number of H atoms needed to bring the number of bonds up to the normal four for C atoms.

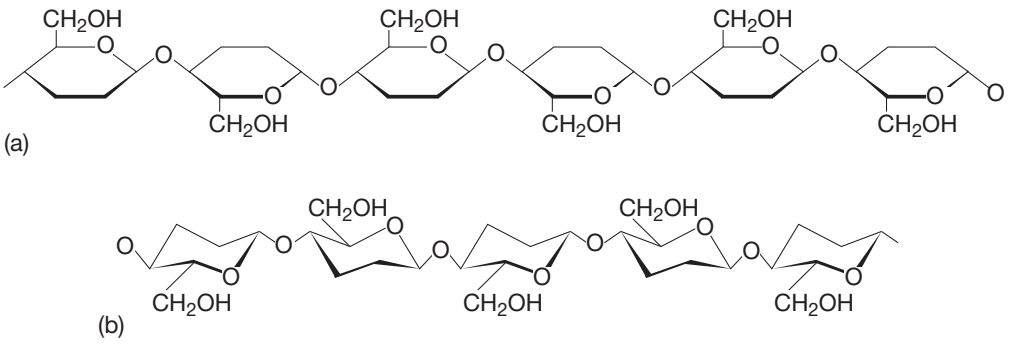


FIGURE 13.6
The structure of cellulose: (a) showing glucose rings as planar, (b) showing the geometry of the glucose rings more accurately. For clarity OH groups on the rings have been omitted

This convention is used in both structures in Figure 13.7.

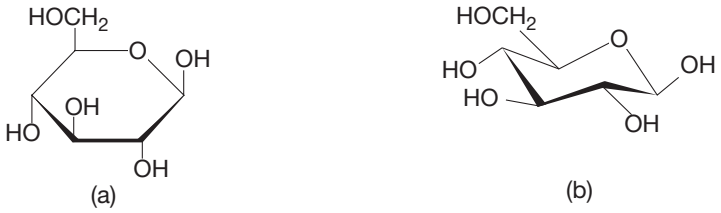


FIGURE 13.7
Different ways of drawing β-glucose: (a) using a planar ring, (b) using more accurate ring geometry

Structure (b) shows the accurate geometry of the β-glucose molecule: note that two OHs are 'below' the ring and two OHs and the CH_2OH are 'above' the ring. This is best seen by handling models: the photographs of molecules in Figure 1.7 (p. 25) may help.

As Figure 13.6 shows, the use of β-glucose results in a fairly linear molecule. Typically there are 2000 to 4000 glucose units per cellulose molecule. Because of all the OH groups and the hydrogen bonding they can produce, these linear chains intertwine with one another and stick together to form fibres. Cellulose is an insoluble fibrous substance. It forms the structural material of cells and plants (including trees).

Starch and *glycogen* consist of α-glucose molecules joined together as shown in Figure 13.8. When we draw planar rings for glucose as in (a), the starch

Rayon (or viscose or viscose rayon) is a common semi-synthetic fibre for fabrics; it is reconstructed cellulose. Wood pulp or low grade cotton is processed chemically then spun into fibres suitable for making textiles

molecule looks linear. This is because we have distorted the geometry around the linkage between pairs of rings. When we draw the proper shape for the glucose rings and use the proper angle for the C−O−C bond (about 109°), we find that our structure curls around as shown in Figure 13.8(b). This is the true situation in starch and glycogen. The molecules are curled up into globular shapes; they are quite different from the linear and fibrous appearance of cellulose. The α linkage in starch and glycogen is the same as that in the maltose model in Figure 13.4.

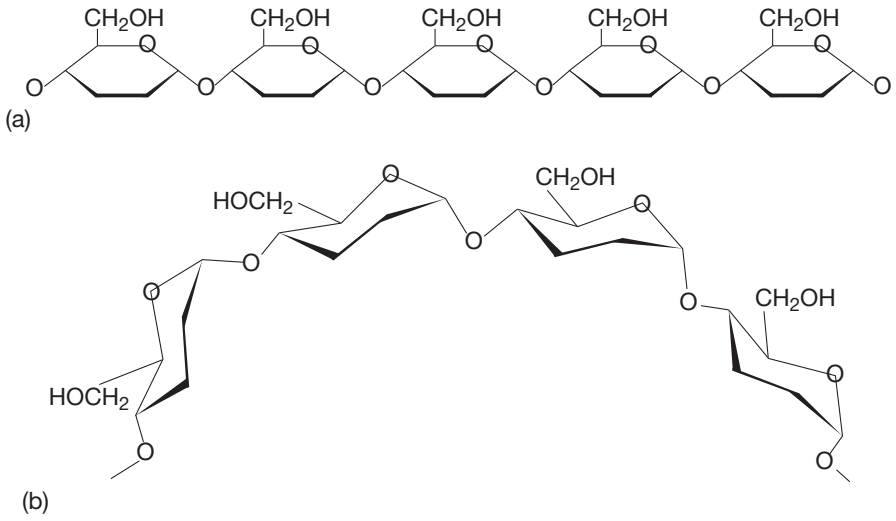


FIGURE 13.8
The structure of starch: (a) using planar rings for glucose, (b) using more accurate geometry. For clarity ring OH groups have been omitted. In this unbranched structure each glucose unit is joined to two other glucose units. Glycogen has these same α-glucose linkages.

There are actually two forms of starch. When starch from plants is warmed in water, about 20% of it dissolves. This soluble starch has a somewhat different structure from the other 80% of insoluble starch. *Soluble starch is called amylose; insoluble starch is called amylopectin.*

Amylose is an unbranched-chain molecule (see Fig. 13.8) with 60 to 300 glucose units per molecule. Amylopectin is a branched-chain molecule with about 300 to 6000 glucose units per molecule. Chain branching occurs when one glucose molecule condenses with three other glucose molecules instead of with the normal two as in Figure 13.8. Condensation with three other glucose units occurs through the 'normal' alcoholic OH to one molecule, through the carbonyl OH to a second molecule and through the terminal OH (that is, the CH_2OH) to a third molecule (Fig. 13.9(a)). Chain branching is shown schematically in Figure 13.9(b).

Glycogen or animal starch is also highly branched. *Glycogen has a lower molecular weight than amylopectin but is more highly branched.*

In summary:

- Cellulose is made of β-glucose units (Fig. 13.6(b)), while starch and glycogen use α-glucose (Fig. 13.8(b)).
- Cellulose and amylose have unbranched molecules while amylopectin and glycogen have branched-chain structures.
- Glycogen and amylose are soluble in water, while cellulose and amylopectin are not.

In order to use polysaccharides as energy sources, animals must first hydrolyse them to glucose. Enzymes are used for doing this. However enzymes are very specific in the bonds that they can break. The human body contains enzymes that can break α-glucose linkages but not β ones. Hence humans can

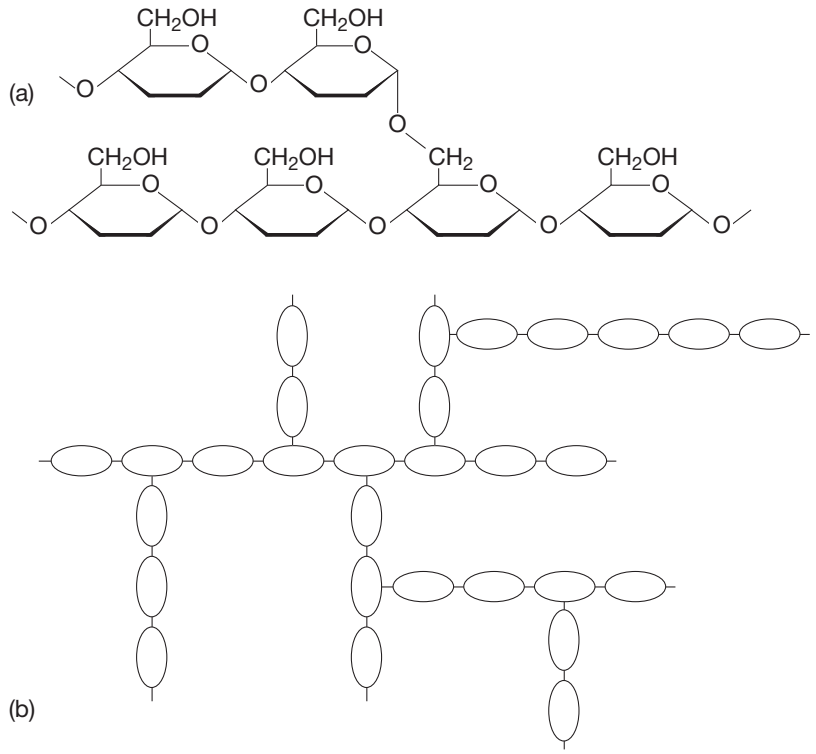


FIGURE 13.9
Chain branching in amylopectin (starch) and in glycogen. (a) shows the actual carbonyl OH–terminal OH union while (b) shows schematically how branching extends through the molecule: the ovals represent glucose units

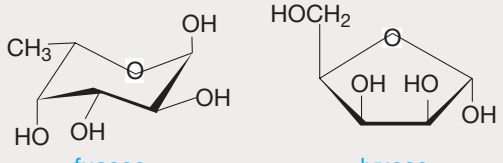
hydrolyse starch and glycogen to glucose but not cellulose. Herbivores (cattle, horses, sheep) have bacteria in one of their stomach chambers that contain enzymes that can break β-linkages; hence these animals are able to use cellulose from grass and leaves as an energy source.

Laboratory hydrolysis of polysaccharides

In the laboratory, polysaccharides can be hydrolysed in the same way as disaccharides; an aqueous suspension or solution of the polysaccharide is refluxed with a strong acid such as hydrochloric or sulfuric until a clear solution (of glucose) results.

Exercises

14 Structures of two less common sugars, fucose and lyxose, are given below in conventional abbreviated form. What is the molecular formula for each of these?



15 **a** Deduce the empirical formula for amylopectin, based upon the molecule of amylopectin being formed from 1000 glucose molecules with the elimination of 999 water molecules.

b Would there be any significant difference if the number of glucose units was 500 or 5000 instead of 1000? Explain.

c Explain why the number of water molecules eliminated in **(a)** is 999.

d Based on **(a)**, **(b)** and **(c)**, what do you expect to be the empirical formula for amylose, glycogen and cellulose?

16 The empirical formula of starch is usually given as $C_6H_{10}O_5$. What would be the molecular weight of a type of starch that contained 300 glucose units?

17 Calculate the percentage oxygen in starch, using the empirical formula given in Exercise 16.

18 Starch can be considered as a polymer of maltose. Explain why. Could a polymer of sucrose be formed by a normal sugar reaction? Explain why or why not.

19 Two molecules of β-glucose condense to form the disaccharide cellobiose using the same C atoms as in maltose. Draw a structure for cellobiose. What relationship exists between cellobiose and cellulose?

20 Using the empirical formula for starch, $C_6H_{10}O_5$, write an equation for its hydrolysis. If the starch contained 100 glucose units, what would this equation become in terms of molecular formulae? Is your equation strictly accurate? Explain.

13.10 RING AND OPEN-CHAIN FORMS: α–β ISOMERISM

The ring structure of monosaccharides (Fig. 13.1) is able to open out into a chain structure as shown in (a) and (b) for glucose and (d) and (e) for fructose in Figure 13.10. Reaction occurs at the carbon atom with two O atoms attached to it. A C=O bond forms on that C atom and an O−H bond on the other C. This C=O group is called a **carbonyl group** and the C atom of it is called the **carbonyl carbon atom** of the molecule (even when the molecule exists as a ring).

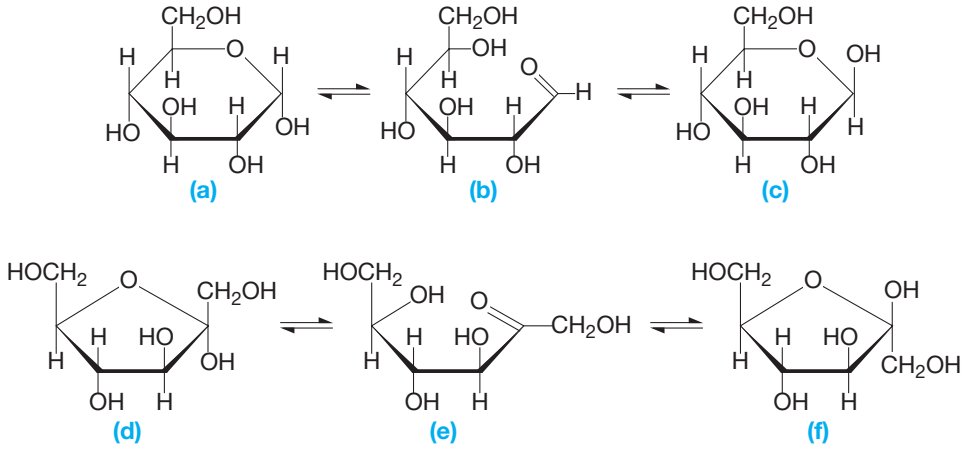


FIGURE 13.10
Ring and open-chain forms of glucose and fructose and how they lead to α and β isomers

That (b) and (e) in Figure 13.10 are really open-chain structures may be better appreciated by drawing them as in Figure 13.11.

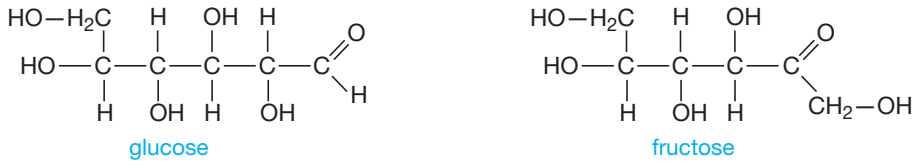


FIGURE 13.11
More conventional drawings of the open-chain structures for glucose and fructose (b) and (e) in Fig. 13.10 respectively)

There is a dynamic equilibrium between the ring structure and the open-chain structure. When the ring closes again, the starting substance may be re-formed, but ring closure may also lead to a slightly different molecule, (c) from glucose and (f) from fructose in Figure 13.10. (a) and (c) are isomers of glucose (called α- and β-glucose respectively) while (d) and (f) are isomers

of fructose (again called α and β respectively). Each individual isomer can be crystallised from solution in pure form under different conditions. However when any one isomer is dissolved in solution, the equilibria in Figure 13.10 lead to a mixture with the other cyclic isomer *and* with a very small amount of the open chain form which is also an isomer.

Although the equilibrium mixture contains predominantly the ring form, the very small proportion present in the open-chain form is sufficient to bring about this isomerisation between α and β forms.

This changing of the arrangement of H and OH about the carbonyl C in carbohydrates is called **α–β isomerism**.

When we draw structures of glucose and fructose, we need not worry too much about the arrangement of H and OH around the carbonyl carbon (that is, whether OH is above or below the ring), because in solution we get a mixture of both forms. However the geometry around the other carbon atoms is fixed, and we must be very careful to get the right groups above and below the ring; otherwise we will have drawn a different compound (for example galactose instead of glucose).

In maltose, in Equation 13.2 and Figure 13.4, the right-hand glucose unit still has an OH on a carbonyl carbon. This means that this ring can still open out into an open-chain form, and hence isomerise into α and β forms. In solution, maltose is an equilibrium mixture of these two isomers.

Sucrose, on the other hand, has no OH on a carbonyl carbon, because the glucose and fructose joined through their carbonyl carbons. Hence sucrose has no α and β forms and so is the one pure compound both as crystals and in solution.

This α–β isomerism is important in forming polysaccharides. As we saw in the previous section it is α-glucose that is involved in forming cellulose and β-glucose for making starch and glycogen.

The presence of the carbonyl carbon in many carbohydrates and the ability of such compounds to isomerise to open-chain forms mean that these compounds are easily oxidised by quite mild oxidising agents. This leads to the classification of carbohydrates into *reducing and non-reducing sugars*.

13.11 REDUCING AND NON-REDUCING SUGARS

We have just seen that many carbohydrates exist as an equilibrium between a ring isomer (the predominant form) and an open-chain form. The open-chain forms of glucose and fructose (from Figures 13.10 and 13.11) can be written as:

$C_5H_{11}O_5-CHO$ (glucose) and $C_4H_9O_4-CO-CH_2OH$ (fructose)

For simplicity we can write these as

$$R-CHO \quad \text{and} \quad R-CO-CH_2OH$$

Both the $-CHO$ and $-CO-CH_2OH$ groups are easily oxidised to carboxylic acids, $R-COOH$ and $R-CHOH-COOH$ respectively.

We call sugars (carbohydrates) that in the open-chain form contain these $-CHO$ and $-CO-CH_2OH$ groups *reducing sugars*, because they are easily oxidised and when that happens they *reduce* the other reagent.

Glucose and fructose are reducing sugars. In fact most monosaccharides are reducing sugars.

Disaccharides may be reducing or non-reducing sugars. Maltose is a reducing sugar because it still has one ring that can isomerise into an open-chain structure (the right-hand ring in the maltose structure in Equation 13.1 on p. 468). On the other hand, sucrose (the right-hand structure in Equation 13.2 on p. 468) is a non-reducing sugar because neither of its rings can convert to an open-chain structure: it is not possible to form a C=O group because the linkage between the glucose and fructose units is through both carbonyl carbons. In summary:

Sugars that have both an OH group and a ring O atom attached to the one C atom are **reducing sugars**. If there is no C atom that has both an O atom and an OH group attached to it, the sugar is a **non-reducing sugar**.

Note that a C atom with two O atoms attached to it as in sucrose in Equation 13.2 and Figure 13.5 is non-reducing; to be reducing one of the Os must be an OH.

A test for reducing sugars

To determine whether a sugar is reducing or non-reducing, we use a reagent that can oxidise a $-CHO$ or $-CO-CH_2OH$ group but which will not oxidise the many ordinary alcohol groups present in all sugars. Three reagents are available:

1. *Tollens' reagent*, a colourless solution of silver nitrate in aqueous ammonia
2. *Benedict's solution*, a deep blue solution of copper sulfate in alkaline citrate
3. *Fehling's solution*, a deep blue solution of copper sulfate in alkaline tartrate.

In all three solutions the cation is present as a complex ion, silver for example as $Ag(NH_3)_2^+$ (Sections 6.10 and 7.6). The reason for this is that alkaline conditions are needed for oxidation of the $-CHO$ and $-CO-CH_2OH$ groups. However Ag^+ and Cu^{2+} normally precipitate out of alkaline solutions as Ag_2O and $Cu(OH)_2$. By forming complex ions of silver with ammonia, and copper with citrate or tartrate, we can keep them in solution even in these alkaline conditions.

When a colourless solution of a reducing sugar is warmed with colourless Tollens' reagent, metallic silver is formed. Depending on conditions, it may appear as a black precipitate of finely divided metal or as a shiny mirror on the walls of a very clean test tube.

When a solution of a reducing sugar is warmed with deep blue Benedict's or Fehling's solution, a reddish-brown precipitate of copper(I) oxide is formed. Cu^{2+} is reduced to Cu^+.

If we write the two types of reducing sugars as $R-CHO$ and $R-CO-CH_2OH$, then equations for these reactions are:

$$R-CHO + 2Ag^+ + 3OH^- \longrightarrow 2Ag(s) + R-COO^- + 2H_2O$$

$$R-CO-CH_2OH + 2Ag^+ + 3OH^- \longrightarrow 2Ag(s) + R-CHOH-COO^- + 2H_2O$$

$$R-CHO + 2Cu^{2+} + 5OH^- \longrightarrow Cu_2O(s) + R-COO^- + 3H_2O$$

$$R-CO-CH_2OH + 2Cu^{2+} + 5OH^- \longrightarrow Cu_2O(s) + R-CHOH-COO^- + 3H_2O$$

Glucose, galactose, fructose and maltose all give positive tests with these reagents. Sucrose gives a negative test. It is a non-reducing sugar (as explained above).

13.12 TEST FOR STARCH

Starch (both amylose and amylopectin) forms a deep blue complex with iodine. If drops of brown iodine solution are added to an aqueous solution or suspension of the substance to be tested, a deep blue colour will form if the substance is starch. At low concentrations this complex is more bluish-purple than pure blue.

Test for glycogen†

Glycogen forms a *reddish-brown* colour when iodine solution is added to it. This colour is quite distinct from the blue to purple colour of the starch–iodine reaction.

Plant or animal sugar?

The glycogen test can establish whether a polysaccharide came from an animal or from a plant, because glycogen is found only in animals, while starch and cellulose are found only in plants.

School laboratory versus forensic laboratory

We have just looked at some simple tests that can be performed in a school laboratory to obtain information about carbohydrates; there are several others that can provide further information but we shall not delve into them. Simple laboratory tests can provide considerable information about the nature and origin of carbohydrates that could be useful to forensic chemists. However they require sample sizes of the order of milligrams, often take considerable time to perform, and the information they provide is not always very specific.

Forensic chemists determine the nature and origin of carbohydrate samples using sophisticated instrumental techniques such as high performance liquid chromatography, HPLC (to be discussed in Section 15.3). Compared with test tube tests performed in a school laboratory, instrumental analysis in a forensic laboratory provides more specific information much more quickly from much smaller samples; however it is much more costly.

† Not required for the NSW HSC; included here because it is a simple test that can give information about the origin of a polysaccharide sample.

Exercises

21 Classify:

a ribose (Structure 1 in Section 13.7) and **b** deoxyribose (Structure 2) as reducing or non-reducing sugars.

22 Is it possible to form a disaccharide between glucose and fructose which is a reducing sugar? If so, draw a structural formula.

23 Which, if any, of the disaccharides you drew in Exercise 12(a) would be reducing sugars?

***24** **a** Describe a chemical test for distinguishing between glucose and sucrose.
b How could you determine whether or not a solution of sucrose had partially hydrolysed to glucose and fructose?

25 Write half equations for the oxidation and reduction processes that are involved in tests for reducing sugars, namely:

a oxidation of R−CHO and R−CO−CH_2OH
b reduction of Ag^+ to Ag and Cu^{2+} to Cu_2O

26 Would it be possible to have a disaccharide that gave a negative test with Tollens' reagent, but which hydrolysed to form only glucose (which gives a positive test)? Explain.

27 A group of students had four aqueous solutions of carbohydrates to identify. They performed the following tests on each solution:

1 add Tollens' reagent
2 add a solution of iodine.

Results are tabulated below.

Sample R was heated with hydrochloric acid for several minutes before being tested.

Sample	Tollens' reagent	Iodine solution
P	no reaction	blue colour
Q	silvery deposit	no reaction
R	silvery deposit	blue colour
S	no reaction	no reaction

a Which is the only sample that could be **(i)** galactose **(ii)** sucrose **(iii)** starch? Explain why.
b What test could you perform to provide further evidence that the sample you identified as sucrose really is sucrose? Explain.
c Which sample could be partially hydrolysed starch? Explain why.

28 Trehalose is a sugar of molecular formula $C_{12}H_{22}O_{11}$. With Tollens' reagent and Benedict's solution it gives negative results. When warmed with aqueous hydrochloric acid one mole of trehalose yields two moles of glucose. Deduce the structure of trehalose.

Important new terms

You should know the meaning of the following terms.

α–β isomerism (in sugars) (p. 475)
alkoxide ion (p. 462)
carbohydrate (p. 465)
carbonyl carbon atom (p. 474)
carbonyl group (p. 474)
condensation reaction (p. 469)
disaccharide (p. 467)
infrared (i.r.) spectroscopy (p. 463)
monosaccharide (p. 467)
non-reducing sugar (p. 476)
organic compound (p. 460)
polysaccharide (p. 470)
reducing sugar (p. 476)
soil (p. 461)
sugars (p. 465)

Test yourself

1 Explain the meaning of each of the items in the 'Important new terms' section above.

2 List three possible sources of contamination of small samples for analysis.

3 Give two reasons why accuracy is important in forensic analyses; give an example to illustrate.

4 How would you decide whether a substance was organic or inorganic?

5 How would you determine the source of some soil found at a crime scene?

6 Describe tests to distinguish between an alkanoic acid, an alkanol, an alkene and an alkane.

7 Draw structures for glucose, fructose, maltose and sucrose.

8 What is meant by ring and open-chain forms of common monosaccharides? Give an example.

9 Write an equation for a condensation reaction involving sugars.

10 Write equations for the hydrolysis of maltose and sucrose. What conditions are needed to bring these reactions about?

11 What is meant by α and β linkages in di- and polysaccharides?

12 What are the structural differences between cellulose, amylose, amylopectin and glycogen?

13 Write an equation, using molecular formulae, for the hydrolysis of starch.

14 What structural features cause a sugar to be reducing?

15 Describe two tests for determining whether a sugar is reducing or non-reducing.

16 How would you determine whether a white powder was starch?

CHAPTER 14

Proteins and DNA

IN THIS CHAPTER

In Chapter 13 we saw how forensic chemists first determine the class of compound that a particular sample belongs to and then identify the actual compound. We looked particularly at sugars in that chapter. Let us now turn to other important biological compounds, proteins and DNA.

14.1 PROTEINS

Proteins are an essential part of all living organisms. There are two general classes of protein—*fibrous proteins* and *globular proteins*.

Fibrous proteins are structural material, such as *keratin* of skin, fingernails, horns, wool and feathers, *collagen* of tendons, cartilage and hides, and *fibroin* of silk. These proteins are tough and insoluble in water and most other solvents. They are stringy in appearance and have a characteristic smell when burnt.

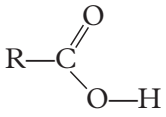
Globular proteins are roughly spherical in shape and are soluble in water. They are involved in the actual functioning of living organisms. Common globular proteins are *insulin* which controls glucose metabolism, *haemoglobin* (in blood) and *myoglobin* (in muscles) which transport and store oxygen and *antibodies* which fight diseases. **Enzymes**, the biological catalysts, are also globular proteins (pp. 298–9 *CCPC*).

Proteins are naturally occurring polymers of amino acids.

Hence to understand the structure of proteins we need to start with the basic monomer units, the amino acids.

14.2 AMINO ACIDS

We saw in Section 5.16 that *alkanoic acids* had the general structure:

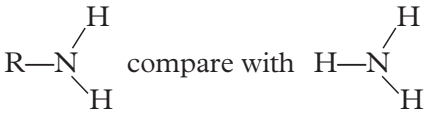


In those chapters R was any alkyl group. More generally R can be any carbon-containing group (not just alkyl) in which case the compound is called a *carboxylic acid*.

Another class of simple carbon compounds is called *amines*. These are derivatives of ammonia.

Amines are compounds in which one or more carbon-containing groups (such as alkyl groups) replace one or more H atoms in an ammonia molecule.

If one H of ammonia is replaced by a carbon-containing group, we get the structure:



Like ammonia, amines are weak bases. They undergo the following equilibrium reaction:

$$R{-}NH_2 + H_2O \rightleftharpoons R{-}NH_3^+ + OH^-$$

(compare with $NH_3 + H_2O \rightleftharpoons NH_4^+ + OH^-$)

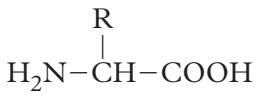
We call $-NH_2$ the **amine functional group**.

Amino acids are compounds that contain both an amine and a carboxylic acid functional group.

The simplest amino acid is glycine:

$$H_2N{-}CH_2{-}COOH$$

There is a whole family of amino acids with the general structure:



where R is a carbon-containing side chain.

Proteins are long-chain molecules with thousands of amino acid molecules of this general structure joined together. (This is discussed below.)

There are 20 amino acids commonly found in proteins. Some of these are shown in Table 14.1. They fall into three broad groups; amino acids with:

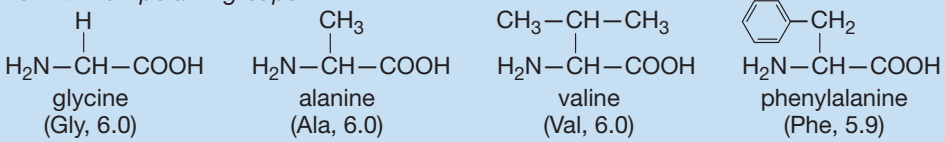
- non-polar R groups (alkyl or similar)
- polar R groups that are capable of forming ions; these have a $-COOH$ or $-NH_2$ group *additional* to the normal ones in the amino acid
- polar R groups that do not generally form ions; the simplest of these contain polar $-OH$ or $-SH$ groups as part of R.

Molecular models of three simple amino acids are shown in Figure 14.1.

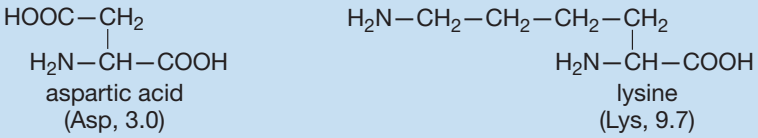
TABLE 14.1 Common amino acids found in proteins

(usual abbreviations and isoelectric points (Section 14.10) are given in brackets under the name)

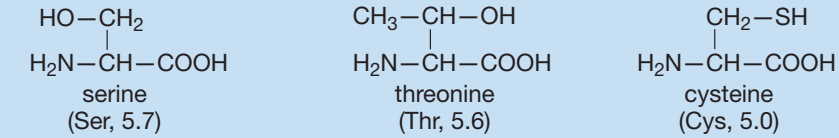
a *with non-polar R groups*



b *with polar R groups that form ions*



c *polar R groups that do not generally form ions*



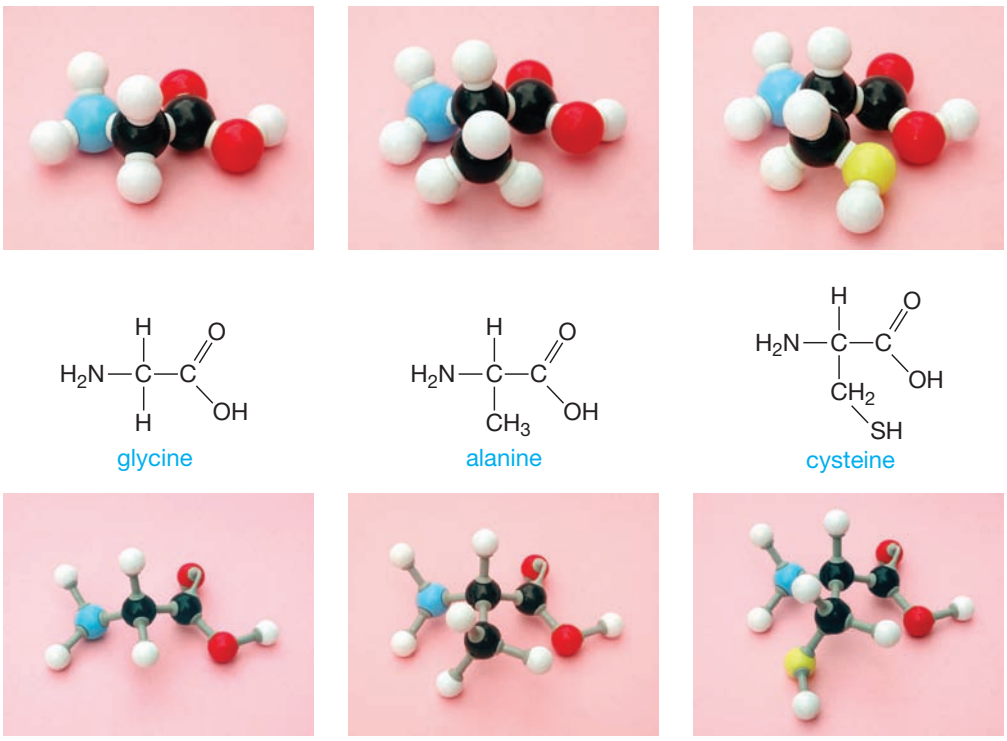


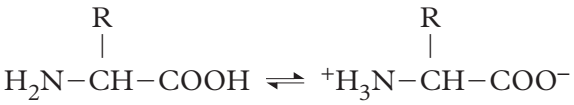
FIGURE 14.1
Space-filling and ball-and-stick models of glycine, alanine and cysteine along with structural formulae drawn with approximately similar orientations. Black is C, white, H, red, O, blue, N and yellow, S

The occurrence of an acidic and a basic group on the one molecule leads to the possibility of an ionic structure for amino acids.

14.3 ZWITTERIONS

So far we have written amino acids as neutral molecules as in Table 14.1. However the $-COOH$ group, being acidic, tends to lose a proton, while the amine group, $-NH_2$, being basic, tends to gain a proton. Hence in solution,

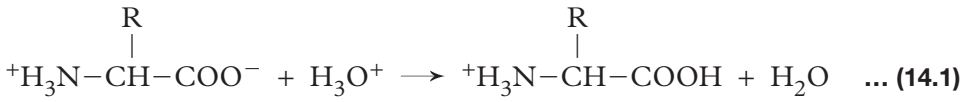
amino acids exist as an equilibrium mixture of neutral molecules with a dipolar ion:



This dipolar ion is called a **zwitterion**. *Most amino acids exist primarily as zwitterions.*

Because of this dipolar structure of amino acids, they are crystalline solids with relatively high melting points, and they are soluble in water.

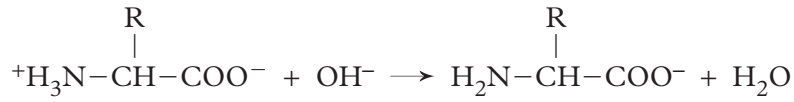
Because the $-COOH$ group is a weak acid and the $-NH_2$ group is a weak base, amino acids take on different forms in solutions of different pH. In acid solution, say pH 2 to 3, the zwitterion changes to a cation:



... (14.1)

(compare with $R-COO^- + H_3O^+ \longrightarrow R-COOH + H_2O$)

In alkaline solution it changes to an anion:



... (14.2)

(compare with $NH_4^+ + OH^- \longrightarrow NH_3 + H_2O$
or with $R-NH_3^+ + OH^- \longrightarrow R-NH_2 + H_2O$)

Exercises

1 A certain amino acid contained 40.4% carbon, 7.9% hydrogen and 15.7% nitrogen. The balance was oxygen.

- **a** Calculate the empirical formula.
- **b** Draw the structural formula of the simplest amino acid that could have this empirical formula.

2 Glycine and alanine are the first two members of a homologous series of 'straight-chain' amino acids. Draw structures for the next two members of the series.

3 MSG (monosodium glutamate) is widely used as a flavour-enhancing agent (often in Chinese food). It is the sodium salt of glutamic acid. Glutamic acid is an amino acid with a propanoic acid side chain. Draw the structure of glutamic acid and hence of MSG. Why is it called 'monosodium'?

4 For each of the amino acids:

a glycine **b** alanine ***c** cysteine ***d** lysine

from Table 14.1, draw the structural formula of the zwitterion form of the molecule.

5 Draw the structures you would expect for **(a)** valine **(b)** aspartic acid (Table 14.1) in solutions of pH **(i)** 2 and **(ii)** 11.

6 **a** Draw the structural formula of the main species present in a solution of serine (Table 14.1) in water.

b Draw the structural formula of the main species present when:

- **i** 1.00 mol of each of serine and sodium hydroxide is dissolved in 2.00 L water
- **ii** 1.00 mol hydrochloric acid is added to the solution in **(i)**
- **iii** another 1.00 mol HCl is added to the solution in **(ii)**.

7 Glycine can be used as a buffer at pHs around 6. Explain why (with equations).

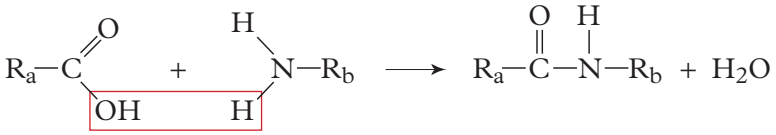
*8 Alanine is amphoteric. Write equations which show this.

14.4 PEPTIDES

As expected from our knowledge of acids and bases, an amine reacts with a carboxylic acid to form a salt:

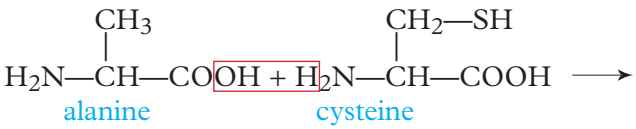
$$R_a{-}COOH + R_b{-}NH_2 \longrightarrow (R_b{-}NH_3^+)(R_a{-}COO^-)$$

However *under certain conditions an amine reacts with a carboxylic acid by eliminating water to form what we call an* **amide**:

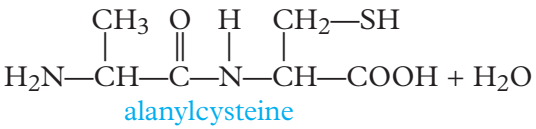


This amide product is a stable covalent compound with no acid or base properties at all.

Pairs of amino acid molecules undergo this same reaction. For example, alanine reacts with cysteine:



... (14.3)



The compound formed when two amino acids react together is called a **dipeptide**. This amide linkage between two amino acids is called a **peptide bond**.

The peptide link which we have drawn out in full in Equation 14.3 is often written more concisely as:

$$-CO-NH-$$

Molecular models of alanylcysteine are shown in Figure 14.2.

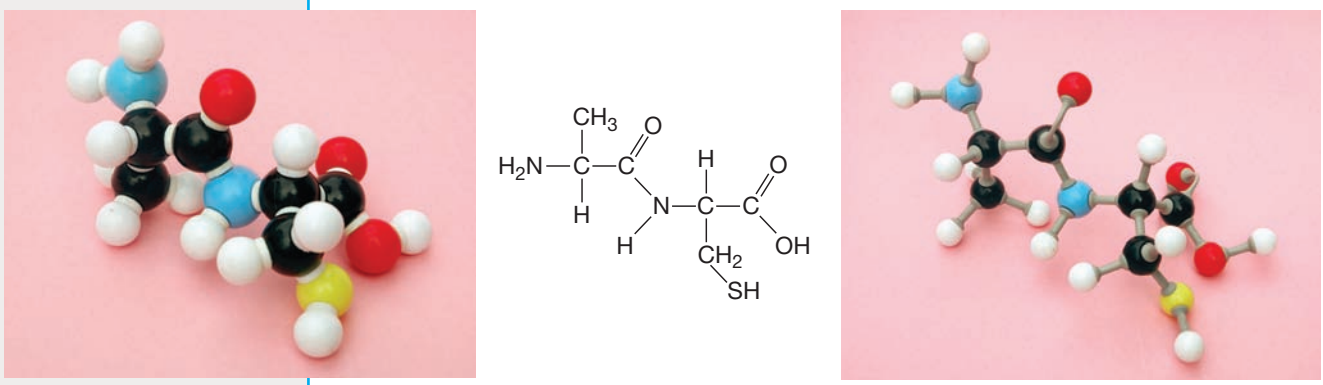
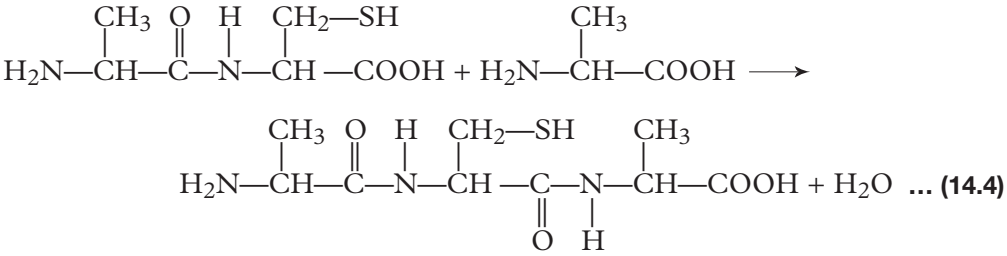


FIGURE 14.2
Space-filling and ball-and-stick models of alanylcysteine, the dipeptide formed between alanine and cysteine. Colours are as in Figure 14.1

Our dipeptide in Equation 14.3 still has terminal COOH and amine groups, so it can join to other amino acids. For example, with another molecule of alanine we get:



... (14.4)

The product is called a **tripeptide**, because it is made from three amino acid units.

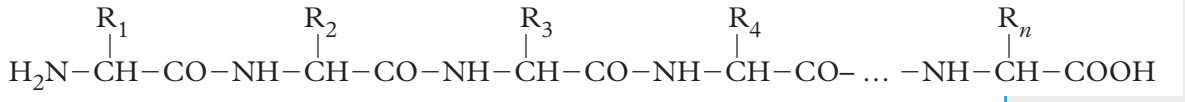
Peptide bonding can extend to form very long chains of amino acid units. We call such long molecules **polypeptides**, many peptide linkages or many amino acids joined together. Polypeptides are polymers. The monomers making up the polymer are amino acids. Formation of a peptide from amino acids is a *condensation reaction* (Section 13.8). Note that the peptide bond or amide linkage formed here is the same as the bond between monomer units in nylon (Section 1.12).

Because formulae of peptides are tedious to draw, biochemists frequently use a shorthand way of drawing them. Each of the twenty common amino acids has a three-letter abbreviation. Some are given in Table 14.1. These are often used instead of complete structural formulae. The dipeptide in Equation 14.3 is written as Ala−Cys. The tripeptide in Equation 14.4 is Ala−Cys−Ala. A possible peptide with ten amino acids is:

Gly−Phe−Cys−Ser−Ala−Gly−Asp−Ala−Lys−Asp

Proteins as polypeptides

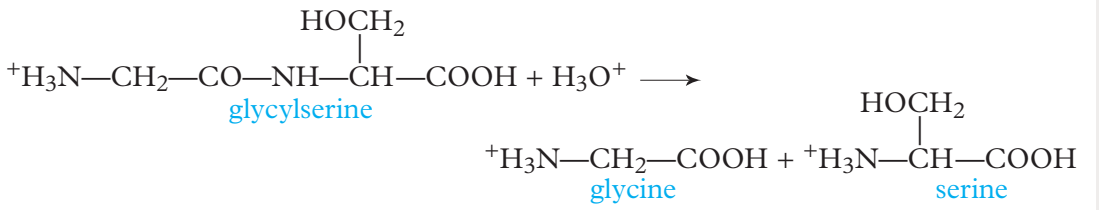
Proteins are polypeptides containing from hundreds to thousands of amino acid units per molecule. The generalised structure of a protein can be represented as:



where R_1, R_2, R_3, … R_n are the side groups of the various amino acids (hundreds to thousands in actual proteins). Some of these will be the same while others will be different. Nine of the twenty R groups (in amino acids) commonly found in proteins are shown in Table 14.1. Notice that one end of the protein molecule has an amino group, NH_2, while the other has a carboxylic acid group, −COOH.

14.5 HYDROLYSIS OF PEPTIDES AND PROTEINS

Peptides and proteins can be hydrolysed to amino acids by warming them with about 6 mol/L hydrochloric acid. For the dipeptide Gly−Ser the reaction is:



If the tripeptide of Equation 14.4 is treated in this way, alanine and cysteine are formed in the molar ratio of 2:1.

$$\text{Ala}-\text{Cys}-\text{Ala} + 2H_2O \longrightarrow 2\text{Ala} + \text{Cys}$$

The reaction occurs with polypeptides and hence with proteins as well.

In humans and other animals, this hydrolysis of polypeptides (proteins) is brought about by certain enzymes. This occurs in digestion, both in the stomach and in the intestine. Some enzymes are very specific as to which peptide bonds they will break. For example trypsin will break a −CO−NH− peptide bond only if lysine or arginine is on the CO side of the bond: chymotrypsin will break the peptide bond only if the amino acid on the CO side is phenylalanine, tyrosine or tryptophan.

Hence by using particular enzymes it is possible to break a protein into several smaller pieces (polypeptides) without completely breaking it into separate amino acids.

In the laboratory it is also possible to break a protein into smaller pieces (rather than completely into individual amino acids) but the breakages occur randomly and not at particular sites as with enzymes.

Exercises

Refer to Table 14.1 for the structures of the amino acids used in the following exercises.

9 Draw the structural formula of a dipeptide formed between:

a glycine and valine
b threonine and alanine
***c** cysteine and aspartic acid
***d** lysine and phenylalanine

10 Can more than one dipeptide be formed between each of the pairs of amino acids in Exercise 9? If so, draw the structures of the ones you did not draw in that exercise.

11 Draw structural formulae for the dipeptides which can be represented by:

a Val−Ser ***b** Ala−Asp ***c** Thr−Lys

12 Write balanced equations (using structural formulae for all carbon-containing compounds) for the reactions that occur when the following compounds are heated with 6 mol/L hydrochloric acid. Make sure that you write each compound in the correctly protonated or deprotonated form:

a $H_2N-CH_2-CO-NH-CH_2-COOH$

***b**
$$H_2N-CH_2-CO-NH-\underset{\displaystyle CH_3}{\underset{|}{CH}}-COOH$$

c
$$HOOC-\underset{\displaystyle HOCH_2}{\underset{|}{CH}}-NH-CO-\underset{\displaystyle CH_3}{\underset{|}{CH}}-NH_2$$

***d**
$$H_2N-\underset{\displaystyle CH_3-CH-CH_3}{\underset{|}{CH}}-CO-NH-\underset{\displaystyle CH_2-SH}{\underset{|}{CH}}-COOH$$

***e**
$$HOOC-\underset{\displaystyle CH_3}{\underset{|}{CH}}-NH-CO-\underset{\displaystyle CH_2-COOH}{\underset{|}{CH}}-NH_2$$

Identify the amino acids formed in each case.

13 Write a molecular equation for the reactions in **(a)** and **(c)** of Exercise 12.

14 a Aspartame is an artificial sweetening agent sold under the trade name, NutraSweet®. It is a derivative of a dipeptide formed from aspartic acid and

phenylalanine. The peptide linkage involves the carboxylic acid group of aspartic acid. Draw the structure of this compound.

b Aspartame is the methyl ester of this dipeptide. The COOH group which has been esterified is the one 'belonging to' phenylalanine. Draw the structure of the ester.

c What is the molecular formula of aspartame?

d NutraSweet® is not recommended for use where cooking is involved. Suggest a reason for this.

14.6 STRUCTURE OF PROTEINS

The chemical composition (number of each type of atom present) and structure (how the atoms are arranged) of proteins determine their biological function; that is, whether they are structural material such as skin, hair, tendon, cartilage, or whether they are involved in the biochemical reactions that constitute life, such as enzymes, haemoglobin, insulin.

There are three aspects to the structure of a protein:

- the nature and sequence of the amino acids that make up the protein, called the *primary structure*
- the way that the amino acid chain is arranged within segments of the molecule (curled into a spiral or folded into pleated sheets and the dimensions of these arrangements), called the *secondary structure*
- the way that these segments are arranged to make up the whole molecule, called the *tertiary structure*.

Covalent bonds (peptide links) are responsible for the primary structure—how the amino acids are joined to one another. The secondary and tertiary structures arise from:

- *hydrogen bonding*, mainly between NH and CO groups on adjacent (but not neighbouring) amino acid units
- *ionic attractions*, between $-COO^-$ and $-NH_3^+$ groups on adjacent (but again not neighbouring) amino acid units
- *sulfur–sulfur links* (covalent bonds) between two nearby cysteine units (−SH HS− converted into −S−S−)
- dispersion forces and dipole–dipole forces.

The sequence of amino acids along the chain determines where these interactions can occur and hence decides the shape that the molecule can have.

If the final structure of the protein is fibrous (many polypeptide chains intertwined) or sheet-like, then its function will be structural—skin, cartilage, hair. If the final structure is globular with many $-COO^-$ and $-NH_3^+$ and −OH groups pointing outside the globule, the protein will be soluble in water and will be involved in the functioning of the organism.

Determination of primary structure

To determine the primary structure of a protein a chemist needs to determine:

- the *identity* of all the amino acids present
- the *relative molar amounts* in which these amino acids are present
- the *sequence* in which the amino acids are linked together in the molecule.

To determine the identity and relative amounts of the amino acids present, the protein is hydrolysed completely (as explained above) into the amino acids

then the mixture is analysed by chromatography (Sections 14.8 and 14.9): this shows which amino acids are present and in what relative amounts.

The sequence of amino acids has to be determined by breaking off amino acids from the protein chain one at a time starting from one end (usually the NH_2 end) and identifying that amino acid by chromatography or electrophoresis (Section 14.10). Then the whole procedure is repeated to split off the next amino acid from the shortened chain. Because of the build-up of by-products, it is generally not possible to cut off more than about twenty amino acids in this way. Typically a protein contains 200 to 400 amino acid units. For full sequencing the protein is cut into smaller pieces by selected enzymes, the pieces separated, then the sequence in each piece determined. By using different enzymes, overlapping pieces can be formed, then after sequencing all the pieces, the chemist can work out the order in which the pieces were joined in the original protein by recognising the overlapping bits, rather like solving a jigsaw puzzle. In this way we get the complete sequence of amino acids in the protein.

Secondary and tertiary structures

The secondary and tertiary structures of proteins are generally worked out by using X-ray studies of crystalline samples of the various proteins.

Forensic identification of proteins

Forensic chemists do not need to work out the structure of proteins from scratch as just described. Rather they have to *identify* protein samples. And generally it is fibrous (structural) proteins that they need to identify, such as hair, skin, cartilage. This is usually done by microscopic examination, because fibrous proteins have quite distinct appearances. If it is necessary to identify a globular (functional) protein, this can generally be done by determining the amino acid composition of the protein (complete hydrolysis followed by electrophoretic or chromatographic analyses, to be described in Sections 14.8 to 14.11).

14.7 SIMPLE TESTS FOR PROTEINS AND/OR AMINO ACIDS

There are two tests that we can easily perform to determine whether or not a sample contains protein or amino acid.

Ninhydrin test

Ninhydrin is a cyclic carbon compound that is able to convert an amino acid (or terminal amino acid group on a protein) into an aldehyde (compound containing a −CHO group). In so doing, ninhydrin is converted to a purple coloured compound.

The **ninhydrin test** is to add colourless ninhydrin solution to the solution suspected of containing protein or amino acid and to heat the mixture for a few minutes. If a purple colour develops, *protein or an amino acid* is present.

Biuret test

The **biuret test** gives a positive result for proteins and peptides (including dipeptides) but not for amino acids.

The solution to be tested is made alkaline with NaOH solution, then a few drops of copper sulfate solution are added to it. A purple colour forms if the sample contains a protein or peptide. This test can be spoilt if too much copper sulfate solution is added, because then a precipitate of copper hydroxide can be formed. The purple colour is due to a complex which forms between Cu^{2+} and two adjacent amide (peptide) linkages. This complex is similar to the copper complexes that were present in Benedict's and Fehling's solutions in Section 13.11.

The name of this test comes from one of the simplest compounds that give a positive result, biuret, $H_2N-CO-NH-CO-NH_2$.

These two tests just tell us that protein or amino acid is present. Two techniques that are widely used to determine which amino acids are present and how much of each there is in a sample of protein are chromatography and electrophoresis.

Exercises

15 A chemist had three solutions, A, B and C. A gave a positive test with ninhydrin and with alkaline copper solution (Biuret test), and B gave negative results with these two tests but formed a deep blue colour when drops of iodine were added to it. C tested positively to ninhydrin and negatively to the Biuret test and to iodine. What do you conclude about the substances present in the three solutions?

16 How would you distinguish between four solutions that contained:

a a protein **b** starch **c** glucose **d** none of these?

Describe carefully what you would observe for a positive result in each of the tests you use.

17 What would you observe if you tested solutions of:

a myoglobin (a protein)
b sucrose
c amylose
d jelly crystals (contain protein and sucrose)
***e** honey
***f** milk
***g** 'diet' lemonade (containing NutraSweet®, see Exercise 14)
***h** flour
***i** apple juice
***j** table salt
***k** egg white

with

i ninhydrin **ii** iodine **iii** Tollens' reagent?

18 How would you establish that:

a Sample (g) in Exercise 17 did not contain sucrose or glucose
b Sample (j) contained chloride but not carbonate or sulfate?

14.8 CHROMATOGRAPHY

Before discussing the separation of amino acids by chromatography, let us take a general look at what is meant by chromatography.

Chromatography is the name given to a group of techniques used to separate substances based upon their differential distribution between two phases, one moving and the other stationary.

The name, chromatography (literally colour writing), arises because the technique was first used to separate coloured substances.

The simplest type of chromatography to demonstrate is *paper chromatography*. It can be used to separate the components of coloured inks. Spots of ink are placed near one end of a strip of filter paper which is then suspended in a suitable liquid as shown in Figure 14.3(a). As the liquid creeps up the paper (by capillary attraction), it 'washes' the components of the ink upwards at different rates. After a few minutes different-coloured, separated spots can be seen as shown in Figure 14.3(b).

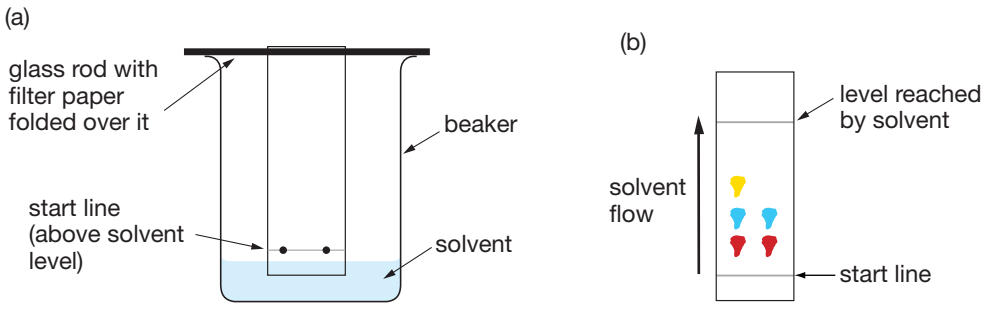


FIGURE 14.3
Separation of inks (from felt-point colouring pens) by paper chromatography

The separation occurs because the substances to be separated have different strengths of attachment to the filter paper (called the **stationary phase**) and different solubilities in the solvent being used (called the **mobile phase**). In paper chromatography the attachment of the substances to the stationary phase is a combination of *adsorption* onto the cellulose fibres of the filter paper and *solubility* in the water trapped in the pores of the paper. Substances that attach only weakly to the stationary phase but dissolve readily in the mobile phase move up the paper more quickly than ones that attach more strongly to the stationary phase and dissolve less readily in the mobile one, and so a separation occurs.

Paper chromatography is a technique for separating substances by using a solvent (mobile phase) to wash the substances in the mixture at different rates up (or down) a piece of absorbent paper such as filter paper (stationary phase).

A mixture of colouring matter extracted from leaves of plants using petroleum ether (a mixture of hydrocarbons) can also be separated by paper chromatography as can the colouring matter from 'Smarties' or similar confectioneries.

Apart from paper chromatography, there are broadly four types of chromatography, depending upon whether the mobile phase is a liquid or a gas, and upon whether the stationary phase is a solid or a liquid (adsorbed on a solid). If the stationary phase is a solid, then separation is based upon the differing strengths of *adsorption* (on the stationary phase) of the substances to be separated. If the stationary phase is a liquid, separation is based upon differing *solubilities* of the substances to be separated in that liquid.

The original form of chromatography was *liquid–solid chromatography*. It was introduced by the Russian botanist Mikhail Tswett in 1902 to separate a mixture of coloured material extracted from plants. A solution of the mixture

to be separated was added to the top of a column of finely divided solid. The solutes adsorbed onto this solid and were then washed down the column by a suitable solvent. Because of different strengths of adsorption of the substances onto the column, they moved down at different rates and so showed up as different coloured bands as in Figure 14.4. With continued addition of solvent to the top of the column these bands were eventually *eluted* from (washed off) the column and collected in separate beakers and so a separation was effected.

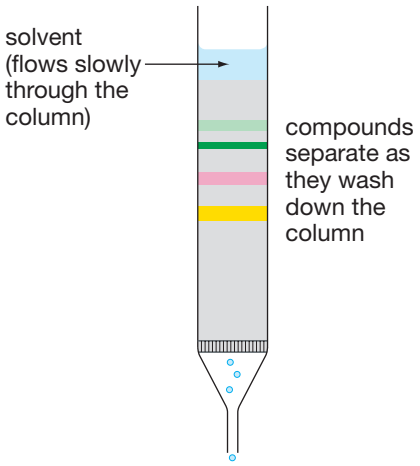


FIGURE 14.4
Separation of a mixture of coloured material from a plant using liquid–solid chromatography. With continued addition of solvent the individual substances can be washed off the column one after the other and collected in separate beakers

In *gas–solid chromatography* the mixture to be analysed (or separated) is vaporised into a flowing gas (mobile phase) and again separated by differential adsorption on a solid (Section 6.2).

For forensic purposes the most useful types of chromatography are liquid–liquid chromatography (sometimes called partition chromatography) and gas–liquid chromatography. These will be considered in Sections 15.2 and 15.3.

14.9 SEPARATION OF AMINO ACIDS BY PAPER CHROMATOGRAPHY

Amino acids can be separated by paper chromatography. The stationary phase is the very polar water–paper combination while the mobile phase is a less polar solvent such as a butanol–acetic acid mixture or a phenol in water solution. The arrangement of Figure 14.3 with the solvent creeping up the paper can be used, but a more efficient separation is effected if the solvent is allowed to run *down* the paper. The arrangement for this is shown in Figure 14.5. In either case the separation is better if the beaker or tank is kept covered so that the paper becomes saturated with solvent vapour.

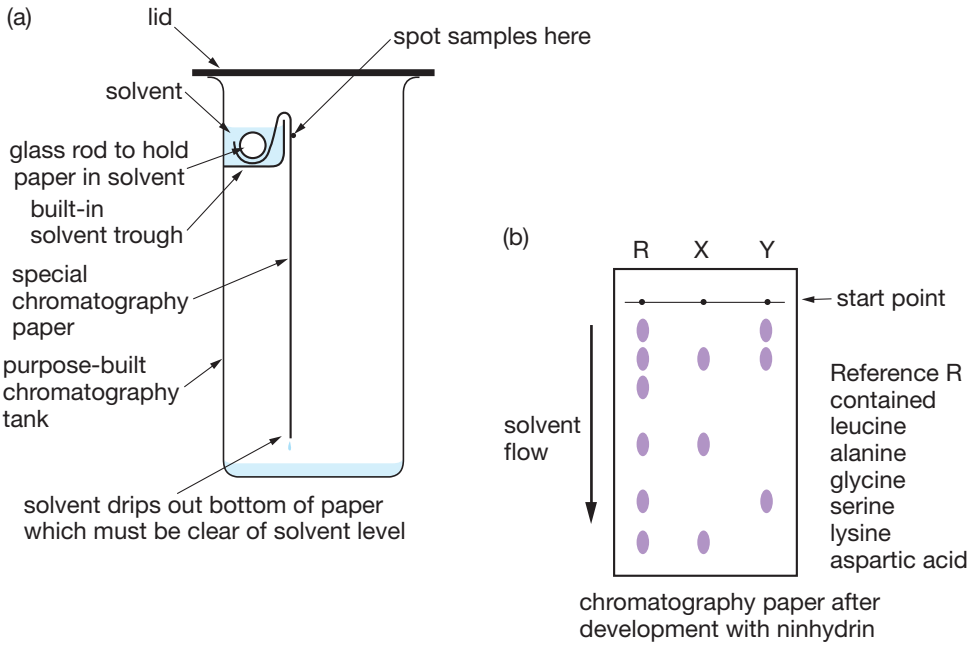


FIGURE 14.5
(a) Paper chromatography for separating amino acids by allowing the mobile phase to flow down the paper which results in a more uniform flow rate and so a better separation.
(b) A developed chromatogram: R is a reference sample containing known amino acids while X and Y are samples being analysed

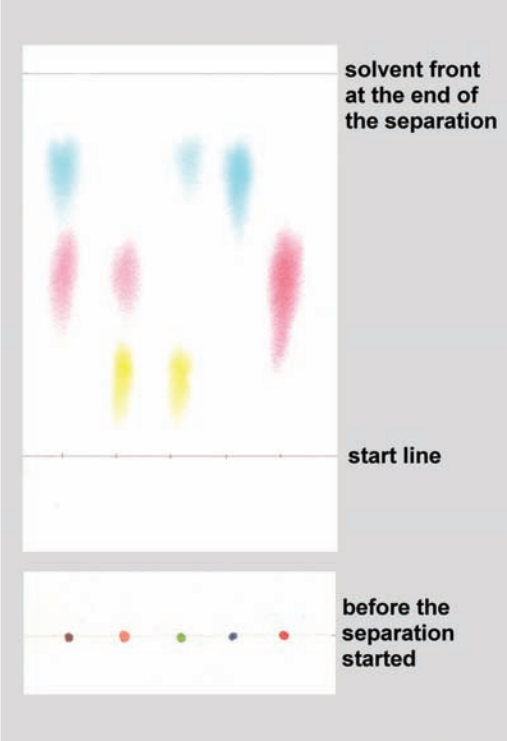
Basis of separation

Amino acids move through the chromatography paper at different rates because of different polarity and therefore different solubilities in the two phases. More polar amino acids attach to the more polar stationary phase (paper) more readily and so move through the paper at a lower speed. Less polar amino acids dissolve more readily in the less polar mobile phase and so move through the paper more rapidly. These differing solubilities are based on the general rule on p. 199 *CCPC*:

Like dissolves like

Making the spots visible

Amino acids are colourless, so we need some way of making the spots visible. The ninhydrin reaction (Section 14.7) is used. After the mobile phase has flowed through the paper for a suitable time, the paper is dried out (by exposing it to a flow of air in a fume cupboard). The dry paper is then sprayed with a ninhydrin solution and warmed in an oven for a few minutes. Ninhydrin reacts with the amino acids and so purple spots develop on the paper corresponding to positions of the amino acids. It is advisable to trace the outline of the spots with a pencil, because the spots can fade with time.



A paper chromatogram showing the separation of the pigments in some common food dyes; the mobile phase is a 1:1:1 mixture of 1-butanol, ethanol and 2 mol/L aqueous ammonia

Identifying the amino acids in the mixture

There are two methods for identifying the spots in the sample to be analysed. In the first method spots of known amino acids are placed beside the sample. The distance from the starting position of each spot in the sample is compared with the distances travelled by the known standards. The unknown spot is identified as being the amino acid that travelled the same distance.

In the second method a standard mixture of several amino acids is used: the order in which these will separate under the conditions of the experiment is known, so just one standard is run beside the sample to be analysed—it produces several spots but these can be identified because of the known order. Again comparison of distances travelled by spots in the sample to be analysed with these standard spots allows identification of the unknown spots as in Figure 14.5(b).

Amino acids can also be separated by ion exchange chromatography but we will not go into that.

14.10 ELECTROPHORESIS AND AMINO ACIDS

Electrophoresis is a method of separating charged substances based upon the different signs of their charges and upon their different mobilities.

Paper electrophoresis uses a strip of filter paper soaked in an electrolyte solution containing a suitable buffer. A voltage is applied across the strip of paper as shown in Figure 14.6 and the sample to be separated is placed on a line

across the middle of the strip. Positively charged particles migrate towards the negative electrode while negative particles move towards the positive electrode. Neutral particles do not migrate at all.

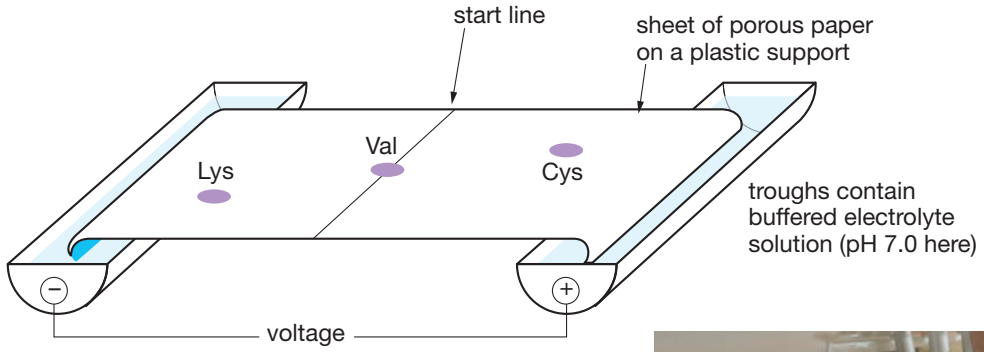


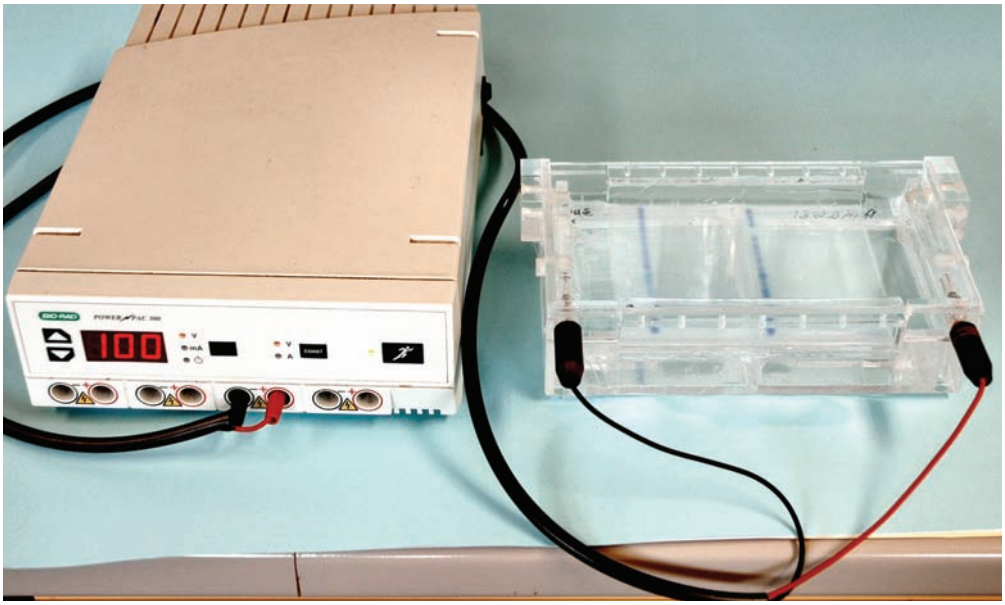
FIGURE 14.6
Electrophoresis: migration of charged particles through a buffered electrolyte when a voltage is applied

The speed at which particles move depends upon their size and charge. If two particles (ions) are singly charged, then the smaller ion will move more rapidly. If two ions are of similar size but one is doubly charged while the other is singly charged, then the doubly charged one will migrate more rapidly.

Electrophoresis is commonly used to identify amino acids or to analyse mixtures of amino acids. Amino acids are colourless, but their positions on an electrophoresis strip can be made visible by spraying the strip with a solution of ninhydrin (as for paper chromatography). This colourless compound turns purple in the presence of an amino acid (Section 14.7), so the sprayed strip shows a series of purple spots that indicate the location of the amino acids.

Instead of a strip of filter paper, electrophoresis can be performed using an agar gel spread on a sheet of glass. This can separate larger quantities of substances than paper.

ABOVE A scientist loads samples into a gel for analysis in the electrophoresis apparatus in the photograph below



LEFT Electrophoresis apparatus using an agarose gel to separate fragments of DNA. The samples, about ten across each blue strip, are placed in holes (wells) in the gel along with blue dye to show where they are; the gel is saturated with and covered with buffer solution. Two sets of samples are being analysed here—the two blue strips. The separation will be stopped (after one to two hours) before the fragments in the left-hand strip reach the right-hand one (the starting line for the second set of samples)

Charges on amino acid molecules or ions

In Section 14.3 we saw that amino acids exist as zwitterions—molecules with both a positive and a negative charge. The overall molecule in this form is neutral and so does not migrate during electrophoresis.

However one end of the molecule is a weak acid while the other end is a weak base and so there can be proton transfer as pH changes as was shown in Equations 14.1 and 14.2. To illustrate further, glycine in near neutral solution exists as the neutral zwitterion, but in acid solution (say pH = 2) the carboxylate ion accepts a proton:

$$\underset{\text{(overall neutral)}}{{}^{+}H_3N-CH_2-COO^{-}} + H^{+} \longrightarrow \underset{\text{(positive)}}{{}^{+}H_3N-CH_2-COOH}$$

This forms a positive ion that will migrate towards the negative electrode in electrophoresis.

If the zwitterion is placed in alkaline solution (say pH = 11), then it gives up a proton and becomes negatively charged:

$$\underset{\text{(overall neutral)}}{{}^{+}H_3N-CH_2-COO^{-}} + OH^{-} \longrightarrow \underset{\text{(negative)}}{H_2N-CH_2-COO^{-}} + H_2O$$

Consequently at pH 11 glycine migrates towards the *positive* electrode, while at pH 2 it migrates towards the *negative* electrode and at pH 6 to 8 it does not migrate at all.

This is true for all amino acids—at low pH they are positively charged, at high pH they are negatively charged and at an intermediate pH they are overall neutral. The actual pHs at which these charges form depend upon the structure of the amino acid in question.

For any given amino acid there is a pH at which it exists as the overall-neutral zwitterion. This is called the **isoelectric point**. At lower pH the amino acid exists as the positively charged species, while at higher pH it exists as the negatively charged species. The isoelectric point of an amino acid can be determined approximately by finding the pH at which it does not migrate during electrophoresis. Isoelectric points for the common amino acids are shown in Table 14.1.

The isoelectric point depends upon the structure of the amino acid. Amino acids with non-polar R groups have isoelectric points close to 6.0. For amino acids such as aspartic acid with an acidic R group (Table 14.1), a lower pH is needed to suppress the ionisation of the acid group in R; hence the isoelectric point is much lower than 6. If R contains a basic group (as in lysine, Table 14.1), a higher pH is needed to suppress the ionisation of this basic group in R; such amino acids have an isoelectric point much greater than 6. Figure 14.7 shows the forms of aspartic acid at different pHs.

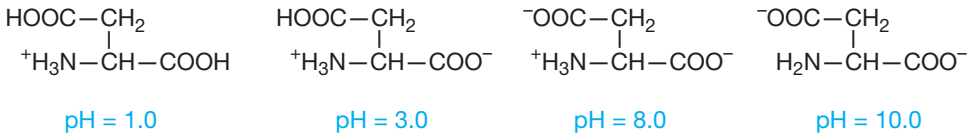


FIGURE 14.7
Forms of aspartic acid at different pHs (isoelectric point 3.0)

Use of electrophoresis for identifying proteins

As mentioned in Section 14.6 forensic scientists do not need to determine the structure of proteins, but rather just establish the identity and/or origin of

particular samples. Electrophoresis can be quite helpful in this regard. Simple acid hydrolysis of the protein sample converts it to a mixture of amino acids and then performing a few electrophoresis experiments in different buffers under standardised conditions allows the amino acids making up the protein to be identified (often by comparison with known samples).

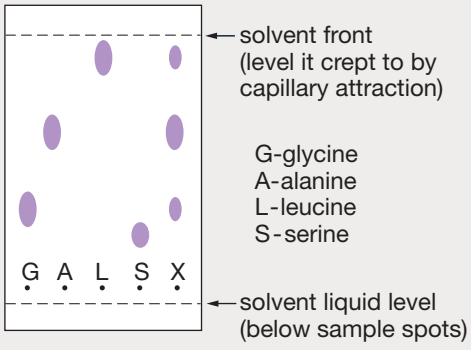
14.11 CHROMATOGRAPHY AND ELECTROPHORESIS COMPARED

The important similarities and differences between the two techniques just described are:

- Both chromatography and electrophoresis separate mixtures of amino acids.
- Chromatography separates them on the basis of their *different solubilities* in polar and non-polar solvents. Electrophoresis separates them on the basis of their c*harges and sizes*.
- In electrophoresis we can vary the charge on a particular amino acid by changing the pH of the solution: this gives considerable flexibility in choosing conditions to get a good separation of particular amino acids. In chromatography, while changing the solvent gives some control over the degree of separation we can get, it is less effective than varying pH in electrophoresis.
- Electrophoresis requires more expensive equipment than does paper chromatography.
- Both methods are powerful tools for forensic chemists in that they allow the easy identification of the amino acids present in a protein and so can identify the protein and possibly its source.

Exercises

19 Paper chromatography was used to identify the amino acids present in a sample, X. The unknown sample and samples of four known amino acids were run under identical conditions, with the spots of sample being placed near the bottom of a sheet of filter paper which was then dipped into a suitable solvent: the solvent crept up the paper by capillary attraction. After development (to make the spots visible) the chromatogram was as shown below. Which amino acids are present in the unknown sample?



20 In Figure 14.5(b) the standard sample contained leucine, alanine, glycine, serine, lysine and aspartic acid. The order of their rates of travel up a paper under the conditions used is the order of listing: leucine is the fastest, aspartic acid the slowest. By comparing the location of spots on the chromatogram, identify the amino acids present in Samples X and Y.

***21** In a paper chromatography experiment using a stationary phase that is more polar than the mobile phase, which of the three amino acids, alanine, serine, aspartic acid, moves the greatest distance and which the least (in a given time)? Explain why.

22 The isoelectric points for some amino acids are given in Table 14.1. For each of the amino acids:

a glycine ***b** alanine **c** cysteine ***d** lysine

draw the structure of the predominant species present in a solution of pH:

i 2.0 **ii** 6.0 **iii** 11.0

23 You may use Table 14.1 for this question.

a In an electrophoresis experiment in a solution buffered at pH 4.0:

i Towards which electrode will each of the amino acids, glycine, valine and phenylalanine, migrate? Why?

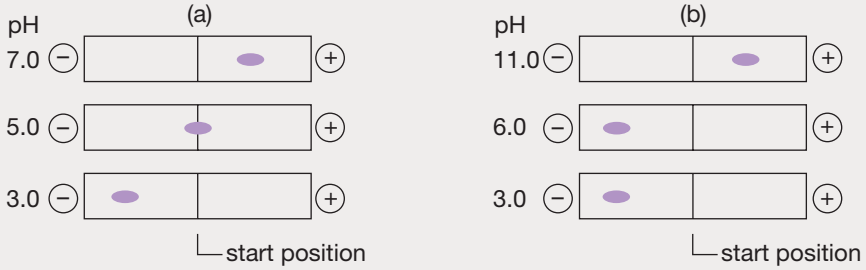
ii Of these three amino acids which will migrate the greatest distance in a given time and which the least? Explain.

b Using Figure 14.7 if necessary, towards which electrode in an electrophoresis experiment will aspartic acid migrate in a solution buffered at pH of:

i 1.0 **ii** 3.0 **iii** 8.0 **iv** 11.0

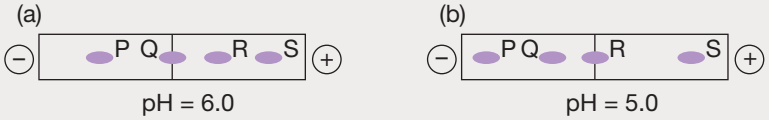
c How would the distance moved (in a given time) by aspartic acid in a buffer at pH 11.0 compare with the distance travelled in a buffer at pH 8.0? Explain.

24 a In order to identify an amino acid Q, a chemist ran electrophoresis experiments on it using buffers at pHs 7.0, 5.0 and 3.0. The developed electrophoresis strips are shown at (a) below. Which of the amino acids in Table 14.1 is Q most likely to be? Explain.



***b** Electrophoresis experiments were run on another amino acid T, using buffers at pH 11.0, 6.0 and 3.0. Results are shown at (b) above. Identify T from the amino acids in Table 14.1. Explain your reasoning.

***25 a** To identify the amino acids present in a mixture, a chemist ran an electrophoresis experiment using a buffer of pH 6.0. The result was as in (a) below. If the only possibilities are the amino acids in Table 14.1, identify as many spots (labelled P, Q, R, S) as you can. (You will not be able to identify all of them from this one experiment.)



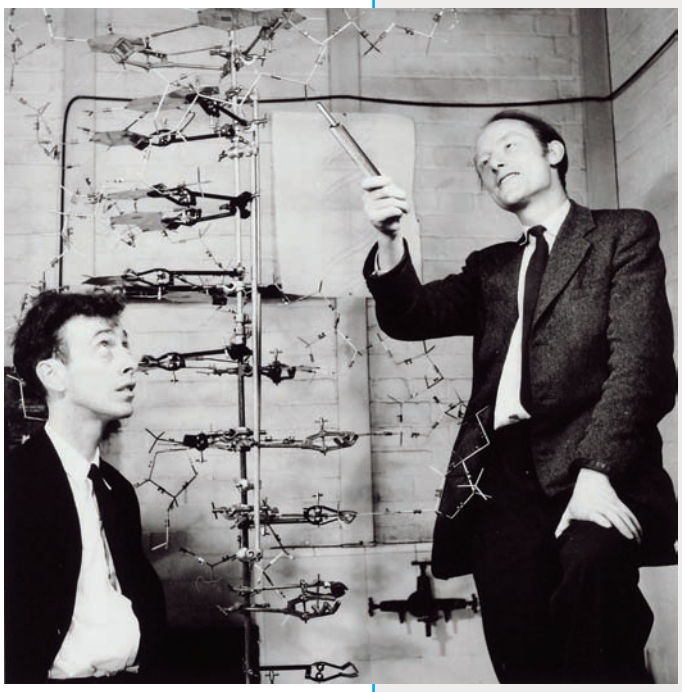
b The chemist repeated the experiment at pH 5.0 and obtained the result in (b) above. If possible identify additional spots and suggest possibilities for spot(s) you are unable to identify. Explain the basis of your identifications.

14.12 DNA

Let us now turn to the substance that in recent years has emerged as the most powerful tool for identifying the person responsible for a disputed activity, namely DNA. DNA was first used in a criminal trial (in Britain in 1987) to convict a rapist, and has been regularly used in court cases ever since. DNA can establish that two biological samples (such as blood, semen, saliva, skin, hair follicle) came from the same person with an error of less than one in ten billion. DNA analysis is also widely used in cases of disputed paternity and to establish other familial relationships between people.

DNA is present in the nucleus of all cells: it carries the genetic information that determines the nature of the cell, regulates its growth and division, and controls the biosynthesis of enzymes and other proteins required for the proper functioning of the cell.

DNA molecules are huge: molecular weights are of the order of 10^{10}. The molecules exist as a double stranded helix coiled around a rod of protein material like cotton around a reel to form what are called chromosomes: the length of a DNA strand (molecule) is over 10 cm though the diameter is only about 2 nm.



The discoverers of the double helix structure of DNA, James Watson (b. 1928) at left and Francis Crick (b. 1916) seen with their model of part of a DNA molecule in 1953

14.13 STRUCTURE OF DNA

DNA is an abbreviation for **deoxyribonucleic acid**—*acid* because the molecule contains many phosphoric acid groups, *nucleic* because the molecule occurs in the nuclei of cells, *deoxyribo* because it also contains the sugar deoxyribose (Structure 2 on page 466).

DNA is a polymeric molecule for which the monomer units are called *nucleotides*.

A **nucleotide** consists of a sugar molecule (deoxyribose in DNA) attached to a phosphate unit and to a base as shown schematically in Figure 14.8(a).

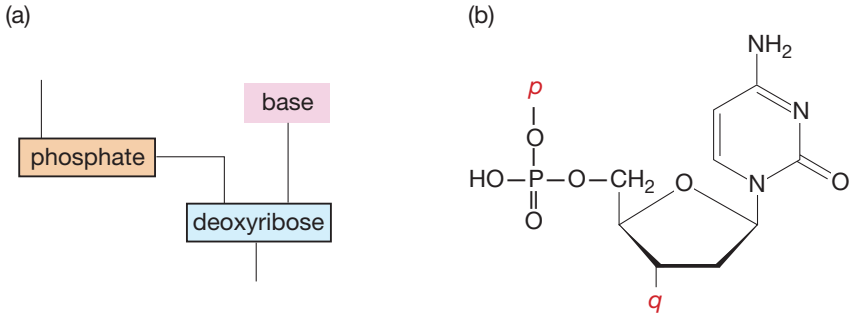
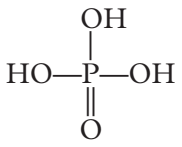


FIGURE 14.8
The structure of a nucleotide, the monomer for DNA: (a) a schematic drawing, (b) a specific example. *p* and *q* are the points where the nucleotide joins on to its neighbouring nucleotides to make the polymer chain

One specific nucleotide is shown in (b). A molecular model of a nucleotide is shown in Figure 14.9.

Phosphoric acid, H_3PO_4, has the structure:



The four bonds around the P atom are arranged tetrahedrally.

The nucleotides of DNA are joined together into long chains as shown in Figure 14.10: (a) is a schematic drawing while (b) is an example of a specific piece of a DNA molecule. There are only four bases used in DNA—adenine, guanine, cytosine and thymine: these are shown in Figure 14.11.

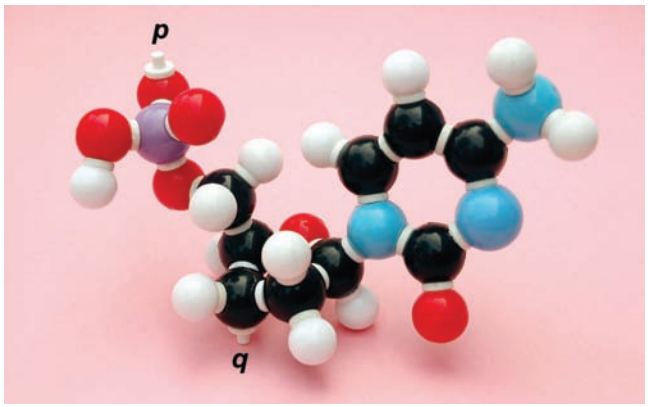


FIGURE 14.9
A space-filling model of the nucleotide shown in Figure 14.8(b). Purple is P, while the other colours are as in Figure 14.1

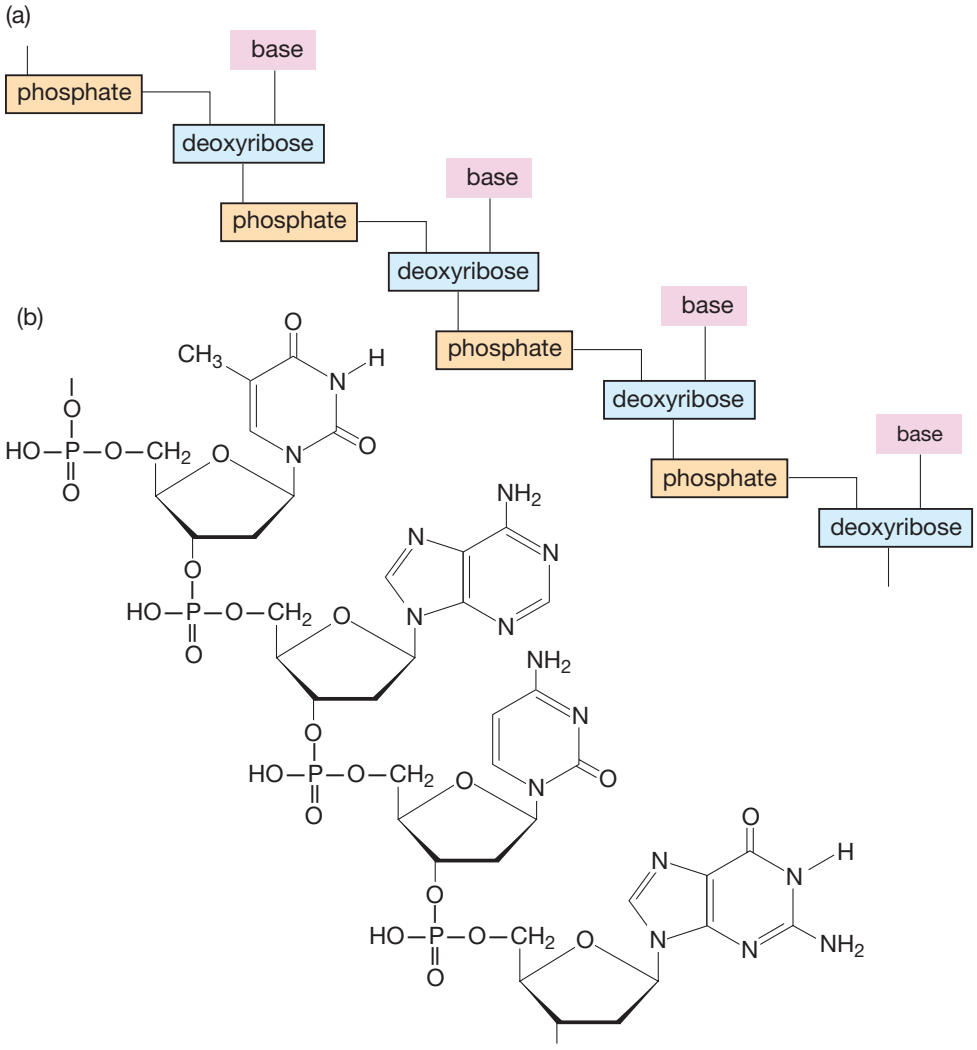


FIGURE 14.10
The structure of DNA: (a) schematic, (b) a section of the actual molecule

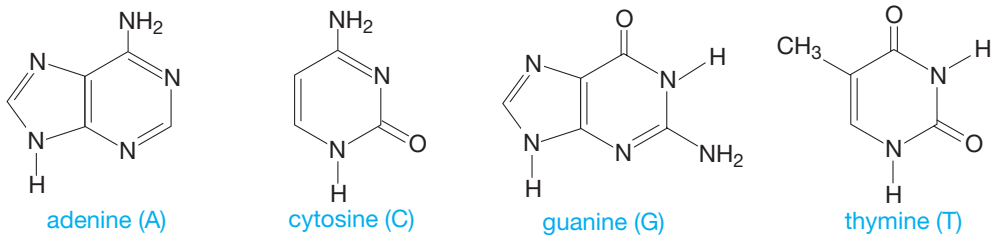


FIGURE 14.11
The four bases that occur in DNA; the letters A, C, G, T are commonly used as abbreviations for these bases

DNA occurs as a double helix as shown in Figure 14.12. Imagine this as two older style telephone hand-set leads intertwined. The sugar–phosphate chains are on the outside of the helixes with the bases pointing inwards. The two helixes are held together by hydrogen bonding between pairs of bases.

This hydrogen bonding occurs only between adenine and thymine and between guanine and cytosine: it is shown schematically in Figure 14.12. Figure 14.13(a) shows the details of this hydrogen bonding while (b) shows schematically how hydrogen bonding holds the two strands together.

A strand of double-helix DNA contains about 10 to 100 million nucleotides and has a molecular weight of the order of 10^9 to 10^{10}.

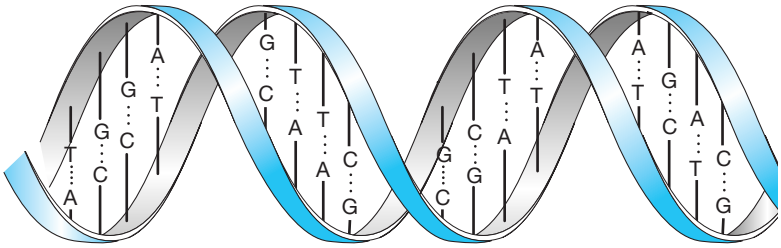


FIGURE 14.12
The double helix structure of DNA: hydrogen bonding is shown schematically

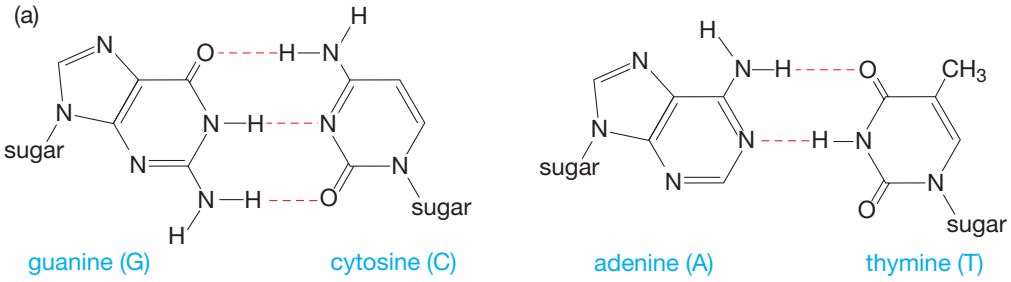
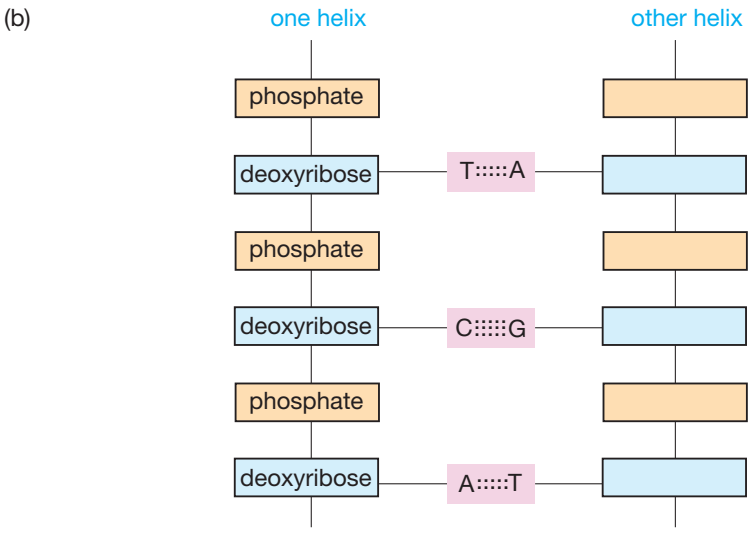


FIGURE 14.13
Hydrogen bonding between the pairs of bases in DNA: (a) its actual chemical nature, (b) schematically how it holds the two strands together



Exercises

26 **a** Why are the four compounds in Figure 14.11 called bases?
b Draw the conjugate acid of cytosine.

27 **a** What is the molecular formula of each of the four bases in Figure 14.11?
b Determine the molecular formulae of the four nucleotide units that occur in DNA. Hence deduce an approximate empirical formula for DNA

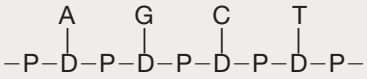
28 Draw the structure of the dinucleotide that contains:

a adenine and cytosine ***b** thymine and guanine

***29** The backbone of DNA molecules is always the same—alternate phosphate and deoxyribose units. Scientists often represent the structure of segments of DNA as the sequence of the bases on deoxyribose units, such as ACT or CGTA (letters from Figure 14.11). Draw the full chemical structure of these segments of DNA.

30 If a sample of DNA was completely hydrolysed (into bases, sugar and phosphoric acid), how many different compounds would be formed? Name them.

31 The diagram below represents a segment of one strand (helix) of a DNA molecule: P is a phosphate unit and D a deoxyribose unit, while A, C, G, T represent the attached bases. Using the same abbreviations, draw the complementary strand (helix) for this: that is, the strand that is hydrogen bonded to it.



32 (*For the mathematically inclined*) If a gene consisted of thirty nucleotides (generally it's far more than this), how many possible genes could be formed from just the four bases in Figure 14.11 (with no restrictions on how many times any one base can be used)?

14.14 THE ROLE OF DNA IN ORGANISMS

DNA carries the genetic code that allows living cells to function and reproduce themselves. Genes are actually sections of DNA molecules. The sequence of bases along the section determines the sequence of amino acids (primary structure) in the protein that the gene controls the synthesis of. By controlling the primary structure of the proteins that are made, genes determine the characteristics of the organism (person). Each DNA molecule contains thousands of genes.

A person has 46 chromosomes, 23 coming from each parent. Each of these chromosomes consists of a long DNA molecule wrapped around (and bonded to) a rod of protein material, so each person has 46 different types of DNA molecules†. It is possible to separate this DNA from the protein material of the chromosomes. *This mixture is called* **a person's DNA**.

Because all humans are very similar organisms—same number of fingers and teeth, same shaped limbs, same metabolism, same breathing mechanism etc—most of the our genetic material is identical (except for that which controls colour of eyes and hair, height and weight, skin toning and so on). This means that there is great similarity between the DNA of one person and another. In fact over 99.9% of genetic material is the same from one person to another. When

† considering each complementary pair of hydrogen-bonded helices as one molecule, though strictly speaking it is two molecules.

talking about genetics we tend to focus on the 0.1% of genes that are different and often overlook the similarities.

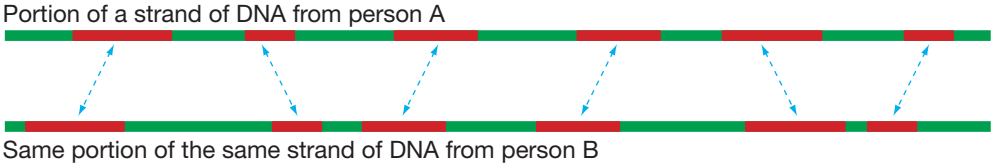
Despite this each person's DNA is unique.

ABOVE Preparing a sample for forensic analysis; care is needed to avoid contamination of the sample and to protect the operator from possible infection or poisoning

14.15 UNIQUENESS OF A PERSON'S DNA

The uniqueness of a person's DNA arises mainly from the non-coding bits of DNA that separate the genes. As already stated, genes are segments of the DNA chain (containing from 200 to 100 000 nucleotides). They are separated by other segments that are not involved in the reproduction process and have no known function. The coding sequences along the DNA strand (the genes) are called **exons** while the non-coding sequences are called **introns**, sometimes called 'junk DNA'. Introns vary significantly from one person to another and it is these regions of the DNA molecules that are used for DNA analysis for forensic purposes. The relation between exons and introns is shown schematically in Figure 14.14.

If people are related, then their introns show some similarity. Fifty per cent of a person's DNA comes from each parent. If an analysis is based upon a certain set of introns, then for a particular person half of those introns will be identical with the corresponding ones of the mother and half with those of the father. Brothers and sisters have 50% of their DNA introns in common, cousins 25%. People who are not related at all have very little of their DNA introns in common. The only people who have the same DNA are identical twins.



A and B have identical or nearly identical exon (coding) regions (red) but quite different intron (non-coding) regions (green)

FIGURE 14.14 Equivalent portions of DNA molecules from two different people showing similarities and differences between exon and intron regions; intron regions are used for forensic DNA analysis

14.16 DNA ANALYSIS FOR FORENSIC PURPOSES

DNA analysis allows forensic scientists to compare the DNA from different samples and to decide whether the samples came from the same person, related persons or completely different persons. This is a powerful tool for solving crime. It has been used to convict murderers and rapists, but it can also unequivocally clear a suspect in such a crime by showing that a sample could *not* have come from that suspect.

The steps in DNA analysis for forensic purposes are:

- separate the DNA from other material in the sample
- make multiple copies of selected segments of the DNA in intron regions (10 segments are currently being used in Australia)

- determine the length (i.e. number of nucleotides or bases) of these copied segments by electrophoresis
- compare samples from different sources or persons to see if they match.

Making multiple copies of selected segments

Portions of a DNA molecule can be copied by using what is called the **polymerase chain reaction**, often abbreviated to PCR. This works in a similar fashion to normal duplication of DNA in living cells.

Part of the process of cell division in living organisms is the duplication of the DNA in the cell. This occurs by the double helix unwinding and by each helix making its partner helix (its complementary helix): the end result is two DNA molecules, each with one strand from the original molecule and one newly synthesised strand.

If a pair of markers, called **primers**, is placed at the beginning and end of the segment of DNA we want to duplicate, the enzyme DNA polymerase can synthesise the complementary chains for both strands of that segment of DNA starting at a primer. The process can be repeated many times and so produce numerous copies of the required segment. This procedure for **PCR amplification** of segments of DNA, as it is called, is shown in Figure 14.15.

If several different primer pairs are used, we can duplicate fragments from several positions along the DNA molecules. Currently in Australia a mixture of ten primer pairs is used to grow fragments simultaneously at ten sites.

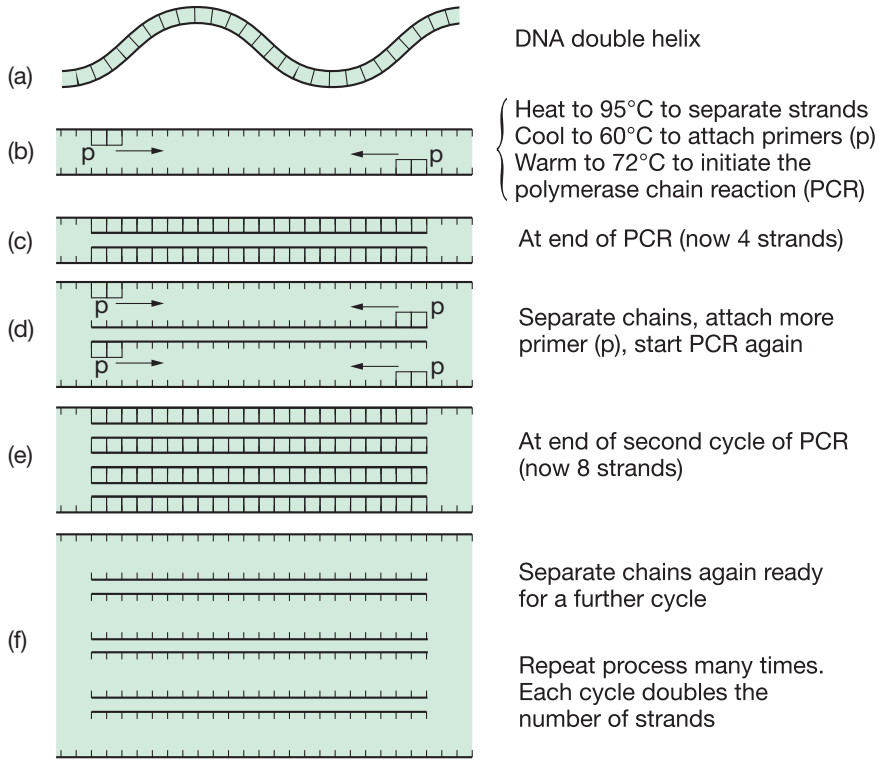


FIGURE 14.15
Making multiple copies of a segment of DNA by using the polymerase chain reaction (PCR). The cycle is commonly repeated 28 times

Because a person has two types of DNA—one from each parent—each primer pair will have grown two DNA fragments—one from each original type of DNA. These fragments may be the same or different.

Determining the length of the synthesised fragments

The synthesised DNA fragments are then separated by electrophoresis using a polyacrylamide gel. The longer the fragment, the more slowly it moves through the gel. A standard sample containing a mixture of polynucleotides of known chain lengths (number of base pairs) is run simultaneously and so the number of base pairs in the synthesised fragments is obtained by comparing the positions of the DNA fragments in the sample with those in the standard.

To detect the bands after electrophoresis and to sort out which bands come from which primer pairs, one member of each primer pair has a fluorescent 'tag' attached to it. This is a portion of a molecule that when hit by u.v. light emits visible light. Different tags emit light of different colours (usually yellow, blue and green). A laser shines u.v. light on the gel and a detector is able to detect the emitted fluorescent light and identify its colour. The detector determines the presence of a band and from the colour identifies the primer pair it came from and by comparison with the standard the operator determines how many base pairs the DNA fragment contains.

The non-genetic (or intron) material that is being amplified (or 'grown') for this analysis generally consists of multiple repeats of certain four-base units, for example −CGGT−CGGT−CGGT− or −ATTC−ATTC−ATTC−ATTC− repeated from 10 to 80 times, where C, G, A, T are the bases in Figure 14.11. Rather than record the number of base pairs in each fragment, the common practice is to record the number of repeat units, called **short tandem repeats (STRs)**. In the two examples just given there are three and four repeat units respectively.

A DNA profile

Each pair of primers attaches to a DNA molecule at a certain known position: the above analysis tells us how many base pairs or how many repeat units there are in the synthesised fragments that start from those positions. Remember that there were two synthesised fragments per pair of primers because there were two types of DNA—one from each parent. This information can be compiled into a **DNA profile** which gives the number of repeats in each of the two fragments that formed from each of the locations used. For example, a possible profile might be:

Location[a]	L	M	N	O	P	Q	R	S	T	U
Number of repeats	[b]	33, 34	23, 25	18, 18	22, 26	18, 23	37, 37	12, 16	32, 32	39, 46

a *Scientists use particular names or 'addresses' for these locations (called loci) that include the identity of the chromosome and the position on it.*
b *one location gives the sex only: xx for female, xy for male.*

At any particular locus (position) the number of repeats that can occur is limited: for example at position M the only numbers of repeats that can be found may be 32, 33, 34, 35 and 36: that is, there are only five possible repeat numbers. For position N, 10 different repeat numbers may be possible. If there were only five repeat numbers possible for each of the 9 positions in the above table (the tenth only tells the sex), then at each position there would be 15 possible combinations (remember 12, 13 is equivalent to 13, 12—order is not significant so by convention the smaller number is always given first) so with nine loci the

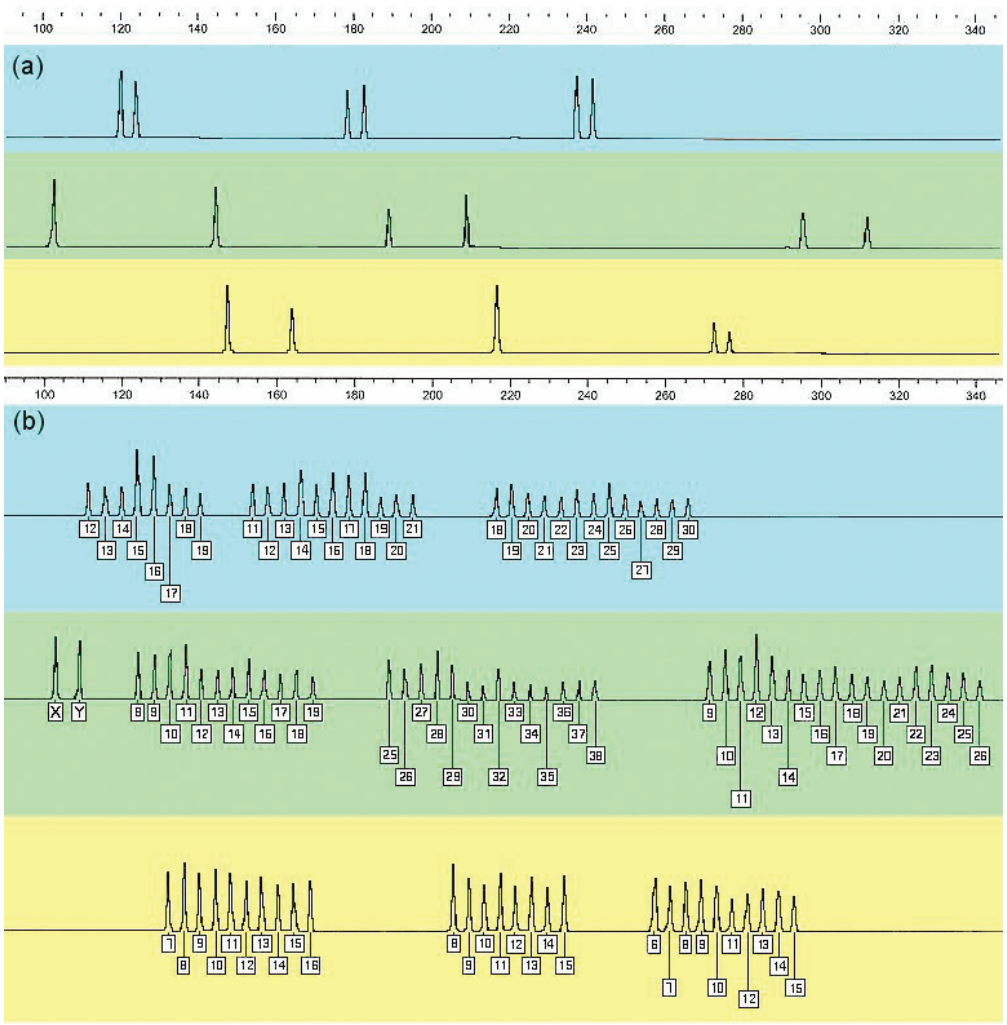


FIGURE 14.16 Electrophoresis results for a forensic DNA analysis; the peaks correspond to bands on the electrophoresis gel. **a** is the sample being analysed while **b** is a standard that contains all possible fragment lengths that can occur at each position used for the analysis. Each peak in the standard is labelled with the number of short tandem repeats that it contains. The blue, green, yellow shading on the chart represents the fluorescent colours emitted by the bands; the different colours allow the peaks (bands) to be assigned to the correct positions (loci) on the DNA. The chart has been simplified to some extent for teaching purposes

number of possible combinations would be $15^9 = 4 \times 10^{10}$ (compare with the world population of 6×10^9). In other words the chance of two people having the same profile would be less than 1 in 4×10^{10}. On average there are more than five repeat numbers possible per position, so this is an overestimate of the likely chance of a coincidental match.

The following example is too simple to be realistic but it illustrates the method used to identify the source of a DNA sample.

Example 1

Suppose that we had a sample from a crime scene and samples from three suspects. Suppose we analysed all four samples using five primer pairs, which set off synthesis at five positions (or loci). Let us call the positions H, I, J, K, L (see the note under the table on p. 503 about labelling). The electrophoresis analysis gives us the number of repeats in the two chains formed from each of the five positions in each of the samples. These results are shown in Table 14.2. To identify the suspect whose DNA matches the sample from the crime scene, we look to see which suspect's sample has produced identical numbers of repeats for each position.

TABLE 14.2 DNA profiles for a crime scene sample and three suspects (greatly simplified)

	Number of repeat units at the position			
Position on DNA	Crime scene sample	Suspect A	Suspect B	Suspect C
H	23, 26	21, 22	23, 26	23, 25
I	32, 34	32, 32	32, 34	31, 36
J	19, 23	19, 24	19, 23	22, 22
K	15, 18	17, 20	15, 18	19, 22
L	12, 15	10, 13	12, 15	10, 18

We see that suspect B's repeat numbers exactly match those of the crime-scene sample. This shows that the sample came from suspect B; it could not have come from suspect A or C.

This illustration used only five loci. Forensic scientists more commonly use ten loci (sex plus nine other). The more loci that are used, the smaller the chance of a coincidental match. The chance of different persons producing eighteen identical repeat numbers (from nine loci) is negligible.

Forensic versus medical/scientific DNA analysis

What has just been described is the method currently being used in Australia for DNA analysis for forensic purposes; that is, for the purpose of identifying the person from whom a biological sample came. It is virtually of no use medically or scientifically. For the medical and scientific study of DNA quite different methods of DNA analysis are used: they concentrate on the exon (genetic coding) regions of DNA, not the intron regions.

There has been intense medical and scientific study of DNA in recent decades, aimed, among other things, at:

- identifying and characterising genes in the DNA molecule
- identifying the variations in genes that lead to genetic abnormalities (such as cystic fibrosis and phenylketonuria) and to susceptibilities to certain diseases (such as breast cancer)
- determining the complete sequence of nucleotides (and the genes) along all the strands of the human DNA molecules, the so-called *human genome project* that was completed in 2000, after an unprecedented cooperative effort by thousands of scientists around the world.

For such studies, DNA molecules are generally isolated from individual chromosomes, then cut into smaller pieces (100 to 200 nucleotides) by what are called *restriction enzymes*. These smaller pieces are cut up further by chemical reactions, then the sequence of nucleotides determined by using electrophoresis and fluorescent or radioactive tags. Because these small pieces overlap, the sequence of nucleotide in the longer fragments can be worked out and ultimately the sequence along the whole DNA molecule. The procedure is similar to that used for sequencing proteins (Section 14.6).

The techniques currently used for forensic DNA analysis are 'spin-offs' from the methods originally developed for scientific research into DNA.

14.17 USES OF DNA ANALYSIS

Forensic scientists use DNA analysis for:

- identifying the person who produced a biological sample found at a crime scene: typical samples are blood, sperm, saliva, skin and hair follicles† with blood or saliva being preferred
- identifying the father of a child in disputed paternity cases
- establishing other familial links when there is a need to verify the claim of one person to be a relative of another person.
- identifying bodies from disasters such as the South-East Asia tsunami of December 2004, aeroplane crashes and terrorist bombings such as the ones in New York in 2001, Bali in 2002 and Madrid in 2004.

In DNA analysis the certainty of the identification of the person from whom the sample came is such that about the only way an accused person can dispute the analysis is to try to establish that the sample did not necessarily identify the actual culprit of the crime or that the crime scene sample had been contaminated, tampered with or substituted for; this is one reason for using great care with the chain of custody as discussed in Section 13.1.

The advantages of DNA analysis are:

- it is *extremely accurate* in identifying the source of the sample (though that may not necessarily be the culprit of the crime)
- it is *extremely sensitive*—only a minute sample is required (because of the PCR amplification).

The disadvantages are that

- it is quite time-consuming to perform
- it requires a specialised laboratory containing expensive equipment with well-trained technical staff and
- it can only be applied to certain biological samples such as blood, semen, saliva or hair follicles.

Some other techniques such as mass spectrometry (Sections 15.4 to 15.7) and infrared spectroscopy (Section 13.6) can be applied to a wider range of forensic samples (that is, not restricted to biological samples), but where DNA analysis is applicable it is much more sensitive than those other techniques.

14.18 DATA BANKS

Because of the great power of DNA analysis to identify persons guilty of criminal offences, moves are afoot in many parts of the world to establish data banks of DNA.

DNA data banks are collections of DNA profiles with which law enforcement authorities can compare samples collected from crime scenes, and where a match is found, from which they can obtain the identity of the person from whom the database sample came.

In Australia in the past decade various states have set up DNA data banks and now these have been combined into a national data bank. What this data bank records is just the DNA profile which consists of the sex and nine pairs of numbers, the number of tandem repeat units at each of the nine selected

† It is the hair follicle (root) that contains DNA, not the hair fibre itself.

intron regions that are used in the analysis; similar data from crime scenes from unidentified people are also stored.

The data bank is used in a similar way to a bank of fingerprints. A newly collected sample can be compared with the stored profiles, and if a match is found then the person from whom the sample came can be conclusively identified; with an extensive database it could be a very powerful tool for solving crimes. To date the only people required to provide DNA samples are convicted criminals in jails.

There is considerable opposition to such DNA data banks from civil liberties groups who claim that the widespread collection of DNA samples is an invasion of an individual's privacy. The argument is similar to the one used against the widespread collection of fingerprints. Often innocent suspects volunteer their DNA in order to clear themselves, but civil libertarians argue that people should not have to prove their innocence. While there is pressure from law enforcement authorities to make it compulsory for suspects to provide DNA samples, at the moment it is voluntary (unless a court orders otherwise) and there is considerable support for the existing situation.

It should be emphasised that the collection of the type of DNA profile currently being used is no more or less invasive of privacy than the collection of fingerprints. Fears that such DNA information could be used by insurance companies to weed out high risk cases or by potential employers to eliminate applicants who may have a tendency to certain diseases are completely unfounded, because the forensic DNA profile is quite useless for any genetic diagnosis purposes; it relates to the intron regions, not the exon regions of DNA. However there is considerable concern within the medical profession that the availability of a wide range of genetic diagnostic tests could lead to pressure from insurance companies and employers for applicants to provide such genetic information, but that is a completely separate issue from the matter of forensic DNA data banks.

WEBSITES

http://www.crimtrac.gov.au/dna.htm
(click on DNA; a simple account of DNA profiling, how it is done, history and data bases)

http://www.industry.gov.au/biotechnologyOnline/human/dnaprofile.cfm
(an introduction to DNA profiling; click on *DNA profiles for forensic use* and *DNA can reveal family relationships* in the left-hand column. You might enjoy solving the two problems in *DNA profiling interactives*.)

http://www.ornl.gov/TechResources/Human_Genome/home.html
(information about the human genome project and DNA generally though not strictly relevant to NSW HSC; click on *Education* for general information. There are some good graphics in the *Picture gallery*, though slow to download by dial-up connection.)

Exercises

***33** To establish which of two males was the father of a baby, DNA analyses were done on samples from the baby, the mother and the two males. The method used was as described in Section 14.16 using six primer pairs for PCR amplification (i.e. 'growing' six fragments from the DNA). The numbers of repeat units in each of these fragments (as determined by electrophoresis) are tabulated below. Which male is the father of the baby? (Use the mother's DNA profile to identify the baby's fragments that came from her, then determine the male from which the baby's other fragments came.)

Position on DNA	Number of repeat units at the position			
	Mother	Baby	Male A	Male B
P	21, 22	19, 22	20, 22	19, 25
Q	37, 40	37, 42	32, 35	40, 42
R	17, 24	17, 24	19, 19	24, 26
S	14, 17	13, 17	14, 18	12, 13
T	25, 29	23, 25	28, 30	23, 28
U	13, 16	10, 13	13, 17	10, 12

34 What is the most likely source of error in using DNA to identify the culprit of a crime?

35 Three men were suspected of raping a woman who suffered abrasions and loss of blood in her struggles during the incident. Two of the suspects had blood on their clothes which they claimed came from scratches they had received in other activities. DNA analyses (using just five loci instead of the usual nine) were performed on saliva swabs taken from the three suspects and from the victim, on the blood on the suspects' clothes and on a vaginal swab from the victim. The results are tabulated below. What can you deduce from the DNA analyses?

Position on DNA	A	B	C	D	E
Victim's vaginal swab	20, 21, 26, 27, 29	32, 34, 35, 37, 38	13, 14, 16, 17, 18, 19	12, 13, 14, 15, 17, 18	21, 23, 25, 28
Victim[a]	20, 26	37, 38	13, 18	14, 15	23, 28
Suspect 1[a]	22, 28	33, 39	15, 17	11, 16	24, 25
Suspect 2[a]	21, 27	34, 35	16, 19	13, 17	21, 28
Suspect 3[a]	26, 29	32, 37	14, 17	12, 18	23, 25
Blood on suspect 1	22, 28	33, 39	15, 17	11, 16	24, 25
Blood on suspect 3	20, 26	37, 38	13, 18	14, 15	23, 28

a from saliva swab

***36 a** Currently on the Internet there are laboratories offering paternity testing by mail. People can mail in DNA samples (saliva swabs, blood spots, hair follicles) of mother, child and suspected father to these laboratories which then send back the results; this can be done without the consent or knowledge of the mother or child. Suppose that a man used such a service to establish that he was not the father of the 'son' for whom he was paying considerable child support. Discuss the ethics of this procedure and the usefulness of this evidence in a possible court case to have his child-support payments discontinued.

b Also currently a mother can refuse to provide DNA samples from herself and her child to allow her divorced husband and alleged father of the child to have paternity testing done (again possibly to have child-support payments cancelled if he is not the father). Discuss the ethics of this situation.

37 Determine the sex and DNA profile of the sample being analysed in Figure 14.16.

Important new terms

You should know the meaning of the following terms.

a person's DNA (p. 500)
amide (p. 484)
amine (p. 481)
amine functional group (p. 481)
amino acid (p. 481)
biuret test (p. 488)
chromatography (p. 490)
dipeptide (p. 484)
DNA data bank (p. 506)
DNA (deoxyribonucleic acid) (p. 497)
DNA profile (p. 503)
electrophoresis (p. 492)
enzymes (p. 480)
exons (p. 501)
fibrous proteins (p. 480)
globular proteins (p. 480)
introns (p. 501)
isoelectric point (p. 494)
mobile phase (p. 490)
ninhydrin test (p. 488)
nucleotide (p. 497)
paper chromatography (p. 490)
PCR amplification (p. 502)
peptide bond (p. 484)
polymerase chain reaction (PCR) (p. 502)
polypeptide (p. 485)
primers (p. 502)
protein (p. 480)
short tandem repeats (STRs) (p. 503)
stationary phase (p. 490)
tripeptide (p. 485)
zwitterion (p. 483)

CHAPTER 14

Test yourself

1 Explain the meaning of each of the items in the 'Important new terms' section above.
2 Describe the difference in structure and function of fibrous and globular proteins. Name examples of each.
3 Draw the neutral molecule and zwitterion structures of two amino acids. Explain why the zwitterion form is the dominant one.
4 Draw structures of and name five amino acids.
5 Draw the structure of a dipeptide and write an equation showing its formation.
6 Draw a general structure for a protein.
7 How would you hydrolyse a dipeptide? Write an equation for the reaction. How would you completely hydrolyse a protein?
8 What are the three steps involved in determining the primary structure of a protein?
9 Describe the ninhydrin and biuret tests, including the results they give. What compounds do they identify?
10 What can you deduce about a compound that gives a positive ninhydrin test but a negative biuret test?
11 Explain how you would use paper chromatography to separate a mixture of amino acids. How would you identify the separated amino acids?
12 What is the basis for a separation of amino acids in paper chromatography? If two amino acids do not separate completely, how would you try to improve the separation?
13 Explain how you would separate a mixture of amino acids by electrophoresis. What is the basis for this separation?

14 Why does changing the pH of the electrolyte being used often improve the separation in electrophoresis?

15 What are the advantages and disadvantages of electrophoresis compared with paper chromatography for separating amino acids?

16 Give the full name for DNA and describe its structure: include a schematic drawing of it.

17 What elements are present in DNA?

18 Draw a structural formula for one nucleotide.

19 If a sample of DNA were completely hydrolysed, how many compounds would be formed? Name them.

20 What is the function of DNA in organisms?

21 Outline the method used for analysing DNA for forensic purposes.

22 Why is forensic analysis of DNA based upon the non-genetic parts of DNA molecules (that is, upon 'junk' DNA)?

23 What is meant by PCR amplification and why is it so important in analysing DNA?

24 For forensic purposes what is meant by a person's DNA profile?

25 Outline the advantages and disadvantages of setting up a national DNA data bank.

CHAPTER 15

Instrumental analysis

IN THIS CHAPTER

Destructive and non-destructive testing
Gas–liquid chromatography
High performance liquid chromatography
Mass spectrometry
Cracking patterns
Forensic use of mass spectrometry
Other designs of mass spectrometers
Electromagnetic radiation
Separating light into its components
Emission of light from atoms
Explaining emission spectra
Atomic emission spectroscopy
Use of emission spectra by forensic chemists
Chemical progress and forensic outcomes

Many of the analyses described so far in this book require samples of at least a few milligrams or more. Often the samples available to forensic chemists are much smaller than this, so forensic chemists often need very sensitive analytical techniques.

In addition there is often the problem that the chemist is not allowed to damage the object to be analysed.

15.1 DESTRUCTIVE AND NON-DESTRUCTIVE TESTING

Most of the analyses discussed so far have essentially destroyed the sample by subjecting it to chemical reaction.

If the original sample is not recoverable, the analysis is called a **destructive analysis**; sometimes the term **destructive testing** of a material is used.

Often the forensic chemist is not allowed to carry out a destructive analysis, for example in the verification that a newly discovered artwork is genuine, establishing the authenticity of historical artefacts or ancient manuscripts, or investigations to prove that the metal in jewellery is a precious one (gold or platinum) and not a cheap alloy. **Non-destructive testing** is required in such cases.

There are some analytical techniques available for non-destructive testing of materials or objects; examples are X-ray analysis, neutron activation and reflectance spectroscopy. However such techniques are beyond the scope of this book.

As an alternative to non-destructive testing, the analyst may be allowed to take a very small sample from the object, such as a fragment of paint from an unobtrusive part of a painting or a scraping of metal from a concealed part of a piece of jewellery. Because such samples are generally very small, the analyst needs to have very sensitive techniques available. In this chapter then we shall consider four such sensitive analytical techniques:

1. gas–liquid chromatography
2. high performance liquid chromatography
3. mass spectrometry and
4. atomic emission spectroscopy.

15.2 GAS–LIQUID CHROMATOGRAPHY

Gas chromatography was briefly described in Section 6.2. Chromatography generally was introduced in Section 14.8. You should re-read those sections before proceeding.

There are two types of gas chromatography:

- *gas–solid chromatography*, which separates substances on the basis of their different strengths of adsorption on the solid and which is widely used for analysis of gases and volatile liquids, and
- *gas–liquid chromatography*, which separates substances on the basis of different solubilities in the stationary phase.

Gas–liquid chromatography, GLC, is the more versatile of the two in that it can separate a broader range of compounds, the main restriction being that the compounds must readily vaporise at temperatures below about 250°C. GLC is widely used for forensic analyses.

Basic arrangement

In a typical gas chromatograph the sample to be analysed (either a mixture of liquids or a solution) is introduced by a syringe into the hot flowing carrier gas (generally helium or argon) as shown in the diagram at the top of p. 197. The sample vaporises immediately and the analysis proceeds by the components repeatedly dissolving into the liquid stationary phase and vaporising back into the gas again. The most frequently used columns today are *capillary columns*; these are long hollow silica tubes typically 0.1 to 0.5 mm internal diameter and from 10 to 100 m in length. The internal walls of the capillary tube are coated with a non-volatile liquid, the stationary phase. Figure 15.1 shows a typical capillary column. The more soluble a substance is in the liquid film, the more slowly it moves through the column.

FIGURE 15.1
A typical capillary column for a gas–liquid chromatograph. This silica column is 30 m long with an internal diameter of 0.15 mm; the internal surface of the column is coated with the non-volatile liquid polydimethylsiloxane. This 'column' has been wound into a coil of about 20 cm diameter

Detecting the separated substances

At the end of the column is a detector which signals when a component is leaving the column (that is, passing through the detector). This is illustrated in Figure 15.2; the recorder trace is called a *chromatogram*. In the last decade or so computer screens have replaced pen recorders. The area under each peak in the trace is a measure of the amount of that substance present.

A *flame ionisation detector* is commonly used to signal when each substance emerges from the column. This consists of a small

flame with a negatively charged electrode located just above it. No current flows through the detector circuit when just pure carrier gas is flowing through the detector, but when a component of the mixture flows into the detector, it is decomposed by the hot flame and produces some positive ions. These are collected by the negative electrode and so a small current flows through the detector circuit. This is amplified and displayed either on a recorder (older instruments) or on a computer screen (newer instruments). There are other detectors in use also, particularly for compounds that do not burn readily.

ABOVE A typical gas chromatograph, showing the oven containing the column and how the sample is injected; during an actual analysis the oven door would be closed

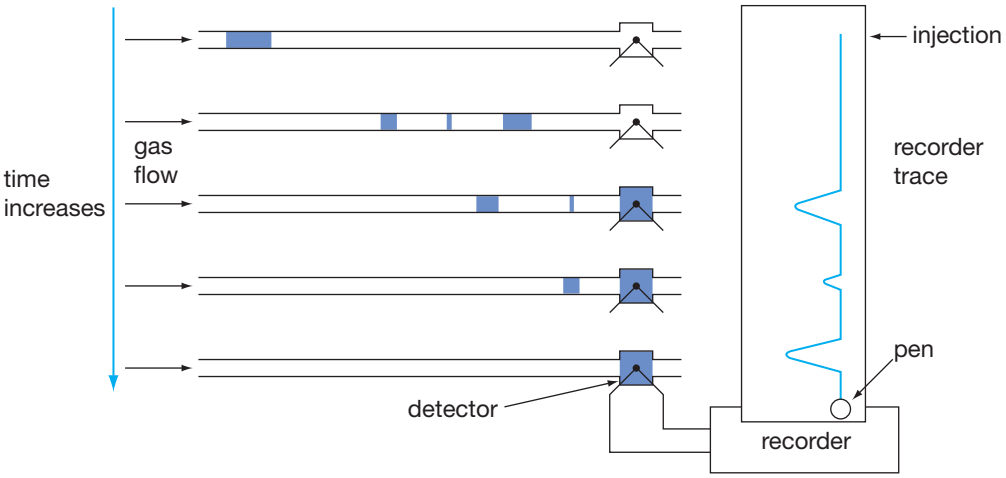


FIGURE 15.2
Schematic separation of a mixture by GLC, showing how the chromatogram (recorder trace or computer display) is generated

Identifying compounds

Compounds are generally identified by the time (after injection) that they take to reach the detector (called the **retention time**) when conditions (such as oven temperature and gas flow rate) are carefully controlled. This means that pure samples of all likely compounds in the mixture must be available and their chromatograms run under the same conditions as to be used in the analyses and their retention times recorded. For quantitative work the instrument must be calibrated by measuring peak areas for known amounts of compound injected. If the peaks are sharp, peak height can be used instead of peak area.

Two typical gas chromatograms are shown in Figure 15.3.

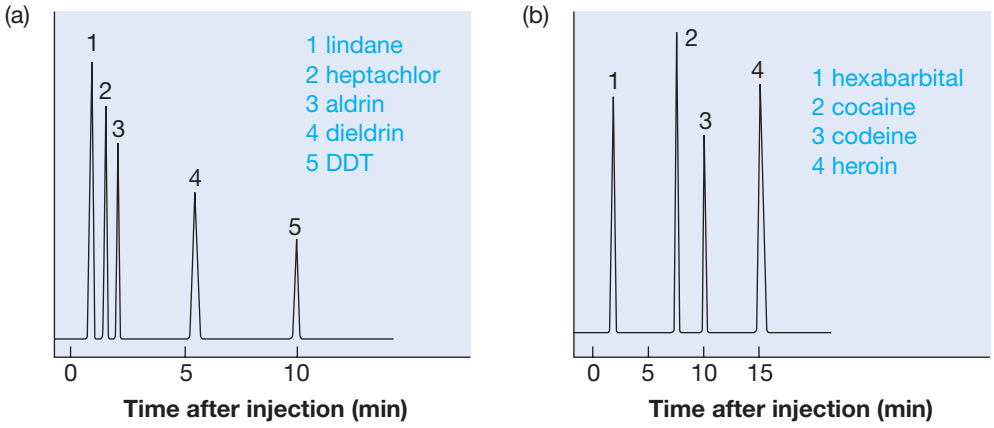


FIGURE 15.3
Capillary column gas–liquid chromatograms for (a) some common pesticides (b) some commonly abused drugs

Because the separation is based upon the different solubilities of the substances being analysed in the stationary phase, the degree of separation

can often be improved by changing the polarity of the stationary phase. Consequently analysts generally have several columns available, with different liquids of different polarities coated on their internal surfaces. If one column does not bring about a complete separation of the components of the mixture, another column is tried.

Forensic use of GLC

Gas–liquid chromatography is very sensitive; it can routinely detect microgram quantities of substances and in many cases quantities in the nanogram range. Hence it is widely used by forensic chemists to identify pesticide residues in foods, illegal drugs and steroids in athletes, trace pollutants in air and water and a wide variety of substances found at crime scenes. It is relatively quick and easy to perform (though great care is needed to avoid sample contamination), but the instruments are quite expensive.

Terminology

The following terminology applies to both GLC and HPLC (next section). The technique is called *chromatography*, the instrument used is a **chromatograph**, and the plot of detector signal versus time is called a **chromatogram**.

15.3 HIGH PERFORMANCE LIQUID CHROMATOGRAPHY

In high performance liquid chromatography, HPLC†, mixtures are separated based upon their differing solubilities in two liquids, one stationary and one moving.

Basic arrangement

The column is usually a metal tube 15 to 30 cm long and about 3 to 10 mm internal diameter. It is packed with very small, uniformly sized synthetic silica spheres (10 to 20 μm diameter) on the surface of which is adsorbed a liquid (the stationary phase). A solution of the sample to be analysed is admitted to the top of the column, then a pump pushes solvent through the column at high pressure to wash the components of the sample down the column. High pressure is needed because of the high resistance to flow of the tightly packed small particles. A suitable detector such as a refractive index monitor or an ultraviolet light absorption cell detects the various components of the mixture as they come out the bottom of the column. It feeds its signal to a computer, which displays the signal as a function of time after the start of the analysis; the chromatogram forms in the same way as shown in Figure 15.2. Figure 15.4 is a schematic arrangement of such a chromatograph; typical HPLC chromatograms are shown in Figure 15.5.

As with GLC, substances are identified by the time taken to wash through the column (the *retention time*) under controlled conditions such as temperature and solvent flow rate or by comparison with standard known samples run under identical conditions. From the size (height or area) of the peak the actual amount of the substance can be estimated.

† Originally HPLC was high *pressure* liquid chromatography, a name that indicated its key feature. The name changed when marketers took over from the scientists.

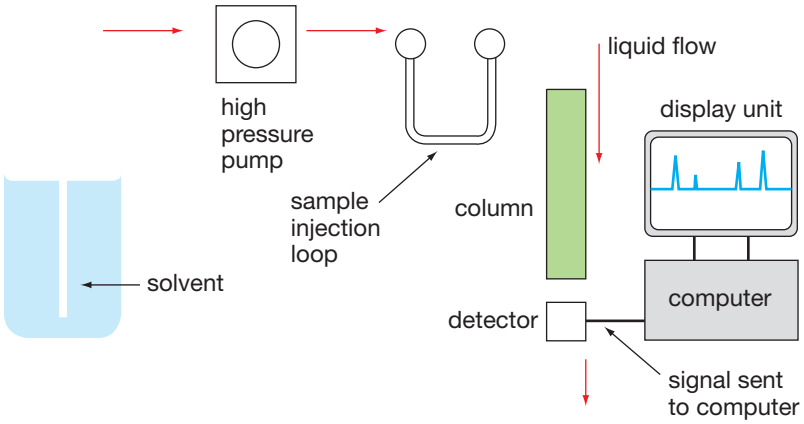


FIGURE 15.4
A schematic arrangement of a high performance liquid chromatograph

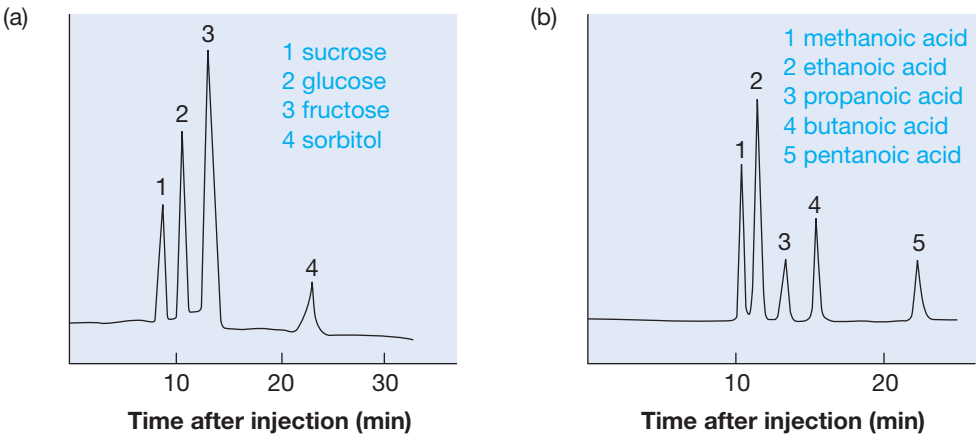


FIGURE 15.5
Two typical HPLC chromatograms: (a) an analysis of apple juice for sugars (b) analysis of a mixture of alkanoic acids

Basis of separation

In HPLC substances are separated on the basis of their *different solubilities in both the stationary and mobile phases*. A substance with a high solubility in the mobile phase and a low solubility in the stationary phase will move through the column at a greater rate than one that has a low solubility in the mobile phase and a high one in the stationary phase. As with GLC, because solubility depends upon polarity of the solvent, separations can often be improved by changing the stationary phase (changing the column). However HPLC has the added flexibility of being able to change the mobile phase also; it is more usual to change the solvent than to change the column.

Forensic use of HPLC

High pressure liquid chromatographic analyses are usually performed at room temperature and, since they do not involve vaporising the sample, can be used for mixtures of non-volatile substances such as carbohydrates, fats and fatty acids.

As with GLC, HPLC is widely used by forensic chemists for a great variety of analyses such as the ones listed for GLC, but also for ones where high temperatures cannot be used such as detecting traces of explosives. HPLC is often used to analyse biological samples such as polypeptides from proteins.

As with GLC, the technique is quick and simple, but great care is needed to avoid contamination, and again instruments are expensive.

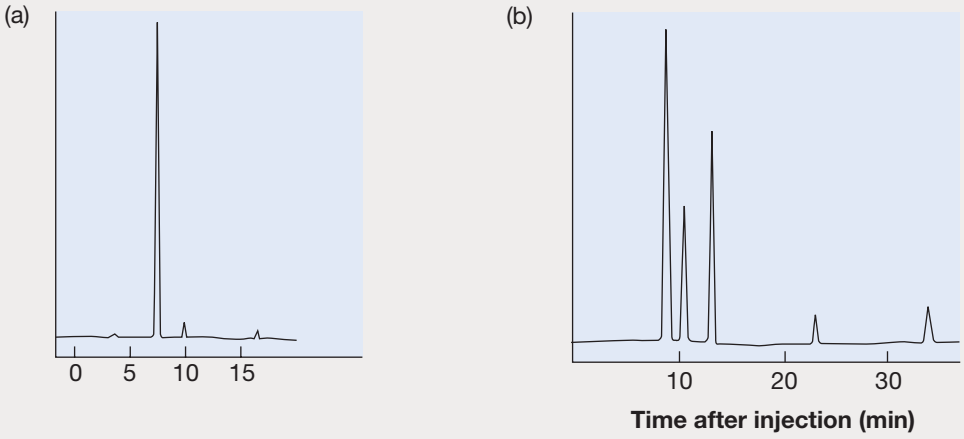
Both GLC and HPLC provide forensic chemists with ways of separating and identifying complex organic and biological samples. A major advantage of these techniques is that they first separate mixtures of complex substances into individual compounds, then provide a method of identifying them (via retention times, provided pure samples of the compounds involved are available). There are no other techniques that can analyse mixtures as effectively. Another advantage is that both techniques are extremely sensitive, being able to identify nanogram quantities. DNA analysis is more sensitive, but it can be applied to only a limited range of samples.

A high performance liquid chromatograph; the column is the white tube being held by the retort clamps. The inlet system is behind door (a); (b) and (c) are high pressure pumps (to supply a mixture of two solvents to the column from the large bottles), (d) is the display unit; the boxes behind the column contain the two detectors available on this instrument while above them is the overall controller for the unit

Exercises

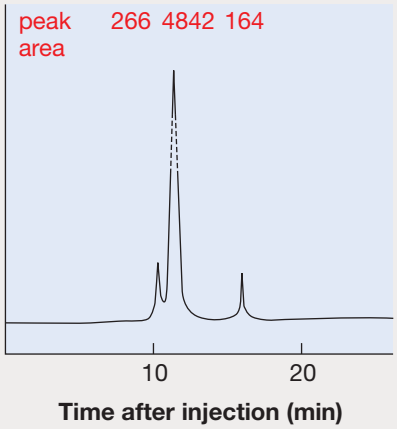
1 A mixture of pesticides was analysed by gas–liquid chromatography. The conditions used were as for Figure 15.3(a). The recorder trace showed a very large peak at a time of 1.6 minutes after sample injection, and small but definite peaks at 2.7 and 5.6 minutes. What do you conclude about the composition of the sample? Explain fully.

2 Solvent extraction was used to extract any likely drugs of addiction from a sample of the blood of a drug addict. The extract was then analysed by gas–liquid chromatography under the same conditions as were used for Figure 15.3(b). The result is shown at (a) below. What drug (if any) is present?



3 HPLC was used to check that a sample of apple juice had not been adulterated by the addition of some other substance(s). The resulting chromatogram, using the conditions of Figure 15.5(a) is shown at (b) above. Is the apple juice 'pure'? If not, how many substances have been added to it? Identify some or all of them (if possible). Explain your reasoning fully.

4 A sample of acetic acid was analysed by HPLC to check for any contamination using the conditions of Figure 15.5(b). The chromatogram is shown below. What if any contaminants are present in the sample and what is their approximate percentage in the sample? Why are your percentages only approximate?



15.4 MASS SPECTROMETRY

Mass spectrometry is a technique that measures the masses and relative intensities of the positive ions formed when a sample is bombarded with electrons or irradiated with u.v. light. The instrument that does this is called a **mass spectrometer**.

One design of mass spectrometer is shown in Figure 15.6. It consists of an evacuated tube, *AA*, with a semicircular section located between the poles of a strong magnet, *B*. There is an ion source at one end and an ion collector at the other. A gaseous sample is admitted to the source and positive ions are produced by electron bombardment (at *EB*): this is shown schematically in Figure 15.7.

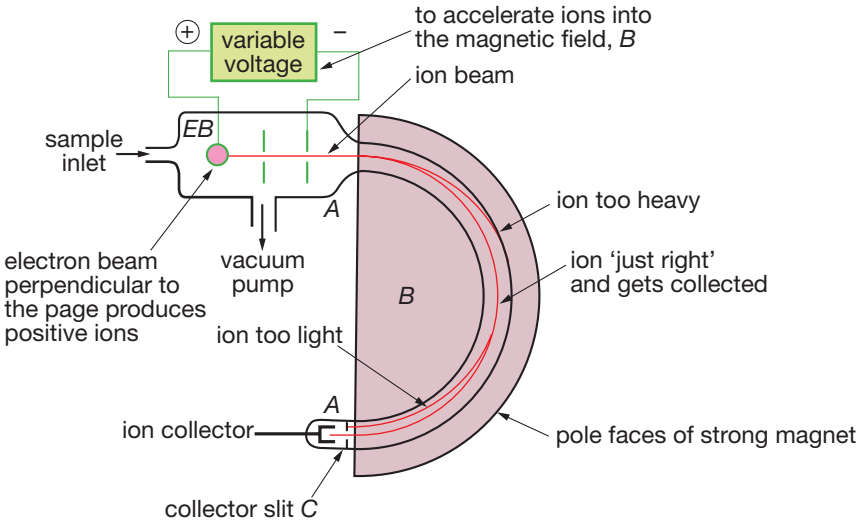


FIGURE 15.6
A simple mass spectrometer

These positive ions are collimated into a fine beam and accelerated through an electric field into the magnetic field. Magnetic fields cause moving charges to follow circular paths. With fixed accelerating voltage, *V*, and magnetic field strength, *B*, the radius of the path is dependent upon the mass-to-charge ratio of the ion. If the values of *V* and *B* are such that the radius, *r*, for argon (say)

corresponds to that of the apparatus, then argon ions will pass through the collector slit, *C*, and reach the collector. Krypton atoms with greater mass will follow a path of greater radius and will not pass through the collector slit; neon, of lesser mass, will follow a path of smaller radius and again will not be collected.

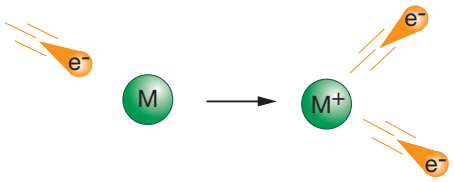


FIGURE 15.7
Formation of a positive ion by electron bombardment: M is an atom or a molecule

Atomic masses and isotopes

By measuring *V*, *B* and *r*, the value of the mass-to-charge ratio of the ion collected can be calculated, and if it is assumed that the charge is equivalent to one electron (the commonest situation), then the mass can be calculated. In this way, accurate atomic or molecular masses are obtained. Determination of atomic masses was the major use of mass spectrometry when the technique was first developed in the 1910s.

Today a mass spectrometer is used to produce what is called a mass spectrum. By varying the accelerating voltage, *V*, different ions can be focused onto the collector slit one at a time. *A record of the ion current (through the detector) as a function of accelerating voltage (that is, as a function of mass-to-charge ratio), as in Figure 15.8, is a* **mass spectrum**. The current through the detector is proportional to the number of ions of the mass being collected. A mass spectrum can be used to identify the components present in a simple gas mixture (such as air as in Figure 15.8(a)), or to determine the isotopes that make up an element.

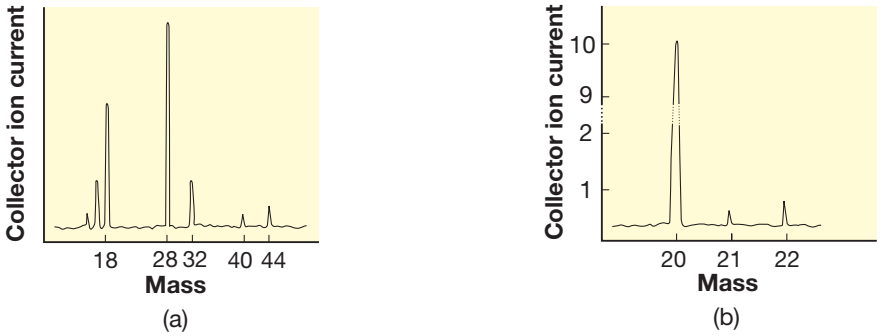


FIGURE 15.8
Typical mass spectra: (a) of residual air, showing peaks due to O^+, OH^+, H_2O^+, N_2^+, O_2^+, Ar^+ and CO_2^+, (b) of neon, showing three isotopes

Francis Aston, in 1919, found that the mass spectrum of neon contained not just the one expected line at mass 20, but three lines (as shown from a modern instrument in Figure 15.8(b)). This directly demonstrated the existence of isotopes. In the next few years many elements were shown to be mixtures of two or more isotopes.

15.5 CRACKING PATTERNS

In the mass spectrum of air in Figure 15.8(a), water produces not just one peak at mass 18, but also peaks at 17 and 16. This occurs because the ion originally formed, H_2O^+, is unstable and some of it falls apart to form OH^+ and then O^+ (masses 17 and 16 respectively).

The mass spectra of molecular substances are generally quite complex: they contain peaks corresponding to many different masses as shown for propane,

acetone (propanone) and propanal in Figure 15.9. Multiple peaks appear from the one compound because the parent ion formed by direct electron bombardment of the molecule is generally unstable and breaks up in a variety of ways to form ions of smaller masses.

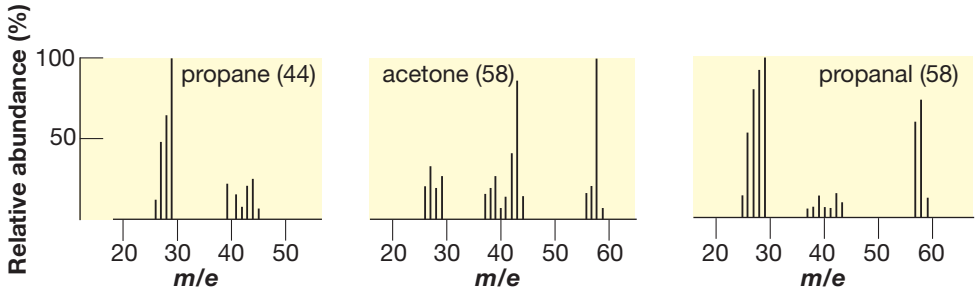


FIGURE 15.9
Mass spectra of propane, acetone (propanone) and propanal; molecular weights are given in brackets

The peak with the largest mass to charge ratio in a mass spectrum is generally from the parent ion; it can give an estimate of the molecular weight of the compound. Often, however, it is not the most intense peak, and for some compounds there is no parent-ion peak at all, because the parent ion is so unstable that it breaks up almost as soon as it is formed.

When propane is admitted to the mass spectrometer, ion bombardment produces the parent ion, $CH_3-CH_2-CH_3^+$ (mass 44, Fig 15.9). This easily loses one or two H atoms to form $CH_3-CH_2-CH_2^+$ and $CH_3-CH_2-CH^+$ (masses 43 and 42). In addition the parent ion can split off a CH_3 group, forming $CH_3-CH_2^+$ (mass 29) and by loss of further H atoms it forms masses 28 and 27.

The **cracking patterns** (as the collections of lines in mass spectra are called) for acetone ($CH_3-CO-CH_3$) and propanal (CH_3-CH_2-CHO) result from the following ion fragmentations:

$$CH_3-CO-CH_3^+ \text{ (mass 58)} \longrightarrow CH_3-CO-CH_2^+ \text{ (57)} \longrightarrow CH_3-CO-CH^+ \text{ (56)}$$
$$\hookrightarrow CH_3-CO^+ \text{ (43)} \longrightarrow CH_2-CO^+ \text{ (42)}$$
$$\longrightarrow CH_3 + CO^+ \text{ (28)}$$
$$CH_3-CH_2-CHO^+ \text{ (mass 58)} \longrightarrow CH_3-CH_2-CO^+ \text{ (57)}$$
$$\hookrightarrow CH_3-CH_2^+ \text{ (29)} \longrightarrow CH_3-CH^+ \text{ (28)} \longrightarrow CH_3-C^+ \text{ (27)}$$

These three compounds are relatively simple ones and lead to fairly simple mass spectra. As molecules become more complex, their mass spectra also become more complex because more and more fragmentation paths are possible. This means that there are more and more lines with differing relative intensities in the mass spectrum. As a result, compounds containing more than about 8 to 10 atoms have unique mass spectra. Consequently mass spectra can be used as 'fingerprints' for compounds.

15.6 FORENSIC USE OF MASS SPECTROMETRY

Because each compound (apart from very simple ones) has a unique mass spectrum, such spectra can be used to identify compounds. If a library of mass spectra is available, then the mass spectrum of an unidentified sample can be compared with those in the library and if an exact match is obtained, then

the sample can be confidently identified. Computers can be used to match unknowns with standards if the spectra are stored electronically. This is the major use of mass spectrometers by forensic chemists. There is one constraint: the operating conditions used to collect the mass spectrum of the unknown compound must be the same as those used to record the library spectra.

Even if the data bank of standard spectra contains no match for the sample being analysed, experienced operators are able to use the cracking pattern to deduce considerable information about the structure of the sample and often to identify it completely. This aspect of mass spectrometry is beyond the scope of this account.

Solid and liquid samples

Although the analysis in a mass spectrometer occurs in the gas phase, solid and liquid samples can also be analysed. A solid or liquid sample is admitted into the high vacuum chamber of the mass spectrometer then slowly heated until some of the sample vaporises and drifts into the ionising electron beam. Microgram samples can be analysed in this way.

Mixtures

Because individual compounds that contain more than about twenty atoms produce quite complex cracking patterns, mass spectra of mixtures (apart from very simple ones) are extremely complex and are generally too difficult to be resolved into individual compounds. Hence mass spectra are not very helpful for identifying the compounds in a mixture.

However if mixtures are first separated into individual compounds, for example by gas chromatography or high performance liquid chromatography, and then the pure compounds admitted to a mass spectrometer, the individual compounds can be identified by their cracking patterns. Today mass spectrometers are routinely coupled to gas chromatographs and high performance liquid chromatographs in order to identify the compounds as they emerge from the chromatographic column. This gives a more definite identification of the compounds than retention times alone: it is possible for two compounds to have the same retention time.

In order to work effectively in this way, the mass spectrometer needs to be able to scan the whole mass range (10 to 1000 or more) very quickly in order to process one chromatograph peak before the next one emerges. This requirement has led to the development of different designs of mass spectrometer.

15.7 OTHER DESIGNS OF MASS SPECTROMETERS

A wide variety of different designs for mass spectrometers is used today. The instrument described in Section 15.4 is a fairly old design: it uses a 180° sector magnetic separator (meaning that the ions travel through 180° in the magnetic field). Other designs use a 60° magnetic sector (ions travel through 60° only).

The part of the instrument that separates the ions according to mass is called the **mass filter**. Today several different types of mass filter are in use. A widely used one, called a quadrupole mass filter, is shown in Figure 15.10. This consists of four parallel metal rods with a direct voltage plus a radio-frequency voltage applied across opposite pairs. The ions oscillate down the space between the rods

and as the voltage is varied ions of different masses are able to get right through to the collector without hitting a rod and being lost.

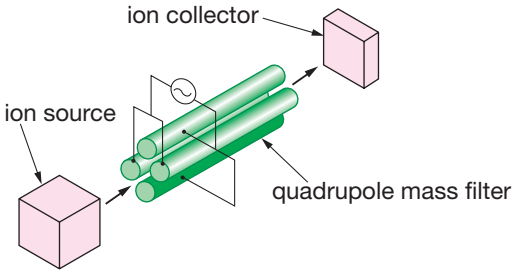


FIGURE 15.10
A quadrupole mass filter

Some mass spectrometers have two mass filters, one after the other. One is often a quadrupole and the other a magnetic sector filter.

Features that users desire are:

- high resolution (ability to separate ions of nearly identical mass)
- high scan speeds (ability to scan from masses 10 to 1000 in a few seconds)
- high sensitivity (ability to detect nanogram quantities).

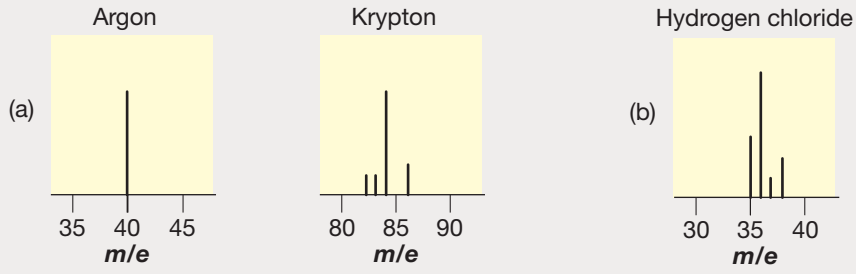
Quadrupole mass spectrometers are able to scan large mass ranges very quickly and so are often used in conjunction with GLC and HPLC.

Mass spectrometry, often after chromatographic separation, is routinely used to analyse a wide range of industrial, environmental and forensic samples. For example prohibited performance-enhancing drugs in the urine of athletes are commonly detected in this way.

A quadrupole mass spectrometer; the black cylinder is the source where the sample is ionised and accelerated into the mass filter (inside the cream box); the filter is similar to the one in the inset

Exercises

5 Portions of the mass spectra of argon and krypton are shown below in part (a). Explain why there is just one peak for argon but four for krypton.



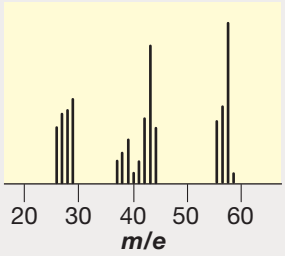
6 **a** Use a table of atomic weights to calculate the relative molecular mass of hydrogen chloride.

b The mass spectrum of hydrogen chloride has four peaks in the mass range 30 to 40 as shown in (b) above. Why is there no peak corresponding to the relative molecular mass calculated in (a)? Explain how this diagram establishes that there

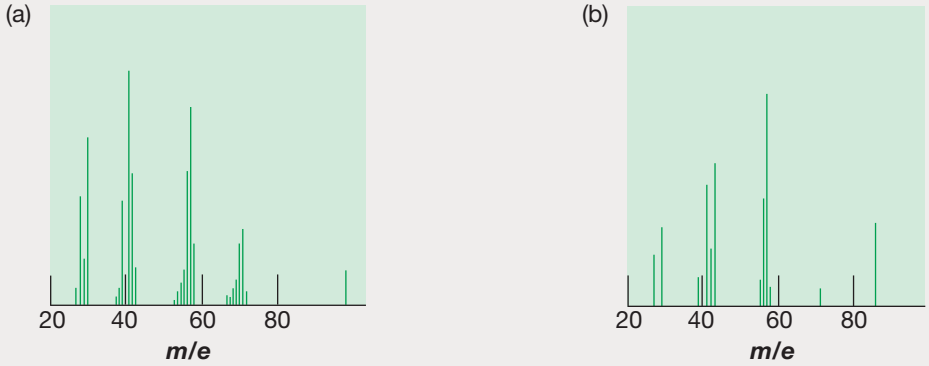
are two isotopes of chlorine. (*Hint*: we might reasonably expect the HCl^+ ion to dissociate into H and Cl^+.) What are the mass numbers of the two isotopes? What is their relative abundance? Explain how you decided this. Reconcile your conclusions here with your answer in (a).

c The mass spectrum of chlorine contains peaks as mass to charge ratios of 70, 72 and 74 (as well as others between 30 and 40). Is this mass spectrum consistent with your conclusions from (b)? Explain. What peaks would you expect between masses 30 and 40?

***7** A chemist suspected that the laboratory supply of acetone was not pure so ran its mass spectrum. The result is shown below. Using the standard spectra in Figure 15.9, is the acetone pure or not? Why? If it is not pure what is the most likely contaminant?



8 A team of forensic chemists measured the mass spectra of two unidentified compounds under the same conditions as were used for recording the spectra in their data bank. The spectra of the unknowns are shown below while some of the spectra from their data bank are given in Figure 15.11. Identify the unknown compounds.



9 Mass spectrometers with low resolution are not able to separate the parent peaks of nitrogen gas, carbon monoxide and ethylene. Explain why. Would you expect a mass spectrometer capable of separating peaks with mass differences as small as 0.002 to be able to separate the parent peaks of these three compounds? Explain.

Before turning to atomic emission spectroscopy, the final instrumental technique to be discussed here, we need to consider electromagnetic radiation which is at the heart of that technique.

15.8 ELECTROMAGNETIC RADIATION

Electromagnetic radiation was briefly introduced in Section 7.9, but let us take a closer look at it.

Visible light is a form of radiation. Three of its properties are:

1 It can transfer energy from one point to another without there being any transfer of mass (for example a light bulb shining on a photoelectric (photovoltaic) cell).

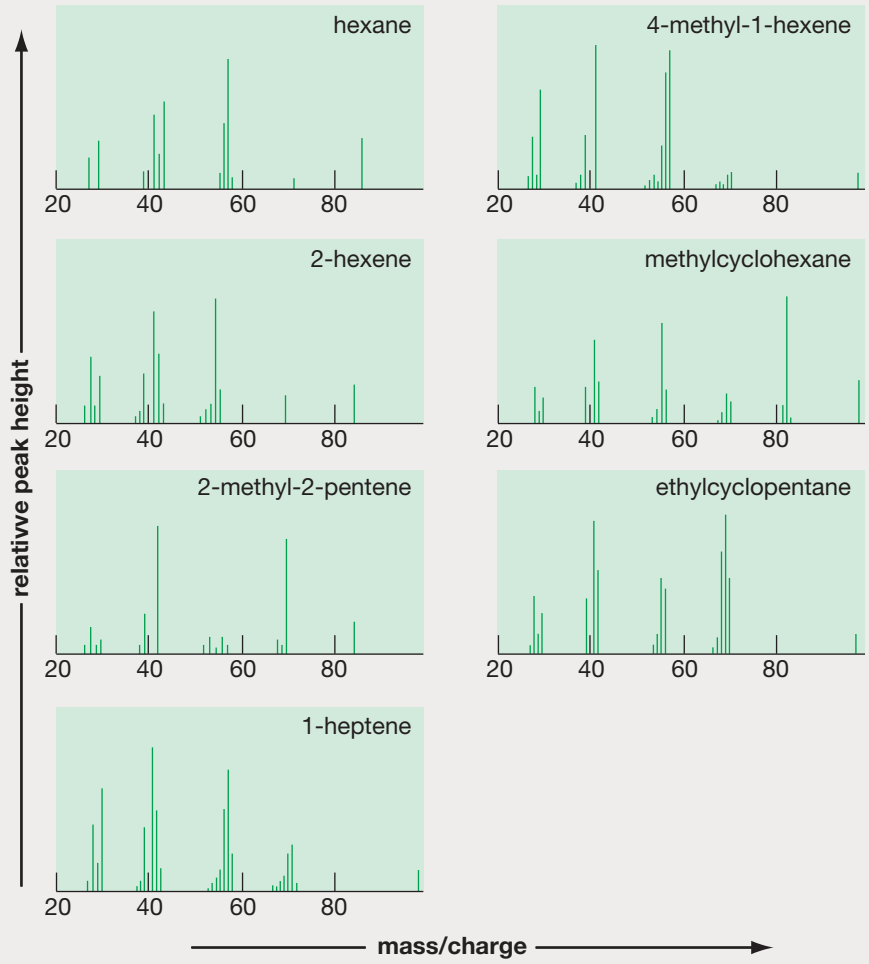


FIGURE 15.11
Some standard mass spectra from a data bank

2 It can be transmitted through empty space (as when sunlight travels from the sun to Earth).

3 It travels with a speed of 3.00×10^8 m s^{-1}.

There is actually a whole family of different types of radiation which have these same three characteristics: radio waves, microwaves, radiant heat or infra-red (i.r.) radiation, visible light, ultraviolet (u.v.) light, X-rays and γ-rays. The general name used for all these types of radiation is **electromagnetic radiation**. The name arises because they all consist of oscillating electric and magnetic fields. Electromagnetic radiation is a type of wave motion. *The essential difference between the various forms of electromagnetic radiation is that they have different wavelengths or frequencies.*

Wavelength and frequency (measured in **hertz**, Hz, or cycles per second) are related by:

$$\lambda\nu = c \qquad \text{... (15.1)}$$

where λ is the wavelength, ν is the frequency and c is the speed of electromagnetic radiation (speed of light).

Each of the types of radiation listed above corresponds to a particular range of wavelengths (frequencies) as shown in Table 7.4 on page 244. Frequency of electromagnetic radiation is commonly used in connection with broadcast radio; each radio station transmits at a particular frequency (wavelength): AM in the range 500 to 1600 kHz and FM in the range 60 to 300 MHz.

15.9 SEPARATING LIGHT INTO ITS COMPONENTS

Isaac Newton showed in 1666 that light could be broken into its various wavelength components by passage through a glass prism as shown in Figure 15.12. The instrument used to do this is called a **spectroscope** (if observations are made by eye) or **spectrograph** (if photographic plates are used).

White light (from an incandescent light bulb) is a mixture of all wavelengths from blue to red (400 to 800 nm). When it passes through a prism, it separates into its various wavelength components and looks like a rainbow. In fact rainbows are formed by this same dispersion of light, brought about in that case by small drops of water.

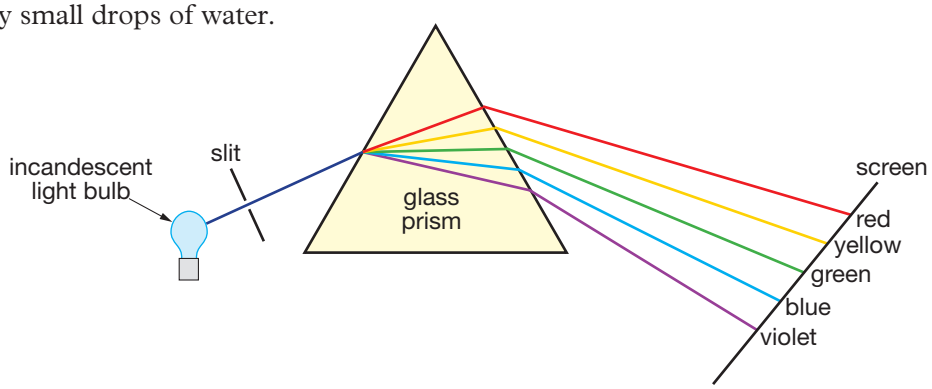


FIGURE 15.12
Dispersion of light by a prism

It is possible to use dispersion of light by such a prism to measure the wavelength of the light. That is, we can use a spectroscope or spectrograph to measure the wavelength(s) of light from a particular source.

In addition if we put a slit at an appropriate place on the dispersed light side of the prism (in addition to the one shown in Figure 15.12), we can select out light of a particular wavelength. An instrument which allows us to select particular wavelengths of light is called a **monochromator**. A prism is one type of monochromator. A monochromator was used in atomic absorption spectroscopy in Section 6.18 to select the required line (wavelength) from the emission lamp being used as the light source.

15.10 EMISSION OF LIGHT FROM ATOMS

Atoms whether in isolation or in compounds do not normally emit light. However if they are given extra energy, either by being heated to a high temperature or by being placed in an electric discharge, they can be made to emit light. This is the basis of atomic emission spectroscopy. It was briefly introduced in Section 6.17. You should re-read that section before proceeding.

If we excite electrons in an atom into higher energy levels, then as those electrons 'fall back' to their normal energy levels—called their ground states—energy is emitted as was shown in Figure 6.11 on page 224. This energy is emitted as electromagnetic radiation. There is a relationship between the energy released, ΔE, and the wavelength of the radiation emitted:

$$\Delta E = h\nu = \frac{hc}{\lambda} \qquad \text{... (15.2)}$$

Therefore the greater the amount of energy to be released, the shorter the wavelength of the radiation emitted. Big jumps in energy release u.v. light, medium jumps visible light and small jumps i.r. radiation.

Emission spectra

If a sample of an element is excited in a gas discharge tube and if the light emitted as the atoms fall back from excited states to ground states is examined (or photographed) through a spectroscope, we observe a series of bright (or coloured) lines on a black background. This is called an **emission spectrum** of the element or more precisely an **atomic emission spectrum**. The wavelengths of the emitted light from one element are different from those from any other element. Several emission spectra were shown in Figure 6.12.

If we look at a sample of an element in which some of the atoms have been excited (for example, in a discharge tube or in a flame), we observe a particular colour—yellow for sodium, red for neon, pink for hydrogen, green for copper†. We are then directly observing the radiation emitted by the excited electrons in atoms of those elements as they fall back to the ground state.

The question arises, then, why do elements emit light at different wavelengths?

15.11 EXPLAINING EMISSION SPECTRA

In order to find an explanation for emission spectra let us start with the simplest element, hydrogen and its spectrum. Figure 6.12 on p. 225 showed the visible portion of the hydrogen spectrum. Actually there are lines in the hydrogen spectrum extending from the u.v. region into the i.r. region. A more extensive form of the hydrogen spectrum is shown schematically at the top of Figure 15.13 on p. 526. Several series of lines have been identified in the figure.

The one electron in the hydrogen atom has available to it a large number of energy levels, labelled $n = 1$ for the lowest energy state, $n = 2$ for the next level up, $n = 3$ for the next one and so on as shown in Figure 15.13. Normally the electron in hydrogen is in the ground state ($n = 1$). When hydrogen is given energy in a discharge tube, the electron is excited into any one of these excited states or levels ($n = 2$ and above).

When the electron falls from one of these excited states back to the ground state, the amount of energy released is large (corresponding to the lengths of the arrows in Figure 15.13(a)) so by Equation 15.2 the wavelength is short. These transitions produce lines in the u.v. region of the spectrum.

The electron does not necessarily fall back to the ground state in one drop. If the electron falls from higher energy states to the first excited energy state ($n = 2$), then the energy differences are less (shorter arrows in the diagram) and so these transitions show up as lines in the visible region of the spectrum: (b) in the diagram. If the excited electron only drops down to the $n = 3$ or $n = 4$ states, then emissions appear in the i.r. region: (c) and (d) in the diagram.

Note that the energy that an electron can have is not continuously variable: there are only certain values of energy that the electron can have. We say that the energy is *quantised*, meaning that the energy that the electron can possess must be discrete whole number multiples of a basic 'parcel' of energy called a **quantum**.

In summary, the bright lines in the emission spectrum of an atom arise because electrons fall from various excited levels to lower energy levels.

† Do not look at discharge tubes without glasses or protective goggles; they emit considerable amounts of u.v. radiation that can damage eyes.

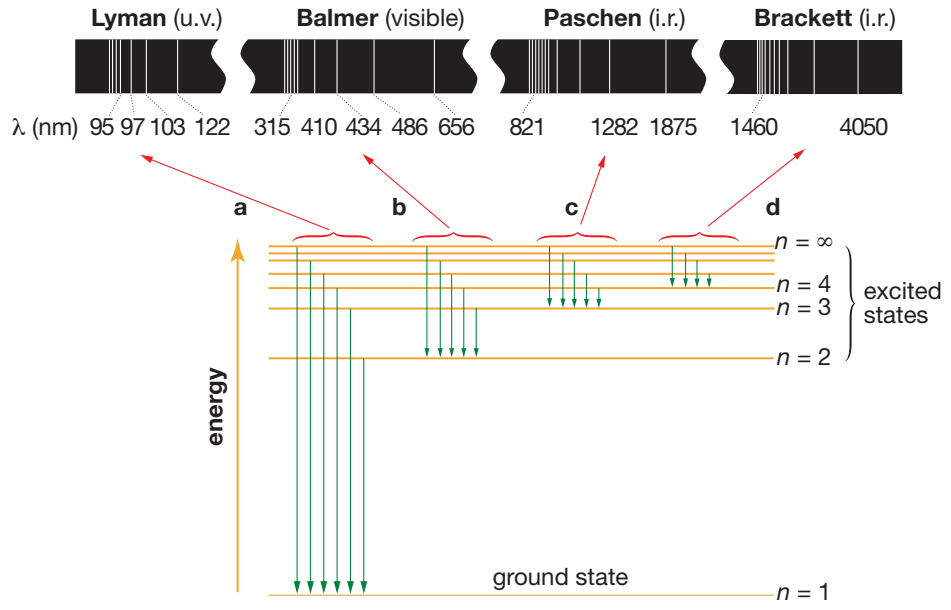


FIGURE 15.13
The hydrogen emission spectrum and its interpretation in terms of energy levels available to the electron

Why different elements have different spectra

The electrons in atoms of all other elements have available to them a set of energy levels that is qualitatively similar to that for electrons in the hydrogen atom. However all the levels have slightly different energy values. Therefore transitions between pairs of levels involve different energy changes and so light emissions are at slightly different wavelengths. This means that the emission spectra of all elements are different from one another. We can use these differences in emission spectra as a method of analysing for different elements.

15.12 ATOMIC EMISSION SPECTROSCOPY

Because the emission spectra of elements are all different from one another, they can be used to identify the elements present in mixtures. This is called **atomic emission spectroscopy**. Emission spectra are used as a type of 'fingerprint' for the element: lines in a sample to be identified are matched with lines in the spectra of known elements.

A typical experimental arrangement for atomic emission spectroscopy is shown in Figure 15.14. The sample to be analysed is heated to vaporise it and excite the electrons either in a flame or in a carbon arc (electric discharge across a pair of carbon electrodes quite close together). Emissions from the excited sample are concentrated with a lens then passed through some form of monochromator such as a prism which separates the emitted radiation into its different wavelengths. In older instruments a photographic plate was used to record the spectrum (as shown in the diagram). Later instruments generally used a fixed slit with a photomultiplier behind it as the detector and rotated the prism so that the different wavelengths of light fall on to the photomultiplier in turn. A photomultiplier is a sensitive device which converts light into an electrical current with the current being proportional to the intensity of the light. A recorder or computer was then used to draw a graph of intensity of light as a function of wavelength and this was the emission spectrum.

If the spectrum is recorded on a photographic plate or film, the instrument is called a *spectrograph*: if the spectrum is produced as a recorder or computer

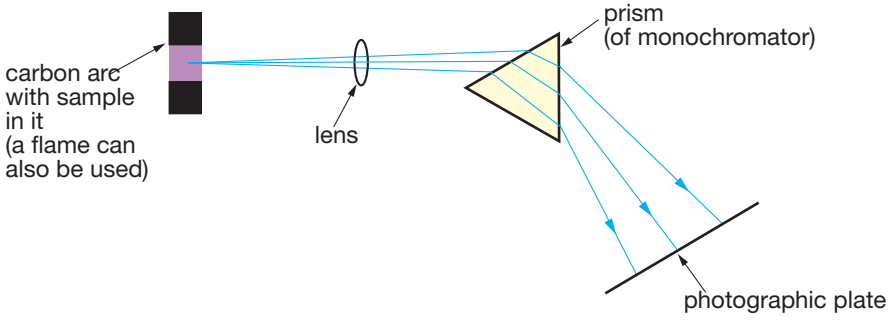


FIGURE 15.14
Schematic layout of a traditional atomic emission spectrograph. In modern instruments, generally called spectrometers, the monochromator is a grating (mirror with very fine lines scratched on it, typically 400 per cm) and the detector is an array of microscopic solid-state photodiodes, called a charge coupled device, CCD; this is similar to the CCDs in digital cameras and video cameras

graph, the instrument is usually called a *spectrometer*. As mentioned earlier, if the spectrum is observed by viewing through an eyepiece the instrument is called a *spectroscope*.

Figure 6.12 showed emission spectra of four elements; these were produced from discharge tubes. Figure 15.15 shows schematic spectra of six further elements; these were produced by exciting suitable compounds of the elements in flames or in carbon arcs. The spectra shown in Figure 15.15 are schematic in that they show the spectral lines as black stripes on a white background—a convenient way of drawing them. However it must be remembered that they appear in a spectroscope as coloured lines on a black background, or in a black and white photograph as white lines on a black background. Spectra in Figure 15.15 are really negatives (that is, reverses) of the true spectra.

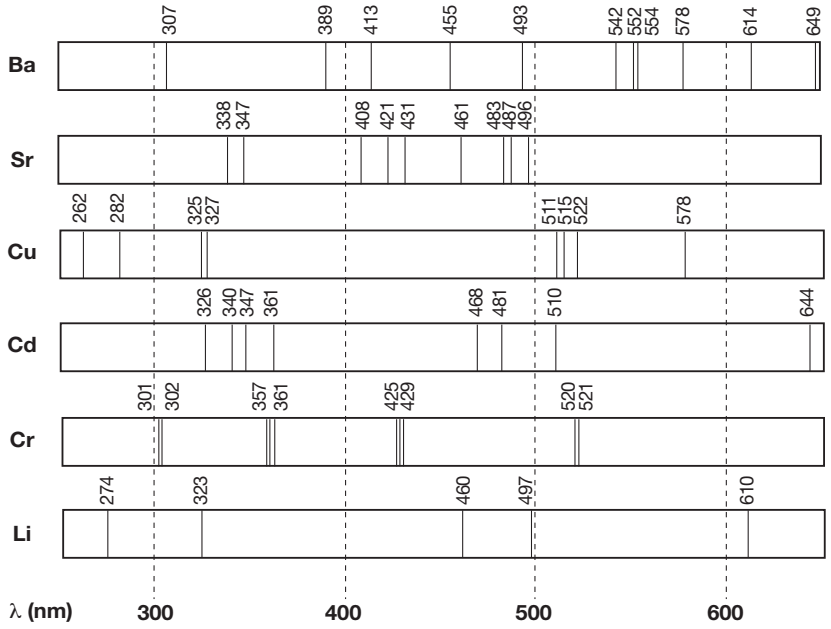


FIGURE 15.15
Some further emission spectra (in reverse form); for others see Fig. 6.12 on p. 225

By measuring the intensity of the emission lines, atomic emission spectroscopy can be used for the *quantitative* estimation of elements in samples, though until recent decades the technique has been more commonly used for *qualitative* analysis. For quantitative measurements atomic absorption spectroscopy has generally been preferred (Section 6.18).

Emission spectra and flame colours

For some elements, there is one particular electron transition which occurs far more readily than any other. This means that the emission spectrum of such an

element has one extremely bright line corresponding to this transition. In such cases, the flame in which the excited atoms are produced takes on the colour of that particular line. This is the origin of the flame colours and the tests based upon them that were described in Section 6.14. You should revise that section before proceeding: flame tests are part of this HSC option.

15.13 USE OF EMISSION SPECTRA BY FORENSIC CHEMISTS

Atomic emission spectroscopy was developed towards the end of the nineteenth century, but was mainly a tool for scientists for a couple of decades. As the twentieth century progressed analytical chemists began to use it for qualitative (and to a lesser extent, quantitative) analysis of samples. One of its earliest applied uses was in the steel industry to monitor compositions of steels as they were being made: it was particularly good for the rapid measurement of relative amounts of various metals in such mixtures.

Forensic chemists have used atomic emission spectroscopy routinely for many decades. During the middle decades of the twentieth century the inconvenience of photographic plates and slowness of recorders attached to photomultipliers limited its use particularly for quantitative work. However in recent decades with the development of faster and more sensitive detectors, such as CCDs (Fig. 15.14), and powerful data processors (computers) its use has become much more common.

It is now widely used to monitor many elements simultaneously and quantitatively. Water authorities routinely use it to measure concentrations of up to twenty elements in the one sample. While atomic absorption spectroscopy measures concentrations more accurately, the need to use a different lamp for each element being measured makes it inconvenient for routine analysis of many elements.

Some forensic uses of atomic emission spectroscopy are:

- determining the origin of small chips of paint at a crime scene by analysing the chip and likely sources of the paint and comparing them
- testing a claim that a painting is by a famous long-dead artist by comparing the atomic emission spectrum of a small sample of a particular coloured paint with a similar sample from a known work of the artist (remember forensic chemistry is not just about law cases—see p. 454)
- determining the origin of small pieces of a metal or alloy or of fragments of coloured glass found at a crime scene by comparing the spectrum of the sample with spectra from likely sources
- determining the source of a soil sample by measuring the concentrations of less common elements in the soil (Section 13.5) and comparing them with known soil compositions from various areas.

In many newer atomic emission spectrometers a plasma is used to excite the atoms. A *plasma* (in physical and chemical contexts) is a highly ionised gas that has a distinct glow: it results from the gas being heated to a very high temperature (typically 5000°C), for example by an electromagnetic induction furnace. Because of its very high temperature a plasma is able to excite a greater proportion of the atoms in a sample than a carbon arc is, and this increases the sensitivity of the instrument. This has made the technique even more useful for forensic purposes.

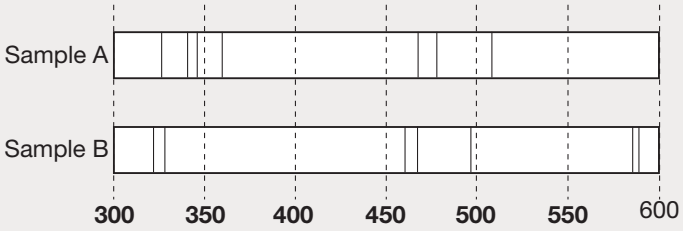
A limitation

Emission spectra can be used to determine the *elements* present in the sample. However, because compounds have to be decomposed into gaseous atoms (by high-temperature flames or electrical discharges), emission spectra tell us nothing about the nature of the compounds in which the elements were present. Consequently atomic emission spectroscopy is of little use in identifying organic compounds, because they contain only five major elements (C, H, O, N, S).

An atomic emission spectrometer with a plasma source (in the top window); below it is the sample handling system (to pump sample solution into the plasma to atomise it and to excite the atoms; the monochromator and detector are behind the grey cover near the operator's face

Exercises

10 A chemist had two samples of metal that were thought to be pure. In order to check the purity and to identify the samples the chemist volatilised a small portion of each sample in turn in a flame and recorded its emission spectrum between 300 and 600 nm. The spectra are shown below. Use Figures 6.12 and 15.15 to determine whether the samples are pure or not. If pure identify the element; if impure what elements are present? (Sometimes closely spaced lines are not completely resolved, and occasionally weak lines do not show up.)



11 The emission spectra in the region 350 to 600 nm of two unknown samples showed lines at:

a 358, 361, 408, 421, 425, 427, 429, 496 and 521 nm
b 361, 468, 481, 511, 515, 522 and 578 nm

Use the emission spectra in Figures 6.12 and 15.15 to determine which elements are present in the unknown samples. Note the warning at the end of Exercise 10.

12 A chemist prepared some lithium metal, and to check its purity determined its emission spectrum between 350 and 650 nm. Emission lines were observed at 610, 588, 497, 460 and 408 nm. Use Figures 6.12 and 15.15 to determine which, if any, impurities are present.

13 The five most intense emission lines for each of eight metals are shown (in nm) in the following table:

Iron	Cobalt	Nickel	Chromium	Manganese	Vanadium	Copper	Tungsten
358	341	341	358	261	309	282	362
372	344	349	359	279	318	325	387
373	345	352	361	280	411	327	401
382	350	357	425	403	438	522	407
386	357	362	427	404	439	578	430

The emission spectra were measured for three different alloys and they showed the following lines:

Alloy A: 318, 358, 372, 382, 386, 411, 438

Alloy B: 341, 345, 350, 358, 359, 372, 382, 386, 425

*Alloy C: 325, 327, 341, 352, 357, 362, 522, 578

Which elements are present in each alloy? Note the warning at the end of Exercise 10.

14 A person had been stabbed to death. The blade of the knife used struck a bone and a small fragment of it broke off and was recovered by forensic scientists. It was possible that the knife could have been resharpened and the damage removed. Two suspects of the murder had knives that could have been the murder weapon. To try to identify the murder weapon the scientists performed atomic emission analyses (in the wavelength range 300 to 420 nm) on the fragment from the body and on the knives from the two suspects. After removal of spectral lines due to iron (all the blades were alloy steels), the remaining emission lines between 300 and 420 nm were as shown below. Using the table in the previous exercise, determine which knife, if either, the fragment found in the victim came from. What conclusion can you draw from this result?

Murder weapon	309, 318, 341, 344, 345, 349, 350, 357, 362, 387, 407, 411 nm
Knife from suspect 1	341, 344, 345, 349, 350, 352, 357, 362, 387, 401, 407 nm
Knife from suspect 2	309, 318, 341, 349, 352, 357, 359, 361, 362, 411 nm

15 Which metals, in addition to iron, were in the alloys used to make the steels in the knives in Exercise 14?

15.14 CHEMICAL PROGRESS AND FORENSIC OUTCOMES

The increasing effectiveness of forensic investigations depends very much on new developments in analytical chemistry. Some examples will illustrate:

- There was little concern about mercury contamination of fish before about 1960, largely because small concentrations of mercury (below 1 ppm) could not easily be detected. After the development of atomic absorption spectroscopy in the 1950s (Section 6.18), it became possible to measure mercury concentrations well below the 1 ppm level and it was found that many species of fish were contaminated with low concentrations of mercury. This led to more stringent controls on releases of mercury to the environment and so to a diminution of the problem. More generally atomic absorption spectroscopy has led to better monitoring and control of all forms of heavy metal pollution.

- Before the development of spectroscopic methods such as those described in this chapter, the identification of an organic substance was a long and tedious process, involving a wide range of tests of which those described in Section 13.6 are the simplest. Unless samples were of gram quantities complete identification was rarely possible. Now with spectroscopic methods, samples as small as micrograms can generally be identified in less than an hour. Spectroscopic methods developed during the 1950s and 60s but have been refined almost continuously over the past four decades.
- Technological development in chromatography, such as capillary columns for gas chromatography and small uniformly sized spherical particles and high pressure pumps for liquid chromatography in the 1970s and 80s, led to more sensitive analyses for substances such as dioxin (a pollutant from garbage incineration and paper manufacture), pesticide residues in foods and illegal drugs in athletes. The development of techniques for coupling mass spectrometers to liquid and gas chromatographs further increased the ability of forensic chemists to detect smaller and smaller quantities of more complex substances. Gas chromatography coupled with mass spectrometry (GC-MS) is now widely used to screen athletes for prohibited performance-enhancing substances.
- The advance in chemistry (or biochemistry) that has most affected forensic investigations in the last decade is the development of methods of analysing DNA for identification of the source of a biological sample. As explained in Section 14.16 it is now possible to identify persons from very small DNA samples with an extremely high degree of certainty. This has greatly increased the ability of forensic scientists to identify the culprits of crimes. In addition it has allowed many persons convicted of crimes to prove their innocence even after ten or more years in jail (over 60 people in the US alone).

As a result of these and other advances in analytical chemistry, forensic chemists are today more likely to be able to solve crimes and other forensic problems than they were twenty to fifty years ago.

As progress is made in analytical methods, forensic chemists are quick to adopt them in order to solve the wide variety of problems they face in fields as diverse as solving crime, tracing sources of environmental pollution, contamination of products, detection of prohibited drugs in athletes and identification of bodies after disasters.

Important new terms

You should know the meaning of the following terms.

atomic emission spectroscopy (p. 526)
atomic emission spectrum (p. 525)
chromatogram (p. 514)
chromatograph (p. 514)
cracking pattern (p. 519)
destructive analysis (testing) (p. 511)
electromagnetic radiation (p. 523)
emission spectrum (p. 525)
hertz (p. 523)
mass filter (p. 520)
mass spectrometer (p. 517)
mass spectrometry (p. 517)
mass spectrum (p. 518)
monochromator (p. 524)
non-destructive analysis (testing) (p. 511)
quantum (p. 525)
retention time (p. 513)
spectrograph (p. 524)
spectroscope (p. 524)

CHAPTER 15

Test yourself

1 Explain the meaning of each of the items in the 'Important new terms' section above.

2 What problems does a requirement for non-destructive testing create for a forensic chemist? When is such a requirement likely to be made?

3 Describe the basic structure of a gas–liquid chromatograph.

4 Explain how high performance liquid chromatography differs from gas–liquid chromatography.

5 Explain how the chromatogram (that is, the graph of signal versus time after injection of the sample) is produced in GLC and HPLC.

6 How are compounds identified in GLC and HPLC?

7 Outline some forensic uses of gas–liquid and high performance liquid chromatography, including mention of situations where one technique is preferable over the other.

8 Explain how a magnetic sector mass spectrometer works (with a diagram).

9 Why is a mass spectrometer able to identify the isotopes in an element?

10 Why are there many lines (peaks) in the mass spectrum of an organic compound?

11 How does mass spectrometry help a forensic chemist identify substances?

12 List five types of electromagnetic radiation and explain why they are different from one another. Give common uses of them.

13 Describe the emission spectrum of an element such as hydrogen or mercury.

14 Explain why the emission spectrum of an element consists of a series of bright lines on a dark background.

15 What do we mean when we say that the energy of an electron is quantised?

16 In what way does the emission spectrum of one element differ from that of another element? Why do they differ in this way?

17 Explain how you can use emission spectra to identify the metals present in an alloy.

18 Why do certain elements impart particular colours to flames? List four elements that do this and state the colours they form.

19 What is the main limitation of atomic emission spectroscopy? As a consequence to what classes of compound is the technique of significant use?

20 Describe two scientific advances in the past thirty years that have been particularly useful for forensic scientists.

EXTENDED RESPONSE EXAM-STYLE QUESTIONS FOR OPTION 3

OPTION 3

Questions in this section are in a similar style to that currently being used in the extended response *questions in the New South Wales HSC Chemistry examination:* *for more information see p. 93.*

Marks (shown at the right-hand end of the question) are assigned to each question in order to indicate how much detail is required in your answer (that is, how much time you should spend on each question). The marks are on the HSC exam basis of 100 marks for three hours work (or 1.8 minutes per mark).

		MARKS
1	Discuss the importance of accuracy in forensic chemistry relative to other branches of chemistry.	3
2	Explain how an atomic emission spectrum is generated and observed.	4
3	Outline the composition and structure of DNA.	6
4	Explain the chemical difference between reducing and non-reducing sugars and describe a test you have performed to determine whether a sugar is reducing or non-reducing.	5
5	Discuss ethical issues that may need to be considered during an analytical investigation.	5
6	Evaluate the usefulness of atomic emission spectroscopy to forensic scientists.	5
7	Compare and contrast the chemical structures of glucose and deoxyribose.	4
8	Identify the chemical features a forensic chemist would concentrate on in order to determine the origin of a sample of soil and assess the usefulness of atomic emission spectroscopy for this purpose.	4
9	Explain the effect upon the isoelectric point of adding an extra $-NH_2$ group or $-COOH$ group to an amino acid.	4
10	Summarise the uses of DNA analysis for forensic purposes.	4
11	Outline an experiment you have performed to separate a mixture of organic substances by chromatography and describe the chemical and/or physical principles underlying the separation.	6
12	Discuss, with at least one example, how developments in analytical chemistry have led to improved outcomes in forensic investigations.	4
13	Evaluate the usefulness of mass spectrometry for forensic investigations.	4
14	Identify the main polysaccharides present in plants and animals and explain how they differ chemically from one another.	5
15	Explain why analysis of a DNA sample allows identification of the person the sample came from and why this identification is considered to be extremely accurate.	4
16	Outline the operating principles of gas–liquid chromatography and assess its usefulness and limitations for forensic purposes.	6
17	Compare and contrast the structures of nylon and proteins.	4

18 Describe two tests you could use to distinguish between classes of carbon compounds and assess the reliability of the conclusions you would draw from these tests. **6**

19 Using sucrose as an example, demonstrate the meaning of the terms condensation reaction and disaccharide. **4**

20 Justify the establishment of DNA data banks by law enforcement agencies and discuss some ethical issues associated with their establishment. **6**

REVISION TESTS FOR OPTION 3: FORENSIC CHEMISTRY

OPTION 3

TEST A

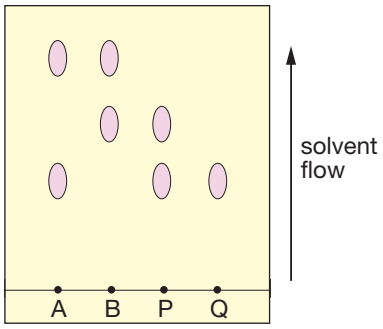
Total marks: 25 **Suggested time: 45 minutes**

MARKS

a i Describe how a mass spectrometer works. **3**

ii Explain why there are many peaks in the mass spectra of relatively simple compounds such as pentane. **1**

b i To identify the monosaccharides making up two disaccharides P and Q, a chemist completely hydrolysed them, then ran a paper chromatogram of the hydrolysed mixtures along with two reference samples. Reference sample A was completely hydrolysed sucrose (a mixture of glucose and fructose). Reference B was a mixture of fructose and galactose. The resulting chromatogram is shown below. Identify the monosaccharides present in P and Q. **2**



ii Explain the fact that the disaccharide Q produced only one monosaccharide. Q was a reducing sugar. Draw a possible structure for Q. What is the relevance of Q being a reducing sugar for doing this? Is this the only structure possible? Explain. **3**

c Describe an experiment you have performed using electrophoresis and explain how it separated the substances in your mixture. Suggest methods for improving the separation of the components in your mixture. **5**

d i Demonstrate the meaning of the terms peptide bond, dipeptide and polypeptide. **3**

ii Identify the difference in the products formed when proteins are hydrolysed by enzymes and by strong acids. **2**

e Outline the basic principles of gas–liquid chromatography and high performance liquid chromatography and evaluate their roles in forensic investigations. **6**

TEST B

Total marks: 25 **Suggested time: 45 minutes**

MARKS

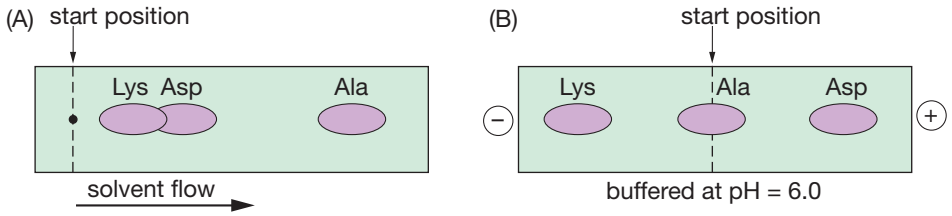
a **i** Distinguish between destructive and non-destructive testing. **1**

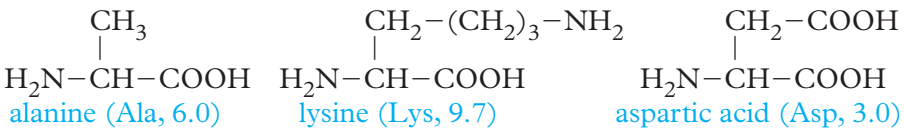
ii Discuss the constraints that a requirement for non-destructive testing or very unobtrusive destructive testing places upon a forensic investigation. **3**

iii Outline steps forensic scientists take to avoid contamination of samples between their collection and the completion of all analyses. **3**

b Demonstrate the meaning of the terms monosaccharide, disaccharide, and polysaccharide. **3**

c In the diagram below, A is a paper chromatogram of a mixture of the three substances, alanine, aspartic acid and lysine. B is the result of an electrophoresis experiment in a buffer with a pH of 6.0 on this same mixture. Structures of the three substances, with their common abbreviations and their isoelectric points given in brackets, are shown below the diagram.





i In A why is there a good separation between alanine and the other two, but only a poor separation between aspartic acid and lysine? **1**

ii Why is there a good separation between aspartic acid and lysine in the electrophoresis experiment? Draw structures to illustrate your explanation. **3**

iii Why has alanine hardly moved in the electrophoresis experiment? **1**

d **i** Describe the procedure you would follow for carrying out flame tests in order to identify certain elements. **3**

ii Explain why only some elements produce characteristic flame colours. **1**

e Assess the usefulness and reliability of DNA analysis in forensic investigations. **6**

OPTION 3 AND THE NEW SOUTH WALES HSC SYLLABUS

This section allows you to check that *Conquering Chemistry* has covered all necessary material for the Forensic Chemistry option of the New South Wales Higher School Certificate HSC Course.

Syllabus content

The following table lists (for Option 3) the items from the *students learn to* column of the HSC syllabus and shows where they are treated in *Conquering Chemistry* (*CCHSC*).

Location of HSC Course material in *Conquering Chemistry* for Option 3 Forensic Chemistry

Syllabus reference *Students learn to*	**Sections where found in *CCHSC***
9.9.1 *The job of the forensic chemist is to identify materials and trace their origin*	
■ outline precautions that may be necessary to ensure accuracy and prevent contamination of samples	13.1, 13.2
■ distinguish between organic and inorganic compounds	13.3
■ explain that there are different classes of carbon compounds including – hydrocarbons – alkanols – alkanoic acids which can be identified by distinguishing tests	13.6
■ explain that the inorganic chemical properties of soils and other materials may be useful evidence	13.5
■ discuss, using a recent example, how progress in analytical chemistry and changes in technology can alter the outcome of a forensic investigation	15.14
9.9.2 *Analysis of organic material can distinguish plant and animal material*	
■ identify that carbohydrates are composed of carbon, hydrogen and oxygen according to the formula: $C_x(H_2O)_y$	13.7
■ identify glucose as a monomer and describe the condensation reactions which produce: – sucrose as an example of a disaccharide – polysaccharides including glycogen, starch and cellulose	13.7 to 13.9
■ describe the chemical difference between reducing and non-reducing sugars	13.10, 13.11
■ distinguish between plant and animal carbohydrates' composition in terms of the presence of: – cellulose – starch – glycogen	13.9

Syllabus reference *Students learn to*	Sections where found in ***CCHSC***
9.9.3 ***Because proteins are a major structural and metabolic component of all living organisms, the analysis of protein samples can be useful in forensic chemistry***	
■ distinguish between protein used for structural purposes and the uses of proteins as enzymes	14.1
■ identify the major functional groups in an amino acid	14.1
■ describe the composition and general formula for amino acids and explain that proteins are chains of amino acids	14.1, 14.4
■ describe the nature of the peptide bond and explain that proteins can be broken at different lengths in the chain by choice of enzymes	14.4, 14.5
■ compare the processes of chromatography and electrophoresis and identify the properties of mixtures that allow them to be separated by either of these processes	14.8 to 14.11
■ discuss the role of electrophoresis in identifying the origins of protein and explain how this could assist the forensic chemist	14.10
9.9.4 ***DNA is an important compound found in all living things and is a most useful identification molecule***	
■ outline the structure and composition of DNA	14.12, 14.13
■ explain why analysis of DNA allows identification of individuals	14.15
■ describe the process used to analyse DNA and account for its use in: – identifying relationships between people – identifying individuals	14.16, 14.17
9.9.5 ***Much forensic evidence consists of very small samples and sensitive analytical techniques are required***	
■ explain what is meant by the destructive testing of material and explain why this may be a problem in forensic investigations	15.1
■ identify, outline and assess the value of the following techniques in the analysis of small samples: – gas–liquid chromatography – high performance liquid chromatography	15.2, 15.3
■ outline how a mass spectrometer operates and clarify its use for forensic chemists	15.4 to 15.7
9.9.6 ***All elements have identifiable emission spectra and this can be used to identify trace elements***	
■ describe the conditions under which atoms will emit light	15.10, 6.17
■ identify that the emission of quanta of energy as electrons move to lower energy levels may be detected by humans as a specific colour	15.10
■ explain why excited atoms in the gas phase emit or absorb only certain wavelengths of light	15.11

Syllabus reference *Students learn to*	Sections where found in *CCHSC*
■ account for the fact that each element produces its signature line emission spectrum	15.11
■ discuss the use of line emission spectra to identify the presence of elements in chemicals	15.12, 15.13

Compulsory experiments

The table below lists the compulsory experiments for this option (from the right-hand column of the syllabus), along with the location of relevant information in this book.

Compulsory experiments for Option 3 Forensic Chemistry

Experiment		Related material in *CCHSC*[a]
1	Perform a sequence of tests to distinguish between organic and inorganic compounds	Section 13.3
2	Perform tests to distinguish between reducing and non-reducing sugars and to identify starch	Sections 13.11 and 13.12 Exam-style Question 4
3	Perform a test to identify protein	Section 14.7 Exam-style Question 3
4	Separate a mixture of organic materials by chromatography	Sections 14.8, 14.9 Exam-style Question 11
5	Identify a range of solvents that may be used for chromatography	Sections 14.8, 14.9
6	Carry out electrophoresis on an appropriate mixture	Section 14.10 Revision Test A Question c
7	Perform flame tests or use a spectroscope to describe the emission spectra of a range of elements including Na and Hg	Sections 15.12, 6.14 Revision Test B Question d(i)

a *Exam-style and Revision Test questions are listed here, because their answers at the back of the book often contain helpful information.*

Location of material for other Option 3 Forensic Chemistry activities Ê

Syllabus item		Relevant material in *CCHSC*
9.9.1	■ solve problems … discuss the importance of accuracy in forensic chemistry	Section 13.2
	■ solve problems … discuss ethical issues … during an analytical investigation	Section 13.2
	■ Experiment 1 (see previous table)	
	■ gather … information …series of distinguishing tests to separate …	Sections 13.4, 13.6

Syllabus item			Relevant material in *CCHSC*
9.9.2	■	Experiment 2 (see previous table)	
	■	… using molecular model kits, computer simulations or other multimedia resources to compare the structure of organic compounds …	Photographs of molecular models in Sections 13.7 and 13.8 may be of help
9.9.3	■	… molecular model kits … composition and generalised structure of proteins	Photographs of molecular models in Sections 14.2 and 14.4 may be of help
	■	Experiment 3 (see previous table)	
	■	Experiment 4 (see previous table)	
	■	Experiment 5 (see previous table)	
	■	Experiment 6 (see previous table)	
9.9.4	■	… discuss the range of uses of DNA analysis… and … ethics of maintenance of data banks of DNA	Sections 14.17 and 14.18
9.9.5	■	analyse … information … to discuss the ways in which analytical techniques may provide evidence about samples	Sections 15.2 and 15.3
9.9.6	■	Experiment 7 (see previous table)	
	■	process … information … to … identify individual elements present in a mixed emission spectrum and … explain how such information can assist analysis of the origins of the mixture	Section 15.13 Chapter 15, Exercises 10 to 15

Prescribed focus areas

Although the syllabus suggests that in this option the emphasis is on prescribed focus areas 3, 4 and 5 (p. 101), namely applications and uses of chemistry, the implications of chemistry for society and the environment, and current issues, research and developments in chemistry, considerable attention to area 2, the nature and practice of chemistry, is also needed; in fact this option introduces more new chemical concepts than either of the other two options treated in this book. For students at this level some quite complex chemistry is introduced, such as reducing and non-reducing sugars, amino acids and proteins and the structure of DNA. The basic principles of gas–liquid chromatography, high performance liquid chromatography, mass spectrometry and atomic emission spectroscopy, though more straight-forward, do constitute a considerable body of new concepts.

Nevertheless the main emphasis is on applications, the use of chemistry for forensic purposes. Bear in mind that forensic chemistry is concerned with a wide range of investigations such as detecting contaminants in products, monitoring content claims of other products and measuring pollutants in the environment, and not just with collecting evidence for criminal prosecutions.

The implications for society probably have more to do with the non-criminal aspects of forensic chemistry—quality control, truth in advertising and safety of products. The use of DNA for paternity testing also has social implications (as a 2005 case involving a senior politician very publicly demonstrated) as well as political ones (issues about paternal child-support payments). For the environment the availability of analytical techniques that are far more sensitive

than those of thirty years ago has far-reaching implications; it is now possible to monitor harmful substances at extremely low concentrations, thus providing better safeguards for people's health and wellbeing.

For current issues there is the question of forensic DNA data banks and the privacy questions involved. For research and developments in chemistry there have been the major improvements to chromatographic techniques. Although chromatography has been in widespread use for about seventy years, the developments in recent decades of capillary columns for gas–liquid chromatography and of high pressure pumps and uniform-sized micro-spheres for high performance liquid chromatography have greatly improved the sensitivity and versatility of these techniques. Added to this, developments in mass spectrometry that have led to high speed scans and direct coupling of chromatographs to mass spectrometers have resulted in even better compound identification methods. Although atomic emission spectroscopy has been in use for over one hundred years, the development in recent decades of plasma sources, CCD detectors and computer data processing have seen a re-emergence of this technique for analytical (and forensic) use.