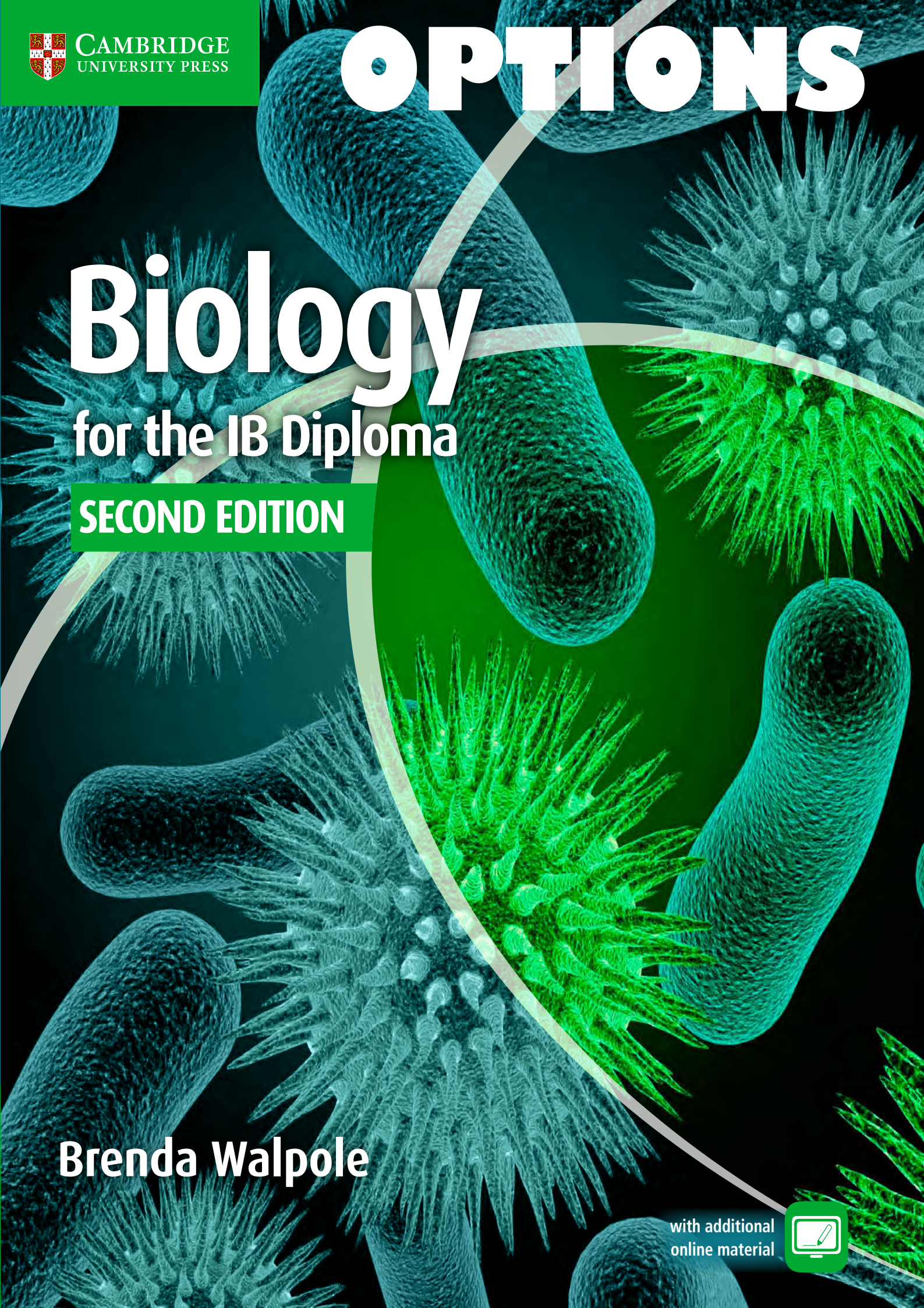

CAMBRIDGE
UNIVERSITY PRESS
OPTIONS
Biology
for the IB Diploma
SECOND EDITION
Brenda Walpole
with additional
online material

Option A Neurobiology and behaviour

The environment of an organism changes throughout its life and it is an advantage for the organism to be able to detect the changes and respond to them. Most responses improve survival chances and all living things – from the simplest unicellular organisms to insects, birds and mammals – respond to stimuli such as light and chemicals. Animals have a nervous system to detect changes and transmit information around the body. Neurobiology is the study of the structure and functioning of the nervous system.

Behaviour is the pattern of responses of an animal to one or more stimuli and the study of animal behaviour is called ethology. Different behaviours enable animals to develop social patterns and mating rituals.

A1 Neural development

Learning objectives

You should understand that:

- During the embryonic development of chordates, the neural tube forms as ectoderm infolds and elongates.
- Neurons are formed by differentiation of cells in the neural tube.
- Immature neurons migrate to their final positions.
- Chemical stimuli cause an axon to grow from each immature neuron.
- Some axons extend beyond the neural tube to connect with other parts of the body.
- As a neuron develops, if forms many synapses.
- Synapses that are not used are lost – only those that are used remain.
- Unused neurons are removed in a process known as 'neural pruning'.
- The nervous system can change with experience and is said to have plasticity.

Development of the neural tube and nervous system

The development of the human central nervous system begins in the embryo (during the first 12 weeks after conception). A similar pattern of development can be observed in all chordates (vertebrates and other animals that possess a supporting dorsal rod called a **notochord**). In the embryonic **ectoderm**, an area of cells called the **neural plate** develops and becomes a region known as the **neural groove**. It is these cells that eventually become the brain and spinal cord. The cells of the neural plate gradually change and develop into folds, which extend up from it and eventually meet and close over to begin the formation of a tube (Figure **A.1**). In this way, as the growth of the folds continues, the neural groove develops into the **neural tube**. The tube elongates and its outer cells form the foundation of the nervous system – cells at the anterior (front) end become the brain (Subtopic **A.2**) and those in the posterior (back) region become the spinal cord.

The process of neural development (neurulation) has been extensively studied in the clawed toad *Xenopus laevis* (Figure **A.2**).

At the time when the neural plate forms, the embryo consists of three layers of cells:

- **ectoderm**, which will form skin and neural tissues
- **mesoderm**, which will become muscle and bone
- **endoderm**, which forms the inner cells of the digestive and respiratory systems.

Spina bifida

Spina bifida is a congenital disorder caused when the neural tube fails to close properly during embryonic development. Some vertebrae do not form completely and remain unfused and open, so that a portion of the spinal cord may pass through the opening in the bones. This occurs most often in the lower back in the lumbar or sacral vertebrae. Spina bifida can be treated by surgery soon after a baby is born. The surgeon places the spinal cord back into the body and closes up the gap between the vertebrae, but the affected part of the spinal cord will not be able to function normally. Spina bifida is one of the most common birth defects and occurs in approximately 1 in 1000 births worldwide. Folic acid, taken as a dietary supplement during the first three months of pregnancy, has been found to greatly reduce the risk of spina bifida and other neural tube defects.

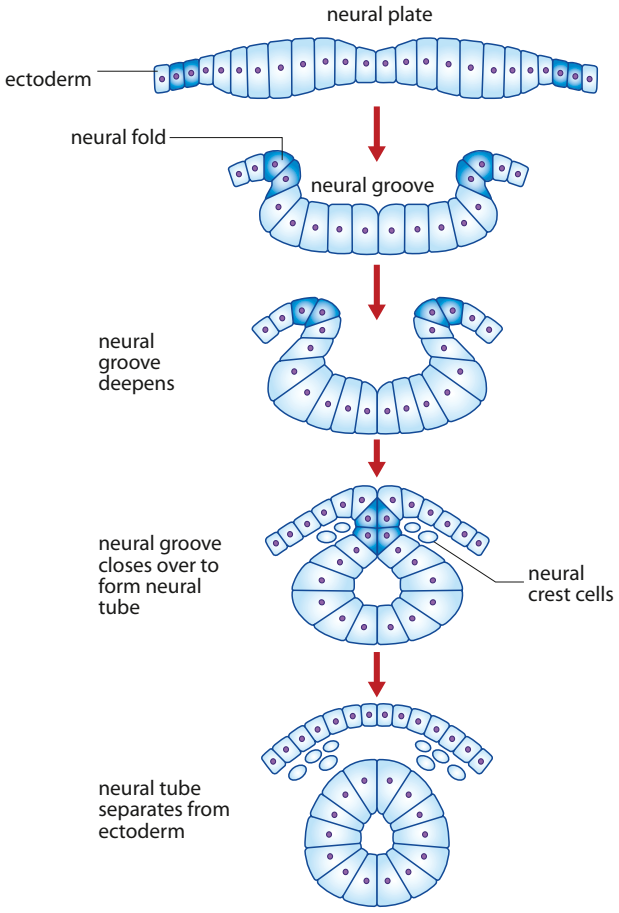


Figure A.1 Neural tube development. Only the ectoderm is shown here – the mesoderm, from which the notochord forms, and the endoderm lie beneath the ectoderm and the neural tube (Figure **A.2**).

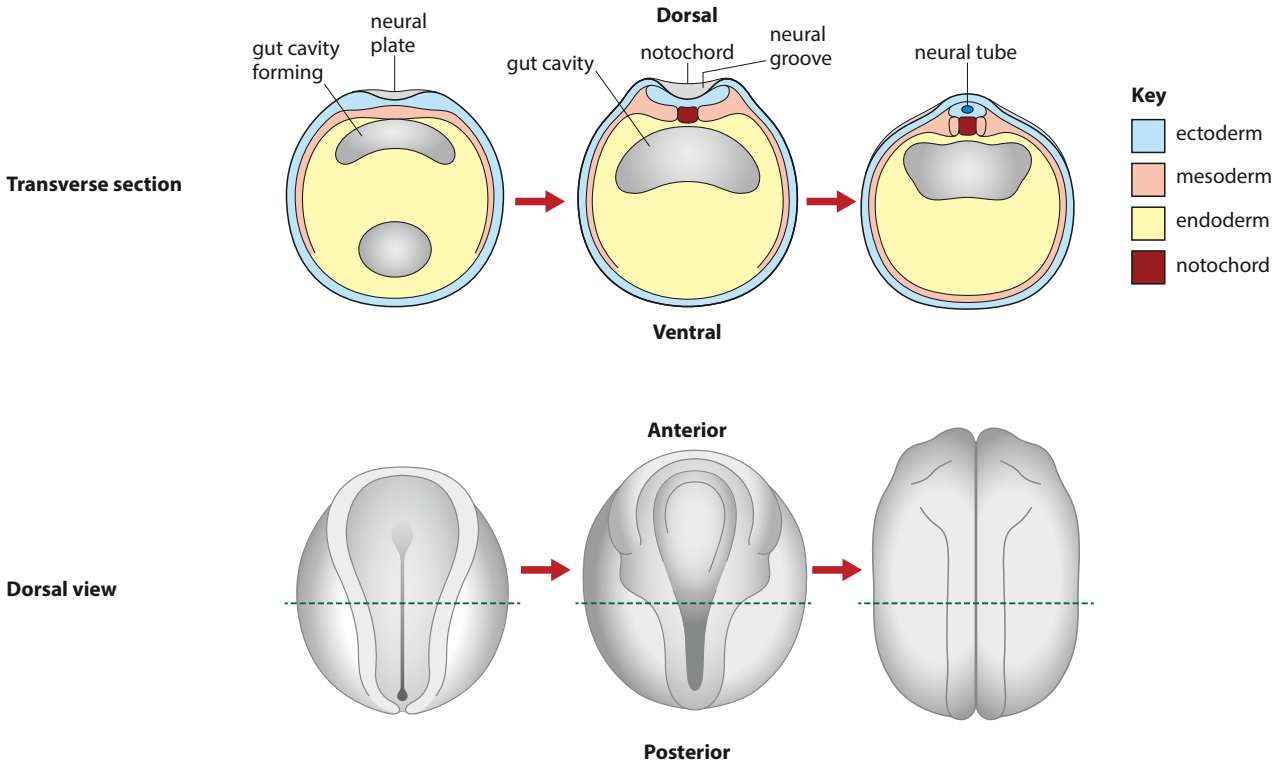


Figure A.2 Formation of the nervous system in the toad *Xenopus*. *Xenopus* is used as a model to study neural development. As an amphibian, it has large eggs with transparent yolk, which enables scientists to study the development of the nervous system as it progresses.

Formation and development of neurons

At the end of the embryonic period of development (12 weeks after fertilisation), the structures that will become the brain and central nervous system (CNS) are established. After this the fetus continues to grow and develop and the fibres of the nervous system extend. Neurons begin to form 42 days (about 6 weeks) after conception and are more or less complete by 18 weeks. Within the developing brain, neurons migrate to different areas and begin to connect with other neurons to make up **neural networks**.

When the neural tube is complete (at about 12 weeks), the cells that will differentiate to become neurons form a single layer lining the tube immediately adjacent to its hollow centre. In the embryo, the hollow centre of the neural tube is cylindrical, like the centre of a straw. But when the brain becomes larger and more complex, the shape of the hollow cavity also changes as the whole embryo changes shape (Figure **A.3**). At this time, a process known as **neural patterning** begins in all parts of the nervous system as neurons take up their positions and start to establish connections. Although this begins in the embryo, the process continues for many years.

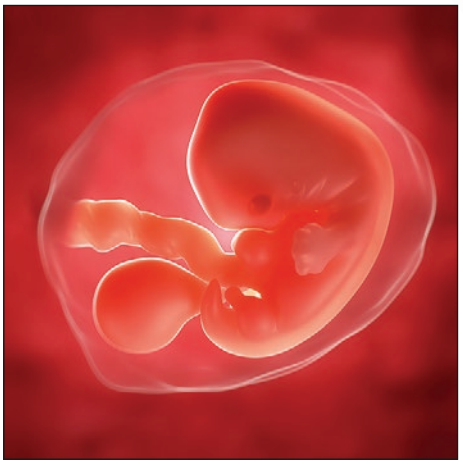


Figure A.3 Illustration of a human embryo at five weeks. The embryo is approximately 2 mm across at this stage.

Neuron migration

As the embryo and then the fetus develop, neurons that are produced in the developing brain migrate to different positions. Some move by a method known as **somal migration** in which a neuron moves by extending a long process from its cell body to the outer region of the brain. This region will later become the cortex (Subtopic **A.2**). The long process attaches itself to the outer surface of the brain and the cell nucleus then travels, through its cytoplasm to take up a new position.

Once young neurons have reached their target region, they must become part of a network in order to be able to process information. Neurons need to develop the **axons** and **dendrites** (extensions of their cell bodies) that enable them to communicate with other neurons. Axons send signals away from the neuron cell body, while dendrites receive signals from other neurons. Each cell develops a network of many dendrites close to the cell body and a single axon that can extend for some distance away from the cell (Figure **A.4**). Every axon extends from an area known as a 'growth cone' at the extreme end of the axon. The growth cone responds to and is guided by molecules of certain chemical substances, which direct it to the correct area. Some of these chemical stimuli are attractive while others repel the growing axon. Axons develop within the areas that will become the brain but also extend beyond the neural tube to reach other parts of the body. When an axon has reached its target area, it begins to develop many **synapses** with other cells. Every developing neuron forms multiple synapses and these synapses allow for communication between the cells of the nervous system via **neurotransmitters**.

Exam tip

Check your definitions of key terms such as: neural pruning, plasticity and a model organism.

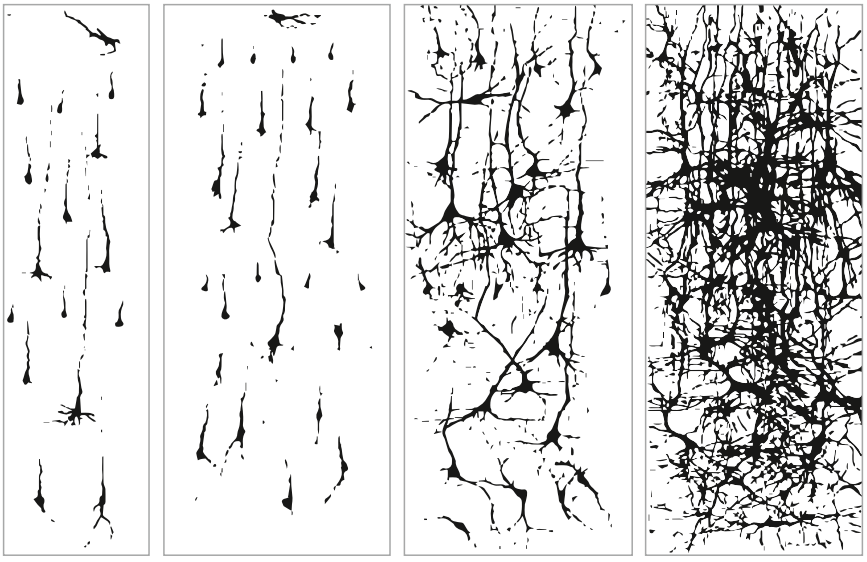


Figure A.4 Connections between nerve cells form during fetal development and become more and more complex in the months and years after birth, establishing intricate neural networks through which information can be passed.

Apoptosis

Apoptosis is cell death that occurs as a result of a regulated sequence of physiological events. A cascade of gene expression results in the breakdown of DNA and histone proteins and eventually the destruction of the cell. All neurons and many other types of cells have this so-called 'suicide' programme. Environmental factors or internal cell events can influence apoptosis. Some factors trigger cell death, while others protect specific cells. Apoptosis occurs in all parts of the developing brain and is particularly high in the cortex, where 70% of cells that form may die as new connections are established or refined.

Cell death in the developing nervous system

During human development there is an enormous amount of cell growth, but the death of neurons is also important. Natural death of cells (**apoptosis**) removes about half of the neurons in certain regions of the brain and, in addition, a second type of cell modification known as 'neural pruning' removes up to half the synapses that have developed between neurons. These two types of cell death are essential to remove unused neurons and thereby help to establish and streamline the complex nerve networks in the brain. The timing of the two types of cell death is different: most cell death (apoptosis) occurs before a baby is born but most neural pruning and synaptic modification occur after birth.

Plasticity of the nervous system

Brain **plasticity** is the ability of the nervous system to change in both structure and function over a person's life, as they react to the changes in their environment.

As a person acquires new knowledge, learns new skills or has new experiences, the brain can establish new neural pathways. Through practice and revision, communication between synapses is enhanced and signals travel more efficiently. Later, if the same neural pathway is used again, the connections between the neurons are re-established and each new attempt means that memory and cognition are made faster. This **synaptic plasticity** is established in a similar way to the way a walker might learn a pathway through a field of corn. If the path is used every day, a clear path will soon become established and the walker will be able to cross the field more quickly and efficiently. Most students can experience synaptic plasticity as they revise for their examinations. On the other hand, if synaptic connections are not used, pathways maybe lost and unused neurons can be removed during neural pruning (see above).

Synaptic plasticity enhances connections between neurons, but a second example of the changes that can occur in the nervous system is **neurogenesis**, the birth and proliferation of new neurons in the brain. For many years it was believed that when neurons died they were never replaced, but stem cells have been found in certain areas of the brain (the hippocampus and the dentate gyrus) that are able to reproduce and migrate to other areas of the brain where they are needed. This may occur after a traumatic event such a stroke, which kills many brain cells. Neurogenesis allows the brain to replace cells that have died and restore functions that have been lost. Alternatively, in some cases, undamaged nerves in different areas of the brain are able to take over the roles of cells that have died and restore some of the functions lost when a person has a stroke.

Language learning

Young children listen to voices around them and soon learn their home language. Infants develop a preference for the vowel sounds of their own language by the age of about six months. The child begins to 'tune in' to their own language and 'tune out' sounds that are not typical of that language. A parallel process takes place in neural development as pathways related to favoured inputs are retained and develop, while unwanted connections are lost by neural pruning. Developmental modifications to the numbers and locations of synapses are probably essential to language learning, both before and after a child is born.

Nature of science

Using models to study the real world – neural development in *Xenopus*

Model organisms are used to study natural events that are difficult to study directly. Species chosen as models are non-human species, used to provide an insight into the way that the biological processes of humans or other organisms work. Different models are used in the study of genetics, reproduction, embryology and biochemistry. Model organisms are chosen for characteristics such as short life span, DNA content or ease of observation. The clawed toad *Xenopus laevis* has been extensively used to study embryonic development in vertebrates because it produces large numbers of eggs, which can be easily seen. Despite the fact that all vertebrates share a common ancestor, and thus share developmental and metabolic pathways, information gained from studies of model organisms such as *Xenopus* must be carefully applied to other species, as there may be important differences as well as similarities.

Test yourself

1 Define 'neural pruning'.
2 Outline the importance of synaptic plasticity in learning and recall.
3 Describe how the neural tube is formed from embryonic ectoderm.

Treating neural tissue injuries

Until recently, it had been assumed that adult brain and spinal cord neurons do not regenerate after injuries, but new research has led to this being reconsidered. Laboratory studies involving investigations of cell signalling and neuro-immunology have increased understanding about how cells may be regenerated.

Experimenters working with rodents placed stem cells from regions of developing and adult brains into damaged areas of the central nervous system. Stem cells that have been grown *in vitro* can be tagged so that they can be monitored after they are grafted into the brain. Experiments like this have indicated that these stem cells are able to adapt to the region in which they have been grafted and differentiate into the appropriate nerve cells. Other studies have shown that these stem cell grafts can also lead to neurogenesis in model animals.

As this knowledge increases, it may one day be possible to treat people with spinal cord and brain injuries by stem cell grafts or by stimulating neurogenesis.

A2 The human brain

Learning objectives

You should understand that:

- The anterior region of the neural tube enlarges to form the brain.
- Different areas of the brain have specific roles.
- The brain stem controls autonomic functions.
- The cerebral hemispheres form the largest part of the human brain.
- Cerebral hemispheres are folded and have a large surface area.
- Higher-order functions are coordinated in the cerebral hemispheres.
- Each cerebral hemisphere interacts with the opposite half of the body.
- A large energy input is needed for brain metabolism.

Basal metabolic rate the amount of energy used by the body when at rest

Brain development

The anterior part of the embryonic neural tube develops to become the brain, and by the end of the embryonic phase of development, at three months after fertilisation, the basic structure of the complete brain is in place. Thereafter, the fetal brain grows rapidly and fibres of the nervous system form. After birth, the brain continues to enlarge and will be four times bigger by the time the child is six years old. At this time it will have become about 90% of its adult volume.

The brain changes in form throughout childhood and adolescence, with both the structure of the brain and behaviour changing as a result. A young child has far more connections in their brain than an adult but these connections are refined (pruned) (Subtopic **A.1**) as the child learns and has new experiences.

A mature human brain contains about 100 billion neurons. Each one can make connections with more than 1000 other neurons so the adult brain may have as many as 60 trillion connections. The metabolism of the brain requires a great deal of energy to maintain its activity – indeed, it has been calculated that the brain requires larger amounts of energy per unit volume than any other tissue. Up to 25% of the energy needed for a human's **basal metabolic rate** is used by the brain, most of which is used to maintain the membrane potential of neurons (Subtopic **6.5**). The majority of the energy comes from the aerobic respiration of glucose. Active regions of the brain use more energy than non-active regions, and this enables fMRI scanning to identify those regions that are active during different activities (see Figure **A.8**).

The structure and function of the brain

The brain is the most complex organ in the human body. Its billions of neurons and connections are responsible for learning, memory and our individual personalities. A human brain has a characteristic folded appearance (Figure **A.5**), which gives it a large surface area, and it is organised so that each part of the brain has a particular function, some regulating automatic processes, such as heart beat and balance, while others control our physical coordination, speech and ability to reason.

- The **cerebral hemispheres** are the largest part of the brain, and form the coordinating centre for learning, memory, language and reasoning. These regions receive information from the sense organs and coordinate and organise motor functions.
- The **hypothalamus** controls the autonomic nervous system (Figure **A.9**). It coordinates the endocrine and nervous systems by regulating the secretions of the pituitary gland.
- The **cerebellum** coordinates movement, posture and balance.
- The **medulla oblongata** (brain stem) controls automatic and homeostatic activities such as breathing, swallowing, digestion and heart rate.

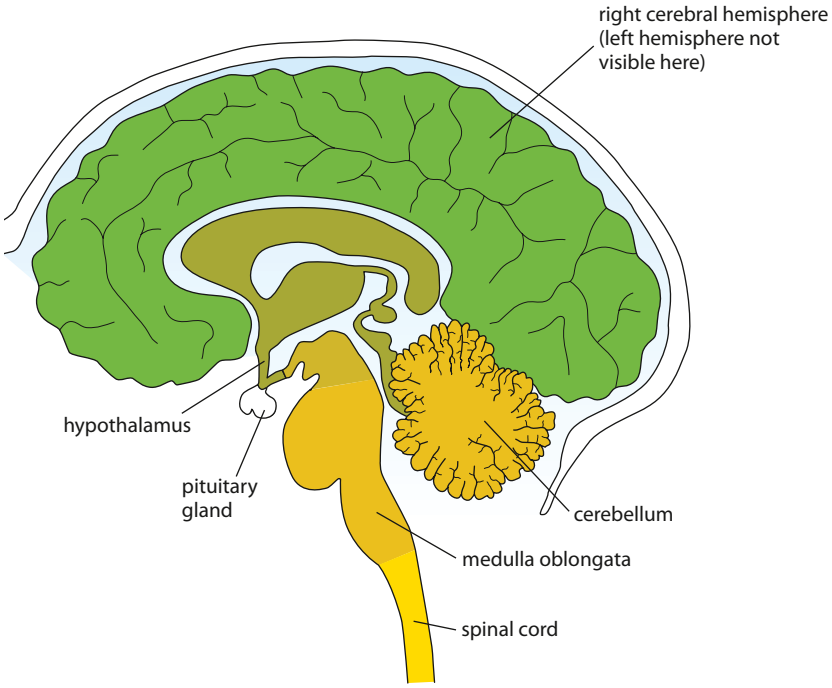


Figure A.5 The human brain.

- The **pituitary gland** has two parts – the posterior lobe stores and releases the hormones oxytocin and ADH from the hypothalamus, while the anterior lobe produces and secretes seven hormones, including FSH and growth hormone, which regulate many of the body's functions (Subtopic **6.6**).

Cerebral hemispheres

The outer, highly folded area of the brain (the **cerebrum**) is made up of two halves, known as the cerebral hemispheres, which are connected by a band of nerve tissue called the **corpus callosum**. The outer layer of each hemisphere is a layer of grey matter called the **cerebral cortex**. The cerebral cortex controls functions such as speech, logic and decision making – the so-called higher-order functions of the human brain.

Although the two cerebral hemispheres look similar, they contain different types of cells and different neurotransmitters. Different areas of the cerebral cortex also have different functions (Figure **A.6**). **Sensory areas** receive impulses from sense organs, **association areas** process the information received and **motor areas** send impulses to effectors in the body.

Information from the left side of the body is received by the right hemisphere and information from the right side of the body is received by the left hemisphere. Likewise, motor signals are sent to each side of the body from the opposite hemisphere. Information is processed by both hemispheres, but there is some division between the functions of each one. For example, association areas of the left hemisphere are important in our use and understanding of language – **Broca's area** is responsible for speaking and writing and **Wernicke's area** is responsible for understanding of language.

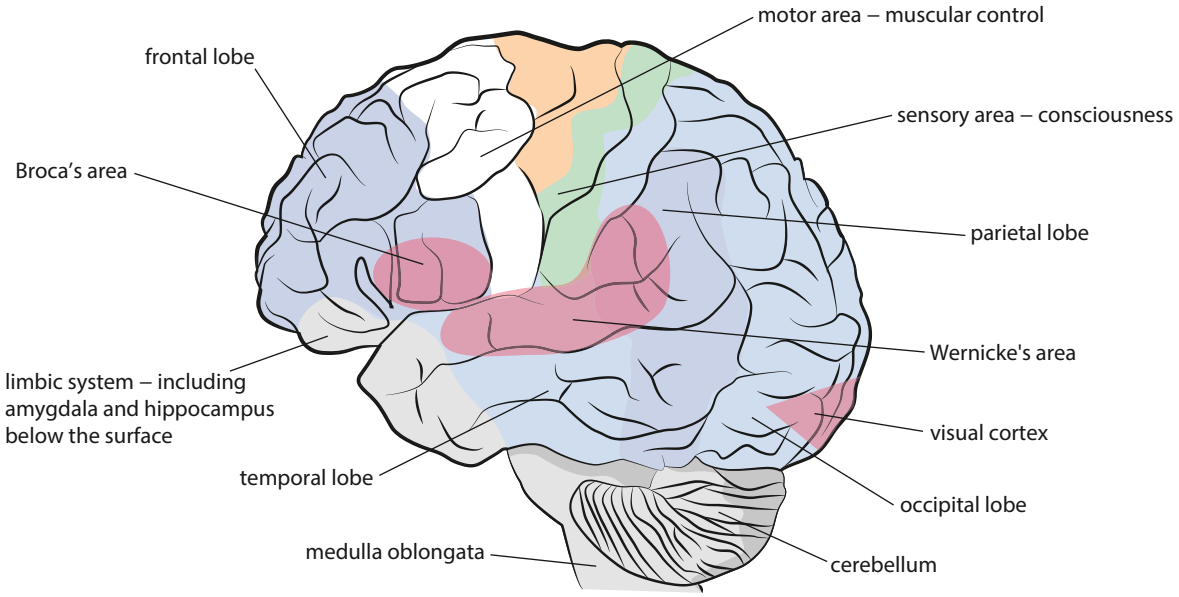


Figure A.6 Approximate locations of important areas of the cerebral hemispheres.

The left side of the brain contains higher levels of the neurotransmitter dopamine while the right has higher levels of norepinephrine. (You may also see norepinephine called noradrenalin in other publications.)

Exam tip
Make sure you do not confuse cerebellum, cerebral cortex and cerebral hemispheres. All are part of the brain but have specific meanings and functions.

Each hemisphere also contains an area known as the **nucleus accumbens**. These regions seem to form part of the brain's 'pleasure centre'. They have been shown to have a role in pleasure, addiction, fear, laughter and reinforcement learning, and research is continuing into their functions.

The **visual cortex** (Figure **A.13**), located at the back of the brain, is the part of the cerebral cortex that is responsible for processing visual information. It receives impulses from the optic nerves and interprets these signals to produce the images we 'see', giving us an understanding of the world around us. Each hemisphere of the brain has a visual cortex. The left visual cortex receives signals from the right visual field and the right visual cortex those from the left visual field.

Brain size and body mass

In general, animals' brain sizes increase with the size of their body so that large animals tend to have larger brains than small animals. But the relationship is not a direct, positive correlation. As the data in Table **A.1** show, mice and humans have a similar brain-size to body-mass ratio, but large animals, including the elephant and hippopotamus, have a relatively small brain for their large body mass even though elephants, in particular, are known to be intelligent animals. For primates, measurements of the size of the whole brain are in fact a better indicator of cognitive abilities than simple brain-size to body-mass ratios, and also remove the need to take account of variations in body mass between underweight and overweight individuals, and differences between adults and young.

Scientists have suggested that a simple size-to-mass ratio is not always helpful in assessing the likely abilities of a species because neurons are very small cells and a large increase in the number of neurons does not cause a huge increase in brain size. A more useful measure than brain-size to body-mass ratios, the encephalisation quotient (EQ), is often used as an alternative. It is calculated from the ratio between actual brain mass and predicted brain mass for an animal of a given size. Scientists propose that this provides a better approximation of the intelligence of the animal and produces a better correlation with the complexity of animal behaviour that we can see.

Species	Brain size : body mass	Encephalisation quotient (EQ)
human	1 : 40	>7
mouse	1 : 40	0.5
elephant	1 : 560	2
dog	1 : 125	1
cat	1 : 100	1
horse	1 : 600	0.9
hippopotamus	1 : 2789	0.9

Table A.1 Brain-to-body-mass ratios and encephalisation quotients for several mammals.

EQ values calculated for mammals, in Table **A.1**, show that carnivores and primates have values greater than 1 while herbivores tend to have lower values. It may be that carnivores with their energy-rich, nutritious diets have greater reserves to support the energy requirements of a large brain but herbivores, which have less energy-rich food, do not. Carnivores may need the greater cognitive ability of a large brain to hunt and kill prey than animals that graze their food.

Investigating brain function

Investigating the brain is a difficult and complex task. Until the arrival of scanning machines, it was difficult to study a living human brain and directly observe its activities.

Animal experiments have yielded valuable information but there are ethical issues involved, particularly when primates, whose brains are most likely to be similar to the human brain, are used. Some procedures involve removing parts of the skull or carrying out experiments on the brain that result in different behaviours. Such experiments may cause distress to the subject animals, which many scientists find objectionable.

Brain lesions – injuries to a specific part of the brain – have provided more direct insights into the functioning of the human brain. Strokes and accidents can damage just one area of the brain and the resulting losses of function give information about what the area controls. One well-documented case of brain injury was that of Phineas Gage in 1848. Gage was a construction worker who survived an accident that sent a large metal pin through his skull and destroyed the left frontal lobe of his brain (Figure **A.7**). Although he physically recovered from the injury and lived for a further 12 years, his personality was completely changed – to the extent that his friends no longer knew him.

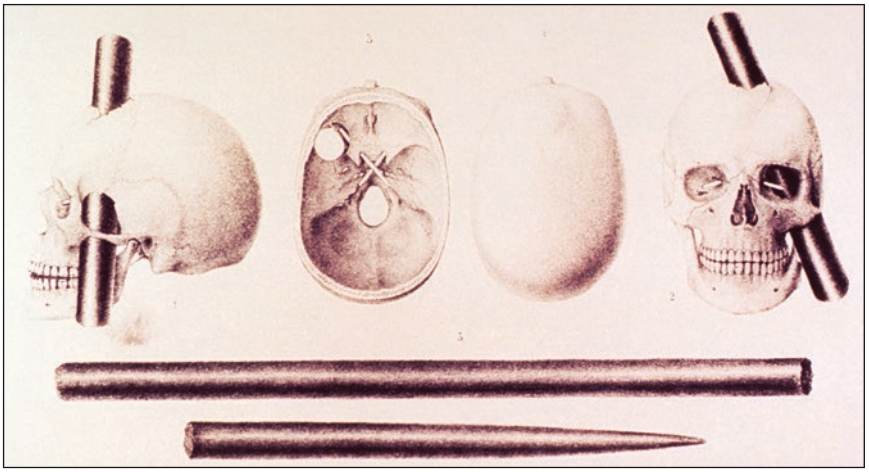


Figure A.7 The skull of Phineas Gage. The fact that his personality changed following the damage to his temporal lobe but he was able to carry on living a fairly normal life, tells us that the temporal lobe is important in coordinating a person's behaviour and reasoning, but not in controlling body functions.

Further information has come from people who have had surgical treatment for epilepsy that involves cutting the corpus callosum, the band of neurons linking the left and right cerebral hemispheres of the brain. Roger Sperry (1913–1994, a psychobiologist who won a Nobel Prize in 1981), and his coworkers carried out 'split-brain' studies on these patients. Sperry discovered that the two sides of the brain can operate almost independently. In a normal brain, the stimulus entering one hemisphere is quickly transferred through the corpus callosum to the other hemisphere so that the brain functions as one. But if the two hemispheres cannot communicate, a person's perception of the outside world is changed. Sperry's work showed that the two halves of the brain have different functions. The left hemisphere of the brain specialises in communication and if a lesion affects this side of the brain, a person may be unable to speak. Damage to the right hemisphere of the brain, which is particularly good at interpreting sensory information and enabling us to understand what we see or hear, may result in a person failing to recognise someone they know well.

Functional magnetic resonance imaging

Since the 1990s, functional magnetic resonance imaging (fMRI) scans have been a key source of new information on brain function. This scanning technique monitors blood flow to different areas of the brain as a subject carries out different tasks. As a region of the brain becomes active, more blood flows to it. Subjects in fMRI experiments are asked to remain still in the scanner as they respond to stimuli or undertake different activities. The scans reveal which areas of the brain are active and help to show how it is working (Figure **A.8**).

Figure A.8 A patient undergoing an fMRI scan.

fMRI scans can reveal Parkinson's disease, tumours and injuries to the brain. But it is important that the images are interpreted carefully. There may be activity in an area of the brain associated with a particular task but correlation does not imply cause. Many brain processes are complex and not confined to one area alone and this explains why, in some cases, different areas of the brain can take over functions of damaged cells following a brain injury or a stroke.

Examination of the brain after death

Brain tissue can also be examined after a person has died and can help our understanding of brain function and disease. Most nucleic acids and proteins are quite stable after death and RNA and protein samples can be used in studies involving RT-PCR (which uses RNA to create DNA profiles), cDNA microarrays (which can identify DNA sequences) and proteomics (the study of proteins and their functions using databases of known sequences). The quantities of nucleic acid and proteins vary from person to person, so it can be difficult to compare different individuals with the same condition or disease, but already researchers have discovered that high levels of certain proteins are associated with brain degeneration and Alzheimer's disease. The rapid development of new techniques available to molecular biologists means that this research is likely to become more and more important in the future.

Sympathetic and parasympathetic control

The **peripheral nervous system** (**PNS**) consists of all the nerves that do not form the central nervous system (CNS; brain and spinal cord). The PNS comprises the sensory neurons, which carry impulses to the CNS, and the **autonomic nervous system (ANS)**, which is involuntary and regulates internal processes (such as activities of the glands and digestive system, and blood flow) without our awareness (Figure **A.9**).

The autonomic nervous system is subdivided into two parts: the **sympathetic nervous system** and the **parasympathetic nervous system**. Both receive impulses from the medulla oblongata in the brain stem but have opposite effects on the body. The sympathetic system causes responses that are important in an emergency – the so-called 'fight or flight' responses. It is excitatory in its effects. The parasympathetic system controls events in non-urgent, relaxed situations and is inhibitory in its effects. Table **A.2** compares the actions of the two systems on some vital functions.

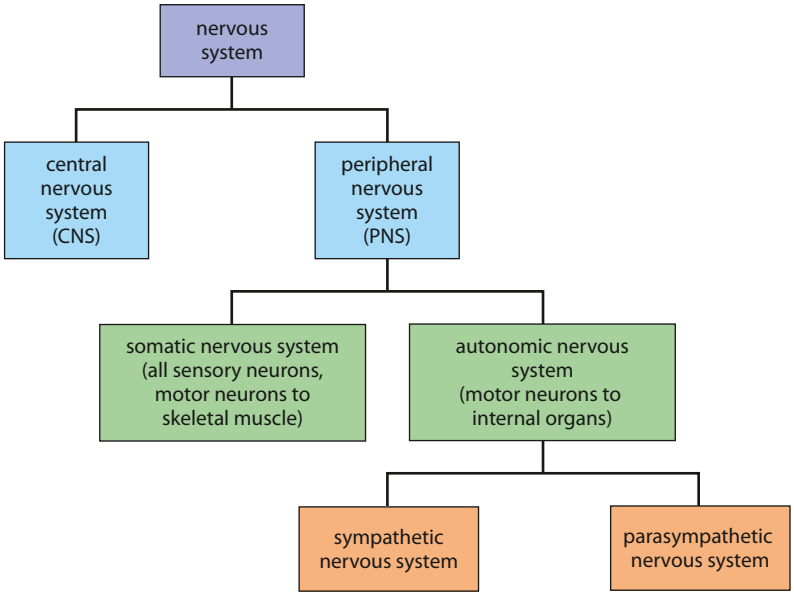


Figure A.9 The components of the human nervous system.

Organ	Effect of parasympathetic system	Effect of sympathetic system
eye	causes contraction of circular muscles of the iris, which constricts the pupil	causes contraction of radial muscles of the iris, dilating the pupil
heart	heart rate is slowed down and stroke volume is reduced as the body is relaxed	heart rate is increased and stroke volume increased so that more blood can be pumped to muscles
digestive system	blood vessels are dilated, increasing blood flow to the digestive system	blood flow to the digestive system is restricted as blood vessels constrict

Table A.2 Comparison of the effects of the parasympathetic and sympathetic nervous systems on three organs.

The pupil reflex

The **pupil reflex** is a constriction of the pupils caused by contraction of the circular muscles in the iris. It occurs when bright light shines into the eye. The rapid, reflex action protects the retina from excess light, which could damage it. Unlike the majority of reflexes, it is controlled by the brain instead of the spinal cord. When light stimulates photoreceptors in the retina, impulses pass along the optic nerve to the medulla oblongata in the brain stem. From here, impulses are sent to the muscles of the iris via parasympathetic nerves. The circular muscles are stimulated and constrict the pupil.

The pupil reflex is used to test the functioning of the brain stem and, along with other tests of reflex actions, is a key indicator of brain death.

Ethics in medicine – the concept of brain death

In medicine, the concept of death is defined in terms of brain functions. Doctors test the activity of the brain stem to determine whether to continue medical treatment. A patient with severe damage to the brain stem (medulla oblongata) is unlikely to recover because this region controls breathing, heart rate and all the automatic, vital functions of life. The legal definition of 'brain death' is based on the activity of the brain stem. If this is permanently damaged, the brain is regarded as having lost neurological function so that consciousness and spontaneous breathing will never be possible.

But life-support machines in modern hospitals can take over the roles of vital organs such as the heart or lungs when a person is seriously ill or injured in an accident. They keep a person's body functioning without the need for impulses from the brain stem and can provide time for an organ to recover. The patient may even be unconscious or in a coma because of damage to the brain, which can recover in time.

Sometimes conflicts can occur when the legal and medical criteria for death do not concur with those of family members.

Questions to consider

- To what extent should the views of family members be given priority in making decisions in medical ethics?
- What are the appropriate criteria to use when making ethical decisions?

Diagnosing death

Bodily death is not clearly definable, and in some cultures diagnosis relies on traditional medical practices. Where loss of brain stem function can be identified, this can be used to diagnose death, based on the understanding that the destruction of the brain permanently removes a person's consciousness so that they cannot interact with others and cannot sustain their own life processes. But sometimes a person with brain damage may enter a 'permanent vegetative state' and this cannot be diagnosed or defined with absolute certainty – there may be a possibility of recovery. At present there is no global consensus on the criteria for brain death and the use of additional tests to confirm death. In an effort to address this, a recent international conference of the European Society of Anaesthesiology in 2013 discussed the criteria used to diagnose death, and the World Health Organization has started work to develop a set of criteria that can be used by doctors all over the world.

Nature of science

Using models to study the real world – sensory and motor homunculi

Sensory and motor homunculi (Figure **A.10**) are models that show the relative space that neurons related to different parts of the body occupy in the sensory cortex and motor cortex of the cerebral hemispheres. Hands, feet, lips and sex organs have more sensory neurons than other parts of the body so, in the model sensory homunculus, these areas are larger. The motor homunculus is similar – again the larger areas are those which have a greater area of the motor cortex is devoted to them. Homunculi are useful for representing the brain to help us visualise the functions of the different areas.

Figure A.10 This model shows a motor homunculus, a similar model can be made to show the relative proportions of sensory neurons.

Test yourself

4 Outline the role of the brain stem.
5 State what is meant by a 'higher-order function'.
6 Outline the role of Broca's area of the brain.

A3 Perception of stimuli

Human sensory receptors

Sense organs supply our brain with the information it needs to keep us in touch with the world around us. Information about changes in our surroundings is detected by sensory **receptors**, which are able to absorb energy of different types from the environment and transform it into nerve impulses.

We have **thermoreceptors** in our skin that respond to temperature, **photoreceptors** in the retina of each eye that respond to light, and **chemoreceptors** in our blood vessels that detect the pH or carbon dioxide concentration of our blood and help to regulate our breathing.

Chemoreceptors in our noses and on our tongues respond to chemical substances in the environment. Our sense of smell is mediated by olfactory chemoreceptors in the nasal passages. These receptors are activated by the presence of small odour molecules (all with a molecular mass less than 350) and send nerve impulses to the brain. It is estimated that mammals have up to 1000 different receptors, and each receptor can be activated by several similar odour molecules. It is the combination of stimulation of different receptors that builds up our sense of smell, and the number of possible combinations is enormous so we are able to distinguish an almost infinite number of odours.

Mechanoreceptors are another group of receptors, stimulated by pressure or forces. Some respond to changes in blood pressure, others to the movement of fluid in the inner ear. Whenever we move any part of our body (for example, a leg to kick a ball, or an arm and fingers to pick up a pen) we need to know exactly where that part of the body is. We receive this information from mechanoreceptors known as stretch receptors, found in muscles. Stretch receptors respond to stretching of the muscles and allow the brain to work out the positions of all parts of the body.

Variation in sense of smell

Sensitivity to smell varies from person to person and is genetically determined. Some people cannot smell the odour of a skunk and others cannot smell the fragrance of freesias. In general, women have a more acute sense of smell than men and their sensitivity varies through the menstrual cycle, peaking at ovulation when it coincides with a surge in oestrogen. The level of this hormone also increases during pregnancy, and some women report an increase in smell sensitivity during pregnancy.

Many animals have a much more acute sense of smell than humans. Dogs can distinguish non-identical twins by smell, but not identical twins. 'Sniffer' dogs, pigs and even rats have been used to locate explosives and illegal drugs.

Learning objectives

You should understand that:

- Receptors are cells that respond to changes in the environment.
- Rods and cones are photoreceptors in the retina of the eye.
- Rods and cones are sensitive to light of different intensities and wavelengths.
- Impulses from rods and cones are carried by bipolar cells to ganglion cells.
- The optic nerve carries impulses from ganglion cells to the brain.
- Information about the left field of vision, from both eyes, is sent to the right part of the visual cortex, while information about the right field of vision is sent to the left side of the visual cortex.
- Sound waves are transmitted and amplified by structures in the middle ear.
- Sounds are detected in the cochlea by sensory hairs sensitive to specific wavelengths.
- Impulses from the sensory hairs, produced in response to sound waves, are transmitted to the brain via the auditory nerve.
- Movement and orientation of the head are detected by sensory hair cells in the semicircular canals.

The human eye

Photoreceptors in the human eye make it a very efficient light-sensitive organ. Light rays entering the eye are bent by the cornea and lens and focused onto the **retina** (Figure **A.11**).

The two types of photoreceptor cells, arranged in a single layer in the retina, are called rods and cones.

- **Cone cells** are not very sensitive to light but three different types of cone are sensitive to three different wavelengths of light and enable us to see in colour.
- **Rod cells** are much more sensitive to light. They absorb all wavelengths of light and function well at low light intensities. In dim light, only rods cause nerve impulses to be transmitted along the optic nerve so we cannot perceive colour and the world appears in shades of grey.

Table **A.3** summarises the characteristics of rod and cone cells in the retina.

Rods	Cones
highly sensitive to light, work in dim light	less sensitive to light, work in bright light
one type of rod can respond to all wavelengths of light	three different cones respond to red, blue and green light so we can detect colour
groups of rods are connected to a single bipolar cell	each cone is connected to its own bipolar cell
not present in the fovea	not present at the very edge of the retina

Table A.3 Comparison of rods and cones in the human retina.

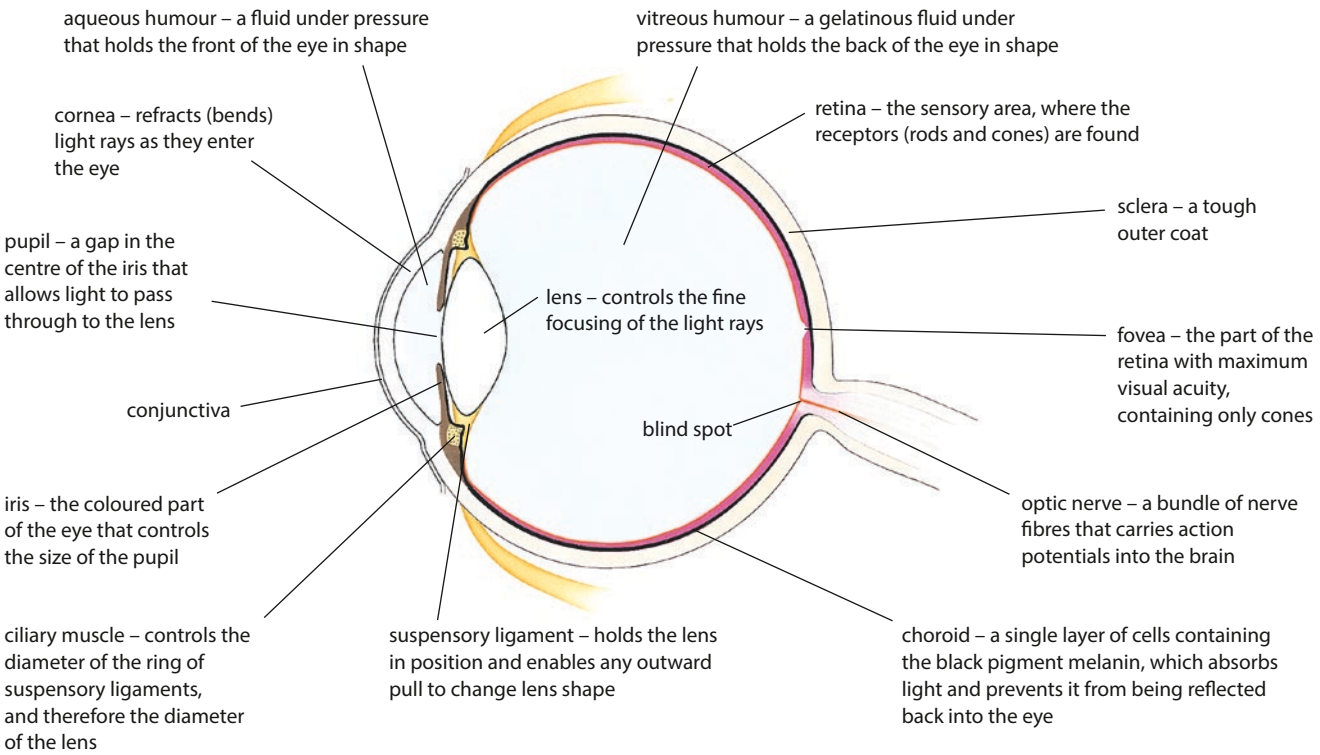


Figure A.11 The structure of the human eye, in transverse section (the eyelids are not shown).

In addition to rods and cones the retina also contains two layers of neurons – **bipolar cells** and **ganglion cells** (Figure **A.12**). These cells conduct the information from rods and cones to the optic nerve.

The **fovea** is an area of the retina directly behind the pupil. It contains the highest concentration of cones in the retina. When you look directly at a small object, its image is focused on the fovea, which is the part of the retina that produces the most visually accurate image.

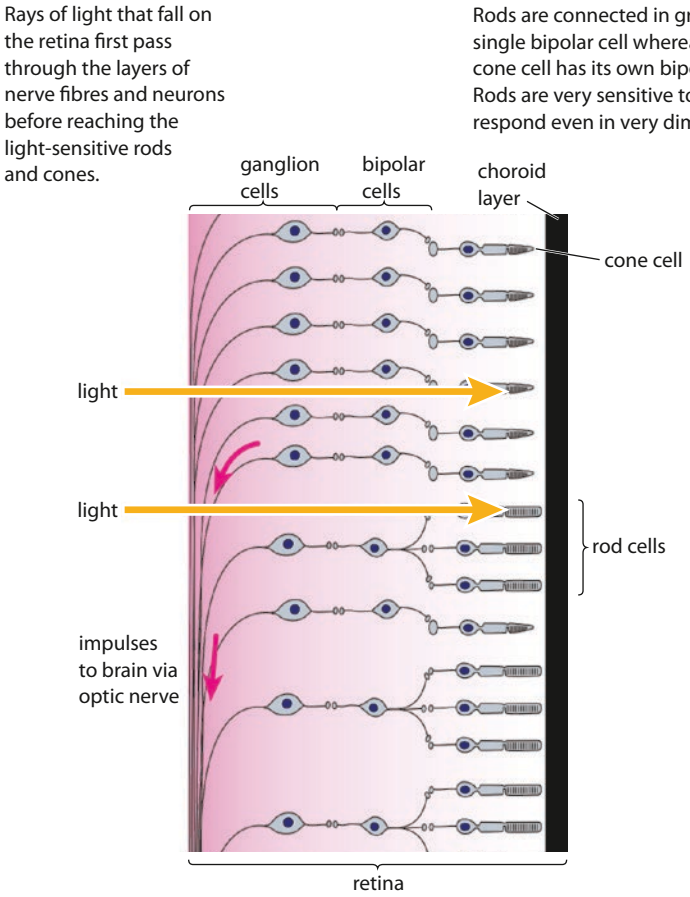


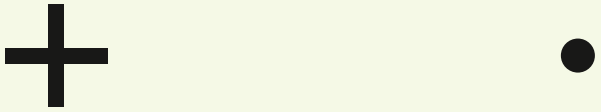
Figure A.12 The retina of the human eye.

Exam tip

Remember that light passes through the bipolar cells before it hits the rods and cones.

Blind spot

At the point where neurons leave the eye in the optic nerve, they pass through the layer of rods and cones, and this creates the 'blind spot'. We do not perceive an image when rays of light fall on the blind spot, because the light does not fall on any rods or cones here. However, the blind spot is in a slightly different position in each eye, which means each eye is able to 'fill in the gap' for the other, and we are not aware of any blank areas in our visual field.



Position your eyes so you are 60 cm away from the cross and dot. Close your left eye and concentrate on the cross with your right eye. Slowly move close to the image. When the image of the dot falls on your blind spot it will disappear.

What do animals see and can they see in colour?

Humans have three types of cone and can see a range of colours that we call the visible spectrum. No one knows exactly what other animals see or to what extent they see in colour. We can only study the physiology of their eyes and the light-sensitive cells they contain and attempt to deduce what their brains may perceive. Colour vision and perception across the animal kingdom is the subject of ongoing research.

Of the species studied so far, the best colour vision appears to be found in birds, aquatic animals and certain insects, especially butterflies and honeybees. Most mammals have weak colour vision; humans and other primates have the most advanced colour perception. Dogs have two types of cones, suggesting that they may view the world in a similar way to red-green colour-blind humans. Cats have three types of cones, but a much lower proportion of cones to rods than humans. They can distinguish blue and green but probably do not perceive red objects well. Many animals can see things that we cannot. Bees can perceive light in the ultraviolet range but do not see red well. This explains why very few wild flowers are pure red.

Questions to consider

- To what extent can the statement 'beauty is in the eye of the beholder' be related to the physiology of the eye and brain?
- We do not know what animals actually see. Could their understanding of what they see be similar to a human's 'understanding' of an abstract impressionist painting?
- Is it ever likely to be possible to answer these questions?
  - Does a bull really get enraged by a red cape?
  - How do bees know which flowers to visit for nectar?

Visual processing

Light rays entering the eye stimulate photoreceptors (rods and cones), which send impulses to bipolar neurons. These neurons combine impulses from groups of rods or from individual cone cells and generate action potentials in the ganglion cells. From here, nerve impulses travel along the axons of neurons in the optic nerve to the visual cortex at the back of the brain. Impulses pass via the **optic chiasma** and relay areas in the thalamus of the brain as shown in Figure **A.13**. When the impulses reach the visual cortex, they must be interpreted to produce the images we 'see'. For example, because of the way light rays pass through the lens of the eye, the image falling on the retina is both inverted and reversed from left to right. However, the images we 'see' are not inverted or reversed. This is because the brain interprets the impulses it receives, so that we perceive the world 'the right way up'. This is **visual processing**.

Contralateral processing

The brain must also coordinate the information it receives from both eyes. As we view an object, each eye receives a slightly different view of the visual field, which is detected by different regions of the retina in each eye. Axons from the region of the retina closest to the nose in each eye cross over in the optic chiasma to go to the opposite side of the brain. This means that all the information from the left visual field goes to the right visual cortex and all the information from the right visual field goes to the left visual cortex (Figure **A.13**). This is called **contralateral processing**. The visual cortex assembles all the information it receives and gives us an understanding of what we are looking at.

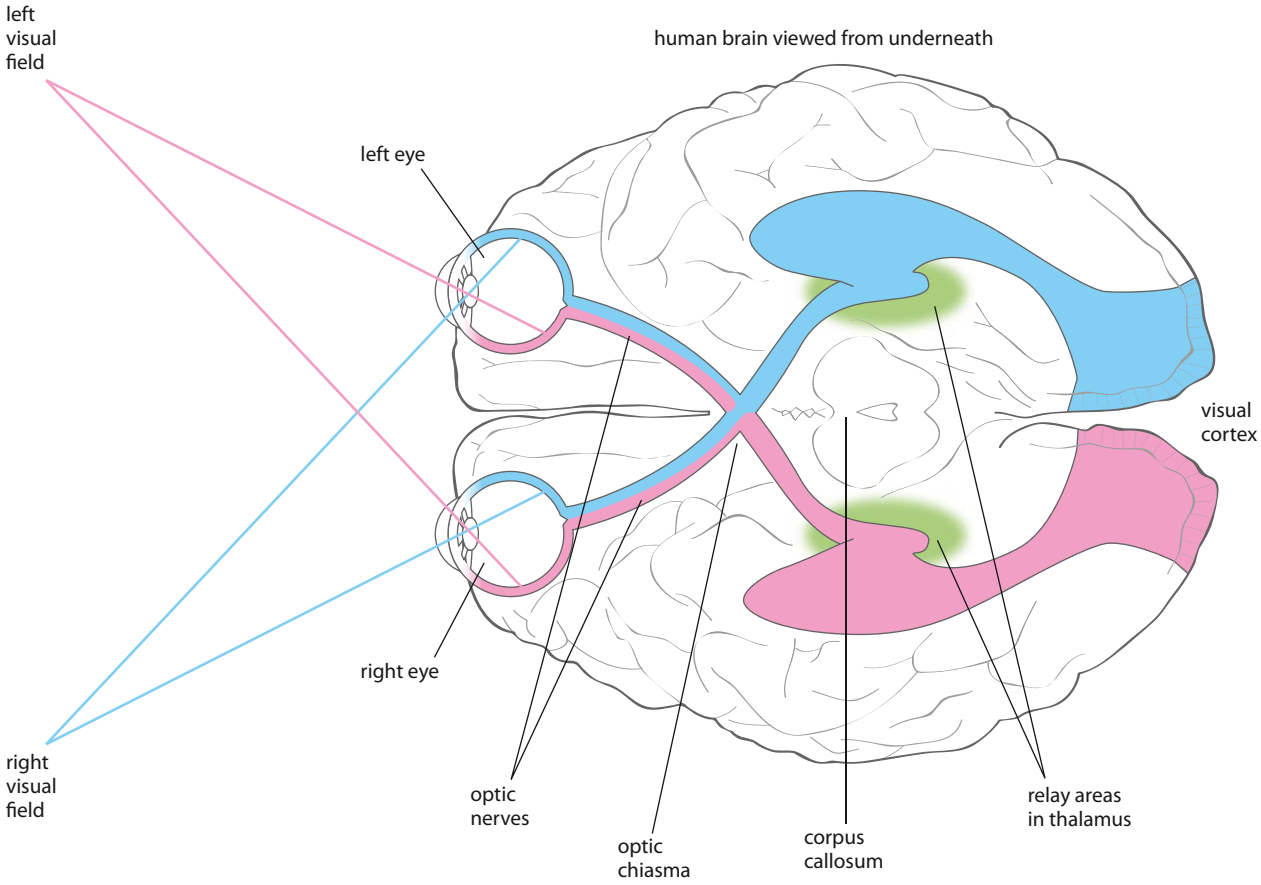


Figure A.13 Both sides of the brain work together to enable us to recognise objects. Contralateral processing allows us to work out the size of an object and its distance from us.

Red-green colour blindness

Red-green colour blindness is a sex-linked condition caused by a faulty gene on the X chromosome. For men to be affected, the recessive 'colour blindness' allele need only be present on their single X chromosome, whereas for women the alleles on both X chromosomes must be recessive (Subtopic **3.4**). This explains why red-green colour blindness is far more common in men than in women. About 8 in 100 men and 1 in 100 women are affected.

Anomalous trichromatism is the medical term for the mildest and most common form of colour blindness. All three types of cone cells are present in the retina but there is a fault in either the red or green cones, giving reduced sensitivity to certain colours. For example, if the red cone is faulty, colours containing red cannot be distinguished clearly. People who are colour blind vary in their ability to distinguish between different colours – some are more affected than others. A very small number of people are unable to see any colour at all and are truly colour blind.

Stereoscopic vision

Each eye captures its own slightly different view of an object and so two separate images are passed to the brain for processing. When the images arrive, they are combined into just one 'picture'. The brain unites the two images by matching up the similarities and adding in the small differences to produce stereoscopic vision. With stereoscopic vision, we can perceive where objects are in relation to us quite accurately. This is especially true when things are moving towards or away from us. We can even perceive and measure 'empty' space with our eyes and brains.

The human ear

Figure **A.14** shows a diagram of a section through the human ear. It is divided into three regions – outer, middle and inner ear.

The outer ear and middle ear are separated by the **ear drum**, and the middle ear is separated from the inner ear by the oval and round windows. The **pinna** is a sound-collecting device and in many animals it can be rotated by muscles to pick up sounds from all directions. Most humans have lost the ability to use these muscles.

The **Eustachian tube** connects the middle ear to the back of the throat via a valve and maintains an equal pressure of air on each side of the ear drum. In the inner ear, the **cochlea** detects sound and the **semicircular canals** detect motion.

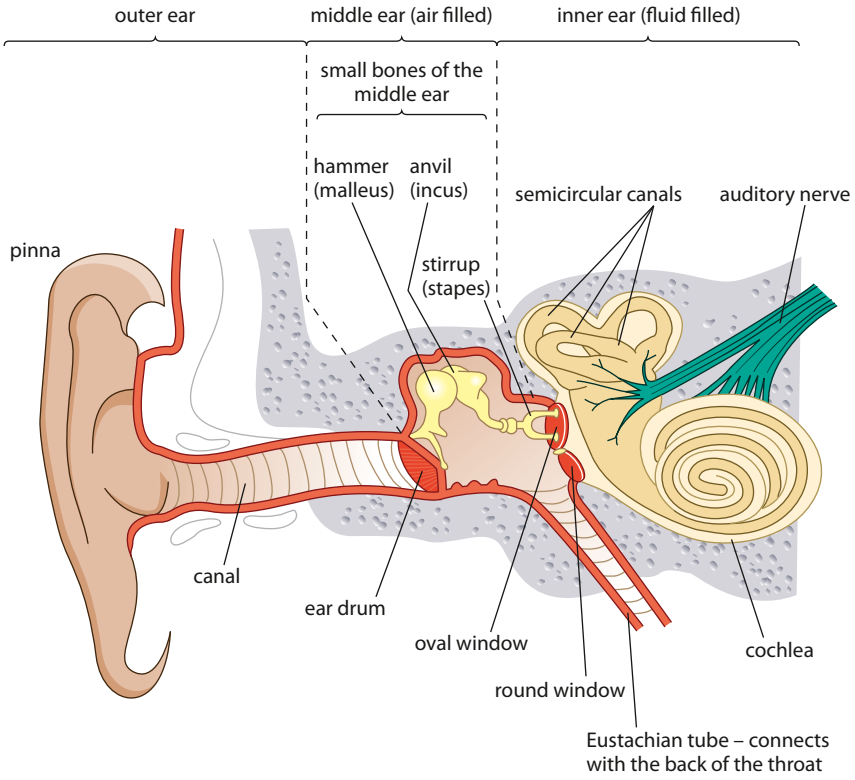


Figure A.14 Section through the human ear. Note that the pinna is not drawn to scale with the internal structures of the ear.

Exam tip

Check that you can label a diagram of the eye and ear – this is often part of an exam question.

How sound is perceived

Sound is created by differences in air pressure, which produce vibrations called sound waves. Sound waves enter the outer ear canal and cause the ear drum to vibrate back and forth (Figure **A.15**, top). These movements are transmitted to the three tiny bones in the middle ear. The ear drum is in contact with the first bone and the third bone touches the oval window. Each bone vibrates in turn so that vibrations pass via the bones to the oval window. By the time the vibrations reach the inner ear they have been amplified up to 20 times because the bones act as levers, increasing the force of the waves, and also because the oval window is much smaller than the ear drum.

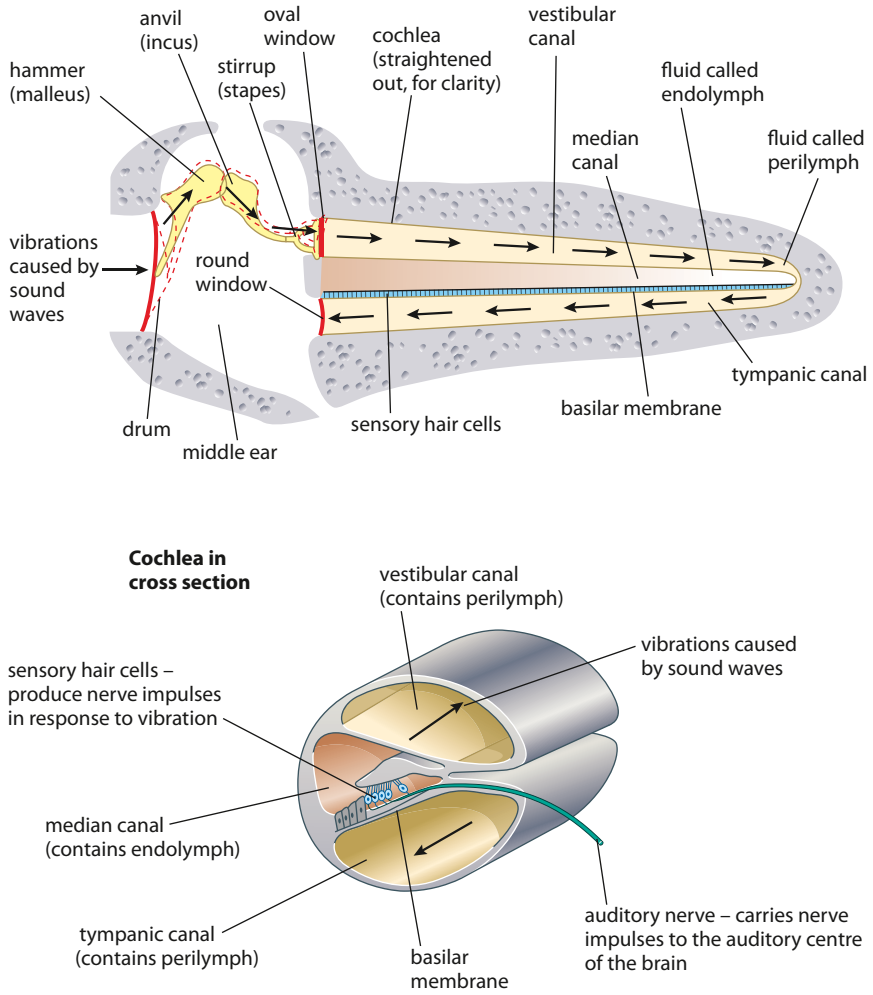


Figure A.15 The structure of the cochlea.

Vibrations of the oval window are passed on to the fluid contained in the cochlea. The fluid cannot be compressed and can only move because the round window at the end of the coiled cochlea absorbs the pressure of the waves of fluid as they arrive.

Inside the cochlea are sensory hair cells attached to membranes, as shown in Figure **A.15**. As the fluid moves, it moves groups of hair cells, which initiate nerve impulses that are passed via the auditory nerve to the auditory centre of the brain. Different regions of the cochlea respond to different frequencies of sound. High frequencies (short wavelengths) are detected nearest to the oval window and the lowest frequencies (longest wavelengths) are picked up further away. Hair cells in any one region vary in their sensitivity and this allows differences in loudness to be detected. A quiet sound stimulates only a few hair cells in a particular region so few nerve impulses are sent to the brain. If the sound is louder, more hair cells are stimulated and more nerve impulses pass to the brain.

Hearing range

A young person can detect sounds of wavelengths between 40 and 20 000 Hz but as we age we lose the ability to detect higher frequencies. Loss of hearing is also common among rock musicians and people who work in very noisy environments without ear protection because excessive noise damages the hair cells in the cochlea. Many species can hear a different range of sounds from humans. Dogs can hear up to 40 000 Hz and bats up to 100 000 Hz.

The semicircular canals

The semicircular canals are organs that give us a sense of position and balance, by enabling us to detect movements of the head. The three fluid-filled canals are arranged so that each one is aligned in a different plane, at right angles to the others (Figure **A.16**). At the end of each canal is a cavity containing receptors, called the **ampulla**, and the canals

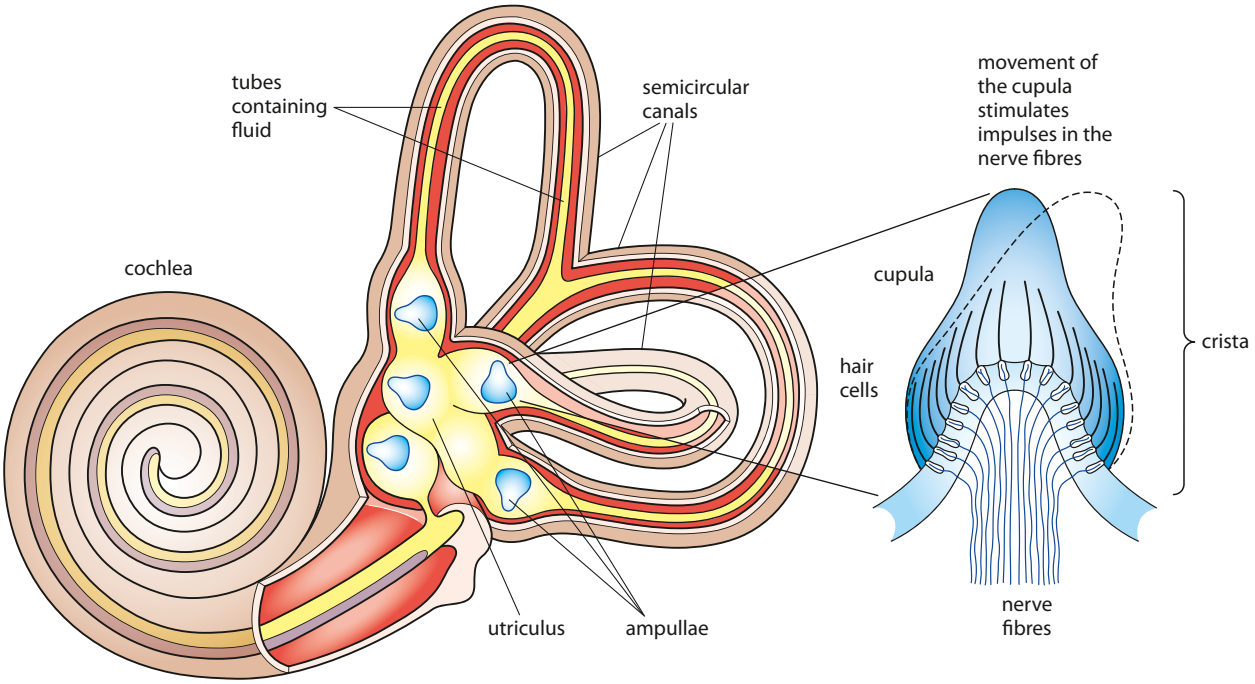


Figure A.16 The semicircular canals and utriculus.

are attached to a central structure known as the **utriculus**, which also contains receptors. The semicircular canals are stimulated by rotation of the head, while the utriculus responds to changes in position. Within each ampulla, sensory hairs are embedded in a jelly-like cap, which is deflected by movements. The inertia of the fluid inside each canal means that the cap is deflected in the opposite direction to the movement of the head. This in turn pulls on the hair cells, which send impulses to the brain.

Nature of science

Scientific understanding drives technical advance – cochlear implants

Understanding the functioning of the ear and how sounds are perceived has led to the development of cochlear implants. A cochlear implant is an electronic device that is surgically placed to provide hearing for people who have damage to the hair cells in the cochlea. Many thousands of people worldwide have benefitted from this technology. Damage to the ear may be caused by diseases such as meningitis, genetic factors or exposure to excessive noise for long periods of time. The cochlear implant bypasses the hair cells in the cochlea and stimulates the cochlear nerves directly using electrical impulses. The brain is able to interpret the frequency of sound as it would if the hair cells were functioning properly. Implants are placed under the skin behind the ear and consist of a microphone that collects sounds, a speech processor that filters and selects important sounds and a transmitter that sends impulses to the internal part of the device. A receiver and stimulator are implanted in bone beneath the skin and these send impulses to electrodes that are woven into the cochlea. The electrodes send impulses to the brain via the auditory nerve.

Test yourself

7 The retina is made up of three layers of cells – photoreceptors, bipolar cells and ganglion cells. State which cell layer light strikes first on entering the eye.
8 Explain the function of bipolar cells.
9 List **three** differences between rods and cones.
10 Outline what is meant by 'contralateral processing'.
11 State the name of the region of the brain where neurons from the left eye and the right eye cross over.

A4 Innate and learned behaviour (HL)

The study of behaviour attempts to understand many aspects of an organism's life, from its instinctive responses to more complex feeding and breeding habits. In a natural environment, two types of animal behaviour can be recognised: innate, instinctive behaviours and learned behaviours that develop as a result of experience.

Innate behaviour

Innate behaviour is very often called 'instinct'. This behaviour is common to all members of a species and is genetically controlled. Innate behaviour occurs independently of the environment and is crucial to survival, helping in activities such as finding food, building a nest or escaping from danger. Short-lived species do not have time to acquire learned behaviours or skills and a high proportion of the behaviour of most invertebrates, which are relatively short lived, is innate. Examples of innate behaviour include the movements of dragonfly nymphs as they prepare to pupate (Figure **A.17**), movements of woodlice towards damp areas to avoid drying out, the dances performed by honeybees to communicate the direction of a food source and the mating behaviour of many bird species. Another example of innate behaviour is seen in the mating behaviour of lions (Subtopic **A.6**).

Experimental study of innate behaviour

Innate behaviour of invertebrates includes the movements they make towards or way from stimuli such as food or light. Two examples of this kind of orientation behaviour that can be investigated are taxis and kinesis.

A **taxis** is when an organism moves towards or away from a directional stimulus. A choice chamber is a simple piece of apparatus that can be used to investigate taxis in small invertebrates (Figure **A.18**). It consists of a circular clear plastic box with four sections in the base, over which a piece of gauze can be stretched as a platform for small invertebrates to walk on.

The innate behaviour of woodlice can be observed and recorded using a choice chamber. The conditions under the gauze can be varied to change the humidity in each section of the chamber and the lid covered to provide different light intensities. To investigate the animals' response to light and dark, one half of the chamber can be covered with black cloth,

Learning objectives

You should understand that:
- Innate behaviour is genetically determined and develops independently of the environment.
- Autonomic and involuntary responses are known as reflex actions.
- A reflex arc is a neural pathway that enables a reflex action to occur.
- Reflex conditioning involves forming new associations between neurons.
- Learned behaviour develops as a result of experience.
- Imprinting is a type of learned behaviour that occurs at a particular stage of life and is not affected by the consequences of the behaviour.
- Operant conditioning is a form of learning that involves trial and error.
- Learning is defined as the acquisition of skills or knowledge.
- Memory is the process of encoding, storing and accessing information.

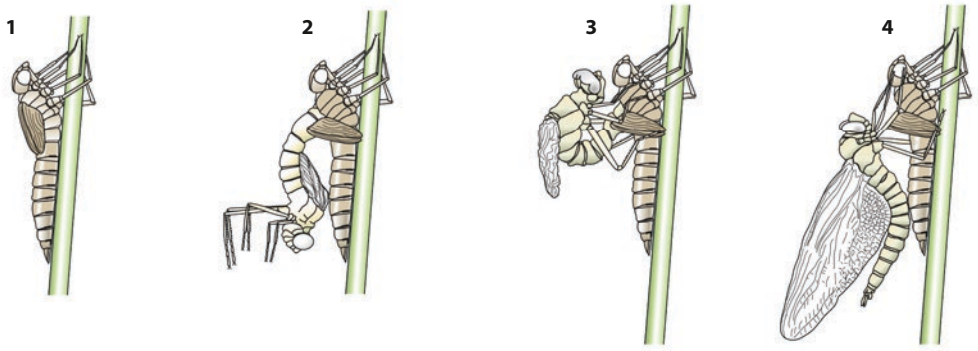


Figure A.17 A dragonfly nymph demonstrates innate behaviour.

while the other is left open. When investigating any one environmental factor, all other variables must remain unchanged.

Woodlice can be introduced to the chamber through a hole in the lid and their movements observed. In an investigation into their response to light, the woodlice begin to move from the light area into the dark side after a few minutes. The numbers of animals in each half can be counted after a fixed period of time to provide quantitative data, and the experiment repeated to obtain more accurate results. Directional movement away from light is known as **negative phototaxis**. It helps woodlice to avoid bright sunny places that are likely to be dry and also ensures that they move to shelter under stones and logs away from predatory birds.

A **kinesis** is a response to a non-directional stimulus. It can be recognised as a change in the level of response to a stimulus.

To investigate the response of woodlice to humidity, a choice chamber can again be used, this time with half the chamber providing damp condition and the other half providing very dry conditions. This can be achieved using a drying agent under the gauze in one side of the choice chamber and damp paper in the other. A single woodlouse put into the chamber moves about on the gauze and behaves differently in the different conditions. Its movements can be recorded as shown in Figure **A.19** by tracing its path on an overlay of clear acetate film. The trace can be marked at 10-second intervals, as shown, so that the speed of the woodlouse can be calculated.

The animal moves further and turns more frequently in the dry conditions. This type of behaviour is a kinesis: the rate of movement and turning depends on the level of the humidity stimulus. Woodlice live in damp environments with a high humidity. In dry conditions, they will keep moving and turn more, searching for a more humid environment.

Both of these investigations show how the innate behaviour of woodlice increases their chances of survival. The longer an animal survives, the greater the likelihood of it being able to reproduce. Since these innate behaviour patterns are genetically controlled and inherited, responses that increase survival are more likely to be passed on to offspring.

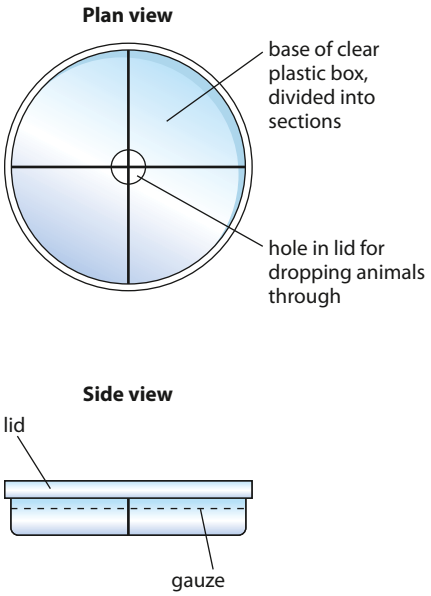


Figure A.18 Plan view and side view of a choice chamber.

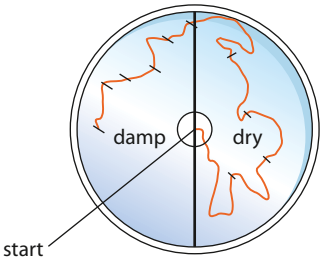


Figure A.19 The distance moved by the woodlouse in 10 seconds is greater in the dry half than in the damp half. The animal also changes direction more frequently.

Intuition in science

It seems intuitive to surmise that the innate behaviour of woodlice in seeking out darker and more humid conditions would increase their chances of survival.

Questions to consider

- Is such intuition sufficient basis upon which to base conclusions?
- In this case, how could this intuition be tested, and would such tests be ethically acceptable?

Test yourself

12 In each of the following, state the type of orientation behaviour that has occurred and suggest why the behaviour helps the animals to survive.

a A sample of 50 *Euglena,* a single-celled photosynthesising organism, was placed in a small, oblong dish and observed under even illumination under a microscope. The distribution of the *Euglena* was seen to be evenly spread over the dish. The illumination was changed so that half of the dish was illuminated and the other half was shaded. After 5 minutes, 48 *Euglena* were in the illuminated side and two remained in the shaded side.

b Newly hatched female silkmoths, *Bombyx mori,* release a pheromone (scented sex hormone) called bombykol, which causes male moths to turn and fly towards them.

c Ten garden snails were placed at the foot of a vertical wall. After 15 minutes the snails had re-orientated themselves and climbed vertically up the wall.

d Three human body lice were placed in a circular chamber that was divided into two parts. One half was kept at a temperature of 35 °C and the other at 30 °C. At the cooler temperature, the insects made few turns but at the warmer temperature, the insects made many random turns and travelled a greater distance in the same period of time.

e *Planaria* are small flatworms that live in lakes and ponds. They have simple light-sensitive eyespots and chemoreceptors at the front of their bodies. Experiments with ten *Planaria* in a choice chamber showed that the animals all moved away from a source of light into a darker area. If a small piece of fish (the natural food of *Planaria*) was introduced into the light section of the choice chamber, five individuals moved towards it.
How should this experiment be modified to include a control?

Reflex actions

Autonomic and involuntary responses (Subtopic **A.2**) are together known as **reflex** actions. A reflex is a specific reaction that is always produced in response to a particular stimulus, and which does not require prior learning. The reflex also forms the basis of Pavlovian conditioning. Sometimes a rapid response to a stimulus is vital for an animal's survival and reflex actions all take place quickly and automatically. Human reflexes include the pupil reflex, which reduces the diameter of the pupil in very bright light to prevent damage to the retina, and the coughing reflex, which occurs when a piece of food enters the trachea.

The **pain withdrawal reflex** takes place if you touch something that causes pain. For example, if you touch a very hot object or are stung by a bee, you pull your hand away quickly, without thinking about it at all. The pain withdrawal reflex is an example of a **reflex action**, mediated by a rapid and simple neural pathway called a **reflex arc** (Figure **A.20**).

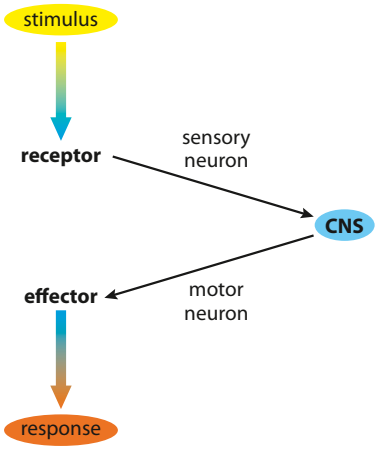


Figure A.20 The basic parts of a reflex pathway.

The pain withdrawal reflex arc involves the receptor cell in your finger, a sensory neuron that carries an impulse to the CNS, a relay neuron in your spinal cord, and a motor neuron that carries an impulse from the spinal cord to the effector – the muscles of your arm that cause you to draw your hand away (Figure **A.21**).

The pathway of a reflex arc is genetically determined so that appropriate responses to different stimuli occur. There are a number of different reflexes that are controlled by the spinal cord, such as the pain withdrawal reflex and the knee jerk reflex. The brain also controls some reflex actions such as the blinking reflex, which happens if something touches the conjunctiva of the eye.

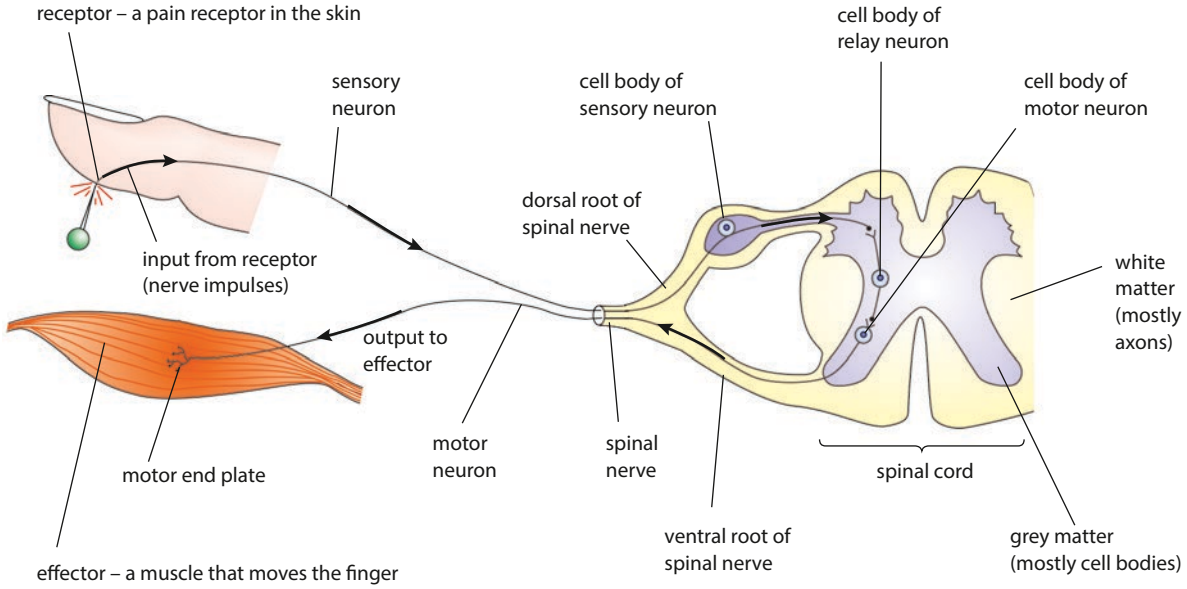


Figure A.21 The spinal reflex arc for the pain withdrawal reflex.

Stimulus a change in the environment (either internal or external) that is detected by a receptor and elicits a response
Response a reaction or change in an organism as a result of a stimulus
Reflex a rapid, unconscious response

Important parts of a reflex arc

Receptors detect a **stimulus** and initiate a nerve impulse. There are many types – e.g. pain, temperature and pressure receptors in the skin.

Effectors are muscles or glands that carry out a **response** to a stimulus.

Reflex arc pathways usually involve the **CNS** – either brain, spinal cord, or both.

Sensory neurons carry impulses from receptors to the CNS.

Motor neurons carry impulses from the CNS to the effector.

Relay neurons inside the CNS connect the sensory and motor neurons via synapses. Relay neurons also connect to neurons going up and down the spinal cord, carrying information to and from the brain. So if you touch something painful, not only do you withdraw your hand rapidly, but information is also sent to the brain so you learn not to do it again.

Learned behaviour and survival

Learned behaviour, unlike innate behaviour and reflex actions, develops as a result of experiences. It is much more adaptable and produces a greater range of behavioural patterns than innate behaviour. Learning is defined as the acquisition of new skills or knowledge, or the modification of existing abilities, which an animal encodes and stores in its **memory**, and can later access as required. Longer-lived organisms with more developed nervous systems are likely to show a higher proportion of behaviour that is learned. Many animals learn from their parents or from older members of their species. Primates, big cats, wolves and many other mammals spend a long time with their parents learning social and hunting skills from them. The matriarch of an elephant herd remembers where water supplies can be found during the dry season and the routes are learned by younger members of the herd.

Primates, in particular, show the ability to acquire new skills that help them to survive. Many monkeys and apes can remember where a particular tree will be fruiting at a certain time of the year and pass on this knowledge to their young. The wild chimpanzees in the Bossou Reserve in Guinea have learned behaviours such as fishing for ants and termites in logs using sticks (Figure **A.22**), and cracking nuts open with a stone hammer and anvil. The young chimps watch other members of the troop and then try to copy them. These behaviours provide a wider range of food sources for the animals that are able to develop the necessary skills.

Many animals learn from experience or by trial and error. Caterpillars of the monarch butterfly in North America feed on a poisonous plant called milkweed. Poison is stored in the caterpillars' bodies and after pupation it is also found in the adult butterflies. If a young toad or bird catches a monarch butterfly, it quickly spits it out and avoids similar prey afterwards. Learning in this way prevents unpleasant and potentially toxic food being taken again.

Raccoons are mammals common throughout North America. Their normal habitat is forest but they have learned that human habitations are excellent sources of food, which they find in pet bowls, garbage cans and even kitchen cupboards. All these new food sources – along with good dens that can be found under houses, in attics and garden sheds – have improved survival rates so much that the animal has become a serious pest in some places.

Figure A.22 These chimpanzees are fishing for termites using tools that they have made.

Imprinting

Imprinting is a type of learning that takes place at a particular stage of life and is independent of the outcome of the behaviour. The best-known work on imprinting was carried out in the 1930s by Konrad Lorenz (1903–1989) who investigated learning in graylag geese. He discovered that in a natural situation newly hatched goslings follow their mother, but that if eggs are removed from a nest and hatched in an incubator, the chicks will 'imprint' on the first moving object they see. When he remained with the eggs as they hatched, and the mother was not present, he found that the newly hatched goslings would follow him (or his boots) instead of the mother bird. He also noticed that this effect only occurred if the exposure was within 13 to 16 hours after hatching – a time he named the **critical period**.

More recently, imprinting has been used to help re-introduce captive-bred birds to the wild. These birds have no adult birds to teach them how to hunt or how to follow the migratory routes of their species. In a novel study in Russia, an Italian hang glider pilot called Angelo D'Arrigo (1961–2006) allowed chicks to hatch under his hang glider so that they imprinted on it. He later encouraged the young birds to follow him into the air so he was able to teach them to fly and hunt. Because the flight path of a hang glider follows air currents and thermals in a similar way to birds, he used his hang glider and knowledge of imprinting behaviour to teach endangered captive-bred Siberian cranes their migration routes from Siberia to the Caspian Sea.

Critical periods for learning

The idea that learning occurs during a critical period is not limited to imprinting in geese. Songbirds have a critical period for song learning and humans also seem to have some critical learning periods. Children below the age of 4 years, for example, seem to learn a language with little effort, without lessons or instructions, while older learners often have much more difficulty and adults seldom learn to speak a new language as well as native speakers.

Operant conditioning

Operant conditioning is behaviour that develops as a result of the association of reinforcement with a particular response. It is a type of trial-and-error learning. The apparatus used to investigate operant conditioning is known as a Skinner box (developed by the behavioural scientist, B. F. Skinner). The box contains a bar or lever that an animal, such as a pigeon or rat, can press or manipulate in order to obtain a reward of food or water (Figure **A.23**). Using these boxes, researchers can study behaviour of animals in a controlled environment. At first, a hungry animal learns by trial and error that pressing the lever causes food to be released. The animal learns to associate the lever with the food reward. The reward is known as **reinforcement**. Experimenters can then condition or train an animal to press the lever in response to specific stimuli such as a sound or pulse of light. If the animal performs the task correctly it receives the food reward. Skinner discovered that operant conditioning develops more quickly if the reward is given sooner after the animal responds, but unexpectedly he found that the response develops more strongly if the reward is not *always* given after the response – the animal presses the lever over and over again, in anticipation of the missing reward. This behaviour could be compared to that of a person in an unmoving elevator who repeatedly presses buttons on the control panel until the expected reward – the movement of the elevator – finally occurs.

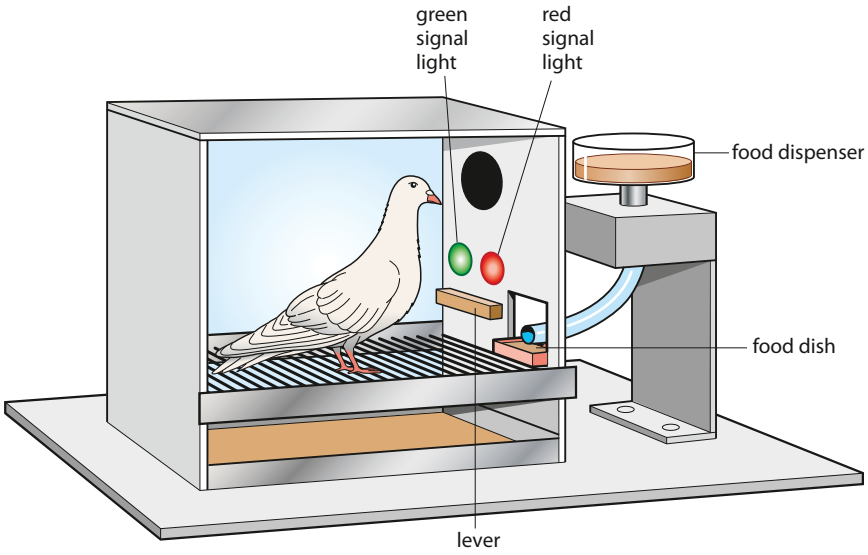


Figure A.23 A pigeon receives a reward of food (reinforcement) for pressing a lever in a Skinner box. Birds can be conditioned to press the lever in response to a signal from a coloured light.

Pavlov's dogs and classical (reflex) conditioning

Ivan Pavlov was a Russian physiologist, psychologist and physician. In the 1890s, he studied the gastric function of dogs and tried to relate the quantity of saliva produced by the dogs' salivary glands to the stimulus of food. Salivation is a reflex response to the presence of food in the mouth but Pavlov noticed that his experimental dogs began to release saliva before they started to eat and he decided to investigate this 'psychic secretion' (Figure **A.24**).

Just before giving the dogs food, and before they could see or smell it, he rang a bell. After repeating his experiments several times he noticed that the dogs salivated as soon as he rang the bell. They had come to associate the sound of the bell with the arrival of food. Even when Pavlov used different sound stimuli, the results were always the same. He called this modification of the dogs' behaviour **classical conditioning** and he used a number of specific terms to explain his results.

- Before training, the normal behaviour involved an **unconditioned stimulus** (the food) producing an **unconditioned response** (the release of saliva).
- After training, the dogs responded to the **conditioned stimulus** (the sound of a bell) and produced the **conditioned response** (the release of saliva without the appearance of food).

Imprinting vs. conditioning

Imprinting and conditioning are two very different modes of learning.

Imprinting usually carries a selective advantage – for example, it causes young birds to stay almost literally 'under the wing' of their parent, where they are likely to benefit from its protection and guidance. But imprinting itself happens 'automatically', regardless of any reward or advantage it brings – so that young birds will imprint on almost anything, however unhelpful.

This type of learning is very different from operant conditioning, for example, which depends on direct rewards. When he carried out his investigations, B. F. Skinner proposed that all behaviour is a reaction to environmental stimuli. Rewards and punishments influence operant conditioning and act as inducements to learning. This fact has many practical advantages in nature and is also used to good effect by teachers in schools and colleges.

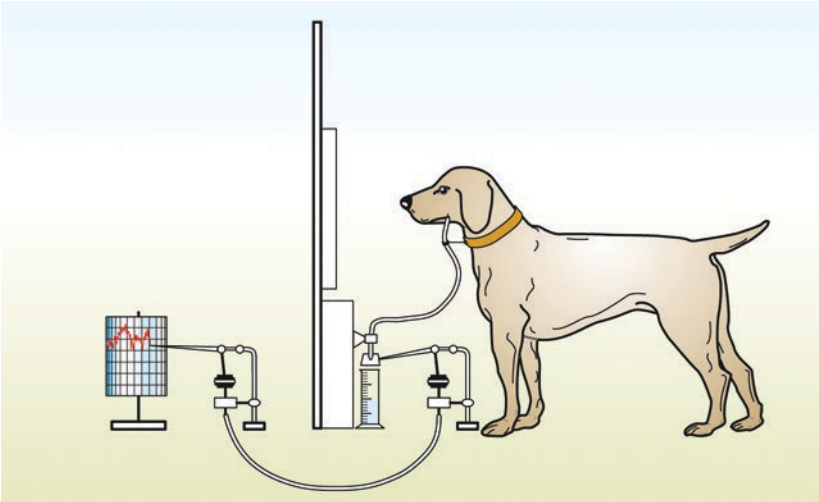


Figure A.24 One of Pavlov's experiments into classical conditioning. A tube in the dog's cheek collected saliva, and the volume collected was recorded on the kymograph drum. Pavlov used this apparatus to investigate the response of the dog's salivary glands to different types and strengths of stimuli.

Behavioural responses involving classical conditioning

Many instances of simple classical conditioning can be observed in animals, including humans.

Questions to consider

Consider the examples below and decide if the behaviour involves Pavlovian classical conditioning or not.

- If you have bells in your school to mark the end of each lesson, watch what the other students do as soon as the bell rings at the end of a class. It is likely that they start to pack up their books and pens, even if the teacher is still talking.
- As you walk past a house, a dog in the garden starts to bark at you.
- A homeowner puts out food for the birds first thing every morning. Early in the morning birds start to gather in the trees near the bird table.
- A sheepdog runs in particular directions or lies down when his owner makes specific whistles.
- A chicken kept in a battery farm cage for a year will start to scratch and peck at the ground when released into a farmyard.

Inheritance and learning in the development of birdsong

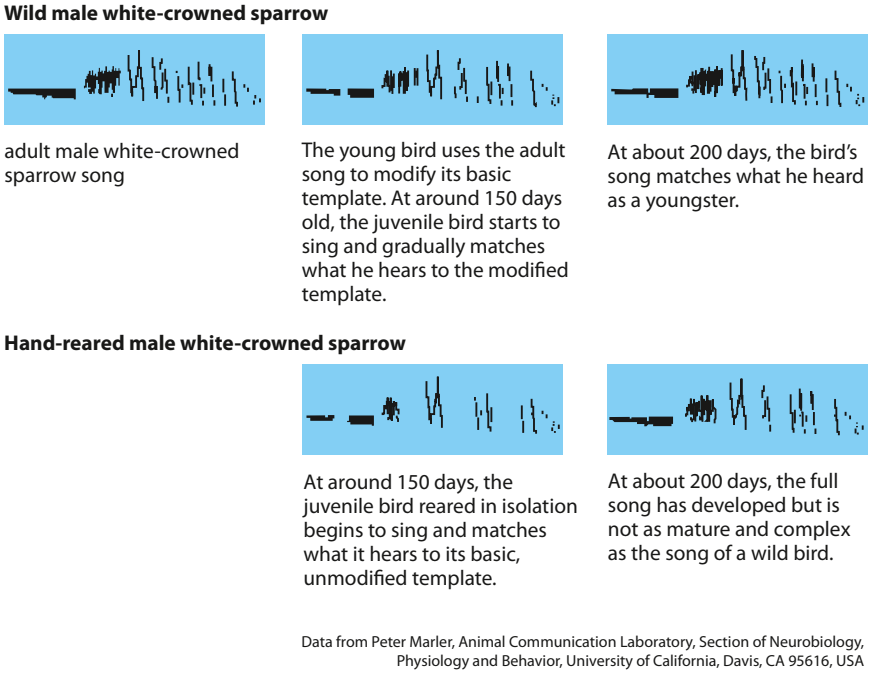
Birds sing to defend their territories and to attract mates in courtship rituals. It is usually male birds that sing so a male bird's song is crucial to both its survival and reproductive success. A bird's song is a long and complex series of notes, which can be analysed using acoustic spectroscopy. Different species have quite different songs, but within a species it appears that, although the basic song is the same for all members of the species, variations do develop. Young birds are born with an innate, *inherited* ability to sing a basic song but *learn* details of their species' song from their fathers. Variations in the song gradually appear and over generations these variations can build up to form local 'dialects'.

One bird that has been extensively studied is the North American white-crowned sparrow (Figure **A.25**). An immature male bird inherits the ability to sing a basic song called a 'template'. However, even before it is able to sing, a young bird listens to its father singing close to the nest and it uses what it hears to upgrade its own basic template. When the young bird starts to sing its own song, it matches what it hears to this upgraded template. A hand-reared bird that never hears adult birds sing is deprived of this learning process and is unable to produce a proper song when it matures. The sonograms of wild and hand-reared birds are shown in Figure **A.26**. A male who does not sing properly will be unable to find a mate.

Figure A.25 A male North American white-crowned sparrow in song.

Memory

Memory is defined as the process of encoding, storing and accessing information. The memory associated with learning a song or learning to respond to the sound of a ringing bell involves changes in an animal's neurons and neural pathways and can be influenced by neurotransmitters (Subtopic **A.5**).



Data from Peter Marler, Animal Communication Laboratory, Section of Neurobiology, Physiology and Behavior, University of California, Davis, CA 95616, USA

Figure A.26 Sonograms of North American white-crowned sparrows.

Nature of science

Looking for trends and discrepancies – laboratory experiments and field investigations

In the study of animal behaviour, both laboratory experiments and field studies have been vital in developing our understanding of different types of behaviour. Skinner and Pavlov both used laboratory studies and apparatus they built themselves to develop theories of operant and reflex conditioning, but other behavioural scientists such as Lorenz have made their observations in the field. Austrian Nobel laureate Karl von Frisch was a contemporary of Lorenz who studied bees and spent long hours tracking the insects and noting the flowers they visited many miles away. Eventually he was able to decode the bees' waggle dance, which they use to communicate the location of flowers to other members of the hive.

In each case, many hours of painstaking observations are needed to collect sufficient data to draw conclusion about the patterns and trends in different animal behaviours.

Test yourself

13 Describe **two** examples of ways in which learning might improve an organism's chances of survival.

14 State **one** example of a taxis and **one** example of a kinesis.

15 Explain how innate behaviour is different from learned behaviour.

A5 Neuropharmacology (HL)

Inhibitory and excitatory synapses

The structure of synapses was discussed in Topic **6**. The most important parts of any synapse are the pre-synaptic membrane, the neurotransmitter it releases and the receptors on the post-synaptic membrane that are stimulated by it (Figure **A.27**).

The synapses discussed in Topic **6** are **excitatory synapses**. When a neurotransmitter is released from the pre-synaptic membrane, the post-synaptic membrane is depolarised as positive ions enter the cell and stimulate an action potential.

But there are many different types of synapses in the body and many different neurotransmitters. Some pre-synaptic neurons release neurotransmitters that inhibit the post-synaptic neuron by increasing the polarisation of its membrane (hyperpolarisation), therefore making it harder to depolarise the membrane and trigger an action potential. Post-synaptic transmission is therefore inhibited at these **inhibitory synapses**.

Neurotransmitters

Neurotransmitters are the chemical substances that are released from the pre-synaptic neuron at a synapse, diffuse across the synaptic cleft, and activate receptors on the post-synaptic membrane. It is these receptors that determine whether the neurotransmitter will cause depolarisation of the post-synaptic membrane or not – that is, whether the stimulus is excitatory or inhibitory. For some neurotransmitters, such as glutamate, the most important receptors all have excitatory effects and increase the probability of an action potential occurring in the post-synaptic cell. For other neurotransmitters, such as GABA, the important receptors all have inhibitory effects. And some neurotransmitters, such as acetylcholine (ACh), are received by both excitatory and inhibitory receptors. Nerve impulses are initiated or inhibited in the post-synaptic cell as a result of the **summation** of all excitatory and inhibitory stimuli they receive from pre-synaptic neurones.

Learning objectives

You should understand that:

- Some neurotransmitters excite post-synaptic membranes but others inhibit them.
- The summation of excitatory and inhibitory effects of neurotransmitters secreted by pre-synaptic membranes results in the initiation or inhibition of nerve impulses in post-synaptic membranes.
- Fast synaptic transmission in the brain is regulated by many different slow-acting neurotransmitters.
- Learning and memory involve changes in neurones caused by slow-acting neurotransmitters.
- Psychoactive drugs work by either increasing or decreasing post-synaptic transmission in the brain.
- Anesthetic drugs interfere with neural transmission between areas of sensory perception and the central nervous system.
- Stimulant drugs mimic the stimulatory effects of the sympathetic nervous system.
- Addiction to drugs can be influenced by genetic predisposition, social environment and dopamine secretion.

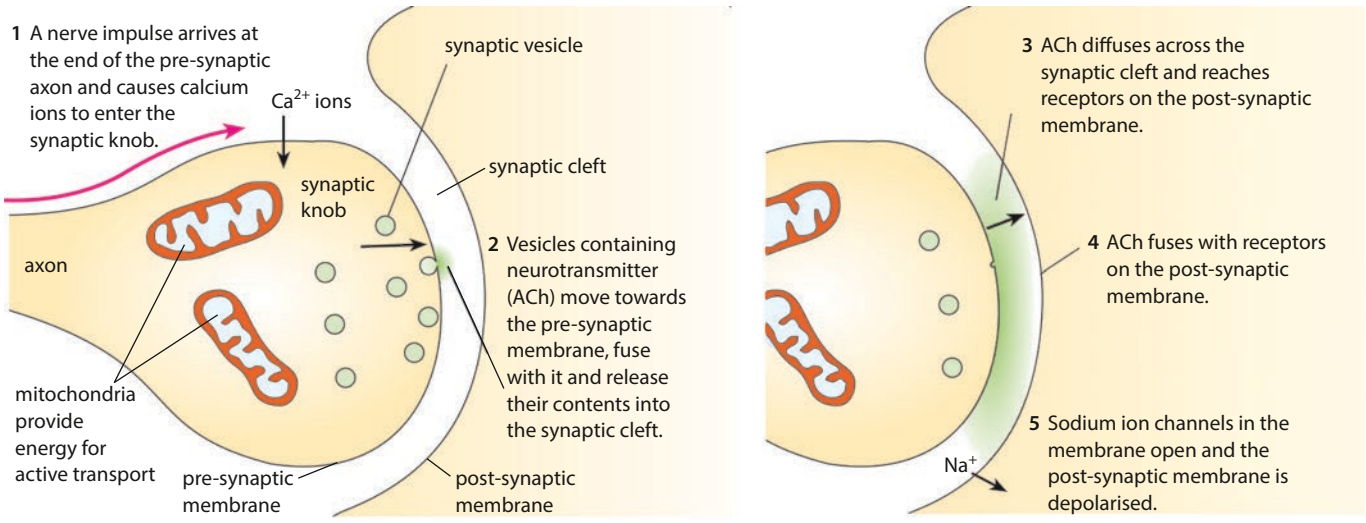


Figure A.27 Synaptic transmission involving ACh (acetylcholine).

Decision making in the CNS

Synapses are the places where action potentials are passed from one neuron to the next. Some post-synaptic neurons are stimulated by many different pre-synaptic neurons, some excitatory and some inhibitory (Figure **A.28**). The balance of stimuli from these many pre-synaptic neurons can either excite or inhibit the post-synaptic neuron, giving a range of possible outcomes. The neuron may receive more stimulatory impulses overall so that it fires an action potential, or it may receive mainly inhibitory impulses so that it does not. The balance of the impulses provides an arrangement that allows us to make decisions about the actions we take.

Many different neurotransmitters are found in the brain. Some of these are listed in Table **A.4**. Memory and learning involve changes in neurons caused by a number of these neurotransmitters. For example, patients suffering memory loss due to Alzheimer's disease have a lower level and concentration of acetylcholine and acetylcholinesterase (the enzyme that removes neurotransmitter from the synapse) as well as a loss of cholinergic neurons. Glutamate, one of a group of amino acids that act as excitatory neurotransmitters, is also thought to play an important role in learning and memory. Glutamate has a key role in synaptic plasticity (the ability of synapses to change their **strength** – that is, the ease with which a pre-synaptic depolarisation can cause an action potential to be stimulated in the post-synaptic membrane). The form of plasticity associated with glutamate is known as long-term potentiation (LTP) and occurs at synapses in several parts of the brain. LTP is a long-lasting enhancement in signal transmission between two neurons and is thought to be an important process in learning and memory. Memories are probably encoded by modifications of synaptic strength. In abnormal conditions, glutamate can also act as a neurotoxin and cause damage to the brain, especially in older people.

Synaptic strength

Synaptic strength is defined as the change in membrane potential in a post-synaptic membrane (measured in millivolts, mV) following a pre-synaptic action potential. This change in the post-synaptic membrane may result from activity either at a single synapse or at a number of connections between pre-synaptic neurons and the post-synaptic neuron, and can also depend on the sizes of these connections.

Synaptic strength can vary as a result of previous activity in the neurons, and can be short-term, lasting a few seconds or minutes, or long term, lasting hours (long-term potentiation, LTP). Learning and memory are thought to result from long-term changes in synaptic strength, which result from this synaptic plasticity.

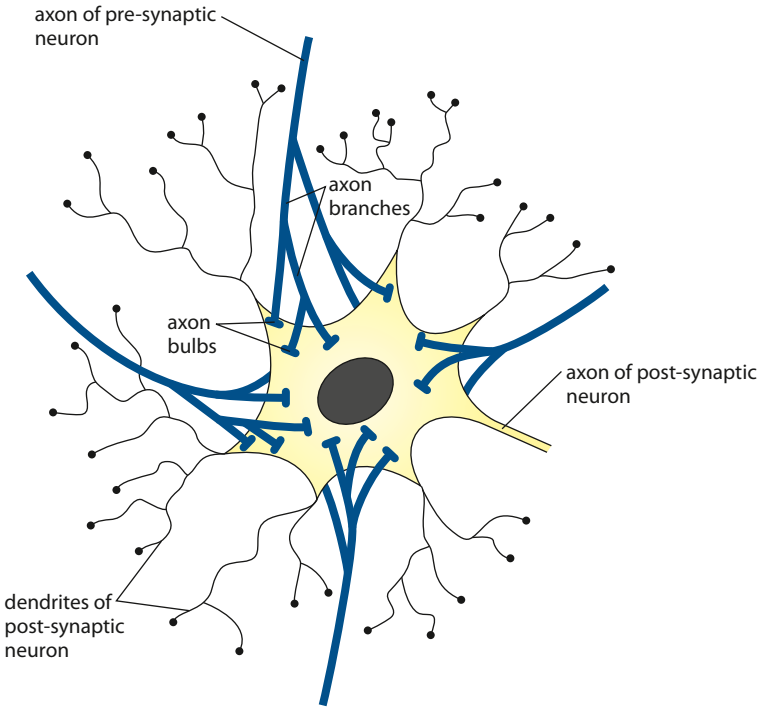


Figure A.28 Some of the neurons that form synapses with the post-synaptic neuron are inhibitory and prevent an action potential being stimulated. Others stimulate the propagation of the impulse.

Neurotransmitter	Effects
acetylcholine	Widely distributed excitatory neurotransmitter that triggers muscle contraction and stimulates the release of certain hormones. It is involved in wakefulness, attentiveness, anger and aggression. Alzheimer's disease is associated with a lack of acetylcholine in certain regions of the brain.
dopamine	Involved in controlling motor actions such as movement and posture. In the brain it modulates mood and pleasure related to motivation. It has a role in positive reinforcement and dependency. If it is lost from certain parts of the brain it can cause muscle rigidity and Parkinson's disease.
GABA (gamma-amino butyric acid)	Found widely distributed in the neurons of the cortex and the majority of fast inhibitory synapses in almost every part of the brain. Many sedative drugs act by enhancing the effects of GABA. GABA contributes to motor control, vision and many other functions, and regulates feelings of anxiety.
glutamate	A major excitatory neurotransmitter that is associated with learning and memory.
norepinephrine (noradrenalin)	Important for attentiveness, emotions, sleeping, dreaming, and learning. It is also released as a hormone into the blood, where it causes blood vessels to contract and heart rate to increase. It plays a role in mood disorders such as manic depression.
serotonin	Contributes to functions such as regulating body temperature, sleep, mood, appetite and pain. Depression, suicide, impulsive behaviour and aggressiveness all appear to involve certain imbalances in serotonin.
substance P	Responsible for transmission of pain from certain sensory neurons to the central nervous system. It is a neurokinin (small peptide) and is released when stressful stimuli are received.
endorphins (opioid peptides)	A group of neurotransmitters found in pain pathways and emotional centres of the brain. They are produced by the pituitary gland and hypothalamus during exercise, excitement or when pain is felt. They are known as 'natural painkillers' because some have an analgesic effect and can cause feelings of pleasure or well-being.

Table A.4 Neurotransmitters and their effects.

Cholinergic and adrenergic synapses

Two of the most important neurotransmitters in the nervous system are acetylcholine and norepinephrine (noradrenalin). Synapses are divided into two types, defined by which of the two neurotransmitters they use.

Cholinergic synapses use acetylcholine and are found in the parasympathetic nervous system. Nicotine increases transmission at these synapses and has a calming effect on mood.

Adrenergic synapses use the neurotransmitter norepinephrine and are found in the sympathetic nervous system. Norepinephrine is crucial to the 'fight or flight' response. Amphetamines stimulate these synapses and produce feelings of alertness and euphoria.

Psychoactive drugs

Psychoactive drugs are chemical substances that affect the way the brain transmits impulses at synapses. They are capable of altering the functioning of the brain and a person's personality.

Drugs act in different ways.

- Some have similar structures to neurotransmitters and so either block receptors, preventing a response, or have the same effect as the neurotransmitter but are not removed so that the response is prolonged.
- Some prevent neurotransmitters being released.
- Some increase the release of neurotransmitters.
- Some prevent neurotransmitters being broken down and so prolong their effects.

Exam tip

Do not forget that psychoactive drugs can either increase or decrease postsynaptic transmission.

Excitatory drugs (stimulants)

Some psychoactive drugs are excitatory – that is, they promote the transmission of impulses at excitatory synapses or inhibit transmission at inhibitory synapses. Excitatory, or stimulant, drugs mimic the stimulation provided by the sympathetic nervous system. Examples of excitatory drugs include:

- cocaine
- amphetamines
- nicotine.

The effects of these substances are summarised in Table **A.5**.

Excitatory drug	Mode of action	Effects
nicotine	• acts at synapses that use the neurotransmitter acetylcholine • is not broken down by the enzyme acetylcholinesterase, which breaks down acetylcholine • remains in the synapse, binding to the same receptors on the post-synaptic membrane as acetylcholine • increases levels of dopamine in the brain, which stimulates synapses in 'reward' pathways, giving feelings of pleasure and well-being	• produces feelings of pleasure, in the same way as cocaine and amphetamines, although to a lesser degree • strongly addictive because effects wear off quickly, so users must dose themselves frequently to maintain pleasurable sensations and prevent withdrawal symptoms • has a calming effect, despite being an excitatory drug, possibly because it reduces agitation caused by cravings and withdrawal symptoms
cocaine	• stimulates transmission at brain synapses that use dopamine • leads to a build-up of dopamine in the synapse by blocking its return to pre-synaptic neurons • causes continuous transmission of impulses in 'reward' pathways, giving feelings of pleasure and well-being	• produces feelings of increased energy, confidence and euphoria often mixed with restlessness and anxiety • highly addictive, with users seeking to maintain feelings and 'highs' induced by dopamine • as it wears off, feelings of euphoria turn into depression, and the user may 'crash', losing all energy and sometimes sleeping for long periods • prolonged use can cause long-lasting mental health problems such as depression, anxiety, paranoia and delusions
amphetamines	• stimulate transmission at synapses that use norepinephrine (noradrenalin) • have similar effects to cocaine but are longer lasting • cause the release of neurotransmitter into the synapse and prevent it being broken down • increase the concentration of dopamine present	• produce feelings of euphoria and high levels of energy and alertness • may cause hyperactivity and aggression in some people

Table A.5 Effects of excitatory drugs.

Ecstasy (MDMA)

Ecstasy or MDMA (methylenedioxymethamphetamine) causes neurons to release serotonin, a neurotransmitter that controls emotions, mood, pain perception and sleep. The drug also alters the sensitivity of the brain to serotonin. Some of the behavioural effects of the drug are due to a massive release followed by depletion of serotonin. Depletion of serotonin can cause depression and memory loss. One serious effect of ingestion of MDMA by humans is hyperthermia (raised body temperature) which can induce other clinical problems and occasionally death. MDMA has also been shown to induce dose-dependent hyperthermia in experimental animals.

Consider the data shown in question **11** at the end of the topic, which shows effects of MDMA given to experimental animals.

Inhibitory drugs (sedatives)

These drugs increase transmission at inhibitory synapses or suppress transmission at excitatory synapses. This class of drugs is also known as sedatives and includes anesthetics used in medicine. Examples of inhibitory drugs include:

- benzodiazepines
- alcohol
- THC.

Their effects are summarised in Table **A.6**.

The effect of anesthetics on the nervous system

An **anesthetic** is defined as a drug that that causes a reversible loss of sensation. **Local anesthetics**, such as novocaine used by dentists, are used to block the transmission of nerve impulses to pain centres in the central nervous system during minor surgical procedures. They work by binding to and inhibiting sodium channels in the cell membranes of neurons. This prevents nerve impulses passing messages to the central nervous system from the region near the site of injection, but the patient will still be awake and will not have any change in perception in other parts of the body.

General anesthetics cause a reversible loss of consciousness. They induce a state of general insensibility to pain. During medical procedures an anesthetised patient loses consciousness but their vital functions, such as breathing and heart beat, continue. Many different substances have been used as general anesthetics for more than 150 years (Figure **A.29**). They are known to cause a reduction in nerve transmission at synapses but even today their exact mechanism of action is not fully understood.

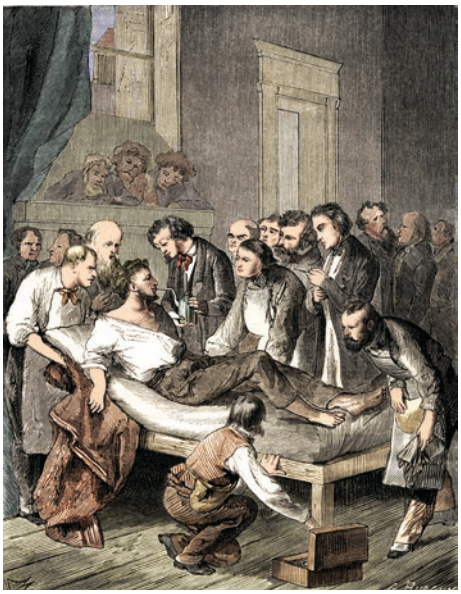


Figure A.29 William Thomas Green Morton (1819–1868), an American dentist from Boston, administering ether to a patient during the public demonstration of an operation by the surgeon John Collins Warren, on 16 October, 1846. This was the first documented successful major surgical intervention under a general anesthetic.

Inhibitory drug	Mode of action	Effects
benzodiazepines	• bind to the same post-synaptic receptors as GABA, the main neurotransmitter at inhibitory synapses • cause hyperpolarisation of post-synaptic membranes so that they are more difficult to stimulate	• reduce anxiety, cause relaxation and can induce sleep • used therapeutically to treat anxiety, insomnia and seizures
THC (tetrahydrocannabinol – the most important psychoactive substance in cannabis)	• affects receptors in cells in the cerebellum and cerebral hemispheres that use the neurotransmitter anandamide • similar in structure to anandamide and binds to the same receptors, known as cannabinoid receptors • causes hyperpolarisation of post-synaptic membranes so that they are more difficult to stimulate	• induces feelings of relaxation and affects coordination • causes panic and paranoia in some users • can interfere with short-term memory and learning, as many cannabinoid receptors are found in areas of the brain concerned with memory
alcohol	• increases the binding of GABA to receptors in post-synaptic membranes • causes hyperpolarisation of post-synaptic membranes so that they are more difficult to stimulate • decreases the action of the neurotransmitter glutamate, which stimulates post-synaptic neurons	• in small quantities, affects behaviour by reducing inhibitions • in larger quantities, can cause a lack of coordination, slurred speech, loss of balance and, in some cases, aggressive behaviour

Table A.6 Effects of inhibitory drugs.

Another way in which patients may be prepared for medical procedures is sedation. During sedation small amounts of anesthetic or similar drugs are used to produce a 'sleep-like' state. Under sedation the patient is physically and mentally relaxed so that procedures such as endoscopy, which may be unpleasant or painful, can be carried out. The patient may have little memory of events and may not recall what has happened to them.

Barbiturates and benzodiazepines are drugs that are used to produce either sedation or anesthesia (loss of sensation, such as touch, for example) but themselves have no pain-relieving properties. **Analgesics** such as aspirin and ibuprofen are painkillers that relieve pain but do not eliminate sensation, such as touch. Combinations of anesthetics and analgesics are sometimes used to produce additive therapeutic effects.

What causes addiction?

Addiction is a chemical dependence on a psychoactive drug. Many different factors are involved in addiction as the body becomes **tolerant** of a drug, needing more and more of it to produce the same effects.

Three factors seem to be common to all addictions, whether drugs have been taken for therapeutic reasons or recreation.

Social factors

Peer pressure can influence young people to experiment and drug-taking behaviour can be associated with a need to belong to a group.
Culture also affects whether drug use is acceptable. In some cultures, cigarette smoking is freely accepted, but in others it is only acceptable for men to smoke in public; in some cultures, alcohol is used to celebrate at social events, while in others it is prohibited entirely. The use of opium, for example, has a long history. In Homer's *Odyssey*, written around 800 BCE, opium is referred to as a '*drug that had the power of robbing grief and anger of their sting*' and in the 1850s Chinese immigrants who helped build the railways in the USA smoked opium as an integral part of their culture to relieve stress and exhaustion. Today, drug addiction is often linked to factors such as poverty or poor family circumstances, which can increase the chances of an individual starting to use drugs. However, certain drugs (such as cocaine) may be used by more affluent members of society.

Dopamine secretion

Most drugs that cause addiction are those involving 'reward' pathways and the release of the neurotransmitter dopamine. Users of addictive drugs find it hard to give them up because of the feelings of well-being that are induced by dopamine. As dopamine receptors are repeatedly stimulated, they become desensitised so that more and more of the drug is required to produce the same feelings. In this way, the user develops a **tolerance** to the drug.

Genetic predisposition

Relatively few people become addicted to drugs although many are exposed to them. The tendency to become addicted has been shown to be more common in some families and groups than others. Research on

identical twins also supports this view. This evidence seems to indicate that some individuals are more likely to carry genes that predispose them to addiction than others.

Nature of science

Assessing risk in science – approving new drugs

The National Institute for Health Care and Excellence (NICE) is an organisation that gathers the available evidence on new medicines in the UK. Similar organisations exist in many other countries. Such organisations are responsible for considering and summarising evidence on new drugs that are developed for use with patients. Before any new medication is used it must be thoroughly tested so that patients are not put at unnecessary risk. But sometimes patients or their families may press for drug approval processes to be speeded up if they feel it may be of benefit to them. This encourages a greater tolerance of risk and may lead to reckless use of new medicines.

Neuro-adaptation

Excessive use of a psychoactive drug can cause the brain to attempt to counteract the effect of the drug in a process known as **neuro-adaptation**. If the brain is frequently exposed to a drug it adapts to compensate for its presence and if the drug is withdrawn, the brain overcompensates and becomes imbalanced in different way. Neuro-adaptation is the basis of addiction and drug tolerance.

Test yourself

16 Describe how THC affects synapses in the brain.
17 List **three** causes of addiction to psychoactive drugs.
18 Outline the effect of inhibitory drugs at a synapse.

A6 Ethology (HL)

Behaviour is the pattern of responses of an animal to one or more stimuli, and the study of animal behaviour in natural conditions is called **ethology**. Some examples of different behaviours, such as patterns of social interaction and mating rituals, are considered in this subtopic.

Learning objectives

You should understand that:
- Ethology is the study of the way animals behave in natural conditions.
- Natural selection can influence animal behaviour.
- Behaviour that increases survival and reproductive chances will tend to become more prevalent in a population.
- Learned behaviour can spread through or be lost from a population more rapidly than innate behaviour.

Natural selection and animal responses

Natural selection acts on the behavioural responses of animals in just the same way as it does on other characteristics. Behaviour that increases the chances of survival and reproduction will tend to become more prevalent in a population. One well-documented example is the migratory behaviour of a small European songbird called the blackcap (*Sylvia atricapilla*).

Blackcaps breed in summer, in Germany and other areas of northern and eastern Europe, and then migrate south about 1600 km to winter feeding grounds in Mediterranean regions of Spain (Figure **A.30**). Since the 1960s, biologists in the UK have recorded increasing numbers of birds travelling northwest from Germany to overwinter in the UK, a distance of only 900 km. At the end of winter, these birds leave the UK up to 10 days earlier than other migrants left Spain.

Modification of their migration pattern has meant that the birds travelling from the UK can quickly return to their summer breeding grounds and occupy the best nest sites. Observations of the birds have shown that they have different-shaped beaks, more suited to the food available in the UK, and also more rounded wings, which are less suitable for long migration. The 'modified migrants' also tend not to interbreed with the birds that migrate back from Spain.

Figure A.30 Changes in migration patterns of the blackcap. Blackcaps tended to breed in Germany in the summer and migrate in winter to feed in Spain. In recent decades, increasing numbers have instead travelled northwest from Germany, to overwinter in the UK.

Figure A.31 There is a selective advantage for great tits in laying eggs early, if the parents can get enough food for egg production, because when the chicks hatch they will be first to take advantage of the springtime abundance.

To study the behaviour of these birds in more detail, eggs were removed from the nests of both types of migrants and the hatchlings hand-reared. In autumn when the young birds migrated, the direction in which each bird flew was carefully observed. It was found that birds reared from eggs of south-migrating parents headed south, and those birds reared from northwest-migrating parents headed northwest. The birds had had no previous migration experience and their behaviour had not been learned from their parents. The migration patterns must be genetically determined, indicating that natural selection is operating on the behaviour of the blackcap.

A second example of how natural selection affects animals' responses can be seen in the breeding behaviour of the great tit (*Parus major*, Figure **A.31**). This European bird lays eggs at a time that is influenced by day length, and the behaviour is genetically determined. In recent years, ornithologists in the Netherlands have noted that many birds are laying their eggs earlier in the year. In general, natural selection favours early breeding because, with changes in climate, trees now tend come into leaf earlier and so small invertebrates that inhabit these trees are also available earlier. Early egg production means that there is abundant food for the birds' offspring at the time when it is needed.

The eventual extent of early breeding may in turn be limited by natural selection, because egg production is very costly and so birds' energy needs may come to restrict their laying behaviour. There may simply not be enough food around very early in the year for parents to produce eggs.

There are many other examples of animal behaviour that increase the chances of survival and reproduction and some of these are summarised below. Consider the examples and notice how learned behaviour, which is not genetically determined, can spread through a population or be lost from it far more quickly than innate behaviour, which is genetically determined and can only be changed by natural selection acting on genetic variation within the population.

Reciprocal altruism in vampire bats

Altruistic behaviour is 'unselfish' behaviour that does not benefit the individual itself but benefits another, which may be genetically related. Altruistic behaviour may decrease the individual's chance of survival and reproduction but increase the number of offspring produced by another animal. Good examples of altruistic behaviour are seen in naked mole rats and in vampire bats.

The common vampire bat (*Desmodus rotundus*) is found in tropical South America, including Mexico, Argentina and Chile, especially where cattle are present in large herds (Figure **A.32**). Bats emerge from their roosts at night to feed on the blood of large warm-blooded animals while they are asleep. Bats have a rapid metabolism and need a regular supply of food; they will begin to starve if they have to go without food for more than 48 hours. Female bats form close long-term associations with small groups of other females and these are the basis of the bats' social structure. Within the groups, a bat that has fed will regurgitate food into the mouth of an individual that has been unable to feed. Food sharing seems to be altruistic; a bat gives up food that could have helped to sustain its own body to another individual whose chances of survival are therefore increased. Zoologists who have studied these animals for many years have suggested that bats share blood only with those who have fed them in the past or with bats that are related to them. This indicates that this is an example of reciprocal altruism. Bats that are newcomers to the colony do not receive donations of food so easily. Altruistic behaviour between related animals increases the chance of genes that are shared by these individuals being passed on to the next generation, and so this behaviour is favoured by natural selection (Subtopic **5.2**).

Figure A.32 Despite their reputation, vampire bats demonstrate altruistic behaviour towards others in their colony by sharing meals.

Synchronised oestrus in female lions – innate behaviour that increases chances of survival

Lions do not breed with the seasons as other mammals do, but within a pride, female lions – which often live in groups of related animals – come into oestrus and are ready to mate at the same time as other females in the group. Synchronised oestrus is also well known in other animals such as rats and hamsters that are kept in laboratories. Cubs that are born into synchronised litters have several advantages. They are suckled and protected by more than one female in the pride and the adult males will remain close by to protect them. When the time comes for the young lions to leave the pride and establish themselves, they are more likely to leave along with other animals of the same age that will help them survive.

In addition there is evidence that oestrus is also stimulated if the pride is taken over by an incoming male lion. Incoming males will attack and kill dependent cubs. Females with cubs will not be ready to mate for 1–2 years but if their cubs are lost, they will come into synchronised oestrus and mate with the new dominant male.

Figure A.33 Blue tits will pierce the foil on a milk bottle to feed on the milk inside.

Blue tits – learned behaviour developed and lost

Until the start of the twenty-first century, most milk in the UK was delivered to customers' doors in milk bottles. At the beginning of the twentieth century some of the milk containers that were delivered had no tops so that birds that perched on the edge of the container had easy access to the cream floating on the top of the milk. The blue tit (*Parus caeruleus*) was one species that quickly learned to feed on the cream.

From the 1920s onwards, dairies started to seal their milk bottles with aluminium foil tops, but blue tits modified their behaviour to cope with the new situation and by the early 1950s all the blue tits in the UK had learned how to pierce the tops with their beaks (Figure **A.33**). Blue tits had not only learned a new behaviour but also passed on the skill to other members of their species. Even though some other species such as the robin occasionally took cream, only the blue tits passed on the behaviour.

This learned behaviour has largely been lost today because many people now buy milk from a supermarket rather than having a delivery to their doorstep, and where milk delivery still occurs, it comes in cartons rather than bottles. Changes in human dietary preferences have also meant that even households that have milk delivered may choose low fat or homogenised milk, which has no cream on the top for the blue tits to feed on.

Foraging behaviour in shore crabs – optimising prey choice

Efficient feeding behaviour is essential for survival and reproduction. But in hunting and foraging there is an energy cost in finding, catching and consuming food. This has to be balanced with the energy the animal gains from the food. Animals are able to change their behaviour so as to ensure the overall benefit is greater than the cost.

The shore crab (*Carcinus maenas*) feeds on mussels. A shore crab will investigate a mussel by handling it in its claws for about two seconds before deciding whether it is a suitable size to feed on. Each differently sized crab will have a size of mussel for which the ratio of energy content to handling time is at its maximum efficiency. This ratio is known as the

prey value and can be calculated for each crab size. The optimal mussel size for an individual increases with the size of the crab.

If there is an abundance of food available, crabs will choose prey that are close to the predicted optimal size. But if this size is scarce, they will select less suitable mussels, which may be either above or below the optimal size. When a crab is forced to eat larger mussels, the prey value decreases because it must use a slower feeding method.

Crabs are able to adapt their feeding habits by eating sub-optimal mussels in proportion to the relative numbers of mussels present. Mussels tend to be distributed in groups of mixed sizes so crabs always take a mussel of optimum size if they find one. They will reject a sub-optimal mussel after one encounter, but take one after a series of encounters with mussels of sub-optimal size.

Similar prey selection behaviour is seen in the bluegill sunfish. The bluegill sunfish (*Lepomis macrochirus*) is a well-studied species that feeds on water fleas (*Daphnia* sp.) and other small pond invertebrates. Table **A.7** shows how the foraging behaviour of the bluegill sunfish varies depending on the amounts of *Daphnia* available. When there is a low density of prey available the fish consume all sizes of *Daphnia*, but at medium densities they consume only middle-sized or larger prey. When food is abundant they actively select only the largest *Daphnia*. Feeding on small numbers of large prey takes less energy than catching large numbers of small prey, if they are nearby. If the density of food is low, the fish will eat whatever they can rather than go hungry.

Density of *Daphnia* in pond habitat	Sizes of *Daphnia* selected as prey by bluegill sunfish
low	all
medium	middle-sized or large
high	largest

Table A.7 The foraging behaviour of the bluegill sunfish depends on the amount of prey available.

Exam tip
Make sure you can describe examples of animal behaviour that increase the chances of survival and reproduction.

Mate selection – courtship in birds of paradise

The elaborate tail of the peacock (*Pavo cristatus*) has fascinated biologists for many years; even Charles Darwin wondered why such an elaborate and impractical structure had evolved and what purpose it served. But research has shown that females of the species (peahens) prefer males with larger tails, so these males have more reproductive success and are more likely to pass on their genes to the next generation.

Similar elaborate displays of plumage and behaviour are seen in many bird species but the birds of paradise of Papua New Guinea are thought to have some of the most complex displays of all. The courtship behaviour of Queen Carola's Six-wired Parotia (*Parotia carolae*) bird of paradise has been extensively studied and its complex behaviour has been observed. Male birds are mostly black and have three ornamental head 'wires' attached above each eye. They also have display feathers that form a skirt during courtship, white flank plumes and iridescent throat and breast feathers (Figure **A.34**). A male bird performs an elaborate courtship ritual involving dance, display and calls. His first strategy is to select a display

area, which he clears of debris and prepares with a mat of fungi. This 'stage' is decorated with fur and brightly coloured leaves. The male may also remove leaves from the branches of surrounding plants so that females can get a clearer view of his display. Females perch on branches above the stage and watch as the male shakes out his feathers to form a fan, dances and shakes his body and head, and pirouettes to impress them. He may also call and rattle his feathers. In some other species, such as the blue bird of paradise, males flip upside-down on branches to perform hanging displays, spreading their breast plumes into a fan as they bounce and wave their tail wires. Other species' displays involve several males displaying together. Females select a mate from the quality of the displays.

It seems that female birds of paradise prefer males with elaborate and attractive displays, but the reasons for this are not fully understood. Females may prefer such males because they are likely to produce attractive (and therefore reproductively successful) male offspring, or because an elaborate dance is a sign of good health.

Figure A.34 Six-wired bird of paradise.

Breeding strategy affects chances of survival – coho salmon

Coho salmon (*Oncorhynchus kisutch*) live in the North Pacific Ocean but return to fresh water to mate and spawn in gravel beds of the streams where they were born. After spawning once, the adult fish die. Young fish live for one or two years in the streams, growing and feeding in fresh water, but travel to the sea for the rest of their life cycle. Coho salmon have blue backs and silver sides in the ocean, but spawning fish are darker with red colouration on their sides (Figure **A.35**). Adults usually spawn when they are about 3 years old but some precocious males, known as 'jacks', spawn early at the age of 2 years. Males adopt one of two irreversible life-history strategies in their mating behaviour; either they mature early at a smaller body size ('jacks') or delay maturation until they are larger, when they are known as 'hooknoses'.

At spawning time, the males develop a hooked snout and large teeth and a distinctive colouration. Male mating strategy depends on body size and their competitive fighting ability. Hooknoses are more brightly coloured with larger hooked jaws and sharp teeth and fight with other males for access to females. Jacks are smaller, less colourful and poor fighters so they must use a 'sneaking' strategy to get close to females. Jacks mate in shallow water and hiding places where their less vivid colouration enables them to remain hidden. The 'fitness' or lifetime reproductive success of jacks and hooknoses depends on the probability that they will survive to maturity. On one hand, the jacks' strategy reduces the exposure of a fish to the dangers (such as predators) in the ocean, by 1 year. But the opportunities for 'sneaking' mating behaviour are limited by the number of shallow water hiding places. If too many males become jacks, the number of opportunities for sneaking is reduced. For a hooknose, survival is influenced by the number of other hooknoses – more hooknoses means more competition.

Human activity and changes to the environment are affecting coho salmon populations, which are declining. It is probable that the number of fish becoming jacks and the behaviour of the species will change as a result of, for example, if larger fish are taken by fishing or their numbers decline due to pollution.

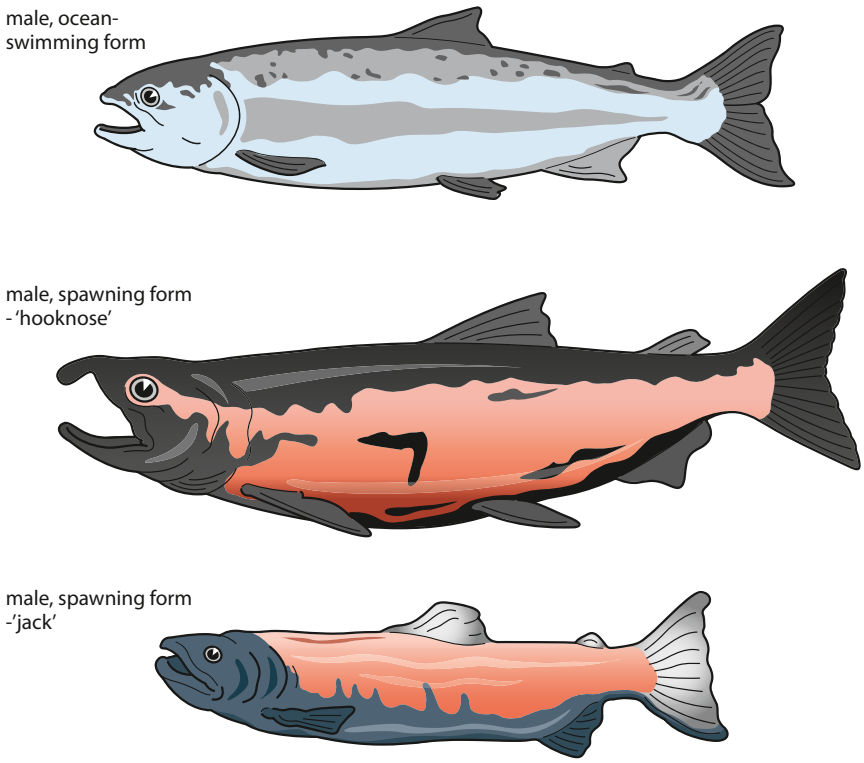


Figure A.35 Male coho salmon in oceanic and spawning phases.

Animal behaviour in literature

British author Rudyard Kipling's 'Just So Stories' were first published in 1902. They tell the stories of different animals and their appearance and behaviour in a series of fanciful descriptions. For example, 'How the Camel got its hump' tells how the camel was given a hump as a punishment for refusing to work, and other stories tell 'How the rhinoceros got his skin' and 'How the leopard got his spots'. The stories are imaginative descriptions, which were written for children, and are very different from scientific observations of animal behaviour and physiology.

Questions to consider

- What are the features of a scientific explanation of animal behaviour?
- How does this differ from a historical or literary view?

Nature of science

Testing a hypothesis – changing migratory patterns

The behaviour of migratory black caps has been extensively monitored since it was proposed that their migration patterns might be affected by climate change. To provide more conclusive evidence, experiments involving relocating their eggs were carried out to test the hypothesis that behaviour was changing.

Researchers have also carried out experiments to test the hypothesis that European starlings migrate using set patterns of distance and direction as aids to navigation. These birds make short migrations in a southwesterly direction from the Netherlands to France and southern England. When populations of starlings were moved to Switzerland by experimenters, they did not fly northwest to their usual winter homes but travelled in the same southwesterly direction and the same distance so that they arrived in Spain, where they were retrieved. Testing their hypothesis enabled the researchers to conclude that direction and distance were cues for the starlings' migration pattern.

Test yourself

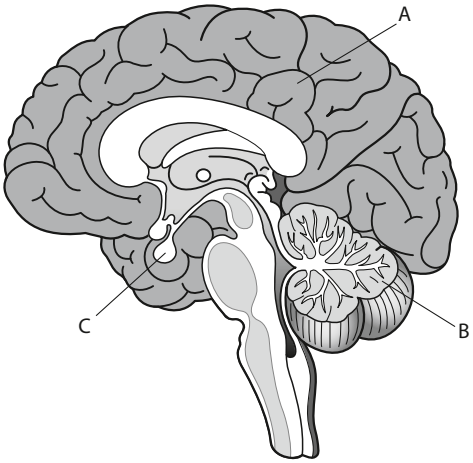
19 Give an example of a learned behaviour that has spread through an animal population.

20 How does synchronised reproduction benefit the offspring that are produced?

Exam-style questions

1 **a** Outline the role and functions of the visual cortex and Broca's area of the brain. **[4]**

b Name the parts labelled A, B and C in the diagram of the brain. **[3]**



2 Discuss how the pupil reflex is used to test for brain death. **[4]**

3 Draw a labelled diagram of a reflex arc and outline the pathway of a withdrawal reflex. **[6]**

4 Outline what is meant by the term 'contralateral processing'. **[2]**

5 Outline the role of hair cells in the cochlea in the processing of sound. **[3]**

6 Distinguish between innate and learned behaviour. **[2]**

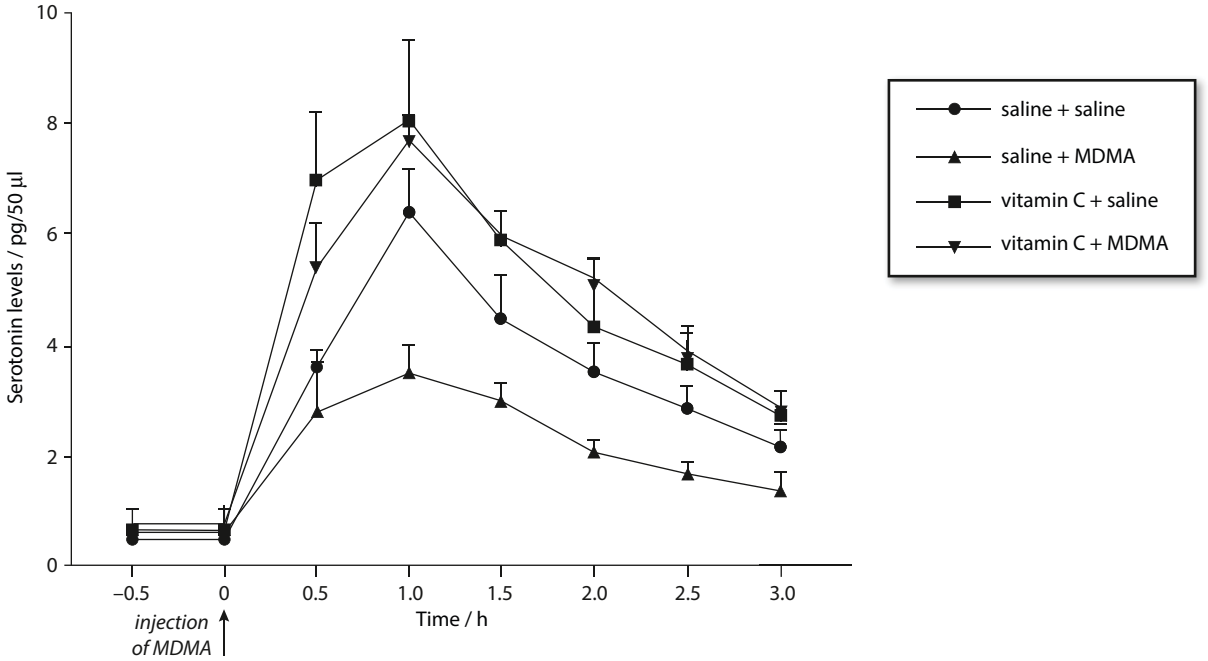
7 Discuss how learning may improve survival chances. **[3]**

8 Outline Pavlov's experiments on conditioning in dogs. **[3]**

9 Explain how a pre-synaptic neuron can inhibit a post-synaptic neuron. **[2]**

10 Outline the reasons why learned behaviour can spread through a population or be lost from it more quickly than innate behaviour. **[4]**

11 The graph shows the effect of an injection of MDMA (ecstasy) on serotonin levels over a period of 3 hours in four groups of experimental rats. Two groups (indicated by the triangles) had been given MDMA either with a saline solution or with vitamin C 1 week prior to the experiment. The two other groups (indicated by the squares and circles) had not been pre-treated with MDMA. Serotonin is a neurotransmitter concerned with the regulation of mood, appetite and sleep.



a i Describe the effect of MDMA on the serotonin levels of rats which had not previously received doses of MDMA. **[2]**

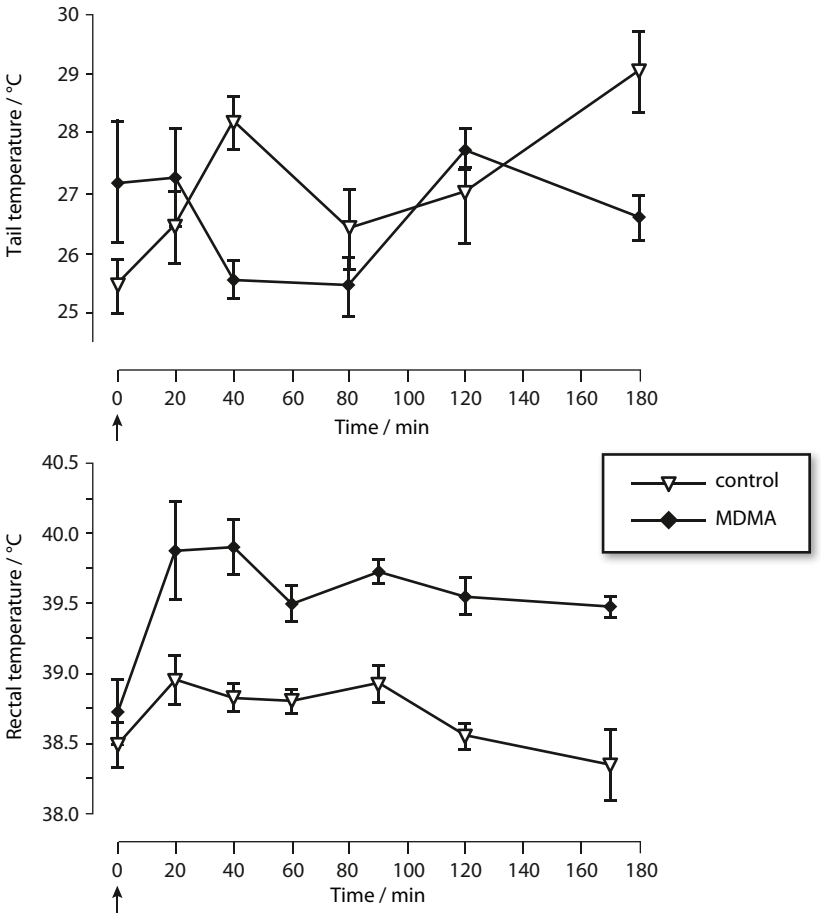
ii Identify the group of rats whose serotonin level was most affected by the MDMA. **[1]**

iii Estimate the time needed for the level of serotonin to return to its normal (pre-injection level) in the control group of rats, indicated by the circles. **[1]**

iv Compare the effect of MDMA on the group of rats which had received doses of vitamin C in the week before the experiment with the group that had not. **[1]**

v Explain the purpose of the saline + saline injection. **[1]**

One serious effect of ingestion of MDMA (ecstasy) by humans is hyperthermia, which can induce other clinical problems and occasionally death. MDMA also induces dose-dependent hyperthermia in experimental animals.



b i The average temperature of control rats was 38.5 °C. Calculate the average rectal temperature of the rats between 20 and 120 mins after administration of MDMA. **[2]**

ii Suggest a reason for the changes in the temperature of the rats' tails over this period. **[1]**

Option B Biotechnology and bioinformatics

Introduction

Microorganisms, or microbes – so called because of their small size – have a vital role to play in all ecosystems. They recycle waste, fix nitrogen and can be used by humans to make foodstuffs such as bread, cheese, yoghurt, tofu and beer. Many microorganisms are used in industry and genetically modified bacteria also produce human proteins such as insulin and growth hormone. It has been estimated that microbes account for almost half the biomass on Earth.

Although the majority of microbes are harmless, a few species do cause disease or serious illness. Influenza, polio and HIV are viral diseases and bacteria cause tuberculosis, cholera and leprosy.

B1 Microbiology: organisms in industry

Classifying microorganisms

Microbes are difficult to classify because of their small size and in the last 30 years our understanding of their similarities and differences has changed as new techniques have helped in our study of their structure and biochemistry.

In Subtopic **5.3**, you learned that all living organisms are classified into a number of kingdoms. In the past, it was commonly agreed by scientists that the organisms in these kingdoms could be separated into two groups, based on their structures:

- **prokaryotes**, with little cellular organisation and no organelles such as a nucleus
- **eukaryotes**, which contain organelles including a nucleus

New data from the sequencing of RNA has shown that the group known as prokaryotes should be split into two distinct groups – the Archaea and the Eubacteria. The eukaryotes still form a separate group, called the Eukarya. These groups are called **domains** (Figure **B.1**). The Archaea and Eubacteria are the groups containing most microorganisms and, as they have been studied further, details about other structures and molecules have been discovered that support this three-domain model.

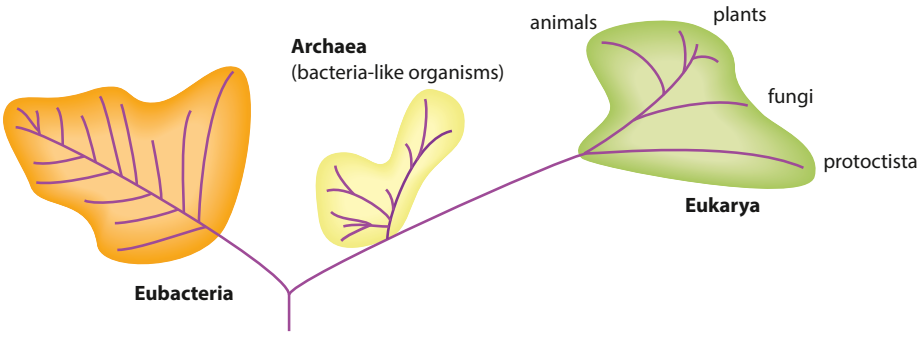


Figure B.1 A phylogenetic tree showing the grouping of organisms into three domains.

Learning objectives

You should understand that:

- There are many different types of microorganism, which have different metabolic reactions.
- Microorganisms are useful in industry because they are small and they grow and reproduce at a rapid rate.
- Pathway engineering is a way of optimising the regulatory and genetic processes that occur in microorganisms.
- In industrial processes, pathway engineering can be used to produce metabolites that are of interest.
- Useful metabolites can be produced on a large scale using fermenters.
- Growth of microorganisms in fermenters is carried out by batch or continuous culture.
- The growth of microorganisms in fermenters is limited by the accumulation of waste products.
- Conditions within fermenters are monitored by probes.
- Microorganisms are cultured by maintaining conditions at optimal levels for the growth of the organisms.

Diversity of metabolism in microorganisms

Microbes have a varied range of metabolic processes that enable them to use different sources of energy and carbon. As Table **B.1** shows, microbes are divided into four groups based on their methods of metabolism: **photoheterotrophs**, **chemoheterotrophs**, **photoautotrophs** and **chemoautotrophs**.

Method of metabolism	Energy source to generate ATP	Carbon source used to obtain organic compounds	Example
photoheterotrophic	light	organic compounds	*Heliobacter* is a photoheterotrophic bacterium found in waterlogged soils and paddy fields. It is able to fix nitrogen so it is probably important in soil fertility.
chemoheterotrophic	chemical reactions	organic compounds	Fungi such as the yeast *Saccharomyces* are chemoheterotrophic microbes. These organisms cannot photosynthesise, so they must use organic material as an energy source. They respire or ferment sugars to make ATP. Most bacteria are chemoheterotrophs.
photoautotrophic	light	inorganic carbon dioxide	An example of a photoautotroph is *Anabaena*, a cyanobacterium (filamentous blue–green bacterium) found among freshwater plankton and on grass. It fixes nitrogen and forms symbiotic relationships with some plants.
chemoautotrophic	chemical reactions	inorganic carbon dioxide	*Nitrobacter*, a nitrifying bacterium found in the soil, is an example of a chemoautotroph. Others include the sulfur-oxidising Archaea that live in hostile environments such as deep sea vents.

Table B.1 Different types of microbe metabolism.

Photoheterotroph microbe that uses light energy to generate ATP but which gets the organic compounds it needs from other organisms

Chemoheterotroph microbe that uses the chemical energy released from chemical reactions to generate ATP and obtains organic compounds from other organisms

Photoautotroph microbe that uses light energy to create ATP and to produce organic material from simple inorganic materials

Chemoautotroph microbe that uses energy released from chemical reactions to generate ATP and which makes its own organic material from simple inorganic materials

Diversity of Archaea and Eubacteria

The **Archaea** (Archaebacteria or ancient bacteria) inhabit some of the most extreme environmental conditions on the planet and those that reflect the conditions present in the early part of the Earth's existence. Three different groups of Archaea are identified from their metabolism and habitats.

- **Thermophilic bacteria** have evolved to survive at temperatures in excess of 70 °C and up to 100 °C in some cases. They inhabit hot, sulfurous springs in volcanic regions and hydrothermal vents on the ocean floor. One species, *Thermus aquaticus*, provides the enzyme DNA polymerase, vital for use in the polymerase chain reaction (PCR) for amplifying copies of DNA.
- **Halophilic bacteria** live in very salty environments like tidal mud flats and inland lakes (such as the Dead Sea) where the Sun has evaporated much of the water. They are also found in salt mines.
- **Methanogenic bacteria** are anaerobes found in the guts of ruminants (cows and sheep) and termites as well as in waste landfills, sewage works, paddy fields and marshland. They produce methane as a waste product of respiration. Methane is a major greenhouse gas.

Eubacteria evolved at the same time as the **Eukarya**, yet still possess primitive features. They have no internal organelle structure and often reproduce by simple binary fission. They are divided into groups based on their shapes, as shown in Figure **B.2**.

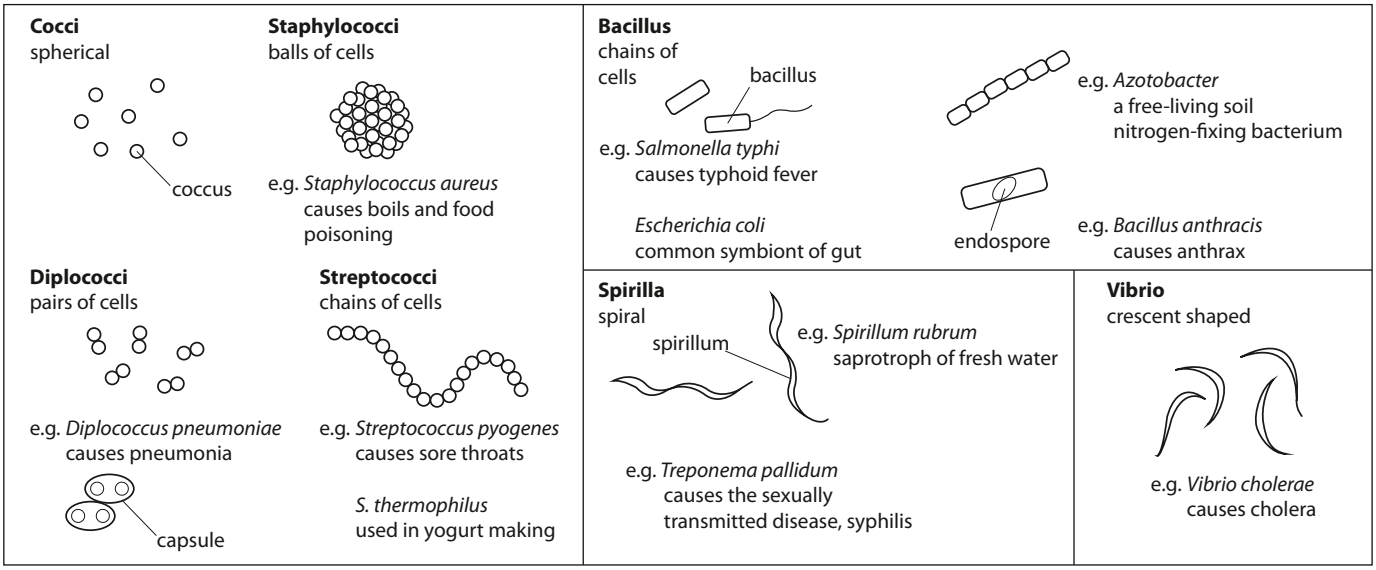


Figure B.2 Shapes of Eubacteria.

Exam tip
Remind yourself of the relationship between the three domains and the different groups of Archaeans by drawing a family tree or cladogram.

Gram-positive and Gram-negative Eubacteria

As well as being recognised by their shape (Figure **B.2**), bacteria can be separated into two main types by the structure of their cell walls. The Gram staining method is a useful way of differentiating these two types:

- Gram-negative bacteria do not retain the colour when the dye crystal violet is added.
- Gram-positive bacteria retain the dye and appear purple, even when washed in a decolourising solution.

These results are explained by the difference in structure of the cell walls – Gram-positive bacteria have large amounts of peptidoglycan, which retains the dye, in their cell walls whereas Gram-negative bacteria do not (Figure **B.3**). The main differences between Gram-positive and Gram-negative bacteria are outlined in Table **B.2**.

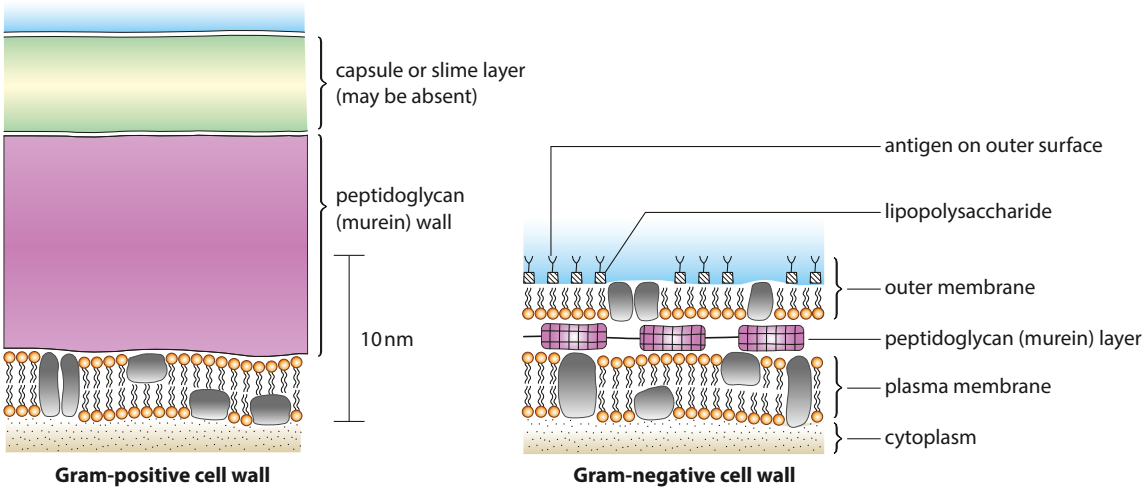


Figure B.3 Peptidoglycan consists of sugar molecules joined to polypeptides, which surround and protect the cell.

Gram-positive	Gram-negative
thick cell wall	thin cell wall
several layers of peptidoglycan connected by peptide bridges	layer of peptidoglycan sandwiched between inner and outer layer
no outer layer	outer layer contains lipopolysaccharide (LPS) and protein

Table B.2 Differences between Gram-positive and Gram-negative bacteria.

Diversity of microscopic eukaryotes

The microscopic eukaryotes are a diverse group of organisms. Using a microscope, it is possible to see the variety in their structure. They also vary in their methods of nutrition and of movement. Table **B.3** summarises the differences found in the group. The structures of the organisms listed in the table are shown in Figure **B.4**.

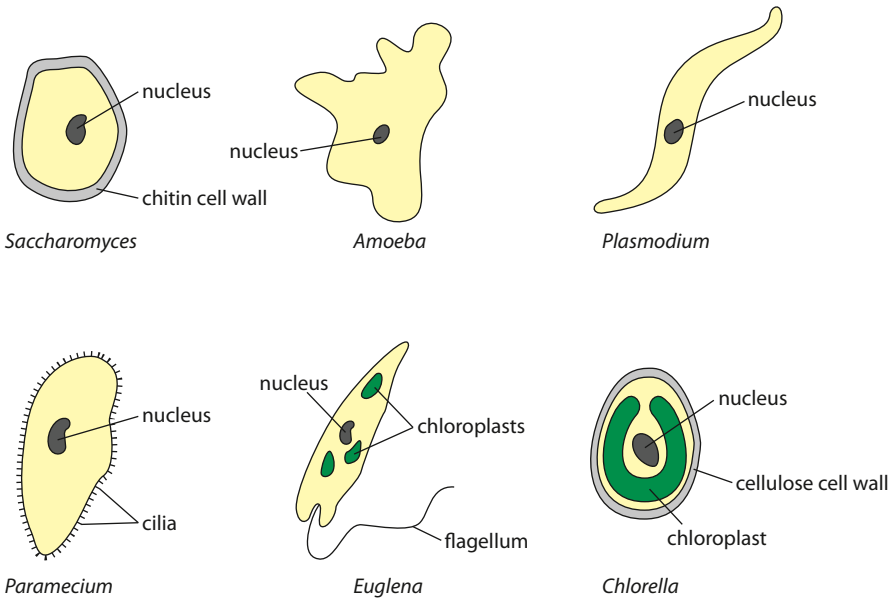


Figure B.4 Basic structures of the microscopic eukaryotic organisms listed in Table **B.3**.

Organism	Cell structure	Nutrition
Saccharomyces sp. (yeast)	cell wall made of chitin	heterotrophic, absorbs small molecules saprotrophically and feeds on sugars
Amoeba	no cell wall	heterotrophic, feeds on other organisms
Plasmodium (malarial parasite)	no cell wall	parasitic, some of its life cycle occurs in human cells
Paramecium	no cell wall	heterotrophic, takes in food by endocytosis
Euglena	no cell wall	autotrophic and heterotrophic, contains chlorophyll
Chlorella (single-celled green alga)	cell wall made of cellulose	autotrophic, photosynthesises using a chloroplast

Table B.3 Differences between types of eukaryotic organisms.

Microorganisms in industry

Microorganisms are used in commercial processes that produce foodstuffs, medicines and industrial products such as enzymes. Their fast growth rate and small size means that suitable amounts of product can be produced relatively quickly and efficiently.

Strict safety precautions are needed to ensure that products are pure and uncontaminated. This is very important where genetically engineered strains are used. Microorganisms may be cultured either on a solid substrate, which is a more traditional method used in cheese and sauerkraut production, or in an aqueous culture. Aqueous cultures may be produced by the batch culture method or the continuous culture method (see below).

Pathway engineering

Pathway engineering is a way of manipulating metabolic pathways in microorganisms so that particular metabolites of interest are produced in useful quantities. It is a developing new technology that has the potential to assist in many production methods based on biological organisms. It has been used in the production of both penicillin and citric acid to improve productivity, efficient use of carbon sources and product purity. Study of **genomics**, **proteomics**, **fluxomics** and **physiomics** are all used in pathway engineering, and can significantly increase the amount of product produced per cell.

Metabolic pathway engineers use four key strategies to achieve their results:

1. **optimisation of the primary metabolic pathway** by removal of factors that limit the rate of the reaction or transcription, as well as removing allosteric regulation, so that more product is produced
2. **genetic blocking** of any competing pathways
3. **maximising the amount of carbon substrate** that is directed to the metabolic pathway of interest and away from other pathways
4. **modification of secondary metabolic pathways** so that energy metabolism and availability of required enzymes and cofactors for the pathway of interest are maximised.

Two key examples of successful pathway engineering have been in the production of penicillin from *Penicillium* sp. bacteria and of citric acid from the fungus *Aspergillus niger*.

Knowledge of the **gene clusters** involved in penicillin production has enabled pathway engineers to increase output by the amplification of certain genes. This development first led to a 40% increase in output in the 1990s and later the introduction of structural genes and a whole gene cluster increased production still further.

Genetic improvement of *Aspergillus niger* used in industry was achieved by the use of mutagens, particularly gamma ray irradiation, so that citric acid yields increased more than three-fold. Further improvements were achieved by genetic engineering of the primary metabolism of *Aspergillus*. This increased the 'metabolic flux' or flow through the pathway forming citric acid, a product that would not normally accumulate. One method used is to decrease the flows through branches off the main pathway so that fewer by-products are formed. Another way is to engineer organisms so that they overproduce the enzymes needed for citric acid production so that the flux through the main pathway increases.

Genomics the study of the genome of an organism
Proteomics the study of the structure and functions of proteins
Fluxomics the study of the flow of fluids and molecules within cells
Physiomics the study of an organism's physiome, the interconnections of aspects of physiology that result from genes and proteins in the organism

Gene clusters sets of two or more genes that code for the same or similar products

Microorganisms naturally produce new genotypes as a result of mutations. Most mutations are harmful to the organism and are eliminated by natural selection but a few are beneficial. Others are not beneficial to the microorganism but are useful to humans. Microbiologists purposely apply **mutagens** (such as nitrosoguanidine, hydroxylamine, ultraviolet light and gamma rays) to different microorganisms, including *Aspergillus*, to modify their genomes – any useful mutations (for example, those that lead to an increase in the production of substances such as antibiotics) can then be detected by selection and screening processes and conserved in the genome. This has led to huge improvements in production in industrial microbiology since the 1960s.

Figure **B.5** shows some useful products that are made from glucose metabolism by fungi such as *Aspergillus* sp. and *Penicillium* sp.

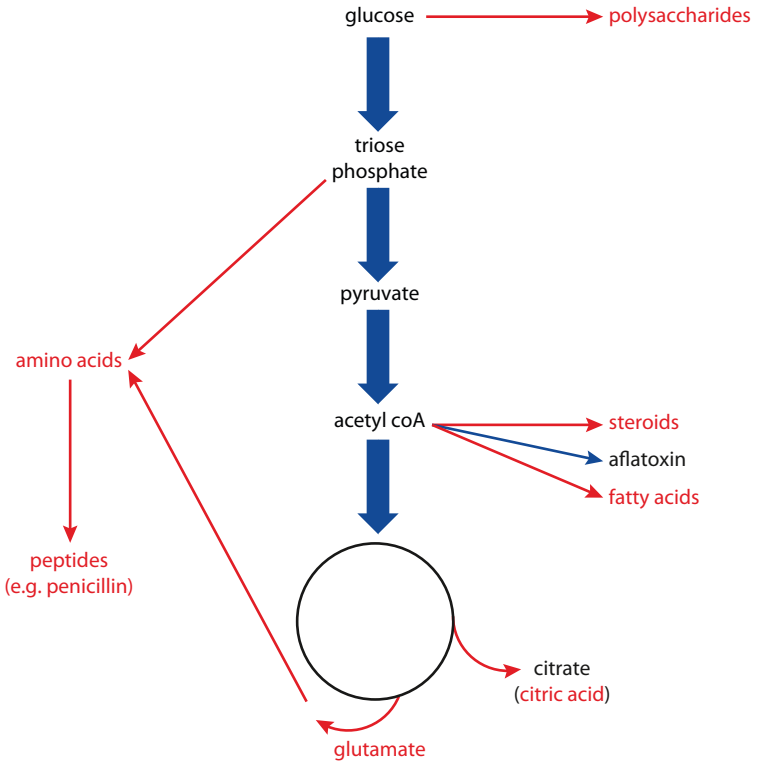


Figure B.5 Diagram to show some useful products of the metabolism of glucose by fungi. Useful substances are shown in red. Pathway engineers optimise genetic or regulatory processes in microorganisms to maximise production of these substances.

Large-scale production of metabolites of microorganisms

In order to produce and harvest the useful metabolites of microorganisms, a large-scale industrial process is needed. Most organisms are cultured in aqueous culture, a method in which the substrate has high water content. Two methods are commonly used: the batch culture method and the continuous culture method. In both cases, cells are grown in **fermenters** (Figure **B.6**), which are large tanks that can hold up to 200 000 litres. Cells are provided with nutrients and the exact environmental conditions they need to grow at their maximum rate. If the microorganisms used are aerobic, air is pumped in, and the culture is also stirred to ensure even mixing of nutrients throughout the culture. Growth in a fermenter can be limited by accumulation of waste products but in a continuous culture these can be removed.

Nutrients enter the fermenter through valve-operated pipes so that exact quantities can be controlled to maximise growth. All fermenters must be sterile before the process begins and this is usually achieved by passing steam through the pipes and tanks. People who work in the biotechnology industry must also take precautions and wear protective clothing to avoid contamination of the process or of themselves.

Although the culture vessels are known as fermenters, the metabolic process that occurs in them does not have to be fermentation (that is, anaerobic respiration).

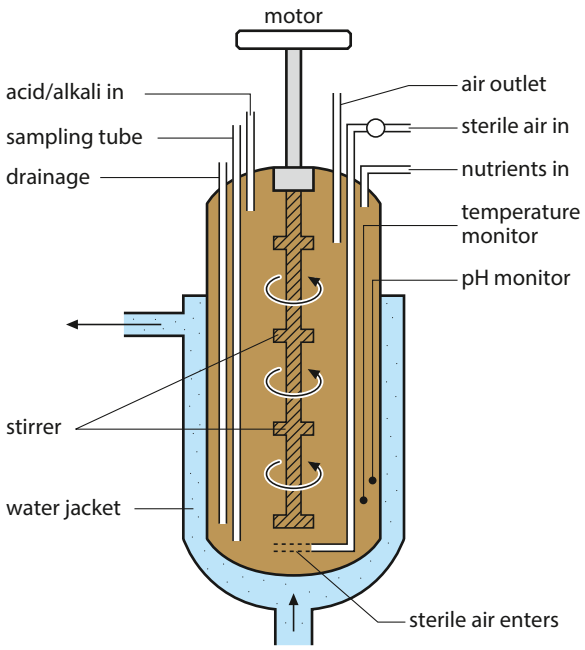


Figure B.6 The inputs and outputs of a typical fermenter.

Batch culture

Batch culture is used in the mass production of penicillin, and is carried out in a closed fermenter. The microorganism (*Penicillium* sp.), from a culture whose genome is known and stable, is added to the fermenter with the necessary nutrients, which have been pre-sterilised. The fermenter is then left for the process to take place. As the microorganisms grow and reproduce, nothing is added to or removed from the fermenter except waste gases, which are allowed to escape. As the curve in Figure **B.7** shows, the microorganisms multiply rapidly at first, but this exponential phase of growth is usually quite short.

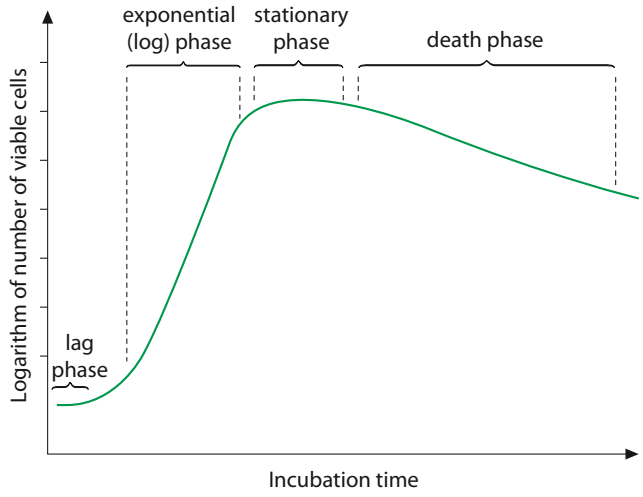


Figure B.7 Growth curve of microorganism in batch culture.

Secondary metabolites substances produced by microorganisms after the main growth phase has been completed, and which are not essential for normal cellular functions; penicillin produced by *Penicillium* sp. is an example of a secondary metabolite

As a batch culture proceeds, the temperature can rise by 1 °C per hour, and because penicillin is a protein and can easily be denatured by high temperatures, the fermenter must be cooled. Cooling coils or a water jacket may be used in some fermenters. Although the temperature is kept constant, other conditions in the fermenter change – up to 0.5 tonnes of glucose per day may be used and waste products begin to accumulate. Penicillin is a **secondary metabolite** so eventually the amount of product made will start to decline as the reproduction of the microorganisms slows down. At this point the fermenter will be shut down for cleaning and restocking.

Continuous culture

Continuous culture is used in the manufacture of citric acid by the fungus *Aspergillus niger*. Citric acid is difficult to produce from inorganic chemicals but is quickly and easily made by microorganisms. In continuous culture, nutrients are supplied continuously to the fermenter at a steady rate (usually the nutrients are molasses or hydrolysed corn starch, which are cheap sources of sugar). Continuous culture matches supply and demand so that the organisms can be kept in an exponential phase of growth. The conditions in the fermenter are carefully monitored by probes that check pH, nutrient concentration, oxygen levels and product levels. Used medium is drained off, together with the citric acid product. The fungi are filtered out of the liquid that leaves the fermenter and citric acid is extracted by precipitation with calcium hydroxide.

A continuous culture method means that an uninterrupted supply of product is made and the equipment can be used without breaks for cleaning and recharging the fermenter. Continuous culture can therefore be more economical than a batch culture method. Continuous and batch culture methods are compared in Table **B.4**.

Batch culture	Continuous culture
product is separated from the mixture at the end	product is harvested continuously
conditions inside the fermenter change (although temperature is monitored)	all environmental factors are monitored and kept as constant as possible (sometimes this is difficult and production can be disrupted)
cells have a relatively short time in the exponential growth phase	cells are kept in the exponential growth phase
larger 'deep tank' fermenters are needed	smaller fermenters can be used – process is more productive
usually less cost effective, but if a batch is contaminated only one batch is lost	continuous process is more economical, but if a fermenter is contaminated losses are greater

Table B.4 Comparing batch and continuous culture methods.

Uses of citric acid

Large volumes of citric acid are produced by fermentation every year – estimates put production at around a million tonnes per year. Citric acid is a natural preservative and is also used to add an acid taste to foods and soft drinks (in Europe it is classified as food additive number E330).

Citric acid is also used:

- as an emulsifying agent to prevent fats from separating in ice cream and other products
- with sodium bicarbonate in effervescent indigestion tablets
- as a grease remover in many cleaning products
- as a pH adjuster in creams and gels.

Biogas production

Biomass in the form of wood and agricultural waste, such as straw and animal manure, already provides a useful source of fuel. Now many countries are looking at the use of **biofuels** to reduce their dependence on fossil fuels. Biofuels are produced by converting biomass into ethanol or methane. This is done in **bioreactors**, either on a large industrial scale or in small fermenters on farms or in villages. Methane produced from animal manure and agricultural waste is known as **biogas**. A simple, small-scale fermenter that can be used to produce biogas is illustrated in Figure **B.8**. Manure and straw are fed into the bioreactor, where they decompose anaerobically as different groups of bacteria present in the manure break down the organic material. The slurry that remains is a useful fertiliser.

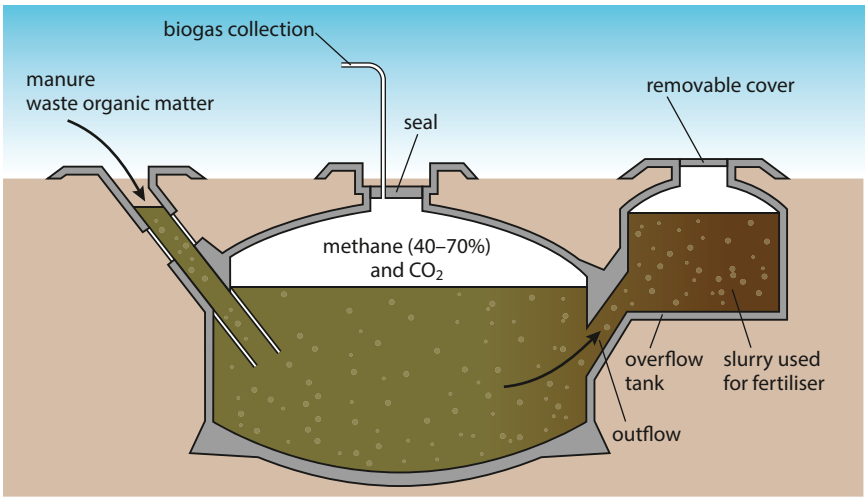


Figure B.8 Cross-section of a biogas reactor.

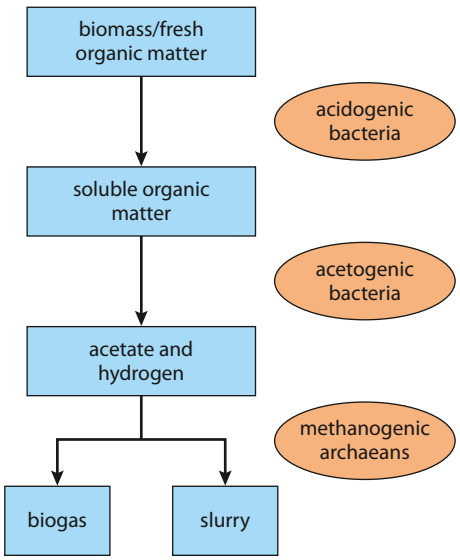


Figure B.9 Processes in biogas production. Three groups of microorganisms digest organic waste to produce methane

Three groups of microorganism produce the enzymes that digest organic material, and each stage breaks down the complex carbohydrates, fats and proteins into simpler compounds.

- Organic material is first converted to organic acids and ethanol by anaerobic, **acidogenic bacteria**, which occur naturally in manure.
- **Acetogenic bacteria** then use the organic acids and ethanol to produce acetate (ethanoate), carbon dioxide and hydrogen.
- Finally, **methanogenic archaeans** produce methane either from carbon dioxide and hydrogen or by breaking down acetate:

carbon dioxide + hydrogen → methane + water

$$CO_2 + 4H_2 \rightarrow CH_4 + 2H_2O$$

acetate (ethanoate) → methane + carbon dioxide

$$CH_3COOH \rightarrow CH_4 + CO_2$$

The sequence of these processes is shown in Figure **B.9**.

The biogas that is produced contains up to 70% methane and about 30% carbon dioxide. Biogas can be used to produce electricity or burned directly as a renewable fuel, and the by-products of the reactions can be used as fertiliser.

The production of biofuels has the advantage of being **sustainable** because plants regrow each season. In addition, it makes use of the methane gas that is naturally produced by the anaerobic digestion of organic matter. Methane is a potent greenhouse gas, partly responsible for global warming. On a small scale, biofuels produced at a local level have many advantages but some people have voiced concerns that problems may arise if large areas of land are used to grow crops for biofuels rather than food.

Worldwide, a range of fuels is now produced from biomass on a large scale. Bioethanol can be used as a fuel for vehicles in its pure form, but it is usually used as a petrol additive to increase a car's performance and improve vehicle emissions. Bioethanol is widely used in the USA and in Brazil. Biodiesel made from oils and fats is most often used as a diesel additive to reduce levels of emissions, and is the most common biofuel used in Europe. In 2008, biofuels accounted for almost 2% of the world's transport fuel.

Nature of science

Serendipity in science – the discovery of penicillin

The discovery of antibiotics began by accident. On 3 September 1928, Professor Alexander Fleming (1881–1955) was examining a batch of culture plates on which he had grown *Staphylococcus* bacteria. He noticed that one of the plates had a green mould growing on it. The mould was *Penicillium notatum*. The mould was circular in shape, and the area around it seemed to be free of *Staphylococcus*. On other areas of the plate, the bacteria were continuing to grow well. This chance event led Fleming to consider the possibility that the bacteria around the circular mould had been killed off by a substance produced by the mould. He discovered that the mould could kill other bacteria and that it could be given to small animals without any harmful effects. However, he moved onto other research and it was not until 10 years later that Howard Florey (1898–1968) and Ernst Chain (1906–1979), working at Oxford University, isolated the bacteria-killing substance, penicillin, produced by the mould. Chain was a German chemist and Florey an Australian pathologist. It was Chain who isolated and purified penicillin and Florey who tested its safety to use on animals and humans. In 1945 Fleming, Chain and Florey shared the Nobel Prize for Medicine.

Luck or judgment?

Alexander Fleming noticed that one agar plate he was about to discard was contaminated by a mould. He was surprised to see that the mould had not only stopped the bacterial growth but had killed the bacteria.

Questions to consider

- To what extent was Fleming's discovery a lucky observation?
- Are we more likely to perceive those events that we are open to?

Test yourself

1 Outline the differences between batch culture and continuous culture of microorganisms.
2 State the name of the organism used in the commercial production of citric acid.
3 Outline the modes of nutrition of *Saccharomyces*, *Euglena* and *Chlorella*

B2 Biotechnology in agriculture

Learning objectives

You should understand that:

- Transgenic organisms have been created to produce proteins that were not part of their species proteome in the past.
- Genetic modification can overcome environmental resistance and therefore increase crop yields.
- Genetically modified crops can be used to produce new products.
- Bioinformatics has helped in identifying target genes.
- A target gene is attached to other sequences, which control its expression.
- An 'open reading frame' is defined as a significant length of DNA beginning at a start codon and ending with a stop codon.
- To create genetically modified plants, recombinant DNA must be inserted into cells and be taken up by chromosomes or chloroplast DNA.
- Successful uptake of introduced DNA is indicated by the use of marker genes.
- Recombinant DNA can be introduced into protoplasts, leaf discs or whole plants.
- Recombinant DNA can be introduced either indirectly by vectors or directly by physical or chemical methods.

Genetic modification in agriculture

Transgenic describes an organism that contains genes from another organism

Gene technology enables scientists to transfer genes from one species to another completely different species and create a **transgenic** organism in just one generation. For example, bacterial genes have been transferred to plants, human genes transferred to bacteria and spider genes transferred to a goat. Gene transfer is possible because the genetic code is universal – no matter what the species, the genetic code spells out the same information and produces an amino acid sequence in one species that is exactly the same in any other species.

Through gene transfer, transgenic organisms can be created that are able to produce proteins that were not previously part of their species' proteome. Animals have been farmed for the products of such inserted genes – for example, Factor XI, a clotting factor used in the treatment of hemophilia, can be produced in the milk of genetically modified sheep and alpha 1-antitrypsin, a protein whose deficiency leads to the breakdown of lung tissue, has also been produced in the same way. Research continues on transgenic goats, bred to produce a growth hormone that may bolster the immune systems of transplant recipients and reduce the side effects of chemotherapy.

Genetic modification in crop plants

By the first decade of the twenty-first century, almost 100 plant species had been genetically modified and many trials have taken place to assess their usefulness. In comparison, there are very few examples of genetically modified animal species. Most genetic engineering has involved commercial crops such as soybeans, maize, potatoes, tomatoes and cotton. Plants have been modified to make them resistant to pests and disease, tolerant to herbicides and able to produce novel products. Some have been modified to extend their ranges so that they are able to tolerate drier conditions and increase crop yields in these areas. Three examples of important genetically modified crops are described below.

Glyphosate-resistant soybeans

Herbicides are used to kill weeds in crop fields but they are expensive and can affect local ecosystems as well as cultivated areas. One commonly sprayed and very powerful herbicide is glyphosate, which is rapidly broken down by soil bacteria. For maximum crop protection, farmers needed to spray several times a year. But now, the genes from soil bacteria have been successfully transferred into soybean plants making them resistant to the herbicide.

Farmers can plant the modified seeds, which germinate along with the competing weeds. Spraying once with glyphosate kills the weeds and leaves the soybean plants unaffected. The crops then grow and out-compete any weeds that grow later when the glyphosate has broken down in the soil. Yields are improved and less herbicide has to be used.

Glyphosate-resistance genes are introduced into soybeans using a vector, the bacterium *Agrobacterium tumefaciens*. This bacterium causes tumours known as galls when it infects plants – it has a large tumour-inducing plasmid known as Ti, which it incorporates into the DNA of infected plants. This Ti plasmid has been used to introduce glyphosate-resistance genes into the genomes of both soybeans and maize (Figure **B.10**), both of which are vital agricultural crop plants. The genes conferring glyphosate resistance are transferred from plant cells of resistant species into the Ti plasmid using **restriction endonucleases** to open the plasmid. These recombinant plasmids are reintroduced into *A. tumefaciens* cells, which then infect plant cells susceptible to glyphosate. Plant cells with included resistance are grown in tissue culture and whole crop plants have been produced from them.

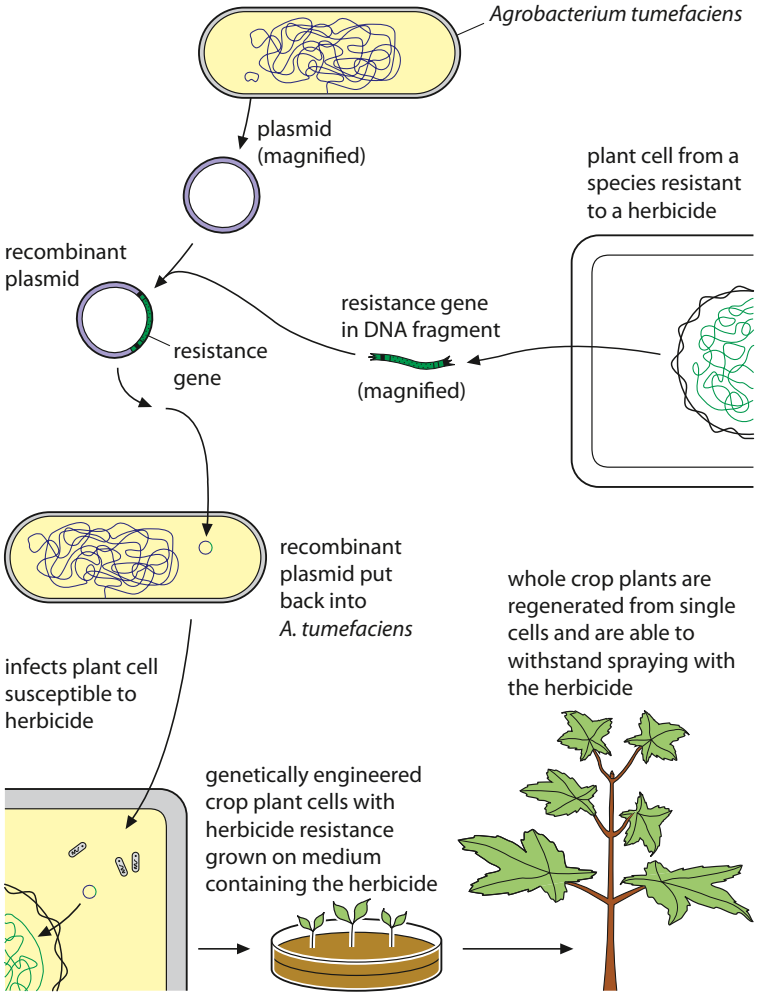


Figure B.10 Transfer of glyphosate-resistance genes by Ti plasmid.

Environmental impact of glyphosate-tolerant soybeans

Transgenic glyphosate-resistant soybeans are grown in many countries, where scientists and environmentalists closely monitor their effects on the environment. The number of transgenic soybean plants grown in the USA increased from approximately one third of all soybean plants in 1998 to more than a half, 2 years later. Studies showed that in the USA, in some areas, there was a 40% decrease in the use of herbicides on the glyphosphate-resistant beans compared with non-transgenic beans, between 1995 when the plants were introduced and 1998. The use of glyphosate-resistant soybeans has also meant that farmers have to dig and turn the soil less frequently, so environmental degradation due to agriculture has been reduced. Some groups have raised concerns that glyphosate-resistance genes in soybeans could be transferred to wild species, but this risk has been found to be low because soybean plants are almost completely self-pollinated. Furthermore, in many countries there are no wild relatives of soybean plants so transmission to other species is very unlikely. One of the possible negative impacts of the soybeans is weed resistance. Weeds that can resist glyphosate will become dominant, as will those that spring up at times when the glyphosate spray is not applied, and thus avoid its application.

Exam tip
If you are asked to discuss a subject, make sure you consider more than one point of view in your answer

Figure B.11 Tobacco leaves infected with TMV.

Producing vaccines in tobacco plants

Tobacco mosaic virus (TMV) is an RNA virus that infects plants. It has recently been adapted for vaccine development, because its simple structure – a strand of RNA encapsulated by just one type of coat protein – means that its genome can be easily manipulated. The virus has been modified to carry genes for several proteins, including those which induce antibody and T-cell responses in animals and so can be used to provide protection against a number of viral diseases such as flu and hepatitis B. This method has already produced a vaccine against Newcastle disease, a viral disease in chickens, and has been extensively investigated for use in the rapid production of flu vaccine in case of a pandemic among humans. Producing vaccines using plants infected with manipulated viruses is known as 'molecular farming'.

As its name suggests, TMV infects tobacco plants, which are fast-growing and have large leaves (Figure **B.11**) – they produce large quantities of biomass in a short period of time so are ideal candidates for vaccine production.

Vaccine production

Vaccines are made by first introducing multiple copies of the genetic information needed to produce the target proteins into the TMV viruses. This genetic information must be introduced into tobacco plants and the TMV is the ideal 'launch vector' to do this. Viruses can only reproduce inside living cells because they do not possess the molecular machinery to do so independently. So as the virus infects a tobacco plant, its genetic material is incorporated into that of the plant cells (Figure **B.12**), which then produce both the normal viral proteins (to make new viral units), and the proteins coded for by the introduced genetic material.

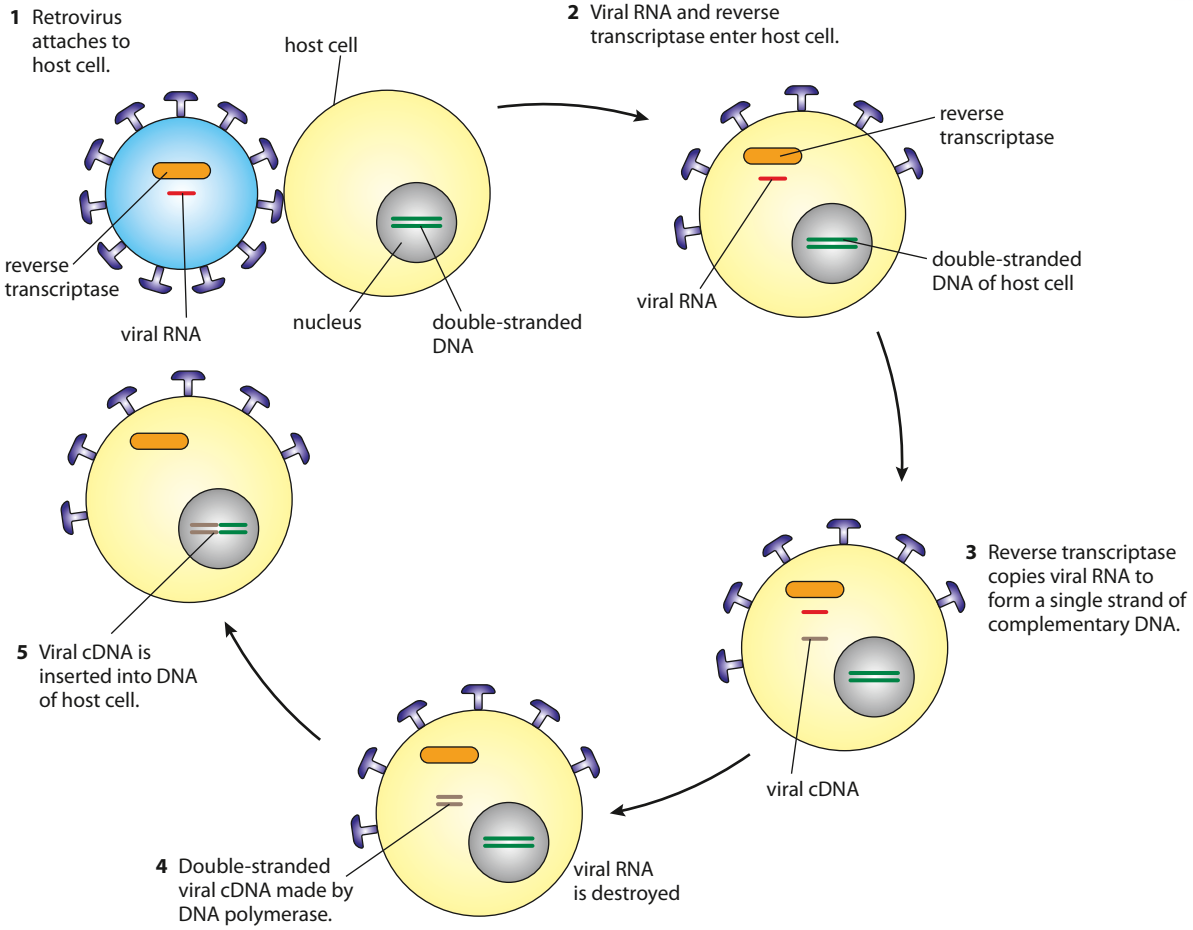


Figure B.12 How the genetic information from a retrovirus becomes part of the DNA of a host cell.

The proteins can then be harvested from the plants to make vaccines. Plants have the potential to make vaccines against any virus in this way, producing high yields of the precise protein that is needed.

The plants can be grown in **hydroponic culture** with carefully controlled conditions to maximise their growth. Although the viruses can infect tobacco plants naturally (through damaged leaves, for example), virus **vectors** are often introduced into mature plants artificially in order to speed up the production process. For example, in some production processes, tobacco plants are grown for 4 weeks before the virus vector is introduced by means of a new technique known as vacuum infiltration. To do this, the plants are turned upside down, and submerged in water containing TMV vectors. A vacuum is then created, drawing air away from the water and plants. As the vacuum is switched off, the water containing the virus vectors is 'sucked' into the plants. The plants are left to grow and in about 7 days they will have produced the target proteins in their leaves and stalks. The plants are harvested, the leaves are cut into small pieces and liquefied, and the proteins are extracted from the liquid.

In the past, vaccines against viral diseases such as influenza and hepatitis have been grown in birds' eggs. But tobacco-based plant vaccines have several advantages:

- vaccines can be produced more quickly
- there are no known allergies to plant-produced vaccines, unlike those from birds' eggs
- large tobacco leaves make harvesting easy
- many doses of vaccine can be made from a relatively low biomass.

Reverse transcriptase and its use in molecular biology

The enzyme reverse transcriptase was discovered in 1970 in a group of viruses known as retroviruses, which includes HIV, TMV and feline leukaemia virus. These viruses contain RNA as their genetic material. Reverse transcriptase enters along with the viral RNA when a retrovirus invades a host cell. It is used to transcribe the virus RNA into a single strand of DNA, using nucleotides from the host cell. The new complementary DNA (cDNA) is then converted to double-stranded DNA using the enzyme DNA polymerase. The original RNA is degraded and the double-stranded DNA is inserted into the host's chromosomes (Figure **B.12**).

Reverse transcriptase is widely used in genetic engineering. Molecular biologists are able to produce therapeutic proteins such as insulin and growth hormone by inserting the genes that code for them into the genetic material of bacteria or other organisms. These organisms then produce large amounts of the required protein as they grow and multiply.

Paper production and potatoes

Potato tubers store starch which, as well as being a good source of food, has a wide variety of uses in the manufacture of textiles, paper and adhesives. Potatoes (*Solanum tuberosum*) produce large amounts of starch and give higher yields per hectare than either corn or wheat. Potato starch consists of approximately 80% amylopectin and 20% amylose (Subtopic **2.3**) but most of the useful properties for industry, such as adhesion, come from amylopectin. In conventional manufacturing processes amylose is not needed and causes production problems because it forms a gel, which makes potato starch unstable.

Recently, researchers at BASF Plant Science succeeded in deactivating the genes for amylose in the potato genome and developed a new variety of genetically modified potato known as Amflora. The Amflora potato produces only amylopectin and so it is ideal for use in the paper-making and the adhesive industries. It enables concrete to stick to walls more effectively, keeps glue liquid for longer and improves the 'glossiness' of paper (Figure **B.13**).

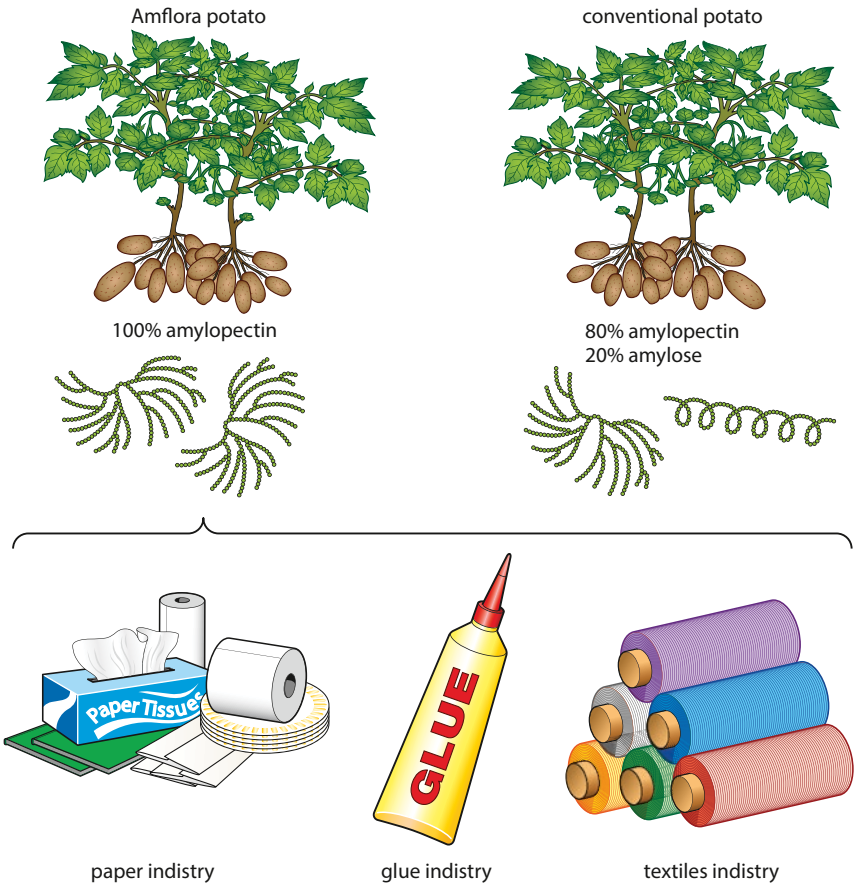


Figure B.13 Amflora potatoes and their uses.

Politics and GM

The modified potato received approval from the European Union in 2010 after an extensive period of review and testing. Amflora potatoes for planting cannot be sold through normal channels but are supplied directly to farmers who have contracts with BASF and who plant the Amflora potatoes at a fixed distance from other potato crops. When harvested, Amflora are taken to separate processing factories. Despite these precautions, and the fact that the European Food Safety Authority (EFSA) has stated that there are no risks to humans or animals from Amflora, the potatoes struggled to gain acceptance – indeed, by 2011, they were only being grown on one small site in Germany. This lead to the company, BASF, deciding to move its biotechnology headquarters to the USA – it has stopped production of genetically modified (GM) products in Europe, citing lack of political support for its work as the main reason. A strain of GM maize, developed by another biotechnology company, Monsanto, is now the only crop approved for use in the EU, apart from the Amflora potato. Other GM crops including maize, rice and tomatoes are widely grown in South America, the USA and Asia.

Support for science

Following the withdrawal of BASF from biotechnology production in Europe, a BBC report said '*No one from the political side supported it. There were no signals from the European Commission that any change was likely*.' Some reports indicated that a new EU health commissioner might freeze approval on new GM crops, while others said that he was trying to clarify the regulations on growing GM crops in Europe.

Question to consider

- To what extent should political and public pressure influence the advance and development of new scientific discoveries?

How recombinant DNA is transferred into plants

In gene technology, manipulated DNA containing genes for the protein or proteins of interest is called **recombinant DNA** (Subtopic **3.5**). The first stage in any gene technology process, such as that used for 'molecular farming' in tobacco plants described above, is to find and isolate the **target genes** for the protein to be produced. These genes are also linked to other genes, which may control the way they are expressed. **Bioinformatics**, which is the use of databases and accessible stores of information, has been crucial in the easy identification of DNA sequences coding for these desired proteins. The most useful sequences are known as 'open reading frames' – these, combined with **marker genes**, can be inserted into target organisms. The newly created transgenic organisms can then produce the novel products.

Open reading frames

An **open reading frame** (ORF) is a sequence of nucleotides in a DNA molecule that has no codons that terminate transcription within it. The site that terminates transcription is found after the ORF and beyond the translation stop codon. This means that, once transcribed and translated, the ORF has the potential to produce a complete polypeptide chain.

ORFs are very useful in helping to predict which sections of a DNA molecule are likely to be genes. Long ORFs can be used to identify regions of DNA likely to code for proteins. Researchers search for a start codon followed by an ORF that is long enough to code for a typical protein in the particular organism they are investigating. But even if ORFs are found, this does not necessarily mean that these regions of DNA are genes that are transcribed and translated. To check for an ORF's potential to be transcribed and translated, an Open Reading Frame Finder (ORF Finder) can be used. This is a graphical analysis tool used by researchers, which finds all open reading frames of a certain, chosen size in the DNA sample being studied. The ORFs may either be in a new sequence that is being investigated or in a sequence that has already been stored in a database. The Finder tool can deduce the amino acid sequences coded by lengths of DNA and these can be compared with known amino acid sequences using the BLAST server. BLASTn is a database that can identify similar nucleotide sequences in different organisms while BLASTp can match amino acid sequences in proteins. This is described in more detail in Subtopic **B.5**.

Introducing genes into plants

In order to modify plants, new DNA sequences must be introduced into their cells and taken up either by the plant's chromosomes or by the chloroplast DNA. Recombinant DNA can be introduced into whole plants, but is often introduced into leaf discs kept in tissue culture or into protoplasts (cells that have had their cell walls removed). When DNA has been taken up by plants, leaf discs or protoplasts they can be grown up to make new transgenic plants. As well as viral vectors, described above, several other methods are used to introduce recombinant DNA into plant cells.

Chemical methods

Sequences of DNA or bacterial plasmids can be encapsulated into tiny, neutral structures known as **liposomes**. Liposomes are vesicles, which have a phospholipid bilayer around them. Encapsulation is achieved by adding a mixture of ethanol and calcium chloride to a mixture of vesicles and plasmids. Liposomes provide an efficient method of transferring DNA to new cells. Cells are 'transfected' by the liposomes, which can actually fuse with the cell membrane of a target cell and release the DNA or plasmid directly into it (Figure **B.14**). In some of the transfected cells the DNA will be taken up by the cell's genome.

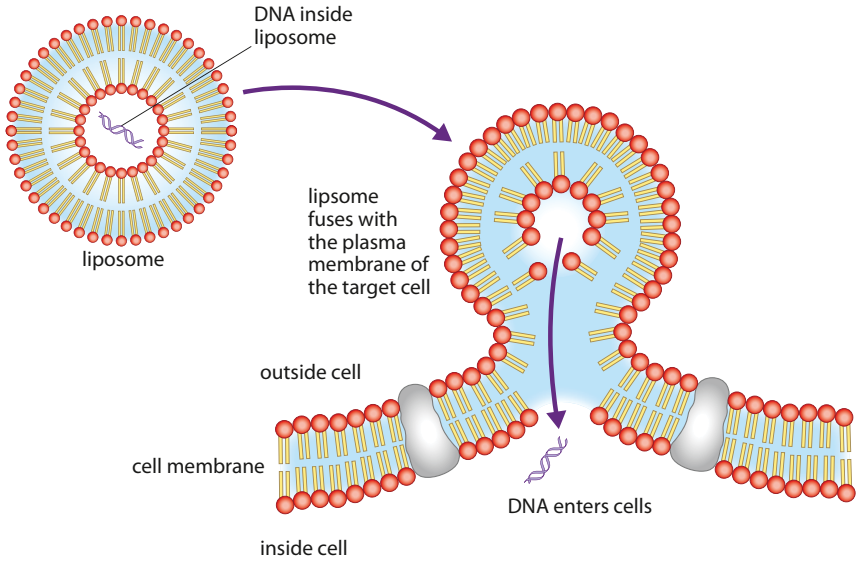


Figure B.14 Liposome structure and transinfection.

An alternative, but less efficient, method uses calcium phosphate and calcium chloride solutions. A solution containing phosphate ions is combined with calcium chloride solution containing the DNA to be transferred. As the two solutions are mixed, a precipitate of positively charged calcium and negatively charged phosphate forms. The DNA that is to be transferred binds to its surface. The transfer is completed by adding a suspension of the precipitate to plant cells grown in a single layer in tissue culture. The cells take some of the precipitate and DNA into their cells. The exact mechanisms of this process are not fully understood.

Physical methods

Three important physical methods of introducing genes into plant cells are electroporation, microinjection and biolistics (the so-called 'gunshot' method).

- **Electroporation** exposes cells to a strong electric field for a very short period of time. This has the effect of changing the permeability of cell membranes for sufficient time for new DNA to enter the cells.
- **Microinjection** involves using a micropipette to inject DNA into a living cell. This is usually carried out under a microscope, and is the same technique as that used to transfer material into animal eggs during *in vitro* fertilisation. New variations of this technique have been developed to bind DNA to large number of nanofibres, which can be inserted into cells or tissues in an automated process.

- **Biolistic** methods involve a 'gene gun', which shoots DNA attached to a nanoparticle of an unreactive material, often gold, directly into the nucleus of the target cell (Figure **B.15**).

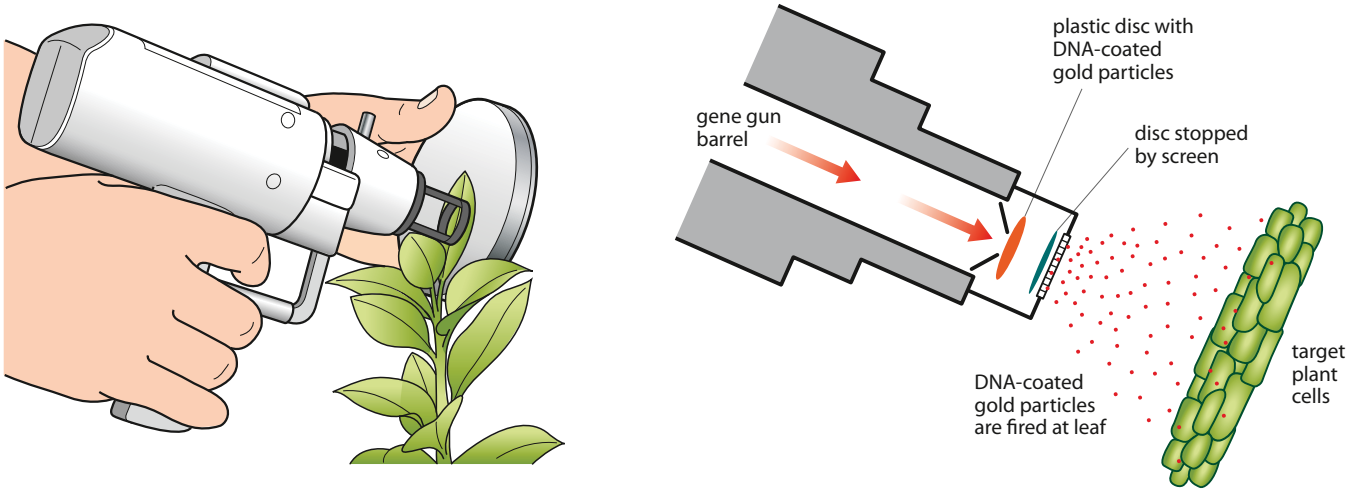


Figure B.15 A gene gun and its method of working.

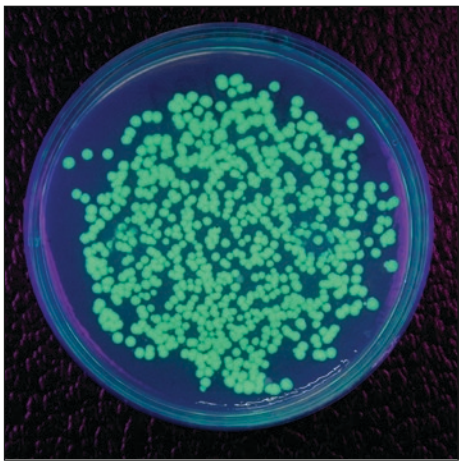


Figure B.16 Transformed bacteria colonies containing a jellyfish gene for GFP protein causing green bioluminescence. This gene can be used as a marker to indicate the successful uptake of a target gene.

Exam tip
Draw a flow diagram to summarise the ways in which DNA is introduced into one organism from another.

Marker genes

Marker genes are linked to sequences of DNA that are being transferred to new organisms so that researchers can check that the insertion has been successful. There are two types of marker genes: screening markers and selectable markers. **Screening markers** identify cells that have taken up new genes by their appearance. The three most common screening markers are:

- a green fluorescent protein (GFP), which makes cells glow under UV light (Figure **B.16**)
- the so-called GUS assay, which stains cells blue and is regularly used in experimental plant science but has limited use in applied work because it kills cells
- the blue–white method, used as both a plant and bacterial marker, which involves adding a bacterial gene coding for beta galactosidase enzyme so that if galactosides are added to a culture medium, cells containing the gene convert it to a blue substance, which can be seen easily.

Selectable markers (usually antibiotic resistance markers or ARMs) protect the cells of organisms that have taken up new DNA from substances that would otherwise kill them. Because very few cells take up new DNA, in most cases it is easier to kill those that have not and use the selectable marker to protect the remainder. Antibiotic markers are often used to kill chloroplasts during the genetic modification of plant cells. Any plant cells that have not taken up the new DNA after they have been treated, and therefore do not carry the resistance marker, can be targeted with suitable antibiotics, which will destroy them.

Nature of science

Assessing risk in science – concerns about genetic modification

Genetic modification of plants and animals has the potential to be very helpful to the human race, but risks and benefits must be evaluated. One of the important risks is the potential of herbicide resistance genes, such as the genes for glyphosate resistance in GM soybeans, escaping into wild populations of plants. If wild species should become resistant to herbicides it might be less easy to clear fields of weeds with the herbicides that we have today. But there are many sides to the discussion – some of the key points to consider are as follows.

- Modifying crop plants and animals to increase yields will provide more food for the growing human population. Plants can also be made tolerant to drought or salt water so that food can be grown in difficult areas.
- Crop plants that are disease resistant not only increase yields but also reduce the need to apply harmful pesticides, which have impacts on the wider ecosystem.
- Many substances – such as human growth hormone, a blood-clotting factor, antibodies, and vitamins – are already being made by genetically modified organisms to improve human health.

On the other hand:

- There is concern that there may be dangers in consuming products from genetically modified plants and animals.
- The long-term effects of genetically modified crops in the environment are not known. Plants or animals could 'escape' into the environment and their genes might become incorporated into wild populations, with unforeseeable effects.
- Human food crops could become controlled by a small number of biotechnology companies, which could make seeds and plants more expensive.
- More genetically modified organisms might lead to a reduction in natural biodiversity.

You can read how pollen from GM maize was alleged to have affected monarch butterflies in Subtopic **3.5**.

Test yourself

4 Define the term 'transgenic organism'.

5 List **three** methods of introducing new DNA into plant cells.

6 Outline the meaning of the term 'open reading frame'.

7 State the key difference between the genetically modified Amflora potato and a normal potato.

B3 Environmental protection

Learning objectives

You should understand that:

- Bioremediation together with physical and chemical procedures can be used in response to pollution incidents.
- Microorganisms are used in bioremediation.
- Microorganisms can metabolise some pollutants.
- Biofilms are cooperative aggregates of microorganisms.
- Biofilms show emergent properties.
- Microorganisms growing in biofilms are highly resistant to antimicrobial agents.
- In biofilms, microorganisms are able to cooperate through quorum sensing.
- Bacteriophages are used to disinfect water systems.

Exam tip

If you are asked to write an essay on bioremediation of an oil spill, make a checklist of all the different things that would need to be done. Don't forget physical and chemical cleaning.

Bioremediation

Bioremediation is the process that uses microbes to treat areas of land or sea that have been contaminated by pesticides, oil or solvents. Bioremediation is usually carried out in conjunction with physical and chemical procedures, such as clearing away contaminated soil or scraping away oil, and treatment with detergents sprays or remedial chemicals to neutralise contamination.

Cleaning up pollution at sea

Crude oil spills from tankers at sea contain many chemicals that harm the marine environment and seashore. Many different microbes are able to oxidise harmful hydrocarbons and break down the oil. To increase the numbers of bacteria and speed up the bioremediation process, nitrate and phosphate fertilisers are sometimes added, to encourage decomposition of the crude oil by the bacteria. Bioremediation like this can halve the time it takes to clean up an oil spill. A number of strains and species of bacteria in the genus *Marinobacter* have been found to be important degraders of hydrocarbons in the sea. These bacteria live in habitats that sometimes have extremes of pH or salt concentrations, and because they are **halophilic** (salt loving) they tolerate marine conditions well. *Marinobacter hydrocarbonoclastus*, for example, is able to degrade petroleum hydrocarbons, such as benzene, which are released during accidental spillage of oil at sea. In 2011, it was discovered that *M. hydrocarbonoclastus* is inhibited by the presence of certain chemicals used as dispersants in cleaning up oil spills, suggesting that dispersants should not be used at the same time as this bacterium in bioremediation treatments (Figure **B.17**).

Figure B.17 Workers using hoses and dispersants to clean up an oil spill. Bacteria may be able to do the job more efficiently in future. Experts from different parts of the world are brought together to deal with serious oil spills. Each one may have experience in a different aspect of the clean up.

Cleaning up pollution on land

On land, bioremediation has been used to remove pesticide residues, crude oil spills, heavy metals and solvents from contaminated soil. There are approximately 20 species of bacteria with the ability to metabolise petroleum hydrocarbons found in oil, which include members of the genera *Pseudomonas*, *Aeromonas*, *Bacillus*, *Flavobacterium* and *Micrococcus*. One of the most efficient hydrocarbon users on land is *Pseudomonas aeruginosa*, which can use crude oil as its sole source of carbon. It can also tolerate and thrive in high concentrations of oil, and if surfactants are used in combination with the bacteria they can degrade the oil more quickly. Bioremediation often happens very slowly when bacteria in the environment break down other toxic substances. To speed up the process, inorganic fertilisers added to the area increase the supply of nutrients and enhance the work of the bacteria. Spreading out polluted soils can also stimulate faster growth of bacteria.

Several species of bacteria have been found to contain the enzyme nitro-reductase, which gives them the ability to break down explosives such as TNT (2,4,6-trinitrotoluene). These species can be very helpful in cleaning up land contaminated with such substances. Over the last century, large quantities of explosives have been manufactured for military and industrial use. Many are highly resistant to biodegradation and large areas of land are contaminated with residues from their manufacture and storage. Incineration of affected soil is very expensive, so bioremediation is an attractive alternative.

In Australia, two species of bacteria, *Pseudomonas* sp. and *Azospirillum* sp., have been isolated from contaminated soils around disused sheep dips. These bacteria have been shown to be capable of breaking down organophosphate pesticides, which remain in the soil and pose a significant threat to the environment and public health. The bacteria are being used to develop a more general bioremediation strategy for the removal of the pesticides.

Removing methyl mercury

Mercury is a very toxic substance and, although all forms of it are harmful, major public health concerns are centred on methyl mercury, a neurotoxin that causes problems ranging from mild numbness to blindness, loss of balance, and in severe cases, death. Methyl mercury can enter the human body through the consumption of contaminated fish. Mercury compounds in sea water accumulate in the bodies of marine organisms, becoming more concentrated as they pass through food chains by a process called **biomagnification**, so that levels in larger fish taken for human food can be highly toxic. Methyl mercury was the compound responsible for the harm caused to residents of Minamata Bay in Japan in the 1950s following industrial pollution of the sea and the shellfish they used as food.

Bioremediation can be used to clean up mercury-contaminated waste water from processes such as gold mining, manufacture of plastics and the processing of fossil fuels, thus preventing methyl mercury from entering the food chain. Methyl mercury can be demethylated by bacteria such as *Pseudomonas balearica*, which carry mercury-resistance genes, and can decompose methyl mercury to produce inorganic mercury. Mercury is a liquid at room temperature and will vaporise on contact with air, but it is almost insoluble in water. If a community of mercury-resistant bacteria such as *Pseudomonas balearica* is established and held under a layer of water

in a bioreactor packed with inert material such as rubber fragments, mercury will precipitate and can be collected and removed.

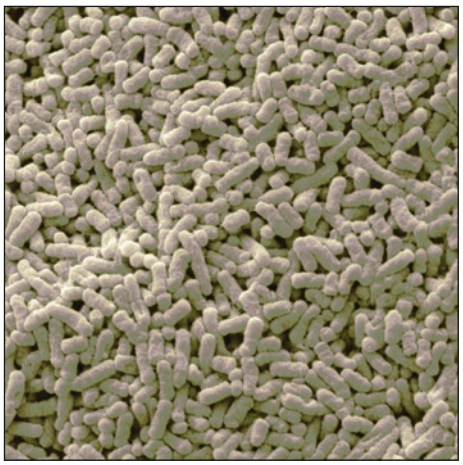


Figure B.18 Biofilm of bacteria forming on a tooth.

Emergent properties

An **aggregate** is a group of bacteria living together that has characteristics not displayed by the individual bacteria. For example, the marine bacterium *Vibrio fischeri*, which lives in the epidermis of sea anemones, does not emit light on its own, but in an aggregate the bacteria become bioluminescent. Other species produce bioluminescence in the light organs of the squid, when they are assembled in large enough groups. These aggregates, and other biofilms that develop different properties from the individual organisms, are said to have **emergent properties** – properties arising as a result of the aggregation.

Bacterial aggregates and biofilms

Most bacteria live as single cells but some (such as *Streptococcus mutans*, which occurs widely in the mouth) form **aggregates** or groups of cells that are connected together. *S. mutans* forms a layer known as a **biofilm** on teeth, at the junction with the gums (Figure **B.18**). The bacteria convert sucrose to a glue-like 'extracellular polysaccharide substance' (EPS), which allows the bacteria to stick to each other and form plaque. Together with the acid produced by bacterial metabolism, plaque leads to tooth decay.

Dental plaque is one example of a biofilm but there are many others, including the persistent slime that forms in a bathroom drain and the coating on the surface of submerged rock, which makes it slippery and dangerous. Some biofilms are useful (one example is their use in bioremediation in sewage treatment) but biofilms can cause major problems. If they form on food-production surfaces they can be a health hazard, and in other circumstances biofilms can cause clogging and corrosion in pipes. In each case, the aggregations are groups of microorganisms living in a matrix of adhesive EPS, which is used to incorporate new cells created by reproduction as well as other cells from the outside.

Biofilms start to form when individual microorganisms attach to a damp surface, which may be either living or non living. The reasons why a biofilm starts to form and attach to particular surface are not fully understood but the rate of water flow across the surface, nutrient availability and debris already present may be key factors. Others include the temperature and pH of the environment, and metabolic interactions between the cells. Keeping a surface perfectly clean is one way to prevent a biofilm forming, so hygiene and cleanliness in hospitals and kitchens are vitally important.

Biofilms are very flexible and are described as being **viscoelastic**, which means they can stretch and change their shape as a flow of liquid pulls or pushes them. Because of this, biofilms cannot be removed by rinsing them away – a flow of liquid over a biofilm may at best cause clumps to disconnect and fall away to settle elsewhere. This causes serious problems in hospitals and industry if biofilms attach to medical equipment or pipes. Some biofilms can be scraped off with gentle scrubbing but others are resistant even to pressure hoses. Biofilms that grow in areas where water flow is strongest tend to modify their growth to become thinner but more firmly attached.

Serious difficulties occur when biofilms become established in areas that are too difficult to reach or too delicate to treat. In hospitals they may form in catheters, nasal tubes or on heart valves. Some medical equipment can be replaced but biofilms that grow on implants can cause life-threatening infections. Biofilms protect the organisms within them from the human immune system and make them highly resistant to treatment with antibiotics and other antimicrobial agents.

Research efforts are being focussed on how to break up different biofilm colonies and make them detach from their surfaces, and on methods to weaken the EPS matrix so that antimicrobial agents can be used to kill the cells. Keeping surfaces clean is the best method of preventing a biofilm from forming in the first place.

Quorum sensing

Microorganisms living close to one another within a biofilm communicate using **quorum sensing**, by releasing signalling molecules into the environment. The accumulation of signalling molecules enables a single cell to detect the density of other cells in the area, and allows cells to coordinate their behaviour. If an environmental factor, such as the availability of nutrients changes, they can respond quickly in order to survive. They are able to defend themselves against competitors that share the same food source and avoid toxic compounds such as antibacterial agents. Quorum sensing systems exist in biofilms, as well as between marine and many pathogenic bacteria, which are known to coordinate their virulence in order to escape the immune system of their host. It may be that quorum sensing also establishes boundaries between different biofilms, and enables cells to reproduce or transfer genetic material between individuals.

About 60% of all hospital infections may be due to biofilms that have formed on medical devices. Catheters, lines and drips are all prone to biofilm infection and on implanted devices, such as valves and replacement joints, biofilms are very resistant to the human immune system and to antimicrobial agents.

Using biofilms in sewage treatment

Biofilms are very useful in sewage treatment. Their tenacity and resilience help them survive in the difficult conditions of a sewage treatment plant. The first stage in the treatment of raw sewage is to pass the material through coarse filters, which remove grit and other items in the sewage (Figure **B.19**). After this, the waste is left to settle in sedimentation tanks so that sludge can be taken away and the remaining liquid effluent given a secondary treatment.

Older methods of secondary treatment use trickle filter beds (Figure **B.20**), which contain biofilms of bacteria and fungi growing on the surface of a porous material known as clinker or on plastic piping, which can be up to 2 m thick. As water trickles through the biofilm, organic matter in it is broken down by the microorganisms in the biofilm, which use it as a source of nutrients. Oxygen levels are kept high by the movement of the trickling water and small invertebrates such as worms and protozoans grow over the biofilm, feeding on it so that it does not grow too thick and impede the filtering process.

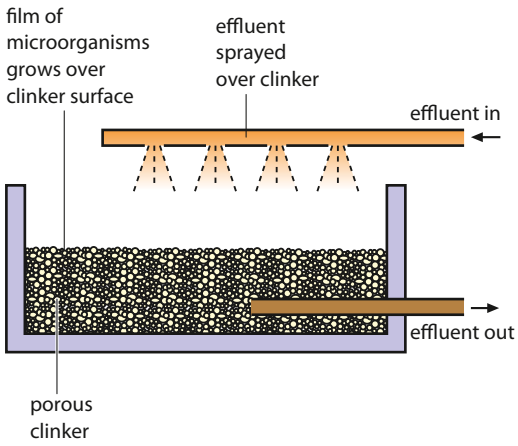


Figure B.20 Trickle bed filter.

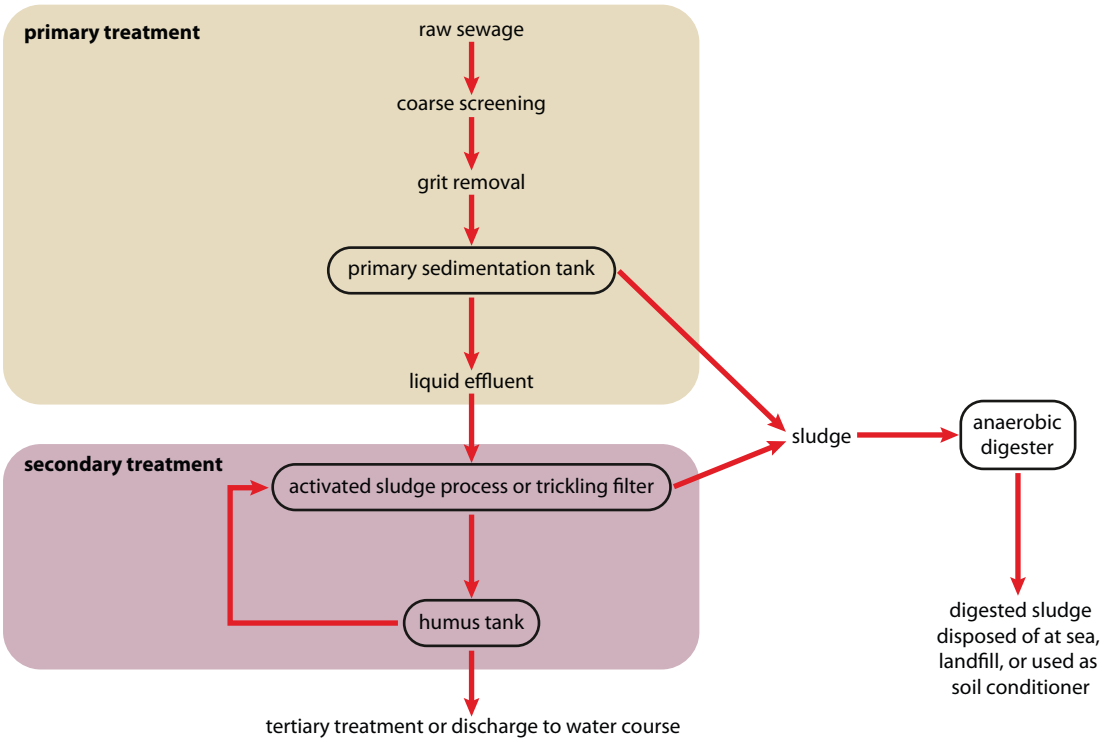


Figure B. 19 Sewage treatment.

Using bacteriophages in disinfection of water systems

Bacteriophages are viruses that infect and kill bacteria (Figure **B.21**) but which have no effect on plant or animal cells. Today, as antibiotic resistance is on the increase (Subtopic **5.2**) and it is becoming more and more difficult to kill bacteria, scientists are examining possible roles for bacteriophages as antibacterial agents.

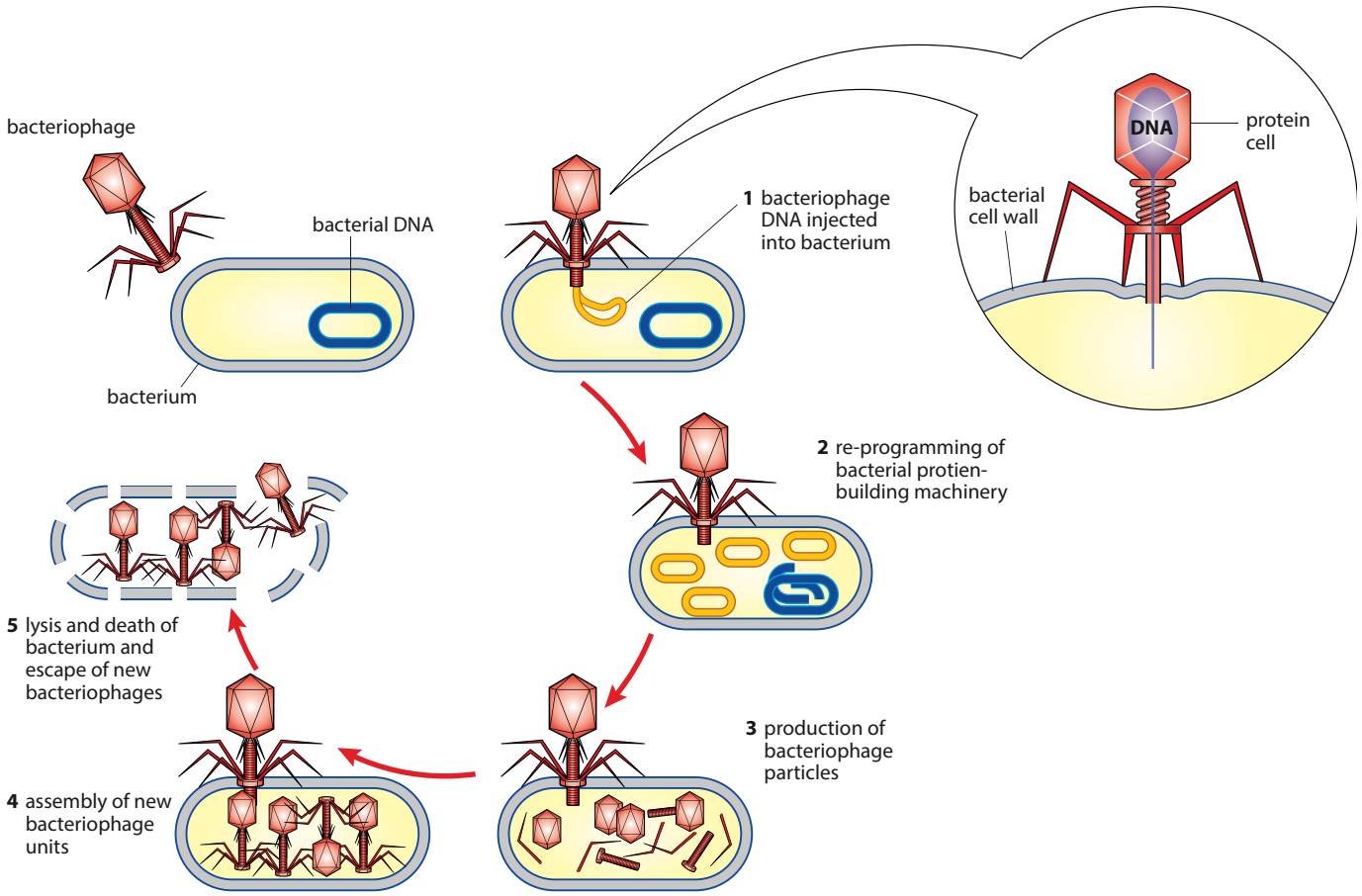


Figure B.21 How bacteriophages kill bacteria.

Water treatment is one area where exciting developments are taking place. Bacteriophages have been successfully used to kill the filamentous bacteria (*Sphaerotilus natans*) that grow in long threads during the settlement stage of sewage treatment. The bacteria interfere with sewage treatment by preventing sludge settling in the settlement tanks so that the remaining liquid cannot be led off for secondary treatment. Disinfectants such as chlorine can be used to kill the bacteria but if bacteriophages are employed, chlorine is no longer needed. After just a day the bacteriophages can remove enough bacteria to improve settlement rates and they have been found to remain active for many months.

Bacteriophages have also been used to remove biofilms of *Pseudomonas aeruginosa* bacteria, which can clog filters at water purification plants. Usually these bacteria must be cleaned away with chlorine and expensive flushing treatments. Studies show that bacteriophages can kill almost 90% of these biofilms and treatment is almost 100% effective if the bacteriophages are followed by a single treatment with chlorine.

Nature of science

Scientific advance follows technical innovation – the laser-scanning microscope

Advances in scanning microscopes have enabled scientists not only to examine the outer structure of bacteria and biofilms but also to build up 3D reconstructions of the complex interactions between individual cells in an aggregate. The **confocal laser-scanning microscope** is able to form images from selected depths within a biofilm in a process known as optical sectioning. The images obtained are built up into a 3D picture using computers. This has enabled researchers to develop a much clearer understanding not only of the interactions between individual bacteria but also the structure of their surrounding matrix. Without such new technology this would not have been possible. This type of microscopy also has uses in medicine and has helped in the diagnosis of keratomycosis, a fungal infection of the cornea. With laser images from the microscope, early diagnosis and treatment have been improved.

Test yourself

8 Outline one physical and one biological strategy used in a response to an oil spill.

9 List **three** properties of biofilms.

10 Outline what is meant by the term 'quorum sensing'.

11 State how bacteriophages can be used in water treatment.

B4 Medicine (HL)

Learning objectives

You should understand that:

- Genetic material or antigens from a pathogen can be used to detect infection by the pathogen.
- Markers can be used to detect predisposition to a genetic disease.
- DNA microarrays can be used to diagnose a disease or test for genetic predisposition to it.
- Metabolites indicating disease can be detected in urine or blood.
- Tracking experiments are used to gather information about the localisation and interaction of a desired protein.
- Genetically modified animals and plants are used in biopharming to produce proteins for therapeutic use.
- Viruses can be used as vectors in gene therapy.

Detecting infection using antigens or genetic material

A doctor may use a blood or urine test in an attempt to diagnose an infection. Antigens or genetic material from pathogens that are present in clinical specimens can be identified and used to pinpoint the cause of infection. Early diagnosis can then lead to faster treatment.

The three most important techniques used are:

- detection of pathogen-specific antibodies in the blood
- detection of the pathogen's antigens
- detection of genetic material from the pathogen.

Detection of antibodies

During an infection, **antibodies** bind to pathogens to disable them and prevent further infection of cells (Subtopic **11.1**). Two antibodies, known as immunoglobulins (IgM and IgG), are always produced as the immune system responds to an infection, and will be present in the blood. IgM is expressed on the surfaces of B cells and kills pathogens in the early stages of infection. These molecules are only produced for a few weeks, so detection of IgM in the blood tells a doctor that an infection is present at the time a blood test is taken. IgG is produced indefinitely by B cells and is used by the immune system to identify and neutralise bacteria and viruses. Presence of IgG provides evidence of a past infection.

Detection of antigens

Antigens (proteins derived from a pathogen) can be detected in blood samples using the ELISA test (Enzyme-Linked ImmunoSorbent Assay), which is one the most commonly used diagnostic tests (Figure **B.22**). It was developed for the detection of HIV but is now used to detect many other pathogens. The ELISA test is highly sensitive and very specific. It has the advantage that it can be used to screen large numbers of specimens at a time, making it invaluable for public health screening and in the screening of blood for transfusions.

ELISA is now a well-established tool for the rapid diagnosis of a wide variety of infectious diseases, helping doctors to choose the appropriate method to treat the infection at an early stage when antibiotics are most effective. Early diagnosis and targeted use of antibiotics may also reduce the development of antibiotic resistance.

The test uses an antibody that is specific to a particular antigen from a known pathogen, and a colour change shows whether or not antigens are present in a sample. The antigens chosen for use in the test are those that appear in blood or urine during an infection – for example, polysaccharides from viral capsules or proteins produced by bacteria.

The procedure involves:

1. immobilising a sample containing an unknown amount of the suspected antigen, usually by attaching it to a well in a plastic microtitre plate
2. attaching the antibodies being used to detect the pathogen to an enzyme and then adding this enzyme–antibody molecule to the plate so that it can form a complex with any antigens present

3 rinsing away any unbound material so that only enzyme–antibody molecules that have formed complexes with antigens remain
4 finally, adding a substrate for the enzyme to the surface of the plate – so that as the substrate binds with enzyme molecules attached to antibody–antigen complexes, a colour change occurs (Figure **B.22**). The degree of change in colour indicates the amount of antigen in the sample. Quantitative results can be obtained by measuring the optical density – the greater the colour change, the greater the quantity of antigen present.

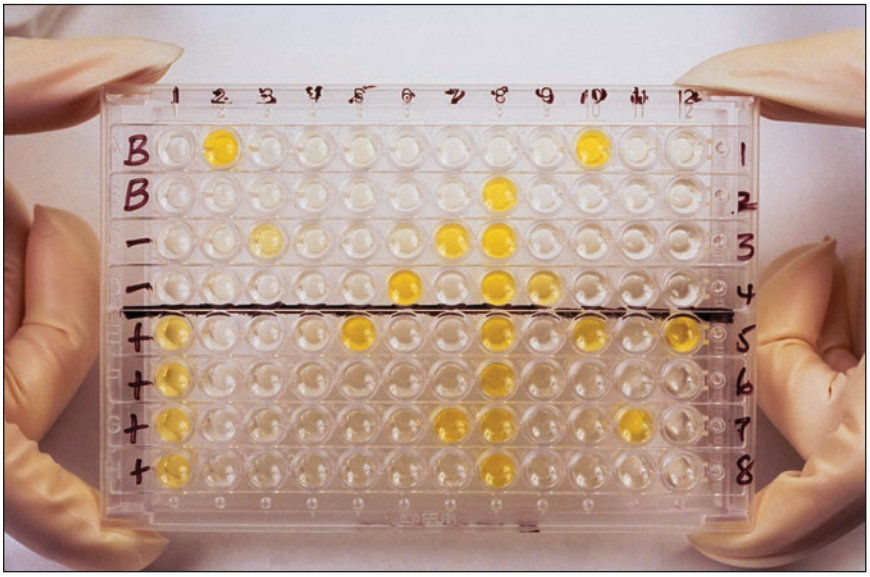


Figure B.22 The ELISA test can be carried out on many specimens at once. Each well in the plate is used for a separate patient's specimen. The yellow colour on this plate indicates a positive result – antigen from the pathogen is present, which shows that the patient is infected with this pathogen.

Detection of genetic material

Genetic material from pathogens such as influenza viruses can be detected using a rapid, sensitive diagnostic technique called the **reverse transcriptase polymerase chain reaction** (RT-PCR). This method uses genetic material from retroviruses such as influenza (which contain only RNA) as a template for reverse transcription. Complementary DNA is produced from the virus RNA. Alternatively, for other pathogens, DNA samples can be used directly. The DNA is amplified using the **polymerase chain reaction** (PCR) (Subtopic **3.5**) and detected either by gel electrophoresis or by using hybridisation with known sequences of DNA to assess the degree of matching and so identify the pathogen.

The sections of genetic material from the pathogen that are selected for amplification in the PCR are carefully chosen. The sections are removed from the genome using restriction endonucleases and by using different endonucleases different regions can be targeted. When trying to identify of bird flu viruses, for example, hemagglutinin (HA) and neuraminidase (NA) genes are targeted because these are unique to strains of bird flu and cannot be confused with material from human flu viruses. By using two different genes in the test a clinician can be more certain that positive results are reliable.

Following the amplification of genetic material, the DNA fragments are separated on agarose gels using electrophoresis. Gel electrophoresis separates the fragments by their size. All DNA fragments are negatively charged but when the electric current is switched on, smaller fragments will move further through the gel than larger ones, which cannot pass easily through it. DNA samples are mixed with a 'loading' dye, which makes them visible and the power supply is turned on. The samples and their dye markers move through the gel and the different molecules separate, within their different lanes. Then the current is switched off and the gel removed. DNA molecules are made visible by staining the gel with ethidium bromide, which binds to them. Ethidium bromide fluoresces under ultraviolet light so the series of bands spread through the gel can be seen if it is illuminated by a UV lamp (Figure **B.23**). Different strains of influenza viruses can be identified from profiles like these, because each genome produces a different characteristic range of DNA fragments when treated with endonucleases.

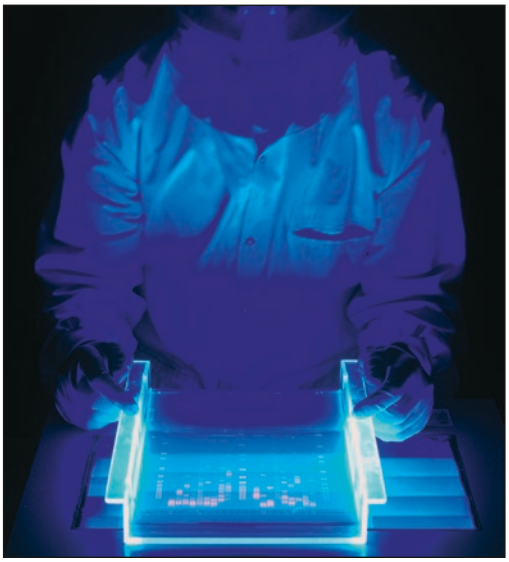


Figure B.23 Agarose gel stained with ethidium bromide. Samples of genetic material from different pathogens produce different banding patterns following electrophoresis, which can be used to identify them.

Genetic predisposition to disease, and diagnosis

Diseases can involve complex interactions between many genes and many are also closely linked to environmental factors. An individual may not be born with a disease but may be at high risk of acquiring it. This is known as **genetic predisposition** or susceptibility. A genetic susceptibility to a particular disease can be identified, in some cases, by **markers** present in a person's body. Recent research has moved the detection of these markers forward so that by understanding their genetic predisposition to disease individuals can make informed choices about lifestyle or medical treatments that could reduce their probability of developing a disease. This might include avoiding certain activities such as smoking, or taking more exercise.

Testing to assess a person's predisposition, or to diagnose the disease itself, is described below for three important diseases – breast cancer, Alzheimer's disease and prostate cancer.

In medical science, a **marker**, or **biomarker**, is any measurable characteristic that indicates the presence or severity of a disease, or a person's susceptibility to it. Biomarkers are often detected in blood or tissue, and may be proteins – such as specific gene products, enzymes or hormones – or other molecules, genes, or cells. High body temperature is a biomarker for fever, while blood pressure is a well-known indicator of the risk of stroke.

Breast cancer

Inherited mutations of two genes known as the *BRCA* genes on chromosomes 13 and 17, account for a small proportion of all breast cancers, but women who have these mutations have a substantial risk, greater than 70%, of developing breast cancer or ovarian cancer. Identification of people who carry the mutations, by genetic testing, allows for preventive measures (including mastectomy), clinical treatment and counselling. *BRCA* analysis is a genetic test that requires only a blood sample to determine whether a patient carries a faulty *BRCA1* or *BRCA2* gene. Genes may be detected using DNA microarrays, which assess the match between genes from the person being tested to known *BRCA* genes.

A **DNA microarray** is a 'biochip' that has thousands of microscopic DNA samples attached to its surface (Figure **B.24**). Each spot contains a specific DNA sequence, known as a **probe**, that is used to hybridise samples of **target DNA**. Hybridisation occurs when complementary nucleotide sequences in probe and target DNA molecules pair up by forming hydrogen bonds. The more complementary sequences present, the greater and stronger the bonding between the two strands. The microarray is washed after hybridisation and weakly or non-bonded strands are flushed away. Hybridisation is usually detected and quantified using luminescent labels attached to target DNA (Figure **B.25**).

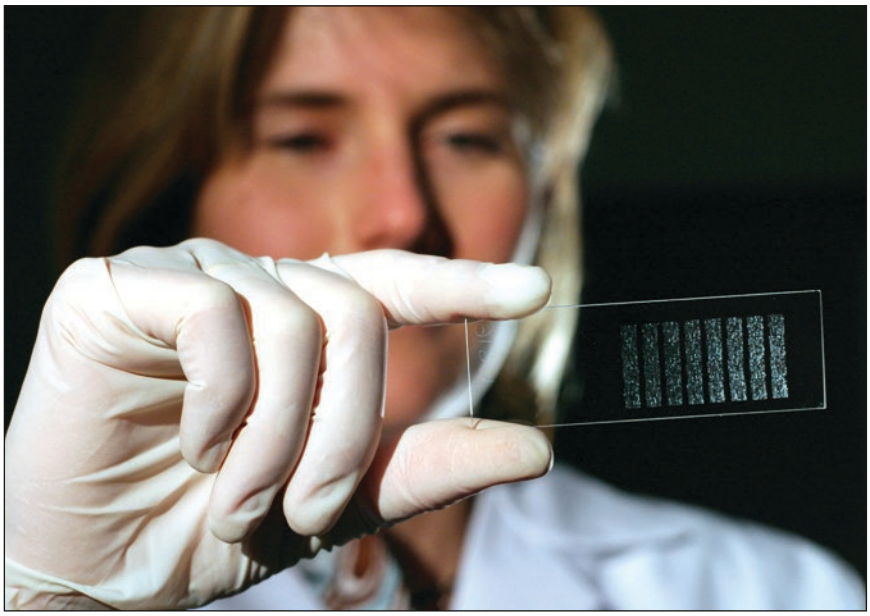


Figure B.24 A microarray contains tens of thousands of microscopic DNA probes.

Microarrays enable scientists to carry out thousands of genetic tests at the same time. They can be used to check for the presence of genes in a genome, to check for genetic predisposition to a disease, or to diagnose a condition. DNA microarrays are used to seek out single nucleotide polymorphisms (SNP) among alleles making them important tools in evaluating faulty genes or locating genetic mutations in cancers.

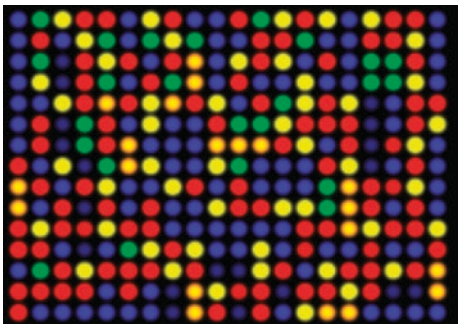


Figure B.25 Results from a microarray. An array of DNA sequences for a particular set of genes is created on the biochip. Each gene's position in the array is known. Samples of unbound DNA or mRNA (messenger RNA) are then labelled with differently coloured fluorescent markers and added to the chip. The genetic material in the samples binds to sites on the array that have a matching (complementary) sequence. The pattern of colours is then analysed.

Tau proteins are found in neurons where they are concerned with the formation of microtubules, tiny structures responsible for internal communication and transport. If they are folded they can become tangled and form insoluble aggregates, which interfere with cell function and are characteristic of Alzheimer's and other neurological diseases.

Alzheimer's disease

Research has revealed that testing either blood or cerebrospinal fluid may allow doctors to detect a person's predisposition to developing Alzheimer's disease, which affects memory and a person's ability carry out everyday tasks. Individuals whose blood contains higher than average levels of the marker peptide beta-amyloid 42 (Aβ42) are at increased risk of developing Alzheimer's disease. Researchers have found that plasma levels of Aβ42 increase before the onset of Alzheimer's disease and decline shortly afterwards. This may indicate that as the disease progresses the decline of Aβ42 in the bloodstream is linked to accumulation of Aβ42 in the brain, which occurs in people who have developed dementia.

More recently, genetic markers that identify mutations affecting the build-up of certain proteins in the brain have also been used to help diagnose Alzheimer's disease. High levels of **tau proteins** and the presence of folded molecules of phosphorylated tau are signs of the disease.

Prostate and other cancers

A **tumour marker** is a substance found in blood, urine or tissue samples that may become elevated if a person has developed cancer. There are many different tumour markers and each one is specific to a particular disease. Tumour markers can be produced either directly by the tumour or by non-tumour cells that are responding to the presence of a tumour. Most markers are tumour antigens.

Today, the most widely used tumour marker in cancer diagnosis is the prostate-specific antigen (PSA), which is detected in a blood test. The test is used to check men for prostate cancer. Men with prostate cancer usually have high PSA levels. But the test is not perfect because high PSA levels can occur in men without cancer, and a normal PSA level does not necessarily mean that no cancer is present. Other tumour markers have been identified which can indicate many other cancers including colon, stomach and breast cancers.

Metabolites in blood and urine

The serotonin production pathway is activated when a person consumes alcoholic drinks so this can be a metabolic marker indicating recent alcohol consumption.

As we have seen, a **biomarker** is a substance that can be measured and which can indicate the presence or severity of a disease. General biomarkers used in medicine include LDL (low density lipoprotein), which can be measured to indicate blood cholesterol levels, and genetic biomarkers for specific cancers or predisposition to them.

Metabolomics is a name given to the analysis of metabolites in a blood or urine sample. In many cases diseases cause disruption of metabolic pathways and the accumulation of, or lack of, metabolites can be biological indicators of certain diseases. Some examples of diseases that can be detected in this way are described below.

Diagnosis of diabetes – ketones in urine

If the body is deprived of carbohydrates, it must rely on the metabolism of fats to supply energy. This can occur if an individual is starving or following a high-protein diet, but it is also a symptom of diabetes. Fat metabolism occurs in several stages and the intermediate products of the process known as ketones can accumulate in the body if they are not metabolised completely. Ketones accumulate in the blood and are excreted in urine. Their presence in urine samples can be an indicator of diabetes.

Diagnosis of PKU – phenylalanine in blood samples

Phenylketonuria (PKU) is a genetic disorder caused by a mutation on chromosome 12. People who suffer from PKU lack the enzyme tyrosine hydroxylase, needed to process the amino acid phenylalanine. If phenylalanine levels in the blood are high, serious mental and physical health problems can arise. Babies are routinely tested for the amino acid with a simple blood test and if it is present suitable treatment can be given.

Glycogen storage disorders – creatine kinase in blood samples

Glycogen storage disorders are a range of conditions that can lead to low blood sugar levels and other symptoms such as fatigue and muscle problems. Diagnosis includes the use of blood tests to check for lipids and urate levels, and for the enzyme creatine kinase, which can indicate whether the glycogen storage disorder is affecting an individual's muscles.

Galactosemia – enzyme absent from blood or urine

Galactosemia is a rare genetic disorder affecting a person's ability to metabolise the sugar galactose (found in milk). It is due to a deficiency in an enzyme responsible for breaking down the galactose. A galactosemia test involves testing blood or urine for three enzymes. A person with galactosemia is lacking one of these enzymes so that high levels of galactose are present in their blood or urine.

Porphyria – accumulation of precursors in the blood or urine

Porphyria is caused by a deficiency in any of the eight enzymes involved in the synthesis of porphyrins and heme. Normally, the body uses porphyrins to produce heme, but if enzymes are missing, porphyrins and precursors in the reaction accumulate and insufficient heme is produced. (Heme is essential for the formation of hemoglobin and healthy red blood cells.) Sometimes the precursors are deposited in the skin and lead to photosensitivity. Some porphyrins may pass out of the body in feces or with urine, colouring it black. In other cases of porphyria, porphyrins such as porphobilinogen (PBG), one of the first substances in the synthesis pathway, can build up. Initial diagnosis of porphyria involves testing urine for the presence of PBG, followed by tests on blood and feces for other porphyrins.

Tracking tumour cells – transferrin and luminescent probes

Transferrins are glycoproteins found in the blood, which bind to iron and control the level of free iron in the plasma. **Transferrin receptor** (TfR) is a glycoprotein found in cell membranes that is involved in the uptake of iron from transferrin and also with the regulation of cell growth. Transferrins carrying iron bind to TfRs and iron can then be transported into the cell by endocytosis.

Recent studies have shown that there are higher levels of TfR on tumour cells than on normal cells. The TfR on normal cells is expressed at low levels but is expressed at greater levels on cells with a high proliferation rate, including cells of the epidermis and intestinal epithelium as well as tumour cells. The high levels of this protein receptor, in an accessible extracellular position, can be used as a target for the

location and treatment of cancerous cells. Tumour cells can be tracked by linking luminescent probes to transferrin molecules, which then attach to the TfRs. In this way, luminescence has been used both to measure levels of TfRs and to identify and track tumour cells in the body.

One new cancer treatment uses this transferrin technology to target and destroy cancerous cells. The technique involves a 'hybrid peptide' composed of two parts: one peptide binds to a target TfR, while a second, lytic peptide then destroys the cell membrane and kills the cancerous cell.

Biopharming

Biopharming refers to the use of genetic engineering to add genes to animals or plants so that they produce useful pharmaceuticals. The GM animals or plants express the added genes and so produce protein from them. As we have seen, proteins produced from recombinant DNA such as human insulin can be produced in bacteria or yeast in bioreactors (Subtopic **B.1**), but one of the main reasons for the increasing interest in animal biopharming is its potential to produce substances that cannot be made using these methods. Transgenic animals have an advantage over both plants and microorganisms because they are able to carry out complex post-translational modifications to the proteins they produce. Animals can therefore make high-quality, biologically active proteins for use as pharmaceuticals.

Harvesting the desired proteins from milk is the most popular way of accessing recombinant proteins from transgenic organisms. Milk production is plentiful and purification of the proteins from milk is straightforward. Goats are one species that is well suited to biopharming because they have a faster breeding cycle than cattle and produce more milk than sheep. Females mature quickly and have a gestation period of just 6 months.

Transgenic goats can be produced in one of two ways. Either by the fusion of cloned fragments of DNA coding for the required proteins with goat genes, which are then injected into eggs or by nuclear transfer (Figure **B.26**). The gene that codes for the chosen protein is attached to a promoter gene, usually the gene for casein, a milk protein, and then microinjected into the nucleus of a newly fertilised egg in the laboratory. Alternatively the DNA is injected along with goat genes into an enucleated egg. In both cases the gene may become incorporated with the genetic material of the embryo as the cell divides. If this happens, the new gene (or transgene) will be incorporated into every cell of the developing goat embryo. The embryo is cultured and then transferred to a goat who will become a surrogate mother. If a female kid is produced, she will produce milk with the new protein, which can be extracted from her milk. When the female kid becomes an adult and breeds, half of her offspring will have the new gene.

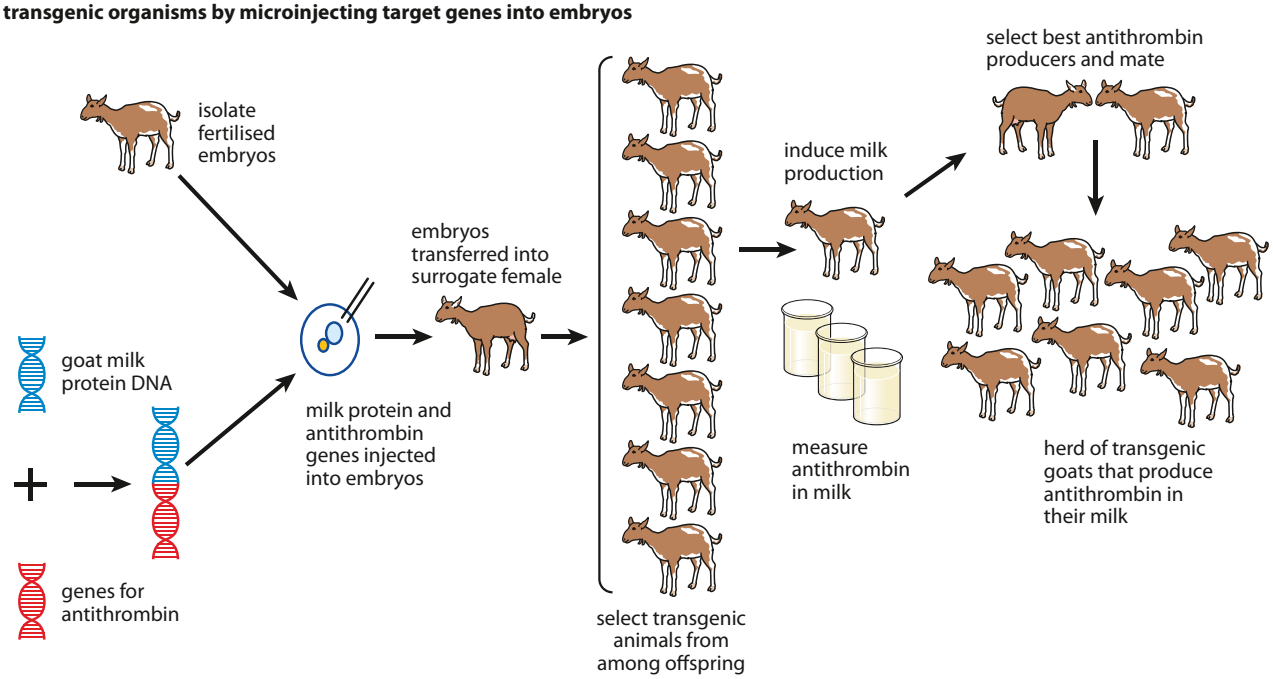
A herd of transgenic goats can be produced using microinjection of genes into enucleated eggs and using the transgenic offspring for reproduction. Cloned early-stage embryos can be used to create additional offspring that carry the modified genome. However, the low survival rate of embryos and animals generated by biopharming is one of the major problems of this technique.

Biotechnologists anticipate that cloning will become far more important as the biopharming of animals develops. Cloning embryos can increase the efficiency rate significantly. The rate of success for microinjection is low. If 1000 fertilised eggs are injected, only 100 embryos are likely to be produced and from these only one transgenic offspring is likely to be born – and there's only a 50% chance of its being female. Cloning embryos produced using adult genetic material and enucleated eggs has a much higher success rate.

In 2009, the US Government gave permission for the sale of the first drug to be produced by biopharming – the drug, called ATryn, is recombinant human antithrombin protein purified from the milk of genetically modified goats. Marketing permission for ATryn was granted in Europe in 2006. (Antithrombin activates enzymes needed for clotting blood so people who inherit antithrombin deficiency have blood that does not clot easily.)

As well as blood clotting factors many other therapeutic substances have been produced by biopharming, including antibodies, growth hormones, fibrinogen, albumin and recombinant enzymes. Treatments for cystic fibrosis using recombinant gastric lipase are also being developed.

Creating transgenic organisms by microinjecting target genes into embryos



Creating transgenic organisms by nuclear transfer

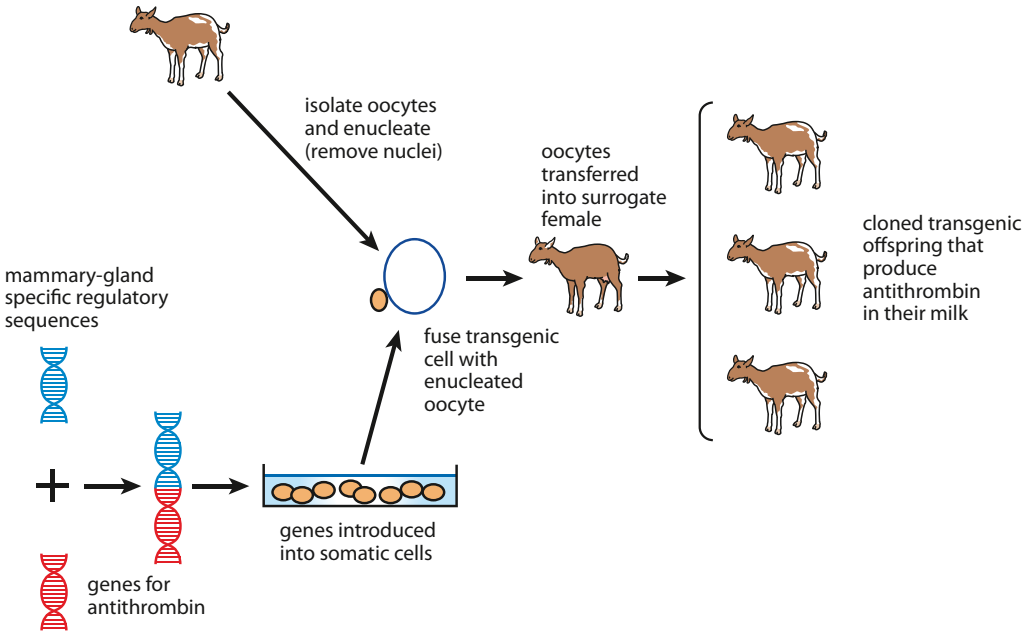


Figure B.26 Creating transgenic goats.

Therapeutic proteins from transgenic plants are also being developed. The proteins can be harvested from the plants' leaves or seeds and used to produce pharmaceuticals. (In Subtopic **B2**, you can read about how the tobacco plant is used to produce vaccine proteins, for example.) New developments in this area include the use of transgenic non-crop plants such as duckweed (*Lemna minor*) and moss. These organisms can be grown in bioreactors, where they secrete proteins produced from introduced recombinant DNA into their growth medium. This makes the purification of proteins for medical use much simpler than for a species grown in a field and then harvested.

Viral vectors in gene therapy

Viruses are very efficient at entering the cells of organisms, and can be used as **vectors** providing a powerful means for the delivery of therapeutic genes into cells. Some viruses are even able to incorporate the genes they carry into the cells they enter.

Before a virus can be used, it must be modified so that it will enter but not replicate inside a target cell, as this would lead to the cell's destruction. Viral genes that are involved in replication are removed or inactivated. Deletion of these genes also allows non-viral genetic material to be inserted and these viruses are then known as vectors.

Retroviruses are the most frequently used vectors. These single-stranded RNA viruses enter target cells via specific receptors and their RNA is converted into DNA and integrated into the genetic material of the cell (Subtopic **B2**), where it remains for the life of the cell. Integrated genes are also passed on when the cell divides.

There has been some success in treating a condition called severe combined immune deficiency (SCID) using retroviruses. Children who suffer from SCID have no immune system because a gene mutation prevents their cells producing the enzyme adenosine deaminase (ADA). Substrates for ADA build up in cells and are very toxic to developing lymphocytes. These cells fail to mature and the patient is left without a working immune system. Stem cells from bone marrow or umbilical cord blood can be taken out of the body and treated with viral vectors that transfer a normal copy of the ADA gene to them. If the treated cells are returned to the bone marrow, the replacement genes can begin to produce ADA. Bone marrow and stem cell transplants now save up to 80% of SCID patients.

Gene therapy

Gene therapy involves modification of genetic material in the cells of a patient in order to bring about a therapeutic effect. Modification is usually achieved by introducing DNA, using viral vectors or other means. Although gene therapy is still in its infancy in medical treatments, discussion of the ethical issues involves principles that apply to all clinical medicine. How are subjects selected for gene therapy trials? How can the safety of individuals who take part in the trials be safeguarded, and do scientists involved in the trials have a conflict of interest between wanting to conduct their research and the best interests of their patients?

In 2007, a 36-year-old woman with rheumatoid arthritis died while participating in a clinical trial for gene therapy. Some experts say she shouldn't have received such an unpredictable, potentially dangerous treatment at all. She was able to lead a full and active life with existing drugs keeping her disease under control. Soon after the experimental treatment a sudden infection caused her organs to fail and there is a suspicion that her death was linked to the therapy. Is it ethical to test unknown therapies on patients whose ailments are not life threatening? Priorities in medical research can raise troubling issues of social ethics.

Questions to consider

- Can 'good' and 'bad' uses of gene therapy be distinguished?
- Who decides which traits are normal and which constitute a disability or disorder worthy of gene therapy treatment?
- Can the safety of patients ever be guaranteed in treatments that are new or experimental?
- Are there conflicts of interest between patient safety and the need to conduct research?

Nature of science

Scientific advance follows technical innovation – technology in medicine

The use of computers to store and collate data, the use of microtechnology to produce DNA chips and the use of robotic plating in ELISA tests are all examples of the way that technology has enabled scientists to collect more, better and more precise information. The use of such technology in medicine is an area that has expanded rapidly in recent decades, allowing huge improvements in the understanding, diagnosis and treatment of disease.

Test yourself

12 List **two** diseases that can be detected by the presence of markers in a blood sample.

13 Describe how a microarray can be used to search for a sequence of nucleotides in a DNA sample.

14 Define the term 'biopharming'.

B5 Bioinformatics (HL)

Learning objectives

You should understand that:

- Databases allow scientists to access information easily.
- The quantity of data in databases is increasing exponentially.
- Similar sequences can be identified in the genomes or proteins of different organisms using BLAST searches.
- Gene function can be investigated by studying model organisms that have similar sequences.
- Sequences from different organisms can be compared using sequence alignment software.
- BLASTp software allows scientists to find alignment between amino acid sequences in proteins and known sequences stored in a database, while BLASTn software allows nucleotide sequence alignment.
- Searches of databases can allow newly identified sequences to be compared with sequences of known function in other organisms.
- Multiple sequence alignment is used in phylogenetic studies.
- Expressed sequence tags (ESTs) can be used to identify potential genes.

Databases

Biological databases are enormous electronic records of information from experiments, scientific papers and computer analyses of data. Rather like a library of books, a database is a collection of structured, searchable and up-to-date data maintained electronically. Information is stored from research in genomics, proteomics, metabolomics, microarray results on genes and phylogenetic studies of evolution, and can be accessed quickly and easily. In genetics, the information provides scientists with easy access to information about gene function, structure and location as well as storing information on the effects of mutations. Similarities between sequences and structures of DNA and proteins in different species can be obtained and compared with new information that is collected. Databases can help our understanding of the interaction of biomolecules or observations of metabolism.

The exponential increase in the quantity of data stored in databases has helped in the development of drugs, our ability to combat diseases and our understanding of evolution. Genome sequencing, in particular, along with many other large-scale research projects, has generated an explosive growth in biological data and there are many specialised databases. Species-specific databases are kept for species such as *Escherichia coli*, *Drosophila* and nematodes, which are model organisms used in many research projects. There are hundreds of public, freely accessible databases available in a matter of minutes to anyone worldwide. But recently some private genome-sequencing companies have started to charge for data that might be of commercial interest – for example, the public genome database for *Saccharomyces cerevisiae* (yeast), a model organism used by many scientists, changed from a free to a chargeable database in 2002.

Data are usually entered in databases and assigned an identifying number for quotation in scientific publications. Most of the data are not directly from the original source – the contents are extracted from other databases by a process of filtering, transformation and manual correction. Because the information is distributed among many general and specialised databases it is sometimes difficult to make sure that the information is consistent.

Model organism non-human species used for study and experimentation in order to understand particular pathways or processes. *Escherichia coli*, *Drosophila melanogaster*, *Saccharomyces cerevisiae* and the plant *Arabidopsis thaliana* are all model organisms. Experiments are done with the expectation that any discoveries made will increase understanding of the same processes in other organisms

Freedom of information

The information in the formerly public database containing information about yeast was taken over by a commercial company in 2000. They now charge US $2000 per lab per year for scientists to use the database. Much of the information in the database has come from scientific publications and personal communication between researchers who have used it.

Question to consider

- Should commercial interests ever be put ahead of free access to scientific information?

BLAST searches

BLAST is an **algorithm** used to compare sequence information, such as the amino acid sequences in different proteins or the nucleotide sequences in DNA from different organisms. The BLAST program was first published in 1990 and now researchers can use a BLAST search to compare a sequence of DNA or amino acids with a database containing known sequences. The search will reveal all the known sequences that have more than a certain proportion of resemblances to the sequence being investigated.

Algorithm a set of rules that defines a sequence of operations for solving a problem in a finite number of steps

The **BLASTp** program allows amino acid sequence alignments to be made and matches to be found between newly discovered sequences and known sequences. Questions it can help to answer include:

- Do analogous proteins from different species share similar amino acid sequences, and if so are the species likely to be related in evolutionary terms?
- Where in a protein is a particular sequence of amino acids located?

The **BLASTn** program allows nucleotide sequence alignments to be made. BLASTn can be used for several purposes such as the identification of species, establishment of evolutionary relationships, and DNA mapping and comparison:

- BLASTn can help to determine the origin of a certain sequence of DNA. It can help to identify a new or known species from a sample of its DNA and search for related species with similar sequences.
- Using the results from BLASTn, a phylogenetic tree can be built up – although phylogenies based on BLASTn alone are not 100% reliable and are always used in conjunction with other phylogenetic analyses.
- In DNA mapping, BLASTn can compare the chromosomal position of a gene sequence of a known species with the positions of other sequences in the database, or locate genes that are common in related species to map differences and similarities between one organism and another.

Studying gene function in model organisms

Knockout mice

Gene function can be studied using model organisms such as the mouse, which is the laboratory animal mostly closely related to humans. The mouse has many genetic sequences that are very similar to those of humans.

A **knockout mouse** is a genetically engineered animal that has had certain genes inactivated or 'knocked out'. The first knockout mouse was created in 1989 and since then thousands of different strains have been made. A gene in a knockout mouse may either be replaced or disrupted by the addition of a section of artificial DNA. Knocking out genes changes the phenotype of the mouse so that it looks or behaves in a way that researchers can see, or it has different biochemical characteristics that can be monitored (Figure **B.27**).

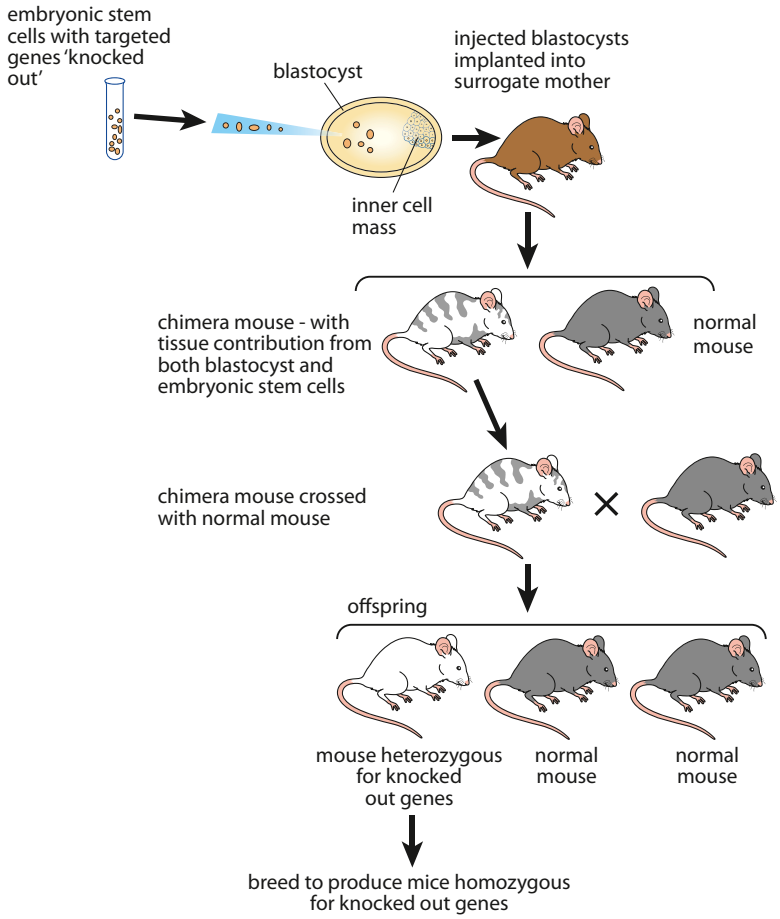


Figure B.27 Genes are replaced or disrupted in a blastocyst, which is implanted into a surrogate mother.

Knockout mice are important animal models for studying the genes that have been sequenced but whose functions are not known. By inactivating a specific mouse gene and observing any differences in behaviour or physiology, researchers can deduce its possible function. Different strains of model mice are named after the gene that has been deactivated. The 'p53 knockout mouse' is named after the gene that codes for a tumour-suppressing protein. The protein stops cell division and induces programmed cell death (apoptosis). In humans, the genetic disease Li–Fraumeni syndrome occurs in people who have a mutation that deactivates their p53 gene. These individuals are at a high risk of developing several types of bone, breast and blood cancers. However, in p53 knockout mice cancers develop in quite different tissues. So, as data is extrapolated from one species to another, scientists must be aware that a change in a gene in a knockout mouse may produce different effects in humans.

Discovery of genes using EST data mining

Improved understanding of normal gene expression can help our understanding of what happens when genes are altered – for example, in disease. Most studies of this type involve investigations of the proteins that are coded for by the genes involved. But finding exactly which gene codes for which protein is not easy and is very time consuming. The process has been improved considerably since the development of new technology to generate **expressed sequence tags** (ESTs). An EST is the nucleotide

sequence of a tiny portion of a known gene, which computers can use to scan databases to help make a match with unknown genes and to map their positions within a genome, assisting in the construction of genome maps. ESTs are short sequences (usually between 300 and 500 nucleotides long) that occur at one or both ends of a gene. ESTs effectively 'represent' the genes expressed in certain cells, and researchers use them as 'tags' to locate a gene in a portion of chromosomal DNA from a different organism, by matching up the base pairs. ESTs provide a route for finding new genes and for obtaining data on gene expression and regulation.

Gene sequences coding for analogous proteins are different in different organisms, varying with the size of the genome and the introns the genes contain. Humans have large numbers of introns and relatively few genes, which makes an individual gene almost impossible to locate. mRNA is used to help find those genes that are expressed and produce protein. But mRNA is an unstable molecule so for research purposes it is converted to complementary DNA (cDNA), which is more stable and contains none of the introns that were present in the original chromosomal DNA sequence. cDNA can be prepared in the laboratory using reverse transcriptase (Subtopic **B2**), which uses mRNA as a template to produce cDNA containing only exons (expressed DNA sequences) (Subtopic **7.1**). ESTs can be sequenced from either end of the cDNA but sequencing at the start of the molecule produces a **5' EST**, which is more useful, because this region tends to be conserved across species and is not significantly different in groups of closely related genes that produce similar proteins.

ESTs are crucial in the search for known genes because they greatly reduce the time needed to find them. For example, suppose a doctor suspects that a person has a disease or condition associated with a missing or defective protein, and wants to check the person's DNA to see if the gene for that protein is defective. ESTs for that known gene could be used to seek out the gene in a DNA sample from the person, and that gene analysed to see if it differs significantly from normal. In humans, genes involved in Alzheimer's disease, colon cancer and other diseases have already been found in this way.

In the 1990s, a new database called dbEST was set up to record data from ESTs. All ESTs that are submitted to the genetic database GenBank are checked and annotated and then lodged in dbEST, which stores a wide variety of genome data.

Exam tip
Make a list of key terms that are important in this and other topics in your course. Some might be **cDNA**, **bioassay**, **cladogram**, **PCR**. Check that you can define all of them.

Nature of science

Cooperation and collaboration – databases allow free access to information

Scientists all over the world use databases to access information on their research organisms, and to contribute to the body of information stored there. The internet has provided the opportunity for interaction in a way that has never been possible before and computers have enabled huge amounts of quantitative data to be stored and accessed. All this information would have taken years to collate and catalogue before the advent of computers.

Reliability of databases

Scientists use databases to provide data, which they use in their research. But how reliable can claims based on such data be? Is it important that the data on which a scientist bases his or her research may have come from researchers using different techniques and producing the data for different reasons?

Test yourself

15 State the type of data that can be accessed in a BLASTn search.

16 Outline what is meant by model organisms and why they are important in research.

17 Describe how ESTs are used to identify potential genes in DNA sequences.

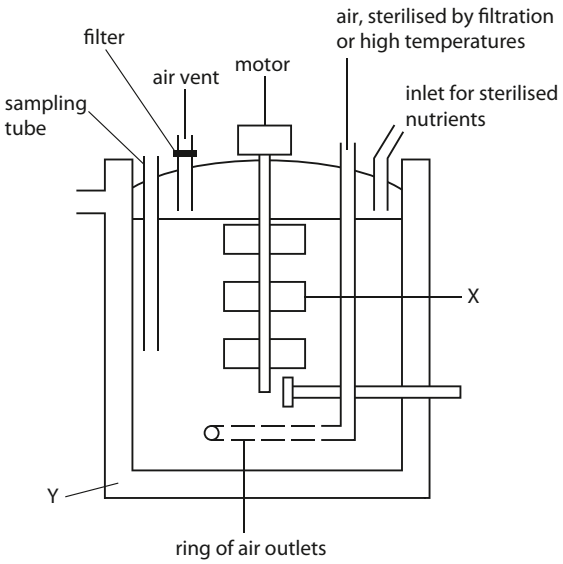
Exam-style questions

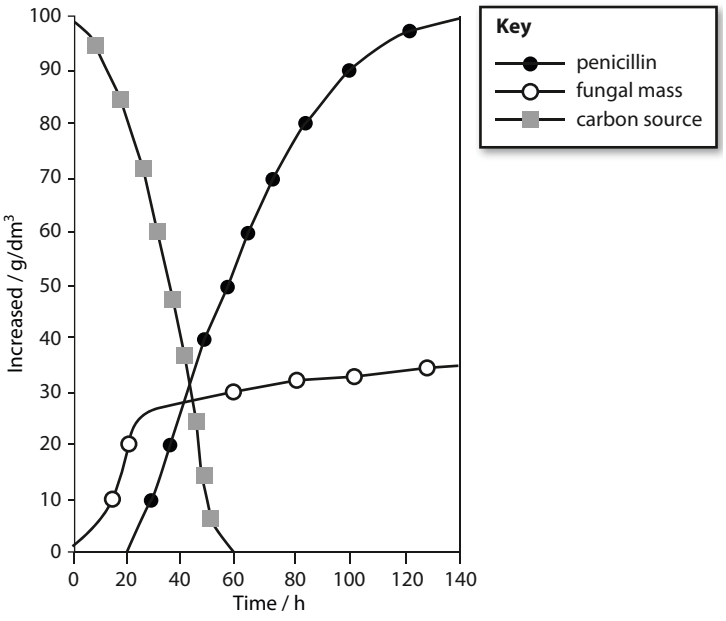
1 Penicillin is manufactured using a process called batch fermentation.

a Outline what is meant by 'batch fermentation'. **[2]**

b State how batch fermentation differs from continuous culture. **[1]**

A fermenter similar to the one shown in the diagram below is used to produce penicillin. The graph shows the rates of penicillin production and use of carbon by the *Penicillium* fungi.





c What is the function of the structure labelled X? **[1]**

d State the function of structure Y. **[1]**

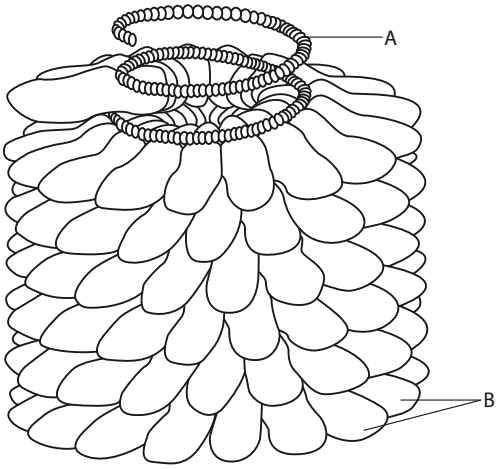
e Define the term 'secondary metabolite'. **[1]**

Using information shown in the graph:

f State the time when the fungal population reaches the plateau stage. **[1]**

g Explain why penicillin production is not increased if further nutrients are added at 30 hours. **[1]**

2 The diagram shows the structure of Tobacco Mosaic Virus (TMV), which infects the leaves of the tobacco plant.



a Name the molecule A. **[1]**

b Name the molecules that make up structure B. **[1]**

c Explain how TMV has been used to allow the bulk production of vaccines from tobacco plants. **[5]**

3 Outline **three** advantages of the Amflora potato in industry. **[3]**

4 Which of the following statements about transgenic plants is true:

A they contain genes from another species
B they are used to produce human antibodies
C they are from different geographical areas
D statements a and b only **[1]**

5 Which of the following microorganism is used for the production of citric acid?

A *Lactobacillus bulgaricus*
B *Saccharomyces cerevisiae*
C *Aspergillus niger*
D *Streptococcus lacis* **[1]**

6 Which of the following statements correctly describes transfection?

A The synthesis of mRNA from a DNA template
B The process by which a cell become malignant
C The introduction of genes from a different species in to a cell
D The synthesis of protein based on mRNA sequence **[1]**

7 Describe what is meant by the term 'open reading frame'. **[2]**

8 Complete the following table inserting a tick (✓) if the structure is present and a cross (✗) if it is not.

Microorganism	Nucleus	Mitochondria	Ribosomes
fungus			
bacterium			
virus			

[3]

9 Biofilms are said to possess emergent properties.

a Define the term biofilm. [1]

b State one emergent property that a biofilm might possess. [1]

c Biofilms co-operate using quorum sensing. Describe what advantages this may confer to the microorganisms. [2]

10 Animal pharming can be described as which of the following:

A Growing animals for farming
B Genetically modifying animals to produce novel products
C Producing transgenic animals for farming
D Creating clones of useful animals **[1]**

11 A micro array is an ordered array of microscopic elements on a flat substrate that allows the specific binding of which of the following:

A A gene or gene products
B A whole genome
C A virus
D Coloured markers **[1]**

12 Which of the following statements about nucleic acids contained in liposomes used for gene transfer in plants are true?

i They are protected from nuclease digestion
ii Nuclei acids are stable in liposomes
iii Nuclei acids are unstable in liposomes
iv They are not protected from nuclease digestion

A i only
B i and ii only
C i and iii only
D iii and iv only **[1]**

Option C Ecology and conservation

Understanding ecology and conservation enables us to appreciate the delicate web of interactions between organisms in different ecosystems and different parts of the world. Climate and other abiotic conditions influence where organisms can live. In addition, it is almost impossible for organisms to live in isolation – species interact to provide one another with food and shelter and each one influences the distribution of others. As human populations increase and encroach further into almost every part of the world, organisms are being lost and the Earth's biodiversity is decreasing. Conservation of threatened species and protection of important ecosystems can be achieved if there is international cooperation and understanding.

C1 Species and communities

Learning objectives

You should understand that:

- Certain factors limit the distribution of species.
- Keystone species can strongly affect the structure of a community.
- Owing to its particular spatial habitat and interactions with other species, every species plays a unique role within a community.
- Interactions between species in a community are classified according to their effect.
- Two different species cannot survive together in a habitat in the long term if they have identical niches.

Factors affecting the distribution of plant species

Organisms are said to live in communities. A **community** may be described by the geographical area it occupies (a lake community, for example), or by the dominant plant species present (coniferous forest, for instance). The organisms present in a community depend on the other organisms living there, as well as on the non-living, **abiotic** aspects, such as soil or climate. The distribution of plants in communities depends on a number of these abiotic factors. Any of the factors described below can limit the chances of survival of an individual or a species. If there is insufficient light or if the temperature is too low, for example, plants will die so these conditions are known as **limiting factors** for the distribution of plant species.

Temperature

No plant can survive freezing conditions for very long because to grow and reproduce plants must carry out chemical reactions within their cells that require enzymes. In arctic climates, plant growth is often very slow because enzymes work slowly at the low temperatures, but it shows seasonal variation as the rate of growth picks up during the relatively short summer. In tropical areas, like rainforests, growth is usually rapid and continuous because temperatures are warm and there is little seasonal variation in temperature.

Water

All plants require water. It is the universal solvent in their cells, the substrate for photosynthesis, and their transport medium. However, plants have evolved a variety of mechanisms to survive periods of drought. Some species remain dormant, some (such as cacti and succulent plants) store water, and others complete their life cycle in a brief rainy season.

Figure C.1 The spines and hairs on a cactus help to deflect harmful ultraviolet rays in sunlight.

Light

Plants need light for photosynthesis. Many use the changing day lengths of the different seasons to trigger flowering. Where light intensity is high, as in a desert, plants have evolved mechanisms to prevent damage to their chlorophyll, such as dense spines or white hair that reflects light (Figure **C.1**).

Where light levels are low, as they are at ground level in a deciduous forest in the northern hemisphere, some plants grow and complete their annual life cycle in the early part of the year, before overshadowing trees have come into leaf.

Soil pH

Most plants prefer a pH of 6.5–7.0 because nutrients are easily available in this range. Some soils are slightly alkaline because they are based on chalk. Chrysanthemum and lavender are two examples of plants that tolerate alkaline soils well and are found in chalky areas. Other soils are acidic; beech, spruce and camellia can grow here. Peat bogs are very acidic because they are composed of decomposing organic material. Very few plants can grow here, although heathers can survive in acid soils.

Salinity

Saline (salty) soils present a particular problem to plants because they make it difficult for them to take up water and minerals. Some plants absorb salt in the soil, secrete it in their leaves and then drop these leaves to remove the salt. A few plants, such as marram grass and lyme grass, can survive in saline conditions.

Mineral nutrients

Soils that are rich in minerals can support a diverse community of plant species, including trees and shrubs. Plants that survive in mineral-poor soils often have special adaptations to supplement their needs. Carnivorous plants such as sundew and Venus flytraps live in very peaty soils that are deficient in nitrogen (Figure **C.2**).

Figure C.2 The sundew (*Drosera rotundifolia*) is a carnivorous plant that attracts, kills and breaks down insects, for their protein. In this way, it can absorb amino acids and use these to make plant protein and other nitrogenous compounds.

Factors affecting the distribution of animal species

Just as for plants, the distribution of animals is affected by the abiotic factors in their environment. If any factor required by an animal is in short supply or is unsuitable for survival, the distribution of the species will be limited by that factor.

Temperature

Animal enzymes are influenced by temperature in much the same way as those of plants. However, animals have the advantage that they can move to avoid the harshest of conditions. In hot, arid areas like deserts, many animals avoid the heat of the day and burrow underground. The jerboa (*Jaculus jaculus*) has long legs that keep its body off the hot sand and its ears have a large surface area, enabling the animal to lose heat efficiently (Figure **C.3**). Birds and mammals can control their internal temperatures but other species use behaviour and other adaptations to maintain theirs.

Some animals, such as the hedgehog, hibernate to overcome the stress of cold winters. Many bird species migrate to warmer climates during wintry seasons.

Figure C.3 The jerboa is a desert rodent with both behavioural and physical adaptations that help modify its temperature in the fierce heat of its habitat.

Water

Most animals need to drink water to survive – very few have evolved to be independent of liquid water. Some desert animals like the jerboa (Figure **C.3**) have done this, however. Jerboas eat seeds and, as the stored carbohydrate is respired in their cells, it produces all the water these animals need – they do not actually drink any liquid water.

Lack of water in certain seasons may change the distribution of animals. Herds of wildebeest and zebra in Africa undertake huge migrations to find new supplies of water and, therefore, vegetation. Carnivorous species often follow these herds, which are their source of food.

Breeding sites

Animals need to find appropriate sites to express mating behaviour and then rear young. These sites may be chosen for safety away from predators, or because they provide rich feeding grounds so the young may benefit. Different species have their own requirements. Many frogs and toads live almost entirely on land but their distribution is limited because they must return to water to breed.

Food supply

Unlike plants, which are autotrophic, animals need a source of food. Herbivores need plants and carnivores need other animals to feed on. The availability of food will determine the distribution of different types of animal. Some animals are restricted to a particular area because it supplies their food – so, for example, rabbits are usually found on grasslands. Others, such as lions, have huge territories and may cover many kilometres searching for food. Animals that have a varied diet are generally more successful and have a wider choice of habitats. If one source of food becomes scarce, they can move on to another.

Territory

Herbivores, such as wildebeest, that live in herds, graze on large areas of grassland and, when the dry season arrives, migrate to find fresh grass. Some birds, such as the European robin, live in smaller numbers and have less need for space but males defend their territories vigorously because they contain food and a nesting area. Carnivores, such as wolves, that live in packs require a large area in which to hunt. They may mark their territory with scent and defend it from other packs. Others, like eagles and buzzards, live solitary lives and have a large hunting territory because their prey is hard to find.

Investigating distribution of species – random sampling

When ecologists want to understand the distribution of a species, or to compare the distribution of one species with another in a different location, it is usually impossible to do so by a direct counting method. In most cases, ecologists take a **sample** of the population and, if the sample is chosen at random, it should provide a good representation of the whole population. Random sampling is used if the area under investigation is large or if time is limited, and it assumes that every organism has an equal chance of being sampled (that is, of being selected as part of the sample).

There are a number of sampling methods used by ecologists to collect data on the distribution of species in relation to one another and to abiotic factors in their environment. Two common methods used are quadrats and transects. They can show not only which species are present, but also how many individuals of each species there are.

Quadrats

One of the simplest and easiest sampling techniques involves using a quadrat (Figure **C.5**). A **quadrat** is a square made of metal or wood that is placed on the ground so that the organisms present inside the square can be counted.

The size of the quadrat will largely be determined by what is being measured. To estimate the number of different trees in a wood may require quadrats of 10 m by 10 m, but a 1 m quadrat would be the best size for studying wild flowers in grassland. Very small 10 cm quadrats might be used for sampling lichens on walls or tree trunks.

Sampling bias

Placing quadrats in order to collect data on the distribution of species may not always be a truly random process. Researchers can introduce personal bias, even without meaning to, by placing a quadrat in a spot that they think will be more interesting or easier to work in. To ensure that the samples within a survey area are made completely randomly, a numbered grid of the area may be drawn up, and random number tables or generators used to select squares on the grid where a quadrat should be placed (Figure **C.4**). Random number tables and random number generators are lists of numbers selected by a computer without any human bias.

Questions to consider

- Is random sampling a useful tool for scientists?
- How significant is the potential for sampling bias and can this ever be completely avoided?

Figure C.5 Using a quadrat to sample a area of grassland.

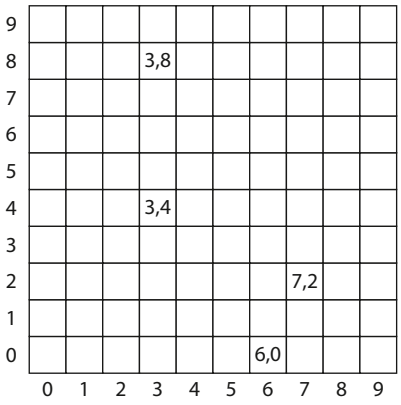


Figure C.4 To select a part of an area to sample with a quadrat, divide the area into a grid of squares, and then select a row and a column number using numbers generated randomly.

Transects

Another commonly used sampling method in ecology is a transect. A **transect** can show the distribution of a species in relation to a particular abiotic factor or it can give an idea of successions or changes in communities of organisms across a habitat (Figure **C.6**). Transects can be used to sample the distributions of plants along a beach or through a field, or to study the changing vegetation as soil or moisture varies. Transects provide a method of systematic, rather than random, sampling.

To take samples along a transect, follow these steps.

1. Stretch a tape or rope from a fixed point for a selected distance across the changing habitat you are interested in. If you are studying a salt marsh or sand dunes above a beach, a distance of 100 m would be appropriate.
2. At intervals of 10 m, or another suitable distance, along the tape put down a quadrat and count the organisms inside it. A series of samples like this provides information about the changes in density and community composition along the transect.
3. Measure the abiotic factor of interest – such as temperature, salinity, soil pH or light intensity – at each quadrat location.

The best type of transect to carry out depends on the terrain and on the organisms present. It may be better to carry out a point transect, where organisms are recorded at specific sampling points along the tape. On the other hand, a continuous 'belt' transect where all species in a 1 m zone along the transect are recorded, might be more helpful in providing a detailed picture of the area.

Figure C.6 These students are using a transect line to survey the plants in a grassy area. A quadrat is placed at measured intervals along the transect line and the plants at each location are counted and recorded. In this way, the plant population can be estimated from a series of samples in a few areas.

Stress zones and limits of tolerance

The distribution of a species is affected by limiting factors, and in the 'zones of stress' at the very edges of their ranges all species struggle to survive and thrive. Various environmental factors can be important in different cases and species are also affected by competition in their habitat. Two examples of the ways in which species survive at the very limits of tolerance are examined here.

Distribution of bristle cone pines in Colorado

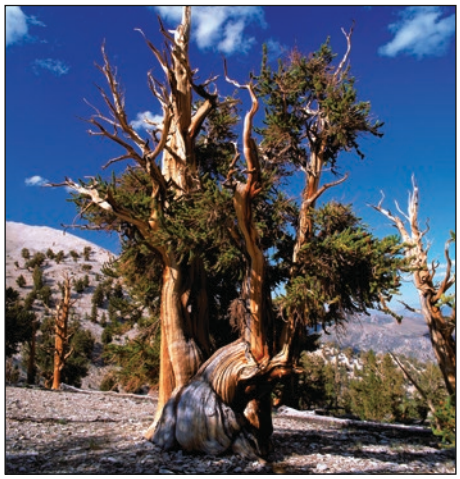


Figure C.7 The bristlecone pine (*Pinus aristata*).

Tree lines at high elevations on mountainsides are easily seen from aerial photographs. For most tree species, distribution is limited by factors that occur as altitude increases. The bristlecone pine (*Pinus aristata*) is a conifer native to the USA, which grows in Colorado and other states at altitudes of between 2100 and 4000 m (7000–13 000 feet) (Figure **C.7**). It is able to live on exposed, cold, dry rocky slopes and high mountain ridges but its limit of tolerance is 4000 m, above which environmental conditions are too extreme for it to survive (Figure **C.8**).

The trees' appearance is determined by climate; at highest altitudes close to the alpine zone it grows as a small tree, while at slightly lower altitudes it grows in a Krumholtz formation with stunted growth caused by exposure to freezing winds. Below the timberline trees can reach a height of 12 m. The growth of the trees is particularly affected by the weather conditions of the previous year. Scientists have identified several key factors that explain how the pattern of the tree line indicates the **limits of tolerance** of the trees to key environmental conditions.

Low temperature and desiccation

In the coldest part of the winter, frost can damage cells by freezing them. This is a particular problem for tall trees, which are more exposed to the atmosphere, and partly explains why the height of trees decreases with altitude. Another reason for this pattern is that when the soil is frozen, water cannot be taken in and tissue becomes desiccated so that trees are unable to reach great heights at high elevations.

High winds and weight of snow

At high elevations, the weight of snow and ice can break tree branches, and in addition high winds above the tree line cause damage to the trees' needles. Wind damage is more of a problem for tall trees than for shorter alpine vegetation, which can survive beyond the tree line.

Insufficient light

Although trees may be adapted to survive at cold temperatures, beyond a certain altitude the combination of low temperature and short growing season means that a tree cannot undergo photosynthesis for long enough during the year to survive. When photosynthesis is limited, the supply of sugars and amino acids needed to make new cells and tissues is low and growth and development cannot take place.

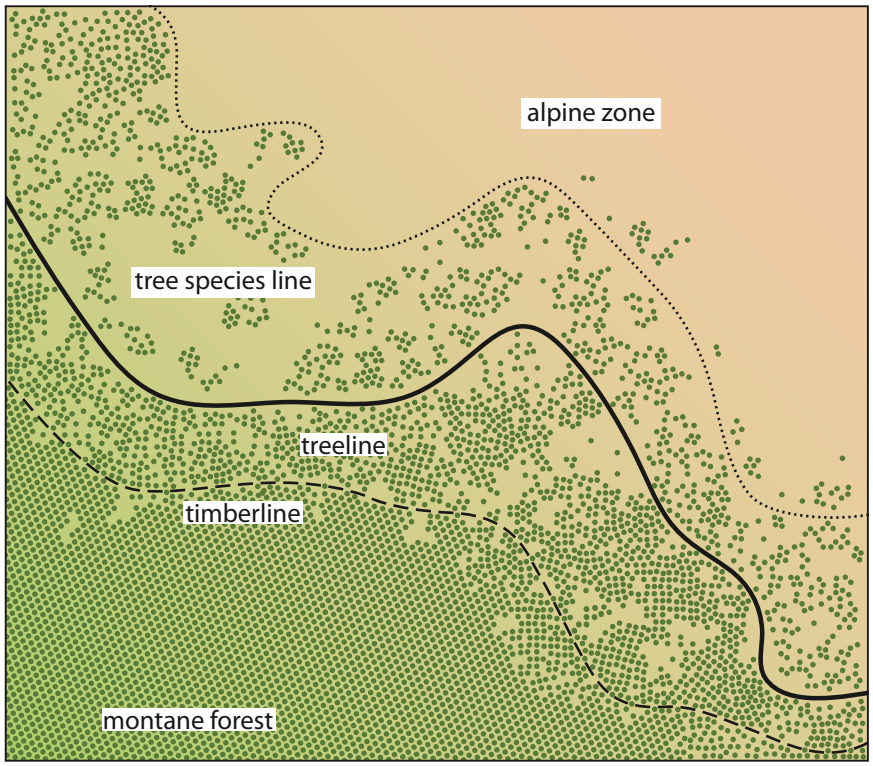


Figure C.8 The different zones of growth in Colorado mountains. The timberline is the limit of tall forest, the treeline is the limit of groups of trees over 3 m high, and the tree species line is the limit of all individual trees.

Distribution of the common limpet on European shores

Common limpets (*Patella vulgata*) are a feature of temperate rocky shores and are found throughout Europe, from Norway to Portugal. Limpets live clinging to rock surfaces where their shells grow to perfectly match their habitat (Figure **C.9**). Limpets are considered to be a **keystone species** of rocky shores because they keep levels of the small fucoid algae (*Fucus* spp.) on which they feed, under control.

Common limpets live in the mid-littoral or **intertidal zone**, an area of the seashore that is covered and uncovered twice a day by the tides. Limpets are adapted to being exposed to air and immersed in sea water but their range is limited by their tolerance of several of the conditions on the seashore, including those described below.

Keystone species a species that is important in maintaining community structure; keystone species may be either herbivores or carnivores that reduce competition in other trophic levels, so that community diversity is sustained

Exposure

In order to live in the intertidal zone, limpets must tolerate extremes of temperature and desiccation. As the tide recedes, limpets are exposed to the sun and their bodies may reach very high temperatures without a cooling covering of water. Limpets also absorb oxygen for respiration from sea water. If they are uncovered for too long they will not be able to survive the heat stress and lack of oxygen.

Limpets have adaptations that help them to tolerate and survive extremes of exposure. Individuals living higher up on the shore, where they are further from the sea and therefore exposed for longer periods between high tides, tend to have higher domes to their shells than those closer to the sea. This shape reduces the ratio of shell aperture (opening) to body size, so the limpet can grow larger without losing more water from the aperture. As well as their primary gills, limpets have a line of secondary gills around the edge of their shell so that they can use water that is trapped on the rock surface under their shells for respiration. They are also able to respire anaerobically while the tide is out and can tolerate an oxygen debt until they are covered with water again.

Figure C.9 Adult limpets usually return to the same area of rock after feeding. They form a small depression, known as a scar, by rubbing against the rock. This scar makes a tight fit for the shell so that the limpet can avoid water loss.

Danger of physical damage

If the sea is rough and waves are pounding the shore, animals can be dislodged from their positions and damaged or killed. Limpets produce sticky mucus, which binds their bodies to the rock surface – they are more tightly bound when the tide is in and conditions are rough. They also have a form of suction, which helps them to hold fast. Nevertheless, they cannot survive in areas where there is excessive disturbance.

Availability of food

Limpets are herbivores and graze on the small, thin fucoid algae that cover rocks. They are active foragers and travel across rocks when the tide is in, using chemical cues to return to their home spot. They sometimes use the edge of their shell to scrape away at rocks and remove the algae.

The distribution of limpets is limited by the availability of small algae for food. These algae cannot survive too far from the sea because they require water to respire and photosynthesise. In areas further down the shore, which are covered by the sea for most of the day, small algae are out-competed by larger species. So limpets do not have a suitable source of food high up the shore or much lower down by the sea – they are limited to an intermediate zone.

Maintaining community structure

A rocky shore contains many different species and limited space. Environmental gradients of temperature, water availability and food allow some species to survive in places where their competitors cannot.

Another factor affecting community structure is the presence of specific predators or grazers, which act as keystone species. For example, limpets are a keystone species, controlling the level of algae on the seashore as they graze and lobsters are a predatory keystone species. Lobsters in the North Atlantic Ocean used to be important predators of sea urchins. When the lobsters were subjected to overfishing, so that their numbers declined significantly, there were not enough lobsters to control the numbers of sea urchins. Sea urchin populations increased significantly and destroyed large areas of kelp, a species of seaweed (*Laminaria* sp.), which they grazed on and was their main source of food. As the kelp was removed, the complex community of molluscs and other small organisms that had lived on it and under it was also destroyed so that overall the diversity of species and the complexity of the food webs that existed in the area were much reduced. Control by predators like the lobster that reduce competition in lower trophic levels is known as **top-down control**. You can read more about this in Subtopic **C5**.

Community structure is also maintained by other interactions between species, including competition, as described in the next sections.

Habitats and niches

In all communities each species plays a unique role. This role is determined by its place in the habitat and the interactions that it has with other species.

A **habitat** is an area offering living space to a number of different types of organism, and includes all the physical and abiotic factors in the environment. An example might be a woodland habitat, whose community includes a huge variety of species, from burrowing invertebrates at ground level to nesting birds in the tree canopy.

Every organism occupies its own space in an ecosystem, which is known as its **spatial habitat**. The surroundings are changed by the presence of the organism – for example, a woodpecker lives inside hollow trees, adapting them to provide nesting places and shelter, while a rabbit burrowing underground affects the soil and plant species growing there.

A **niche** is the particular environment and 'lifestyle' that is adopted by a species. It includes the place where the organism lives and breeds – its spatial habitat – as well as its food and feeding method, and its interactions with other species. As an organism feeds within its niche, it affects the other organisms that are present. For example, an owl feeding on mice in woodland helps to keep the population of mice at a stable level, and rock limpets grazing on small algae control the degree of algal cover. A habitat comprises a number of niches, each of which is unique to its particular species because it offers the exact conditions that the species needs or has become adapted to.

Interactions between organisms

Organisms interact with other organisms living in the same community. The interactions include competition, herbivory, predation, parasitism and mutualism. Almost all organisms influence the lives of others and their interactions can be classified according to their effect.

Competition

Competition occurs when two organisms require the same limited resource. For example, if a pride of lions kills an antelope, they must protect this source of food from scavenging hyenas and vultures that will compete with them for the prey. In most cases, competition will lead to the exclusion of one species by another – as one uses the resource, less is available to the other, so that the less successful species may have to adapt to use some other resource if it is going to survive.

Plants also compete for resources such as light and space. Fast-growing birch trees quickly become established in areas of cleared land, but they require high light levels. Slower-growing species such as oak begin to grow up around them and for a while, they form a mixed woodland. Eventually the birch trees are over-shadowed and out-competed by the more dominant oaks.

Competitive exclusion

Loss of habitat, often caused by human activities such as farming or deforestation, severely limits vital resources such as food, water and breeding sites for the species that live there. When two different species require the same limited resources in the same area, they may find themselves in competition for the same niche. If they are prey species, they may become susceptible to the same predators as well. The principle of **competitive exclusion** states that no two species can occupy the same niche. The species cannot exist together because one will come to dominate and exclude the other. The oak and birch trees described above are an example of competitive exclusion. Both compete for soil resources and light but eventually the oak shades out the light and the birches die off.

In 1934, a classical study on competition was conducted by G. F. Gause (1910–86), a Russian ecologist. He experimented with two species of *Paramecium*, a large protozoan that is common in fresh water – *P. aurelia*

and *P. caudatum*. If the two species were allowed to grow in separate cultures on a food source of bacteria, both species grew well. When the two species were cultured together with an identical food source, *P. aurelia* survived while *P. caudatum* died out (Figure **C.10**). Both species had similar needs in the culture but *P. aurelia* had an advantage that enabled it to outgrow *P. caudatum*.

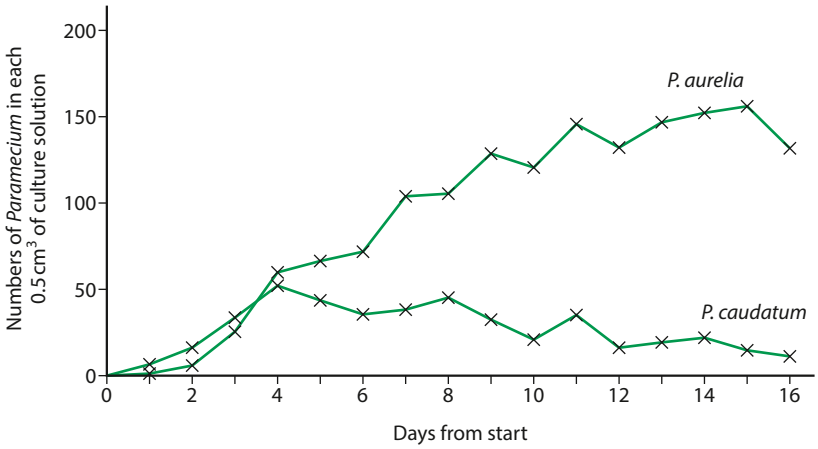


Figure C.10 Over the 16-day culture period, the population of *P. aurelia* increased while *P. caudatum* declined. *P. caudatum* was competitively excluded by *P. aurelia*.

Fundamental and realised niches

We have described a niche as the special space and 'lifestyle' inhabited by a particular species. This is the **fundamental niche** for that species. It is the *potential* mode of existence of the species, given its adaptations.

Often the environment will change through natural phenomena, competition or human intervention. So a species may find that its niche becomes more restricted or begins to overlap with that of another species. This more restricted life pattern is known as the **realised niche**. The realised niche is the *actual* mode of existence of a species resulting from its adaptations as well as from competition with other species. A realised niche can only be the same size as or smaller than the fundamental niche.

In Gausses' study with *Paramecium*, the fundamental niche of both *P. aurelia* and *P. caudatum* was the tank in which they grew alone. However, in a tank together each occupied a more restricted, realised niche where *P. caudatum* was outcompeted and failed to thrive as it became limited by *P. aurelia*. In an urban situation, normally wild animals like raccoons and foxes, whose fundamental niche is living in open countryside and hunting as predators, instead occupy a realised niche in which they scavenge on the waste left by humans.

Herbivory

Herbivory affects both the plants that are food providers and lose parts of their structure, as well as the herbivores that are able to grow and thrive as a result of the food they gain.

A single plant may provide leaves for browsing animals, fruits and seeds for birds, and roots for burrowing animals. The horse chestnut leafminer (*Cameraria ohridella*) is a moth that lays its eggs on horse chestnut leaves. As the larvae hatch, they burrow inside to feed on the tissues of the leaf. The nuts from the horse chestnut tree also provide food for squirrels and deer.

Other leafminer species feed on different tree species around the world, such as oak, birch and holly.

Predation

A well-studied example of the effects of **predation** is that of the Canadian lynx, which feeds on the arctic hare. The numbers of predator and prey fluctuate over the years with changes in the hare population being followed by corresponding changes in the numbers of lynx (Figure **C.11**).

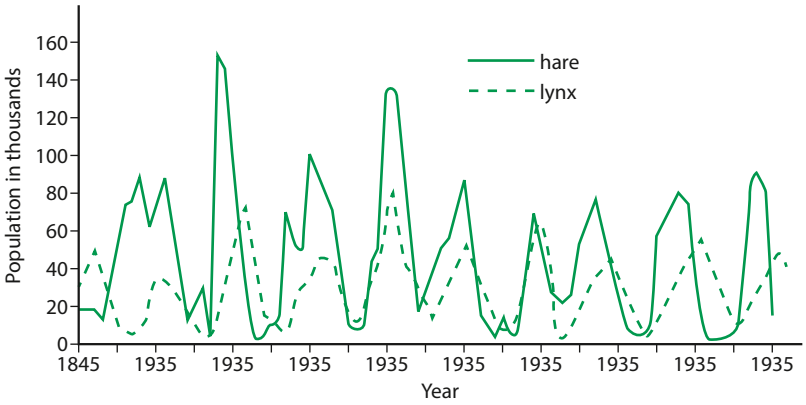


Figure C.11 Changes in the populations of the Canadian lynx and the arctic hare over time.

Parasitism

Parasites are organisms that live entirely on or in a **host** species and cannot survive without it.

Exoparasites, such as fleas and ticks, live on the outside of a host. One economically important example is the southern cattle tick (*Boophilus microplus*), which lives on cattle, feeding on their blood and weakening the animals. It causes significant losses to farmers all over the world.

Endoparasites, such as tapeworms, roundworms and malarial parasites, live inside their host. One example, the barber's pole worm (*Haemonchus contortus*) is a roundworm that lives in the stomachs of sheep in warm, humid climates. It causes anemia and progressive weakness as it feeds on blood in the sheep's stomach. If present in large numbers, this parasite can kill young animals.

Mutualism

Sometimes two organisms coexist and benefit each other, forming what is known as a **mutualistic** relationship.

Lichens such as common orange lichen (*Xanthoria parietina*), which grows on twigs and branches, are the result of a union between a fungus and an alga. The alga carries out photosynthesis and provides sugars for both organisms. The fungus protects the alga from intense sunlight and drying out and absorbs minerals for the benefit of both organisms.

Another mutualistic relationship occurs between the Egyptian plover (*Pluvianus aegyptius*) and the Nile crocodile. The bird feeds on parasites and food particles left around the crocodile's mouth, keeping its teeth clean and healthy. The crocodile openly invites the birds to hunt on its body, even allowing them to enter its mouth.

The relationship between zooxanthellae and reef-building corals

Most reef-building corals contain photosynthetic algae, called zooxanthellae, which live within their tissues. The corals and algae have a symbiotic, mutualistic relationship, in which both the coral and algae benefit from their association. The coral provides the algae with a protected environment and inorganic nutrients, while the algae produce oxygen and help the coral to remove waste products. Zooxanthellae also supply the coral with essential compounds such as glycerol, amino acids and glucose, produced from photosynthesis. The coral uses these products to make fats, proteins and carbohydrates, and to produce calcium carbonate. It has been estimated that up to 90% of the organic material produced by photosynthesis in zooxanthellae is transferred to the coral's tissues, enabling it to grow.

Zooxanthellae also produce the wide range of colours that are seen in corals. If corals are stressed by environmental factors, such as changes in pH or salinity of the water, they may expel the algae so that the colony becomes white. This phenomenon is known as 'coral bleaching' and can result in the corals' death.

Because of their relationship with zooxanthellae, reef-building corals must live in clear water so that their algae can receive sufficient light to photosynthesise, and so coral is usually found in quite shallow water with small amounts of suspended material. Such waters in the tropics are poor in nutrients, but the relationship between the algae and coral allows nutrient recycling, and so provides another mutual benefit to both organisms.

Mutualism and symbiosis – the importance of definitions and communication in science

In order for scientists to be able to communicate ideas and discoveries effectively, it is vital that agreed definitions for important terms are used. However, in practice definitions sometimes slide, diverge or even overlap, which can cause confusion. For example, **symbiosis** is usually defined as a close, long-term interaction between two species. But this definition is the subject of debate among scientists. Some believe symbiosis should only refer to 'persistent mutualisms', such as the interaction between reef-building corals and zooxanthellae, in which both organisms benefit. In this case, **mutualism** and **symbiosis** would mean the same thing. Others believe the term symbiosis should apply to any type of close interaction between two species from which one benefits, but not necessarily both. Under this broader definition, symbiosis would include not only mutualistic relationships, but also **commensalism**, in which the host organism is not affected by the presence of the 'guest', and even **parasitism**, in which the host is harmed while the parasite benefits. Mutualism would be just one type of symbiosis.

Questions to consider

- Do you think it matters if scientists define terms differently, as long as they state clearly the definitions they have used in their writing?
- To what extent is it more important for scientists to use agreed, precise definitions for terms than it is for poets, novelists or journalists?

Nature of science

Using models to study the real world – limits of tolerance graphs

Ecologists often use models to predict events in the natural world. Graphs such as the one shown in Figure **C.12** can be drawn to indicate the likely ranges of different species in different situations. Consider the graph and try to identify the environmental conditions that apply in the case of the limpet and the bristle cone pine. For each species, try to describe what conditions might be like in the 'zone of intolerance' where species are absent.

Knowledge of the stresses that apply to different species can enable ecologists to predict whether another species might be able to survive in a habitat or whether the species being studied could survive in a different location with different pressures.

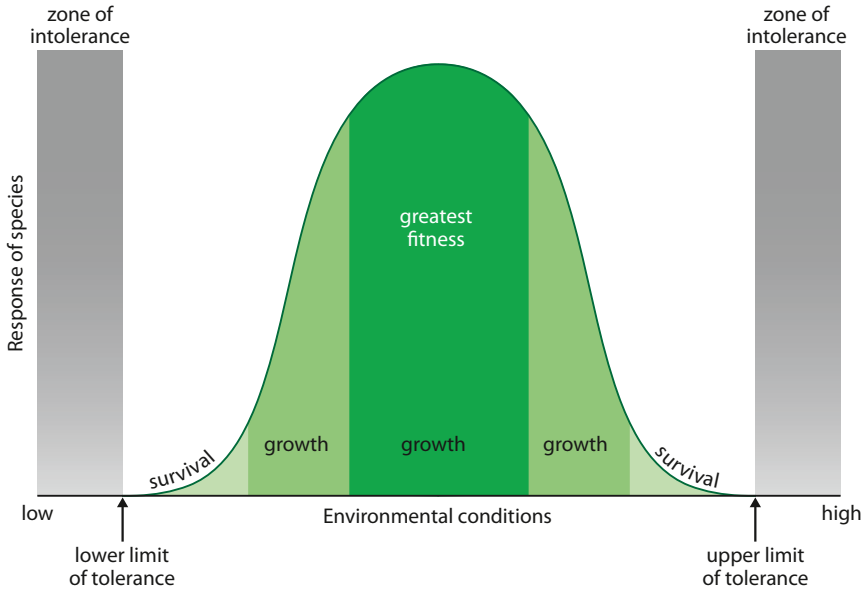


Figure C.12 The graph shows how the range of a species is limited at the upper and lower environmental extremes by zones of intolerance in which the organism cannot survive.

Test yourself

1 Outline **three** factors that affect the distribution of animals and why they are important.
2 Define the term 'competitive exclusion'.
3 Outline the difference between a fundamental and realised niche.

C2 Communities and ecosystems

Learning objectives

You should understand that:

- Most species occupy different trophic levels in a number of food chains.
- A food web is a diagram showing many possible food chains in a community.
- The respiration rate of an animal determines the percentage of its ingested food that is converted to biomass.
- The type of stable ecosystem that forms in an area depends upon the climate.
- In closed ecosystems, energy is exchanged with the surroundings, but matter is not.
- The structure of an ecosystem, and rate of change within it, is influenced by disturbance.

The difficulty of defining a trophic level

grass → rabbit → fox

In a simple food chain, such as the one shown above, grass is the primary producer, the rabbit is the primary consumer and the fox is the secondary consumer, so each organism is said to occupy a separate trophic level. In practice, simple food chains rarely exist – foxes do not feed exclusively on rabbits – so more complicated food webs must be constructed.

In the example of a woodland food web shown in Figure **C.13**, several of the organisms do not occupy a single trophic level because they have a varied diet. The fox could be said to be a primary consumer because it eats fruit, from the crab apple tree, it could also be classed as a secondary or tertiary consumer because it eats both rabbits (primary consumers) and great tits (secondary consumers). In addition, food chains and webs usually contain organisms that feed on dead material. These are the detritivores and saprotrophs, which do not fit into a particular trophic level. A food web shows many possible food chains in a community, and if analysed most show that most species occupy more than one trophic level in many different food chains.

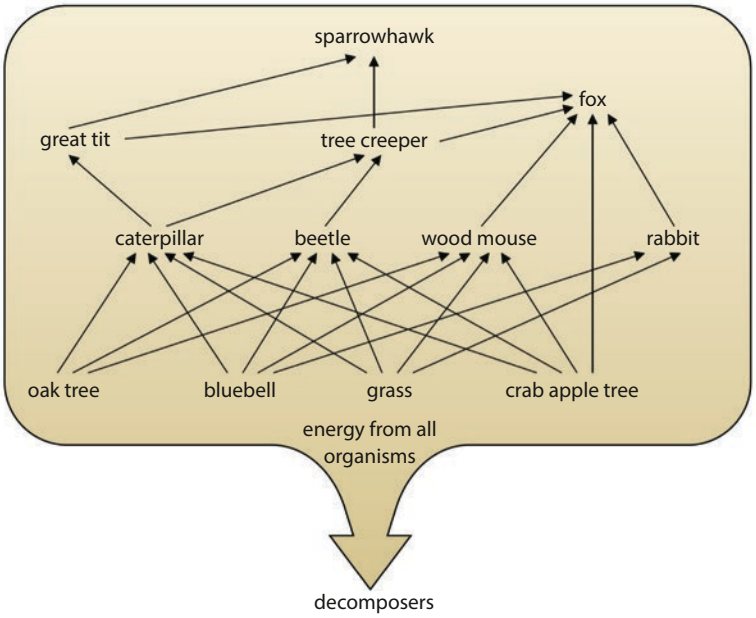


Figure C.13 In a food web such as this, species may feed at different trophic levels depending on what they eat.

Energy in ecosystems

Plants are the primary source of energy in nearly all ecosystems. They carry out photosynthesis and use energy in sunlight to build carbohydrates. They use these carbohydrates, and minerals from the soil, to make all the proteins, lipids and nucleic acids they need to grow. These processes are not 100% efficient and not all of the energy of the sunlight is used. Some is reflected at the leaf surface, some goes right through leaves without being used, and some is lost when plants respire carbohydrate for energy.

When herbivores feed, the energy transferred from plant to herbivore is also not 100% efficient. Not all of the plant material is eaten, not all the material is absorbed in the gut, and some energy is lost in movement and respiration. The same is true for carnivores eating prey animals. Only about 10% of the energy in producers is passed to herbivores and a similar low percentage of energy is passed from herbivores to carnivores (Figure **C.14**).

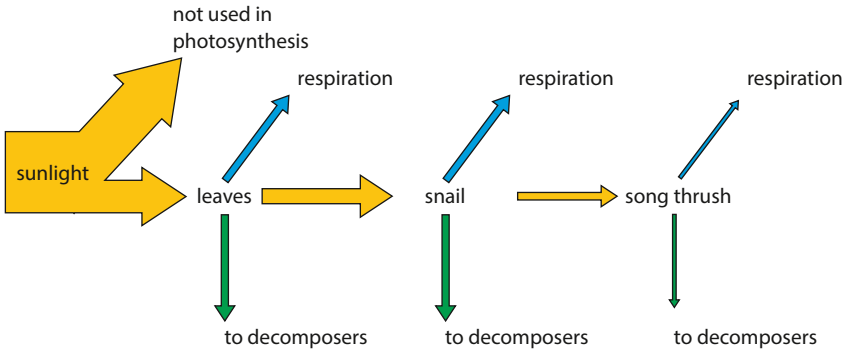


Figure C.14 Energy losses in a food chain (not to scale).

Ecologists show the availability of energy in an ecosystem in diagrams known as **energy pyramids**. Each layer of the pyramid represents the organisms at each trophic level, so layer 1 includes all the primary producers, layer 2 all the primary consumers and so on (Figure **C.15**). It is also possible to construct pyramids of numbers and biomass.

Pyramids of energy

To overcome the difficulty of categorising organisms which occupy multiple trophic levels, animals are often classified according to their main food source.

As one moves up a food chain or food web, energy is lost at each trophic level through respiration and waste. The efficiency of transfer from one level to the next is only about 10%. This is why ecosystems rarely contain more than four or five trophic levels. There is simply not enough energy to support another level. This is true in all ecosystems. The percentage of ingested energy which is converted to biomass in the bodies of animals in the food chain varies in different ecosystems. Pyramids of energy can be compared to demonstrate this (Figure **C.15**).

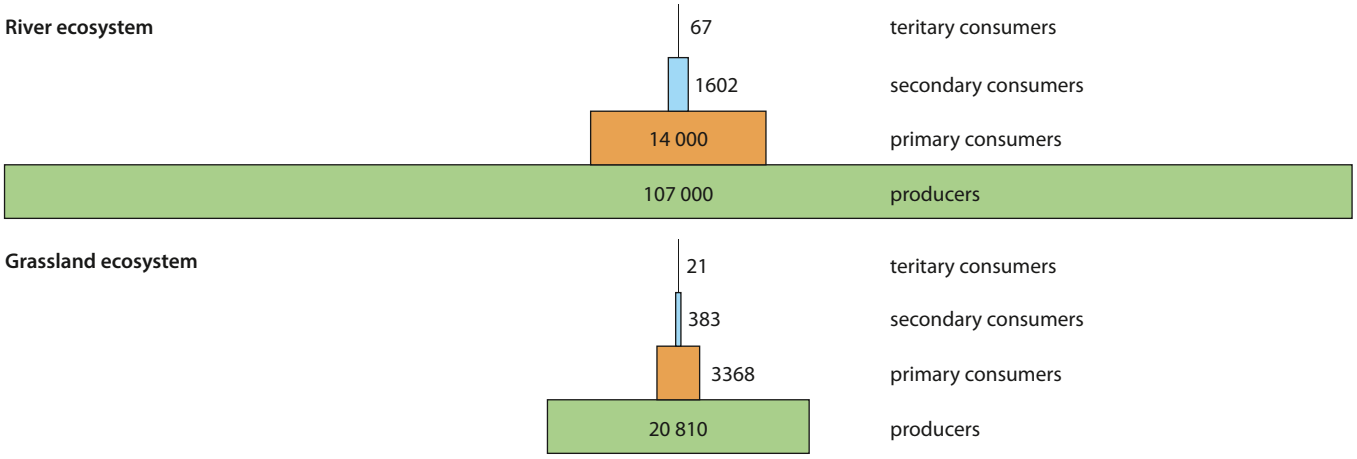


Figure C.15 Pyramids of energy for a river ecosystem and a grassland ecosystem. Each bar represents a trophic level and the width of the bar indicates how much energy it contains. Energy is measured in kJ m^{-2} y^{-1}. Only a small percentage of the energy in each level is transferred to the next.

Biomass

Biomass is biological material, living or dead, that can be used as an energy source. Since living material also contains water, which is not organic and does not contain energy, the biomass in a habitat or sample is usually measured as dry mass of organic matter in organisms, per unit area of land or unit volume of water.

Biomass does not include biological material that has been changed over time into coal or oil. There is much interest currently in using biomass as fuel in place of fossil fuels, because it is a renewable resource. Plants such as perennial grasses, hemp and sugar cane are undergoing trials as sources of industrial biomass.

Gross primary production the total amount of energy used by plants to make carbohydrates during photosynthesis

Net primary production the amount of energy in plants that is available to herbivores, per square metre, per year, after some energy has been lost through respiration

Gross production and net production

A pyramid of energy shows energy flow in an ecosystem. The lowest bar of the pyramid represents **gross primary production**, the total amount of energy that flows through the producers. It is measured in kilojoules of energy per square metre per year (kJ m^{-2} y^{-1}).

Net primary production is the amount of energy available to herbivores from producers after subtracting the energy used by the plants for respiration. This can be represented as:

net production = gross production – energy lost in respiration

Similar calculations can be carried out for each successive trophic level and the data used to construct a pyramid of energy like those shown in Figure **C.15**.

The percentage of ingested energy that can be converted to biomass depends on the respiration rate of the ingesting organism – the greater the amount of energy the animal uses for its own needs, the less there is available to convert to biomass. Animals that are poikilotherms (and have a variable body temperature) are more efficient producers of body mass than homeotherms, which use a lot of energy to maintain a regulated body temperature.

Food conversion ratios

In commercial food production, farmers measure the food conversion ratio (FCR) of their animals. The FCR is a measure of an animal's efficiency in converting food mass into increased body mass (biomass). It is calculated by dividing the mass of food eaten by the gain in body mass over a period of time.

Animals that have a low FCR are considered efficient users of the feed they are given. Pigs have an FCR of about 3.5 while farmed salmon have an FCR of 1.2. It is not possible to compare animals directly because there are differences in the composition and energy content of their foods, but approximate values for other livestock are shown in Table **C.1**.

Animal	Approximate FCR
farmed salmon	1.2
poultry	2
pigs	3.5
sheep	8
cattle	8

Table C.1 The food conversion ratios of some farmed animals.

The low FCR for farmed salmon is due to a number of factors:

- Their (commercially produced) food has a high energy content.
- Salmon use and retain the protein in their food very much more efficiently than other farmed animals.
- Salmon live in water and use very little energy to support their bodies compared with land animals.
- Salmon are poikilotherms (that is, they have a variable body temperature) and need less energy than homeotherms (which maintain a regulated body temperature) to sustain body temperature and functions; their respiration rate is lower.

In natural systems too, energy conversion along aquatic food chains tends to be more efficient than in terrestrial food chains. But because the absorption of light in water is less efficient than on land, the initial capture of light energy by aquatic primary producers tends to be lower than for land plants.

Succession and stability

The processes of succession

Succession is the process of change to communities in a particular area over a period of time so that the appearance of the whole area evolves and changes. Succession involves interactions between both the biotic and abiotic components of the area. If an area of land is left bare as a result of an event such as a fire or land clearance, early 'pioneer' communities modify the physical environment, which, in turn, modifies the biotic community. This enables more species to move in and modify the physical environment still more, and so on until a stable situation is reached.

The different stages of succession are known as **seral stages** and the final stable community, which remains unless there is further disturbance, is called a **climax community**.

A **primary succession** begins when an area of bare ground or rock, with no existing soil, is colonised for the first time. Two well-studied examples are the area on the Indonesian island of Krakatau, which was left bare when the volcano Krakatoa erupted in 1883, and the newly formed volcanic island of Surtsey off the coast of Iceland, which formed in 1993.

The first organisms to colonise bare rock are lichens and mosses, which can settle on the rock surface. Lichens gradually break up the rocks and use dissolved minerals for growth. As lichens die they decompose, leaving debris, which begins the formation of humus and soil. Low-growing lichens and mosses modify the environment sufficiently for seeds of grasses and small shrubs to start growing and these plants modify the land still further. A deeper layer of soil develops as plants die and decompose, and this soil can hold more moisture and contains more organic matter. Later, fast-growing trees such as rowan and birch begin to grow and, as they extend their roots, the soil is bound together and protected from erosion. Eventually these trees will be replaced by slower-growing species, which form a climax community, usually after a period of about 100–200 years.

A typical succession in the northern hemisphere might be:

bare rock → lichens → mosses → grass and small shrubs → fast-growing trees → slow-growing trees

Approx. years:	2–3	3–10	10	15–20	25–50	100–200

Each stage is known as a **sere** or **seral stage**. Seres of particular environments tend to follow similar succession and can be classified according to the environment. For example, a hydrosere develops in water and a halosere in a salt marsh.

Secondary succession occurs where there has been a land clearance, perhaps by fire or landslip. An ecosystem has been established but is replaced as conditions have changed. Soil is already present so secondary succession is usually much quicker than primary succession and a variety of plants such as annual grasses and low-growing perennials can colonise rapidly.

Energy flow and productivity at different stages of a succession

As a succession develops in an ecosystem, it is not only the species present that change. The productivity of the system also changes.

In the first stages of a succession when there are few producers present, gross primary productivity is low but the proportion of energy lost in respiration by these organisms is also low. This means that net primary productivity is relatively high, and the ecosystem grows and accumulates biomass.

In the later stages of succession, there are more consumers present and gross productivity may be high. With more consumers there are more complex feeding interactions and food webs so that net productivity also increases.

Stability

The type of stable ecosystem, known as a climax community, that emerges following a succession depends on the local climate and, to a certain extent, this can be predicted. Rainfall and temperature are the two most important factors that determine the appearance of a stable ecosystem. Figure **C.16** shows a climograph – a diagram relating the prevailing type of ecosystem to conditions of temperature and rainfall. A climograph can be used to predict the type of stable ecosystem that will emerge in an area, from information about the mean annual temperature and mean annual precipitation in the region.

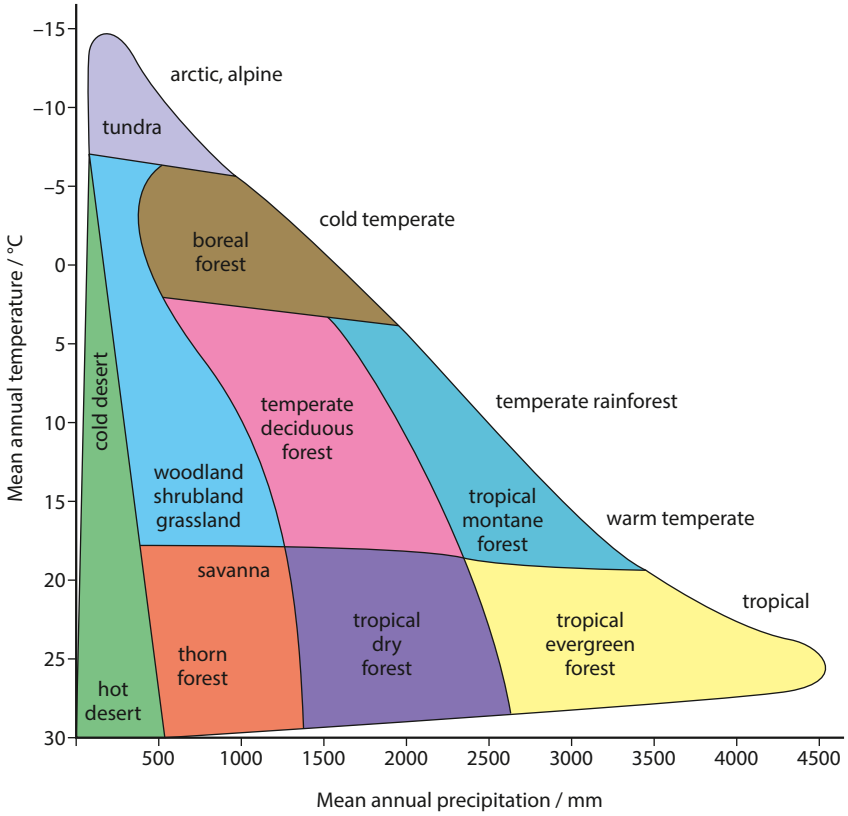


Figure C.16 A climograph shows the differences in vegetation, and therefore in the type of ecosystem, in different climatic regions.

The systems approach

Many ecologists now study ecosystems using the **systems approach**. A system is defined as an assemblage of parts and the relations between them that enable them to work together to form a functioning whole (Subtopic **1.1**). The systems approach studies an ecosystem as a whole, rather than examining individual parts such as a food chain within it. Systems are divided into three types: open, closed and isolated. Living systems may be either open or closed.

An **open system** exchanges both matter and energy with its surroundings across the boundaries of the system. Most living systems and all ecosystems are open systems, which exchange energy and matter with their environment. These open systems and the exchanges that take place can be seen in any living environment.

In a woodland ecosystem, the main inputs include light and carbon dioxide, which plants convert during photosynthesis. Further inputs come from woodland herbivores, which return mineral nutrients to the soil, and bacteria in the soil, which fix nitrogen from the air. Outputs may include water, which is lost during respiration and transpiration, nutrients, which flow away in waterways, and heat, which is exchanged with the environment around the woodland.

In a **closed system** energy, but not matter, is exchanged across the boundaries of the system. These systems are very rare in nature. Most examples are set up for experiments and are artificial. A bottle garden or an aquarium can be set up so that light and heat are exchanged across their boundaries but matter cannot be exchanged or leave the system. In most cases these systems do not survive because they become unbalanced.

Sometimes organisms die as oxygen is depleted or food runs out, or waste matter builds up to toxic levels. Figure **C.17** shows how open and closed systems are represented as diagrams.

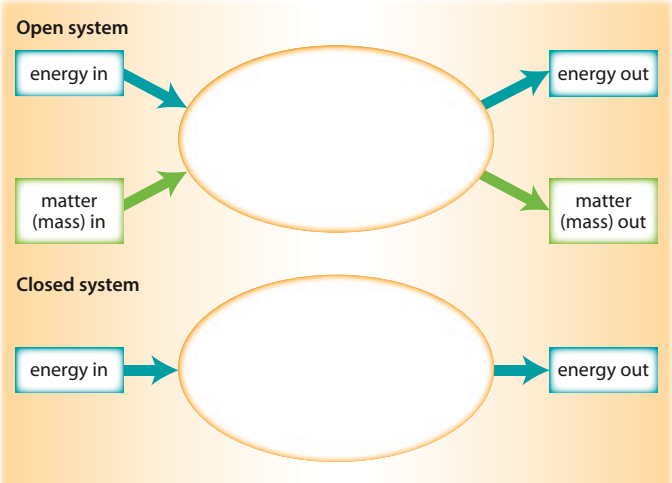


Figure C.17 Open and closed systems.

Ecosystems and disturbance

Disturbance in an ecosystem influences both its structure and the rate of change within it. If there is a disturbance, either a natural event such as a flood or drought, or an event caused by human interference, a decline in the numbers of top carnivores may be one of the first observable signs that something has changed. Extreme climatic events can destroy all or part of a food web. If plants die in a drought, all the consumers in the food chain, right up to the top carnivores, will be affected and the system may take time to recover. The nutrient cycles within the system (Subtopic **C6**) will also be disrupted.

As the human population grows and disturbs or destroys natural ecosystems, many large carnivores have come under severe threat. Carnivores require territories with enough space to hunt and their numbers fall significantly if human populations expand and take over their habitats. In Borneo, large areas of tropical rainforest have been cleared for palm oil plantations and the areas that remain are often separated from one another by roads or housing. This fragmentation of the natural forest has interrupted food chains and led to significant changes. Numbers of clouded leopards (*Neofelis nebulosa*) in north-eastern Borneo (Sabah) have fallen rapidly as their territories and prey have been destroyed so that now the leopard is classified as a 'vulnerable' species. Natural cycling of nutrients is also interrupted when land is cleared and to maintain soil fertility and crop yields more artificial fertilisers must be used.

Pesticides can also disrupt ecosystems. They are used to improve human food production but they also affect the way ecosystems function. The first pesticides made in the 1940s and 1950s were non-specific and so killed many different species indiscriminately, including both pests and useful pollinating insects. One well-documented case is that of DDT, an insecticide that was used by farmers to reduce losses and maximise crop yields. When it was first used, no one understood that DDT was not biodegradeable and remained poisonous in the environment for a long time. Although only low concentrations of DDT were used each time, small amounts soon accumulated in the environment and in organisms' bodies, leading to their deaths and to disruption of ecosystems. This process is known as

biomagnification (Subtopic **C.3**). Once again, humans were first alerted to the problem by the deaths of predatory birds, the top carnivores in the ecosystems.

Nature of science

Using models to study the real world – modelling energy flow in ecosystems

Gershmehl diagrams and pyramids of energy model the flow of energy through ecosystems. Although both have inaccuracies, the models and others like them provide a useful representation of the natural world that can be used to promote discussion between scientists. Models are also very useful in making hypotheses and predictions about future events. Understanding energy flow can enable computer models to be built to predict the effect of a disturbance to an ecosystem or the possible consequences of climate change.

Gersmehl diagrams – do the entities in scientists' models actually exist?

In 1976, the geographer and scientist P. F. Gersmehl developed a model of **nutrient cycling** to highlight differences between ecosystems. His diagrams (Figure **C.18**) show how nutrients are transferred and stored between three different parts of an ecosystem: the litter, biomass of organisms and the soil. As nutrients cycle in an ecosystem, there are interactions between the atmosphere and soil, and many food chains are involved. Nutrient cycles are different in different ecosystems and the rate of nutrient transfer is dependent on the amount of moisture, heat, vegetation and the length of the growing season. Diagrams can be drawn for different ecosystems and provide insights into systems that have high levels of nutrients in the soil or a large biomass of organisms.

Questions to consider

- Are Gersmehl's diagrams real representations of the world or are they useful inventions to explain the natural world?
- Do trophic levels actually exist or are they simply a human strategy to predict and explain feeding relationships?

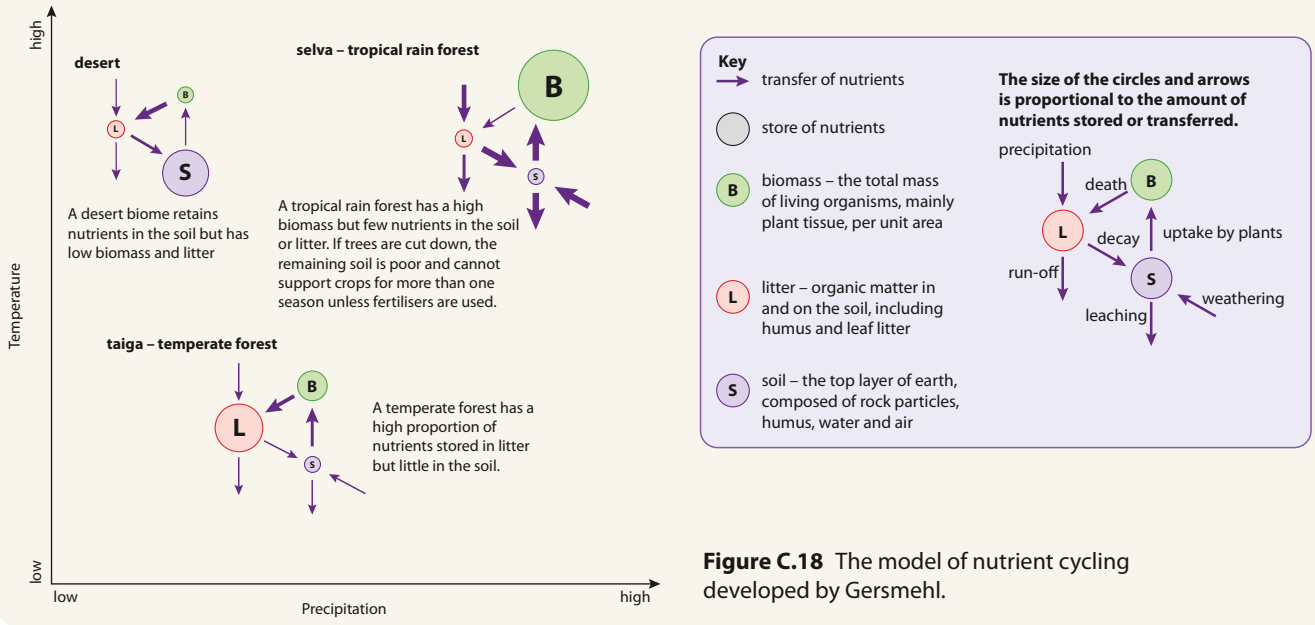


Figure C.18 The model of nutrient cycling developed by Gersmehl.

Test yourself

4 Outline the difference between a closed and open ecosystem.

5 State what is meant by the term 'conversion ratio'.

6 State **two** factors that influence the type of stable ecosystem that develops in a given area.

C3 Impacts of humans on ecosystems

Learning objectives

You should understand that:

- Introduced alien species can become invasive if they escape into local ecosystems.
- The numbers of endemic species can be reduced by invasive alien species, which can cause competitive exclusion, especially if they have no natural predators.
- Biomagnification can lead to the accumulation of pollutants in the tissues of organisms at higher trophic levels.
- Both macroplastic and microplastic debris has accumulated in marine ecosystems.

Interfering with ecosystems – introducing alien species

An **alien** species is one that is not native to the region in which it is found. There have been many occasions throughout history when an organism has been introduced from one ecosystem to another, either:

- accidentally
- deliberately
- or for biological control of a pest organism.

Accidental introduction

The zebra mussel (*Dreissena polymorpha*) is a small freshwater species, originally native to lakes in southeast Russia. It has been accidentally released in many other areas, probably carried in ballast water of cargo ships. It has become an invasive species in many different countries. Zebra mussels are now found in the Great Lakes of the USA where they grow on docks and boats. They have spread into streams and rivers and block water pipes and interfere with water supplies (Figure **C.19**). In some areas, they have out-competed all other freshwater mussels because they grow in dense clumps. Zebra mussels are also believed to be the source of deadly avian botulism poisoning that has killed tens of thousands of birds in the Great Lakes since the late 1990s. On the other hand, zebra mussels are thought to be partly responsible for the increase in the population of bass and yellow perch in the lakes. Zebra mussels are filter feeders and remove pollutants from lake water, which becomes clearer as a result. Algae deep under the water receive more light and grow more vigorously, providing habitats and food for the fish.

Figure C.19 Masked workers use a water jet to clear zebra mussels clogging the walls of the pump room of Detroit Edison's power station in Michigan, USA. Not only do zebra mussels encrust water pipes and pump rooms, but they also excrete a corrosive substance.

Deliberate introduction

Many plants, collected in distant regions, have been deliberately introduced to domestic gardens because of their attractive flowers or exotic foliage. Orchids, bamboos and rhododendrons are now seen all over the world but most were introduced following plant-collecting expeditions in the 19th and 20th centuries.

Much of the time, introduced species create no problems. However, in some cases, an introduced species finds the new conditions so advantageous that it becomes **invasive**. It grows rapidly and becomes a threat to **endemic** (native) species, which it out-competes and eventually eliminates. One such example is Japanese knotweed (*Fallopia japonica*), which was deliberately introduced into European gardens in the 19th century for its attractive flowers. It reproduces vegetatively and even short sections of root can re-grow to become whole new plants. This plant now covers huge areas of land in Europe. It can be controlled with herbicides, but there is a problem using these chemicals near rivers, as the herbicide gets into the waterway and upsets its ecological balance, harming plant and animal life.

The principle of **competitive exclusion** states that no two species can occupy the same niche indefinitely. The species cannot exist together because one will come to dominate and exclude the other. Both rhododendron and Japanese knotweed have competitively excluded native species.

Introduction for biological control

Another example of a deliberately introduced species is the prickly pear cactus (*Opuntia* sp), which was introduced to Australia as a source of cattle feed. The prickly pear rapidly grew out of control. At its height, it was spreading at a rate of 400 000 hectares per year. The dry, hot climate of Australia was ideal for this plant and there were no native animals that would eat it. Scientists conducted research to find a natural consumer for the prickly pear – they found that in its homelands of the USA and Mexico the prickly pear is eaten by the caterpillar of the cactus moth (*Cactoblastis cactorum*). This caterpillar was therefore also deliberately introduced into Australia and now keeps the plant under control. This is an example of successful **biological control**.

A far less successful attempt at biological control has proved disastrous for much of the wildlife of Australia. The Puerto Rican cane toad (Figure **C.20**) was introduced into Queensland in 1935 in an attempt to control sugar cane beetles, which were causing huge losses to cane growers in the north of Australia. In their native regions of Central and South America, cane toads are controlled by a number of predators, particularly snakes. In Australia, potential predators were not adapted to deal with the cane toad's skin, which produces dangerous toxins, so the toad population has grown out of control – so much so that they have spread from Queensland to Northern Territory and New South Wales, wiping out the endemic amphibians, which can live for an average of more than ten years, and who breed more slowly and later in the season than the cane toad. The toads also failed to control the sugar cane beetle, preferring to eat small rodents, insects and even dog food.

The introductions of alien species described above are summarised in Table **C.2**.

Figure C.20 The cane toad's (*Bufo marinus*) large size (up to 15cm in length) and mating behaviour have enabled it to out-compete native amphibians.

Species	Reason for introduction
Japanese knotweed	deliberately planted in European gardens
zebra mussels	accidentally introduced into USA
cane toad	deliberately introduced to control sugar cane beetles in Australia (unsuccessful)
cactus moth	deliberately introduced to control prickly pear cactus in Australia (successful)

Table C.2 Examples of species that have been introduced into new ecosystems.

Restoration of invaded areas

Restoring an area of land invaded by an alien species to its natural state may be extremely time-consuming and expensive. In 2004, a 6-year restoration programme was started on Montague Island in New South Wales, Australia. The island had become covered with kikuyu grass (*Pennisetum clandestinum*) and other non-endemic plants that had been planted in the 1900s to help stabilise the sandy soil and provide food for grazing animals. The kikuyu grass had spread to such an extent that it had displaced seabird nesting areas and was responsible for the death of significant numbers of the native little penguins, which became trapped or strangled in the grass. The grass was also a significant threat to other bird species, such as the shearwaters and crested terns that nest on the island. Management techniques included clearing the grass by controlled burning and spraying with herbicide, followed by re-vegetation of the island with endemic plant species.

In the 20th century, large areas of Snowdonia National Park in North Wales, UK, had become overgrown with rhododendron, which flourished in the wet climate. This plant, which is native to China, had been introduced widely in Britain in the 19th century as a garden shrub because of its very showy flowers. To restore the land, the thick branches had to be cut and the roots pulled out to prevent re-growth. Rhododendron forms an association with certain soil fungi, which prevent the germination of seeds of many other plants. So, even when the ground had been cleared, it had to be left for some time until these fungi died.

DDT and biomagnification

Biomagnification the process that leads to accumulation of chemical substances in food chains; the chemical substances become more concentrated at each trophic level

Some chemicals used in the environment as pesticides are taken into living organisms but then accumulate in their body tissues because the organism cannot break them down and excrete them very well. Insecticides such as DDT and dieldrin are well-studied examples of the way toxic chemicals can accumulate in the environment – in a process called **biomagnification**.

Small quantities of these substances, used to control insect pests, may be taken up by plants, or deposited on the surface of their leaves. The plants may be unaffected, but when primary consumers feed on the sprayed plants they take in a far greater quantity of the toxin. The chemical remains in the bodies of the primary consumers and if a secondary consumer feeds on a number of these animals, it accumulates an even greater amount of the chemical.

Rachel Louise Carson (1907–1964)

Rachel Carson was a writer and ecologist, born in the rural town of Springdale in Pennsylvania, USA. Having graduated from Pennsylvania College for Women in 1929, she went on to receive an MA in zoology from Johns Hopkins University in 1932, and then began a career as a writer and scientist working for the government.

Carson was concerned about the over-use of synthetic chemical pesticides and in her book *Silent Spring*, published in 1962, she challenged modern agricultural practices and called for a change in the way we view the natural world. She was one of the earliest writers to highlight the effects of the biomagnification of pesticides on the populations of predatory birds, such as the American bald eagle.

Some dismissed Carson as an alarmist, but she continued to speak out, reminding us that we too are part of the natural world and potentially subject to the same damage as the rest of the ecosystem. She called for new policies to protect human health and the environment.

DDT is an organochlorine (OC) insecticide that was widely used to kill mosquitoes that carry the malarial parasite. It is stored in the fatty tissues of animals that have ingested it. It is now known that it is not readily biodegradable and can remain in the environment for up to 15 years.

A survey of numbers of peregrine falcons in Europe in the early 1960s showed that they were in decline. Their bodies contained high levels of DDT, which was causing the shells of the bird's eggs to be thinner than normal. As a female tried to incubate the eggs, they broke under her body. This effect was also reported in many other parts of the world in a variety of wild bird populations. Even penguins in Antarctic regions were found to have the chemical in their bodies.

Although the original concentration of DDT used in insecticide sprays was low, at about 3×10^{-6} ppm (parts per million), the chemical was running into waterways and being taken up by microscopic plants in rivers and lakes. As these plants were eaten by microscopic animals, the DDT became more concentrated. It was found that small fish feeding on the microscopic animals had accumulated about 0.5 ppm in their body fat and fish-eating water birds, such as the osprey, had about 25 ppm of DDT in their bodies (Figure **C.21**).

DDT was a successful insecticide because it remained effective for a long time without breaking down, but its damage to the environment was considerable. Since the 1970s, it has been banned in many countries and wild bird populations are recovering. Heavy metals and industrial chemicals such as PCBs (polychlorinated biphenyls) that are also released into the environment remain a problem for living organisms as these accumulate in a similar way.

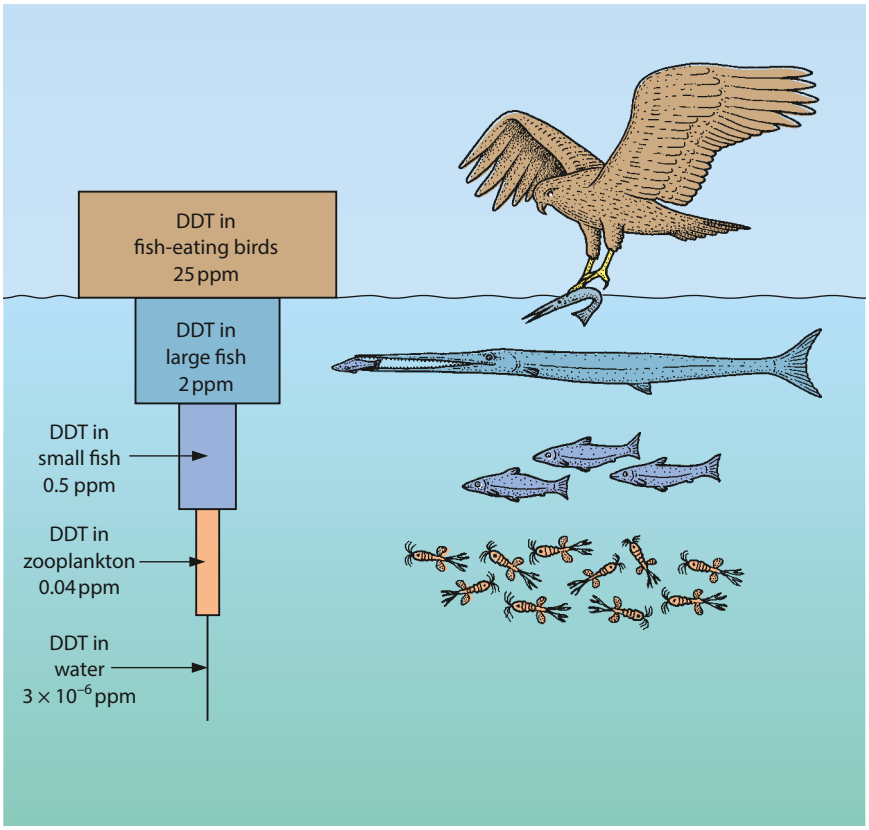


Figure C.21 An example of how DDT concentrations increase up the trophic levels of an estuarine food chain.

The control of malaria versus the use of DDT

DDT was banned for use in agriculture under the Stockholm Convention in 2004, an international agreement that was signed by 170 countries. But DDT is still used today in countries that have high levels of malaria to control mosquitoes, the vectors that transmit malaria to humans. At present there are few effective alternatives, so DDT is used for spraying internal surfaces of homes to kill mosquitoes.

There is considerable debate about the use of DDT in malaria prevention. Consider the following points in favour and against the use of the insecticide:

- Health of individuals is greatly improved if they do not suffer from malaria.
- There are few affordable or effective alternatives to DDT for killing mosquitoes.
- DDT remains active so that it can kill mosquitoes.
- People who live in sprayed homes and the workers who spray the homes are exposed to DDT for long periods of time.
- All the members of a household are exposed, including babies, pregnant women and old people.
- The evidence of human health problems due to DDT is increasing.

Taking into account these points, should the use of DDT be reduced as a precaution?

Those against the use of DDT say that its use and production should be halted because of potential health concerns and damage to the environment.

Those in favour of DDT say that it is safe to use if it is applied correctly. They argue that even if human health is found to be affected by the insecticide, the harm caused would be far less than that caused by malaria.

Others take a pragmatic approach and argue that there is still a need for DDT to fight the transmission of malaria but recognise that there are risks involved in spraying the homes of millions of people.

Material	Time to dissolve
paper	2–4 weeks
cotton cloth	1–5 months
woollen cloth	1 year
tin can	100 years
aluminium can	200 years
plastic bottle	450 years

Table C.3 Time taken for objects to dissolve at sea (Hellenic Marine Environment Protection Association, 2009).

Plastic debris in marine environments

Over recent years the production and use of plastics has increased enormously. It is estimated that over 250 million tonnes of plastic are used annually and that its production requires approximately 8% of the world's annual oil production. Plastic litter degrades very slowly (Table **C.3**), so it builds up in landfill and has serious implications in ocean environments, where it makes up between 60 and 80% of marine debris and as much as 90% of floating debris.

Macroplastic debris is defined as plastic fragments which are greater than 1 mm across, and **microplastic debris** is defined as fragments that are less than 1 mm. Macroplastics include items such as plastic bottles and bags, detergent containers and food wrapping. These items accumulate in marine habitats worldwide and may persist for centuries. Microplastics account for more than 65% of marine debris and mainly comprise PVC, polyester, acrylic and polyamide particles. Researchers have traced much of the microplastic back to synthetic clothes, which can release up to 2000 tiny fibres per garment every time they are washed.

Both types of plastic are ingested by many marine organisms, which mistake them for food. This plastic may enter the food chain or cause blockages of the intestine. Residues of the plastic can also accumulate in organisms' cells. Other problems include entanglement of organisms

in plastic, suffocation and general health problems. Some scientists have suggested that an animal whose stomach is full of plastic fragments feels full and stops feeding which may lead to starvation and death.

The Laysan albatross, which lives on Midway Atoll in the North Pacific Ocean, thousands of kilometres from both mainland Asia and North America, is one tragic example of this (Figure **C.22**). Albatrosses skim the water surface to feed and pick up plastic as they do so. Adults feed the plastic to their chicks and while adults are able to regurgitate some plastic, the chicks cannot and can be killed by its effects. As well as making the chick feel falsely full, sharp plastic pieces can cut through the stomach and cause infections and death.

Figure C.22 Laysan albatross chick which died with a stomach full of plastic debris on Midway Atoll.

Harbour seals (*Phoca vitulina*) in the Netherlands and in the Wadden Sea in Germany have also been shown to be affected by plastics. Samples taken from more than 100 seals revealed that more than 11% had plastic in their stomachs and 1% had plastic fragments in their intestines. Animals younger than 3 years were most affected. Researchers have also analysed fecal samples from both harbour and grey seals and found that all of them contained between a few milligrams and a few grams of granular or fibrous microplastics per sample.

The death of another mammal, a sperm whale, which was found in the Mediterranean Sea in 2013, has also been attributed to the ingestion of large amounts of plastic debris. The debris included several metres of plastic sheeting. The plastic is used as a cover in greenhouses and may have been torn off by the wind or not disposed of properly.

As well as the direct physical effects of plastic debris, researchers are also investigating the importance of so called 'hitch hikers' – species that attach to floating debris, which is then dispersed to other areas or sinks to the sea floor. It is possible that aggressive alien and invasive species could be dispersed in this way and travel to sensitive areas or coastal environments far away from their native habitats. Here they could compete with or endanger sensitive or at-risk species.

Nature of science

Assessing risk in science – biological control

Any programme to introduce an alien species as a means of biological control must weigh very carefully the risks involved in the introduction against the benefits. In the light of past experience, tightly controlled experiments should be carried out before a species is approved for release.

Risks associated with the introduction of an alien species include the following.

- The new species may compete with endemic (native) organisms and reduce their populations.
- This in turn may affect other species within the ecosystem.
- The introduced species may feed on endemic organisms, affecting local food chains and webs.
- The combined effect may be that an endemic species becomes extinct.

Consider the example of the cane toad, described above. What experiments should have been carried out before the species was introduced into Australia?

Test yourself

7 Outline what is meant by the term 'alien species'.

8 Define the term 'biomagnification'.

9 List **three** ways in which plastic debris can harm marine organisms.

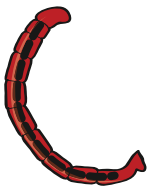
C4 Conservation of biodiversity

Biotic indices and indicator species

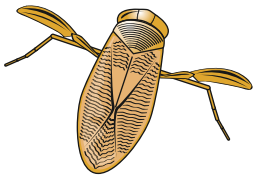
Certain species are very sensitive to environmental changes such as pollution from toxic gases in the atmosphere or chemicals in water. These organisms are called **indicator species** because their presence or absence tells us about environmental conditions in a way that direct measurements of abiotic factors cannot.

Lichens and bryophytes (mosses) are very sensitive to air pollution. Leafy species of lichen can only survive in areas with the highest quality of clean air, so by studying the lichens present in an area we can obtain a measure of air quality. Lichens vary considerably in their ability to tolerate pollutants such as sulfur dioxide in their environment because they have no waxy cuticle as the majority of terrestrial plants do. Without this protection, lichens absorb and accumulate various pollutants, including metal ions in airborne dust. Large, branching lichens grow in clean air, but in polluted air only small flat lichens can survive.

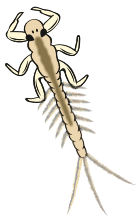
Water quality in rivers and lakes can be measured in a similar way. Some invertebrates can survive in polluted water, while others cannot (Figure **C.23**). Water that is polluted by sewage effluent is usually low in oxygen because bacteria in the water feed on the organic material and respire aerobically, using up the oxygen. Active invertebrates such as stonefly nymphs, mayfly larvae and flatworms are very sensitive to this kind of pollution because they require a lot of oxygen. So if these organisms are abundant, it is a good indication that water is clean. On the other hand, bloodworms (midge larvae of *Chironimus* sp.), sludge worms and leeches are more tolerant of low oxygen levels, and the presence of these organisms in large numbers is an indication of polluted water.



Chironimus larvae can tolerate low levels of oxygen and so indicate high levels of pollution.



The presence of water boatmen indicates moderately polluted water.



Mayfly nymphs are an indicator of clean, unpolluted water.

Figure C.23 The presence of particular organisms can indicate how polluted a body of water is.

Using a biotic index for a freshwater habitat

To gather data for a **biotic index** of a river or lake, the habitat is surveyed and samples of the organisms present are collected. Different indices use slightly different methods of calculation but the species present are recorded and usually counted. The number of each organism found, or simply the presence ('1') or absence ('0') of them, is multiplied by a 'sensitivity factor', which indicates that species' ability to tolerate pollution. A greater value is given to intolerant species, which require

Learning objectives

You should understand that:

- An indicator species is one whose abundance can be used to assess a specific environmental condition.
- The value of a biotic index can be calculated using relative numbers of indicator species.
- *In situ* conservation involves active management of nature reserves or national parks.
- *Ex situ* conservation involves the preservation of species away from their natural habitat.
- Biogeographic factors affect the diversity of species.
- Richness and evenness are two components of biodiversity.

clean water, and a calculation is performed resulting in a figure for the overall cleanliness of the water. A high number of sensitive species gives a high biotic index score but if many species that are very tolerant to pollution are found the index will be low. For example, one index, the Trent Biotic Index for the River Trent in the UK, gives values between 0 and 15 at different points along its length. Zero indicates very polluted water, whereas 15 indicates very clean water.

The Trent Biotic Index was developed for the River Trent in the UK, and most modern systems have evolved from this. Two such systems that are commonly used are the Biological Monitoring Working Party (BMWP) system (described here), used in many countries, and the Belgian Biotic Index Method.

The Biological Monitoring Working Party system

The advantages of this system are that the invertebrate organisms need to be identified only to family level, and that it can be used internationally. Each family is assigned a score from 1 to 10 depending on their pollution tolerance, 10 being the most intolerant. Some example organisms are listed in Table **C.4**.

Example organisms	Score
caddis flies	10
freshwater crayfish, stoneflies	9
dragonflies, damselflies	8
mayfly, tube-making caddis flies	7
small air-breathing snails, amphipod crustaceans	6
pondskaters, water striders, creeping water bugs	5
small mayflies, freshwater leeches, alderflies	4
valve snails, bladder snails	3
non-biting midges	2
segmented worms	1

Table C.4 Some of the families in the ten different classes of pollution tolerance (scores are approximate and common names have been used for simplicity).

All parts of the stream, river or lake are sampled. The sides, centre and areas among vegetation are all included and the invertebrates are collected and sorted into their families. Each is assigned a score from the table. The number of individuals in each family is not important. The results are added together to give a BMWP score.

The efficiency of sampling and sample size are taken into account using the Average Score Per Taxon (ASPT). ASPT is the BMWP score divided by the number of families (taxa) in the sample, which gives an idea of the diversity of the community. The overall water quality is assessed by looking at the BWMP and ASPT scores, as summarised in Table **C.5**.

BMWP	
Score	**Water quality**
>150	very good biological quality
101–150	good biological quality
51–100	fair biological quality
16–50	poor biological quality
0–15	very poor biological quality

ASPT	
Score	**Water quality**
>4.4	very good
4.81–4.4	good
4.21–4.8	fair
3.61–4.2	poor
<3.61	very poor

Table C.5 BMWP and ASPT scores related to water quality.

Worked example

Samples taken from two streams are recorded and given scores as shown in Table **C.6**.

Stream 1		Stream 2	
Families present	**Score**	**Families present**	**Score**
pondskaters	5	bladder snails	3
pulmonate snails	6	midge larvae	2
water striders	5	marsh snails	3
creeping water bugs	5	valve snails	3
leeches	4		
air-breathing snails	3		
alder flies	4		
midge larvae	2		

Table C.6 Species recorded in two streams.

The total BMWP scores are 34 for stream 1 and 11 for stream 2. The ASPT scores are calculated using:

$$\text{ASPT} = \frac{\text{BMWP}}{\text{number of families}}$$

So for stream 1:

$$\text{ASPT} = \frac{34}{8} = 4.3$$

And for stream 2:

$$\text{ASPT} = \frac{11}{4} = 2.8$$

Exam tip
Always include your working out when you are calculating numerical answers. It helps you to find any mistakes when you are checking your results.

Looking up these scores in Table **C.5**, we can see that the water quality of stream 1 is poor by the BMWP score, but fair by the ASPT. Stream 2 scores very poor on both scales.

Conserving threatened species

Conserving and protecting a natural habitat as a nature reserve should benefit all species. However, if population numbers are very low and a species is at risk, more active intervention may be required. Each nature reserve will have its own unique solutions to conservation problems. At Belsize Wood Nature Reserve, a small woodland reserve near the centre of London in the UK, nesting boxes for birds and bats have been put in place, because the number of mature trees providing suitable natural nesting sites is low. In a wetland nature reserve, nesting platforms that float on lakes can be beneficial and offer some protection against predators for nesting birds. At Sungei Buloh Wetland Reserve in Singapore, sluice management allows the control of water levels in the ponds. At any one time, the water level in at least one pond is kept low to expose the mudflats for shorebirds to feed and roost (Figure **C.24**).

Members of the public may question the funding and existence of a nature reserve if access is denied to them. This is a difficult issue, as the more people that visit a nature reserve, the more chance there is of habitats being damaged or destroyed. On the other hand, visitor access can have positive outcomes, if public awareness and knowledge of wildlife is improved. Usually, special trails or walkways are built at reserves to ensure that observers can visit safely without compromising the surrounding habitats (Figure **C.24**). Legislation can also protect nature reserves from development and industrial activities.

Figure C.24 The habitat is carefully managed at Sungei Buloh Wetland Reserve in Singapore. There are a number of trails through the wetland, but the highlight is this 500 m boardwalk that takes visitors right to the centre of the reserve.

In situ conservation

***In situ* conservation** protects species within their normal habitat. This makes sense because each species has evolved to adapt to a particular environment. *In situ* conservation protects species in their own habitats by maintaining the environment, often within nature reserves or national parks. In situ conservation work can involve removal of invasive species, such as the kikuyu grass on Montague Island or rhododendron plants in North Wales (described in Subtopic **C3** earlier), or protecting certain species from predators. Provided there are sufficient numbers in the population, *in situ* conservation should provide sufficient genetic diversity for a population to be sustained.

Ex situ conservation

***Ex situ* conservation** involves preserving a species whose numbers are very low in a **captive-breeding programme** in a zoo or botanic garden to prevent it dying out.

In situations where *in situ* conservation is difficult or inadequate, *ex situ* conservation must be used. This is not ideal, because an organism behaves differently outside its natural habitat. However, it does give rise to the opportunity for captive breeding using scientific knowledge and modern technology. Techniques such as artificial insemination and embryo transfer may be used if animals fail to breed normally, and embryos can be preserved for later use. Difficult pregnancies can be monitored and the young cared for by staff.

An *ex situ* breeding programme has proved invaluable for the Arabian oryx. This animal, once almost extinct in the wild, has been successfully bred in a number of zoos in the USA and Europe. The DNA from the few remaining animals was compared and animals specially selected for breeding so that genetic diversity was maintained as far as possible. Studying the behaviour of captive animals is key to breeding programmes. Some species with complex behaviours such as the giant panda from China are highly challenging to breed in captivity, but the centre at Chengdu in China has been very successful.

Plants are more straightforward to maintain in an *ex situ* situation. Botanic gardens can supply the correct environmental conditions for different plants and computer-controlled glasshouses can maintain the temperature and humidity that each requires. Many countries maintain 'national collections' of a variety of species including endemic plants, exotic genera and important food plants.

There are also **seed banks** for many of the world's staple crops such as rice and maize. These preserve varieties of important crops, called **landraces**, which may be useful in the future to produce new varieties of food plants. At the Millennium Seed Bank at Wakehurst Place in England, seeds are kept in cool, dark conditions, which prevent germination, and can be stored for many decades. The Svalbard Global Seed Vault, on the Norwegian island of Spitsbergen, holds duplicate samples of seeds held in gene banks worldwide, in an underground cavern.

Case study – the Arabian oryx

In 1986, the Arabian oryx (*Oryx leucoryx*) was classified as 'endangered' on the IUCN (International Union for Conservation of Nature) Red List and in 2011 it was the first animal to receive 'vulnerable' status again after having been listed as 'extinct in the wild'.

The oryx is a grazing antelope which is adapted to survive in the extreme conditions of hot, dry deserts. The animals live in small herds of 10–30 animals with a hierarchy of dominance amongst both males and females. They defend their territories using their horns and have keen eyesight to maintain the group. They also use their horns to dig shallow pits to rest from the heat of the day. The oryx lives across the Arabian and Sinai peninsulas and has been reintroduced into Oman, Israel, Saudi Arabia and Jordan (Figure **C.25**).

By the early 1970s, the Arabian oryx was extinct in the wild as a result of hunting by poachers who chased them across the desert in four-wheel drive vehicles. A new population was established by breeding animals that were held in zoos in different parts of the world. A captive-breeding programme began in 1962 at the Phoenix Zoo and was supported by the Fauna and Flora Preservation Society of London and the World Wildlife Fund. It began with nine animals and soon oryx were sent to other zoos and parks to start new herds. The pedigree of the captive-bred animals was monitored to ensure that a sufficiently large gene pool was maintained. Animals were reintroduced into the wild in 1982 and the population thrived for about 15 years. But poaching began again in 1996. However, this time laws were changed to put a stop to the practice and a second reintroduction took place in Saudi Arabia. So far this population has survived successfully. The total reintroduced population now stands at about 1000 animals and is well over the threshold number of 250 mature individuals, below which a species qualifies for 'endangered' status.

Figure C.25 Arabian oryx.

Biogeographic features and species diversity

The study of spatial distribution of organisms, species and ecosystems is known as **biogeography**. Biogeographic regions are distinguished by virtue of their particular climatic conditions, physical characteristics such as altitude and soil type, and the presence of particular species of organisms – areas with similar biogeography have similar distributions of organisms. This means that species with similar lifestyles and adaptations can survive in different parts of the world, if the biogeographic conditions are comparable. For example, tropical rain forests are found at various places on Earth where the climate and terrain is similar. Each area of forest contains similar communities of organisms, even though the actual species may be different in different parts of the world. Biodiversity varies across the world and in different biogeographic regions – it is much greater in rain forests, for example, than in desert ecosystems.

As people have become more aware of the need to conserve species, many governments have set aside land or protected areas of the country to provide regions where organisms are safeguarded. When new reserves of this kind are planned, many factors need to be taken into account to ensure they are successful in promoting conservation of diversity. The chosen areas must have the correct climate and terrain – that is, the correct biogeography – to support the species that are to be conserved. In addition, both the size of the area and the total length of its boundary are important (Figure **C.26**).

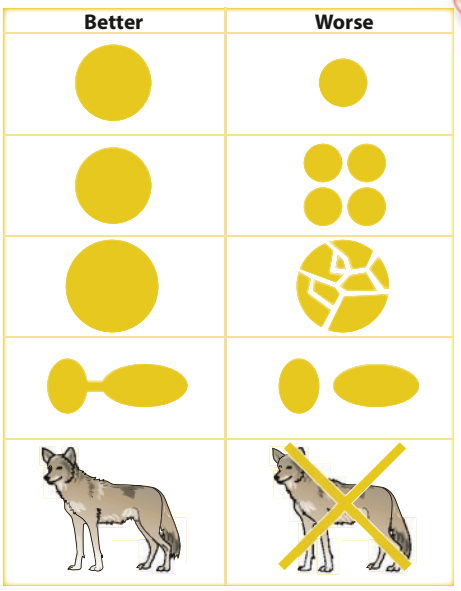


Figure C.26 When designing a conservation area, a large area is better than a small one; a single large area is better than several small areas of the same total size; an intact area is better than a fragmented or disturbed one; areas connected by corridors are better than separate, isolated areas and it is better to have large native carnivores present in the area than not.

Size

Large areas reserved for conservation of biodiversity work better than small ones. Small reserves can only support small population numbers, so there is a risk that inbreeding will occur and the genetic diversity of species will diminish. In a small reserve, there is always a risk that a natural disaster such as flooding or a forest fire will wipe out all the organisms of a species. This is less likely to happen in a large reserve. Edge effects are also less significant in large reserves than in small ones.

Edge effects

The centre of a nature reserve is likely to have different features from the areas at the edge. A woodland reserve has more light, more wind and less moisture at the edge than at the centre. Organisms that live in the centre of the wood will be protected from the influence of other organisms, such as farm animals or human activity, outside the reserve. This is not so for organisms living close to the edge, which may be disturbed by or even compete with organisms outside the reserve. Small reserves have more edge per hectare than large ones, so edge effects have a greater impact on the overall ecosystem in smaller reserves.

One well-studied example of an edge effect involves the brown-headed cowbird of northern and western USA. This bird is a brood parasite, laying its eggs in the nests of other birds at the edge of forests. It feeds in open areas where insects are abundant. As forests have become fragmented, due to urbanisation and farming, more forest edges have become available. The brown-headed cowbird population has increased so much that, in recent decades, many land managers and conservationists have argued that brown-headed cowbirds are a major threat to North American songbird populations.

Wildlife corridors

If it is impossible to create a large nature reserve, good planning may make it possible to link two smaller areas through a corridor. These are often built under busy roads or railway lines, so that organisms have a larger area to move about in and colonise. A corridor is not ideal because animals using it may be exposed to dangers from outside the reserve and corridors can act as conduits for the spread of disease or make certain species easy targets for poachers. On the other hand, the benefits of corridors include the fact that gene flow between two otherwise isolated areas can take place and promote diversity.

Measuring biodiversity – the Simpson diversity index

Biodiversity is a relatively modern term that simply means 'the variety of life on Earth'. One of the best ways to assess the health of an ecosystem is to measure the variety of species living there. The Simpson diversity index allows us to quantify the biodiversity of a habitat. It takes into account both the number of different species present (the species '**richness**' of the habitat) and the abundance of each species. If a habitat has similar population sizes for each species present, the habitat is said to have '**evenness**'.

The value of the Simpson diversity index is best illustrated by comparing two habitats. Two ponds might contain species of invertebrates in the numbers shown in Table **C.7**.

Simpson's diversity index gives us a measure of both richness and evenness. It is calculated using the formula:

$$D = \frac{N(N-1)}{\Sigma n(n-1)}$$

where

D is the diversity index
N is the total number of organisms in the habitat
n is the number of individuals of each species

	Species					Total number of organisms
	Water boatmen	Water measurers	Pond skaters	Whirligig beetles	Water spiders	
Number of organisms in pond A	43	18	38	3	1	103
Number of organisms in pond B	26	18	29	11	5	89

Table C.7 Numbers of different invertebrate species found in two separate ponds.

Using the formula, we can calculate that for pond A:

$$\text{Simpson diversity index } D = \frac{(103 \times 102)}{43(43-1) + 18(18-1) + 38(38-1) + 3(3-1) + 1(1-1)}$$

$$= \frac{10506}{3524}$$

$$= 2.98$$

For pond B:

$$\text{Simpson diversity index } D = \frac{(89 \times 88)}{26(26-1) + 18(18-1) + 29(29-1) + 11(11-1) + 5(5-1)}$$

$$= \frac{7832}{1898}$$

$$= 4.13$$

Although there are fewer organisms in pond B, the individual populations are more even, so the community is not dominated by one or two species. We conclude that pond B is more biodiverse. It is instructive to alter some of the figures and see what effect this has on the value of *D*. An advantage of the index is that you do not need to know the name of every different species – it must simply be distinguished as a separate species. Calculating the Simpson diversity index at intervals over time can give a good indication of the health of an ecosystem and whether conservation measures might be valuable.

Nature of science

Cooperation and collaboration – conserving biodiversity requires international cooperation between scientists, organisations and politicians

In the last 50 years, the importance of biodiversity has come to the forefront of science. Species are not evenly distributed on Earth – biodiversity is far richer around the tropics, and areas containing rainforest are among the most diverse on the planet. People have come to realise that there are many compelling reasons for conserving the biodiversity of habitats such as the rainforests where as yet undiscovered species may provide valuable medicines and other resources for future generations. Conservation in one part of the world may depend on cooperation and collaboration in another and international organisations such as World Wide Fund for Nature (WWF) and the United Nations Environment Programme (UNEP) coordinate such work in many countries. The key objective of all conservation organisations is to preserve species and their habitats. Some work at a local level while others are global. Some organisations, such as UNEP, are funded by governments while others, such as WWF, are non-governmental organisations (NGOs), which are funded by individuals or groups. Organisations such as WWF work with businesses, governments and local communities to create solutions that take account of the needs of both people and nature.

Conservation programmes must select which species are to be protected, but it is often difficult to decide which species most merit conservation efforts. On what basis should one species be chosen over another? For example, is a large mammal such as a tiger or panda more important than a small, seemingly insignificant mollusc? A striking or endearing mammal may encourage people to support a conservation programme but smaller, less appealing species may, in fact, be more important and play a pivotal role in an ecosystem. Should endangered animals be given priority over other species whose numbers are not yet so low?

The choice of species for *ex situ* conservation can also be difficult, and many factors must be taken into account. For example, when zoos select animals for captive breeding programmes, certain animals with aesthetic appeal are likely to increase visitor numbers and therefore raise public awareness and attract greater financial support for conservation. If these animals are returned to the wild, they may engage local people who could benefit from ecotourism. On the other hand, choosing a species for ecological reasons is more likely to benefit a whole ecosystem – assuming the programme does not fail through lack of funding and support.

Ecotourism involves developing a conservation area to make it suitable and attractive for visitors, who pay for local goods and services thus providing economic support for the area and its people.

Science can support conservation efforts by providing the expertise needed to ensure breeding programmes are successful. Different zoos have different areas of expertise and are likely to be more successful at *ex situ* conservation with some species than with others, so this factor too will influence the organisms whose preservation is prioritised.

Test yourself

10 Outline the differences between *in situ* and *ex situ* conservation programmes.

11 Define the term 'indicator species' and give one example.

12 Compare the terms 'richness' and 'evenness' as components of biodiversity.

C5 Population ecology (HL)

Learning objectives

You should understand that:

- Population size is estimated using sampling techniques.
- In an ideal, unlimited environment, exponential population growth can take place.
- As the carrying capacity of the environment is reached, population growth slows.
- The shape of a sigmoid growth curve can be explained by relative rates of natality, mortality, immigration and emigration.
- Limiting factors can act on population size from the top down or from the bottom up.

Estimating numbers in populations

A **population** is a group of individuals of the same species that live in the same area. Population numbers can and do change over time and are affected by a number of factors in the environment. In order to assess the size of a population, **sampling** techniques are used (Subtopic **C1**).

The most common method of estimating population size of animals is the 'capture–mark–release–recapture' technique (Figure **C.27**). It is used for populations where individuals are mobile and move freely in their habitat.

1. A sample of the population is collected by netting or trapping or another suitable method. The sample must be as large as possible and the trapping method must not harm the animals.
2. The number of organisms in the sample is counted and recorded.
3. Each of the captured animals is inconspicuously marked in some way – for example, with non-toxic paint for invertebrates or by trimming a concealed area of fur for small mammals.
4. The animals are returned to the wild and left for long enough to mix with the rest of the population.
5. A second sample of the population is collected after this time.
6. The number of marked and unmarked individuals in the second sample is counted.

The population size is calculated using the Lincoln Index formula:

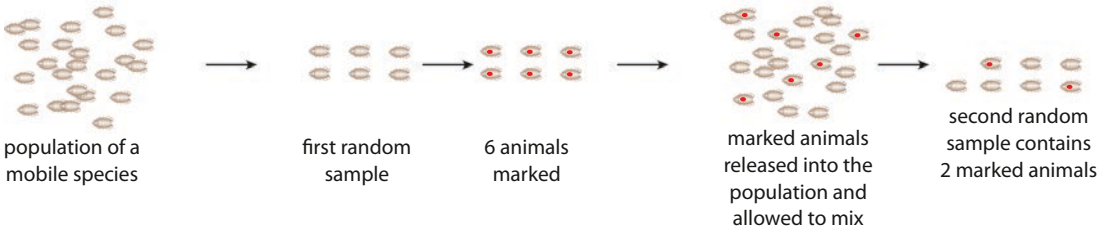
$$\text{total population } p = \frac{\text{number of animals in first sample} \times \text{number of animals in second sample}}{\text{number of marked animals in second sample}}$$

or

$$p = \frac{(n_1 \times n_2)}{n_3}$$

where

P is the total population
n_1 is the number of organisms caught originally
n_2 is the number caught in the second sample
n_3 is the number of marked individuals in the second sample



$$\text{estimated population size} = \frac{\text{number in first sample} \times \text{number in second sample}}{\text{number of marked animals in second sample}}$$

$$\text{estimated population size} = \frac{6 \times 7}{2}$$

$$\text{estimated population size} = 21$$

Note: This method only produces results of acceptable accuracy if the numbers in the samples are larger than shown here. At least 20 animals should be sampled.

Figure C.27 Capture–mark–release–recapture technique for estimating population size.

This method depends on a number of factors, which need to be taken into account.

- Marking the organisms must not harm them or cause them to be conspicuous to predators. That is, the marking itself must have no effect on the population size.
- There should be minimal **immigration** into or **emigration** from the population.
- The measurements must be conducted within a single life cycle, so there are no changes to the population through births or deaths.

The capture–mark–release–recapture technique is most appropriate for invertebrates such as woodlice, snails and ladybirds or small mammals, such as mice, with a limited territory. Sampling organisms with a large territory, or those where the population is small, is not accurate using this method.

Population size and growth

Consider what happens if a few individuals of a species enter an unoccupied area. Perhaps a few rabbits arrive on an uninhabited island covered by lush grassland or some fish are washed into a newly established pond. Assuming there is enough food and there are few predators, the newcomers will reproduce and the population will increase rapidly. After a time, when there are large numbers of individuals, the food supply will start to be used up faster than it can be replaced. The population will be unable to increase any further and the population numbers will stabilise.

This typical pattern of population growth can be represented on a graph, like that shown in Figure **C.28**.

1. As reproduction gets underway, the population shows exponential growth (the steepest part of the curve). At this time, the population inhabits an ideal, unlimited environment – there is abundant food, little competition for space and the effects of predation and disease are minimal.
2. After a time, the exponential phase ceases and one or more of the resources individuals need become limited. The shape of the curve shows that the rate of population growth is declining at this point and the population is said to be in the transitional phase. Individuals must compete with one another for the resources they need, which may include space, light, food, nutrients and water. Competition for resources between members of the same species is known as **intraspecific competition**. This increases as population numbers increase. When the rate of demand for a particular resource is greater than the rate of supply we say that the resource has become a **limiting factor**. Predation, disease, and in some cases the accumulation of toxic wastes, such as carbon dioxide, can also limit a population.
3. Eventually population numbers become more or less constant and the curve levels off, in the plateau phase. The ecosystem has reached its **carrying capacity**, which is the number of individuals in a population that the resources in the environment can support for an extended period of time. Once the carrying capacity is reached, the population growth rate will slow down either because organisms die through lack of an essential resource, or because they fail to breed and their birth rate falls. In the plateau phase, the population remains more or less stable because rates of natality and mortality are balanced, as are rates of emigration and immigration.

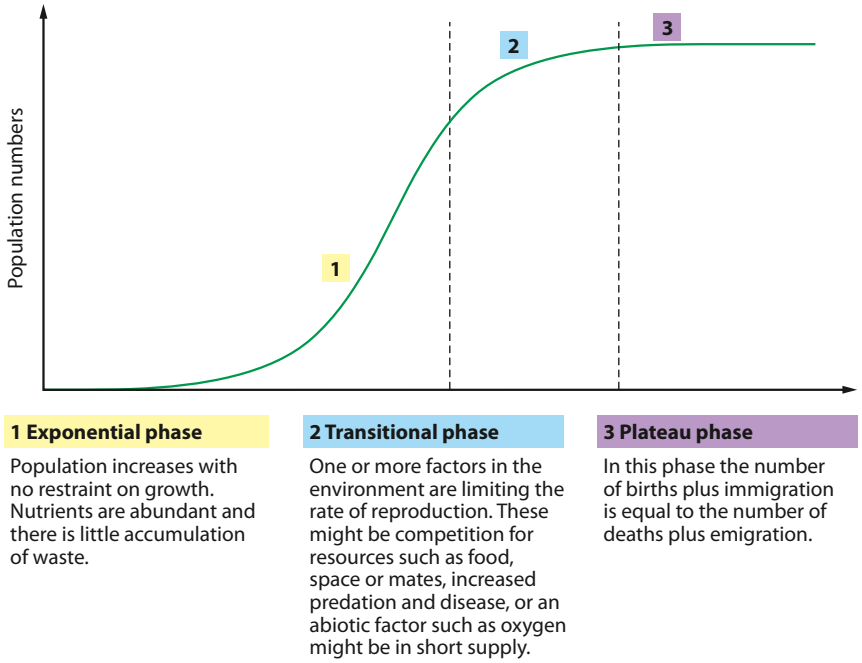


Figure C.28 The sigmoid population growth curve for a model species such as duckweed (*Lemna* spp.), growing in a stable environment.

The human population

The human population growth graph has not followed the sigmoid pattern shown in Figure **C.28**. Since the evolution of humans, no plateau has been reached, and instead the global population continues to rise exponentially. The natural carrying capacity of Earth has been manipulated as humans have found ever more technologically advanced ways to produce food and extract resources from the environment. Humans have also overcome other limiting factors like disease through improved medicine, and colonised almost every part of the planet. But over-population has led to poor living conditions and environmental degradation in many parts of the world.

Why do population sizes vary?

There are a number of important reasons why a population may change in size:

- **natality** – the birth rate may change (the number of new individuals joining the population due to reproduction)
- **mortality** – the number of deaths may change
- **emigration** – members of the population may move away to new habitats
- **immigration** – new members of the species may arrive from elsewhere.

Factors that limit population increases

There are certain key factors that affect a population, no matter what species is considered. These include:

- availability of key resources such as food, water, oxygen, light, space, mates and shelter

- levels of waste products, such as carbon dioxide or nitrogenous waste
- disease
- predation or herbivory.

Ecologists divide limiting factors into two categories: 'top down' or 'bottom up'.

Top-down limiting factors

Top-down limiting factors are those that involve an organism higher up the food chain limiting the numbers of a species at a lower trophic level, usually through predation or herbivory. One example of this is the control of the small algae (*Fucus* spp.) on a rocky shore by grazing limpets (Subtopic **C1**). Another example is the control of kelp forests due to the impact of sea otters. Otters feed on sea urchins, which use kelp as a source of food so if sea otter numbers fall, the sea urchin populations expand and reduce the kelp forest. Ecosystems such as those where sea otters and limpets are found are not controlled by the productivity of the primary producer but rather by a top predator or major herbivore acting as a **keystone species**.

Bottom-up limiting factors

Bottom-up control by limiting factors occurs where the nutrient supply and productivity of primary producers (plants and phytoplankton) control the structure of the ecosystem. In marine coastal ecosystems, plankton populations depend on and are controlled by the availability of nutrients.

Phytoplankton populations increase so that large growths known as **algal blooms** appear when nutrients are abundant. This happens when sea currents cause upwelling, which brings nutrients to the surface where they are accessible to phytoplankton. The abundant growth of phytoplankton is then controlled by top-down control by herbivores, which use it as food. Algal blooms are also controlled by bottom-up control at times when nutrients are in short supply or in places where currents do not bring nutrients to the surface.

Limiting factors and ecosystem stability

Bottom-up and top-down control tends to keep a stable population at the carrying capacity of the ecosystem. The bottom-up resources set the limit for the maximum sustainable population, while top-down control removes individuals from a large population, with the result that resources are not over-exploited. The concept of internal control of populations by interactions between them is a key argument for the conservation of ecosystems.

Test yourself

13 Suggest factors that might lead to an increase in a bird population in a woodland.

14 Complete this equation for the plateau phase of a growth curve:

natality + _____ = _____ + _____

Estimating the size of commercial fish stocks

The commercial fishing industry is of enormous importance worldwide. Fish provide what should be a renewable source of food, but catching fish has become an industrial process, involving technology, such as the use of sound waves to track shoals of fish, and large-scale machinery, including huge trawling nets. Many species are in danger of being over-fished, as their populations are reduced to unsustainable levels. In some species, the numbers of adult fish available to breed is too low to replace the animals removed by fishing. There is a pressing need to monitor populations so that fish species, and the fishing industry as a whole, can survive.

The International Council for Exploration of the Sea (ICES) is an organisation that monitors harvests in the North Atlantic. Fish are not easy to count because they move over long distances. The usual method of estimating a population involves collecting data from landings at fish markets, from the numbers of fish discarded from fishing boats, and from targeted surveys with research vessels.

- The numbers of fish of different ages are recorded to give an idea of the age distribution in the population. The ages of individual fish are a useful indicator of fish stocks. Too few young ones indicate that the fish are not spawning sufficiently to replace caught fish, and too few older, larger fish indicates that over-fishing is occurring.
- Fish age can be estimated by the length and weight of individuals. A more accurate method is to measure the rings in the ear bones. As fish grow, the number of rings increases and these can be measured using a microscope.
- The data collected from catches and age estimation can be used to deduce spawning rates and survival of different species.
- Research vessels can use echo sounding to estimate the sizes of fish shoals in some locations.

ICES offers advice on over 130 species of fish and shellfish. Using the advice from this and other similar organisations, scientists can work out the health of a particular fish population and whether it is being over-fished.

Maximum sustainable yield

The **maximum sustainable yield** is the largest proportion of fish that can be caught without endangering the population. Setting this figure is hotly debated and countries have different views on the issue, often influenced by local interests. At extremes, if the fish population is very small, there will be few adults to produce young, and if the population is large, competition for food will slow growth. The ideal, then, is to fish at a level that maintains the maximum yield by allowing fish stocks to replenish at the optimum rate. Fish are a renewable resource and can always be available for food if they are only taken in a way that allows them to survive and replenish their numbers.

International measures to conserve fish

In recent years, there have been several alarming reports on declining fish populations worldwide. In 2003, 29% of open-sea fisheries were in a state of collapse, defined as a decline to less than 10% of their original yield.

Populations of fish must be monitored and quotas or closed seasons put in place to reduce fishing during the breeding seasons. If net sizes are monitored and controlled, smaller fish can be left in the water to mature and breed. Today bigger vessels, bigger nets and new technology for locating fish are not improving catches, simply because there are fewer fish to catch. Where fishing is banned or regulated, biodiversity can improve and fish populations may be restored relatively rapidly. This means protecting not only fish populations but also other organisms within a marine ecosystem.

International cooperation is essential if measures like these are to be successful. In 1995, the United Nations Fish Stocks Agreement became one of the first international treaties to protect fish. It was the first attempt made to develop a long-term, sustainable fishing strategy and has already had an impact on the regulation of fishing since it came into force in 2001. Now 77 countries including most of the major fishing nations and the European Union have signed up to the Agreement. It has encouraged countries to adopt responsible fishing policies and manage fisheries with more care. In 2006, a Review Conference was held to strengthen the Agreement still further.

The ICES has noted that the situation in the North Sea has been gradually improving since 2006. It is estimated that there are now 21 million mature cod (65 000 tonnes) in the North Sea and fish have been found to be reproducing at a younger age – at 4 years of age, 60% of fish are mature and all are mature at 6 years old.

Cod discard rates have fallen from 62% in 2007 to 24% in 2011 so fewer fish are dying needlessly, but the situation is still precarious. As the graph in Figure **C.29** shows, fishing mortality was estimated at 0.67 in 2010, which means that 49% of all fish between the ages of 2 and 4 years were caught. Although the biomass of spawning fish has increased since its lowest point in 2006 (Figure **C.29**), it is still below the limit of what is healthy for a sustainable population.

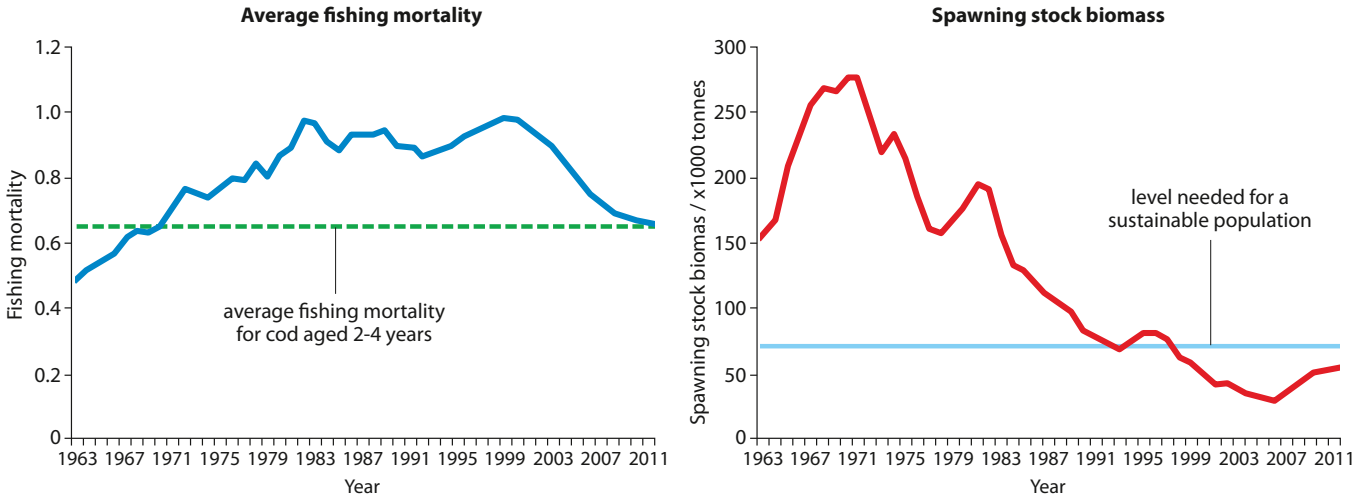


Figure C.29 Fishing mortality and spawning stock biomass for North Sea cod between 1963 and 2011. The population of North Sea cod has fallen greatly since the end of the 1960s, almost certainly as a direct result of over-fishing. The horizontal line in the right-hand graph shows the minimum stock size that has been calculated will allow the cod population to be maintained at a viable level.

Nature of science

Cooperation and collaboration – international efforts to conserve fish stocks

Why is it difficult to ensure that fishing is regulated to the benefit of fishermen, consumers and the environment? Consider the following statements and discuss the problems that arise as international negotiators attempt to interpret and apply scientific data.

- Data on fish stocks is difficult to obtain and there is no common view on what is a sustainable population.
- It is difficult to enforce fishing regulations. Authorities may be active in one region but unable to control actions of other countries that ignore the rules.
- Politicians are under pressure from fishing communities not to limit fish catches.
- As fisheries go out of business due to declining stocks, it is not easy to limit the activities of those that remain.
- Fish are mobile animals and can be caught thousands of miles from where they are bought. Can ethical consumers be sure that the fish they buy is from a sustainable stock?

Exam tip
Remember that if an examination question asks you to 'discuss', it is important to present alternative points of view.

Test yourself

15 **a** Describe how the 'capture–mark–release–recapture' method is used to estimate the size of a population of small invertebrates.

b Explain the limitations of this method.

16 Outline what is meant by 'maximum sustainable yield' of fish stocks.

Exam tip
If you are asked to 'explain' a concept or situation, remember to include the steps in the process and write about them in some detail.

Learning objectives

You should understand that:

- Nitrogen-fixing bacteria convert nitrogen from the atmosphere into ammonia.
- *Rhizobium* bacteria form a mutualistic relationship with roots.
- Denitrifying bacteria reduce soil nitrate levels in anaerobic conditions.
- Phosphorus can be added to the phosphorus cycle in the form of fertilisers or removed as crops are harvested.
- In the phosphorus cycle the rate of turnover is much slower than in the nitrogen cycle.
- Phosphate availability may limit agriculture in the future.
- Eutrophication is caused by the leaching of mineral nutrients from agricultural land into rivers. Increased levels of nutrients in the water lead to an increased biochemical oxygen demand (BOD).

C6 Nitrogen and phosphorus cycles (HL)

The nitrogen cycle

Nitrogen is a vital element for the formation of proteins and nucleic acids in the bodies of plants and animals. However, although almost 80% of the Earth's atmosphere is nitrogen gas, it is so stable that it cannot be used directly by living organisms, and nitrogen is often in short supply as a nutrient. It is recycled in ecosystems through the actions of many organisms (Figure **C.30**).

Nitrogen fixation

Nitrogen is made available to an ecosystem by bacteria that are crucial in transferring nitrogen compounds from the abiotic to the biotic environment. Two types of bacterium, *Azotobacter* and *Rhizobium*, are able to 'fix' nitrogen from the air and convert it into ammonia. Ammonia formed by both organisms reacts with organic acids to form amino acids.

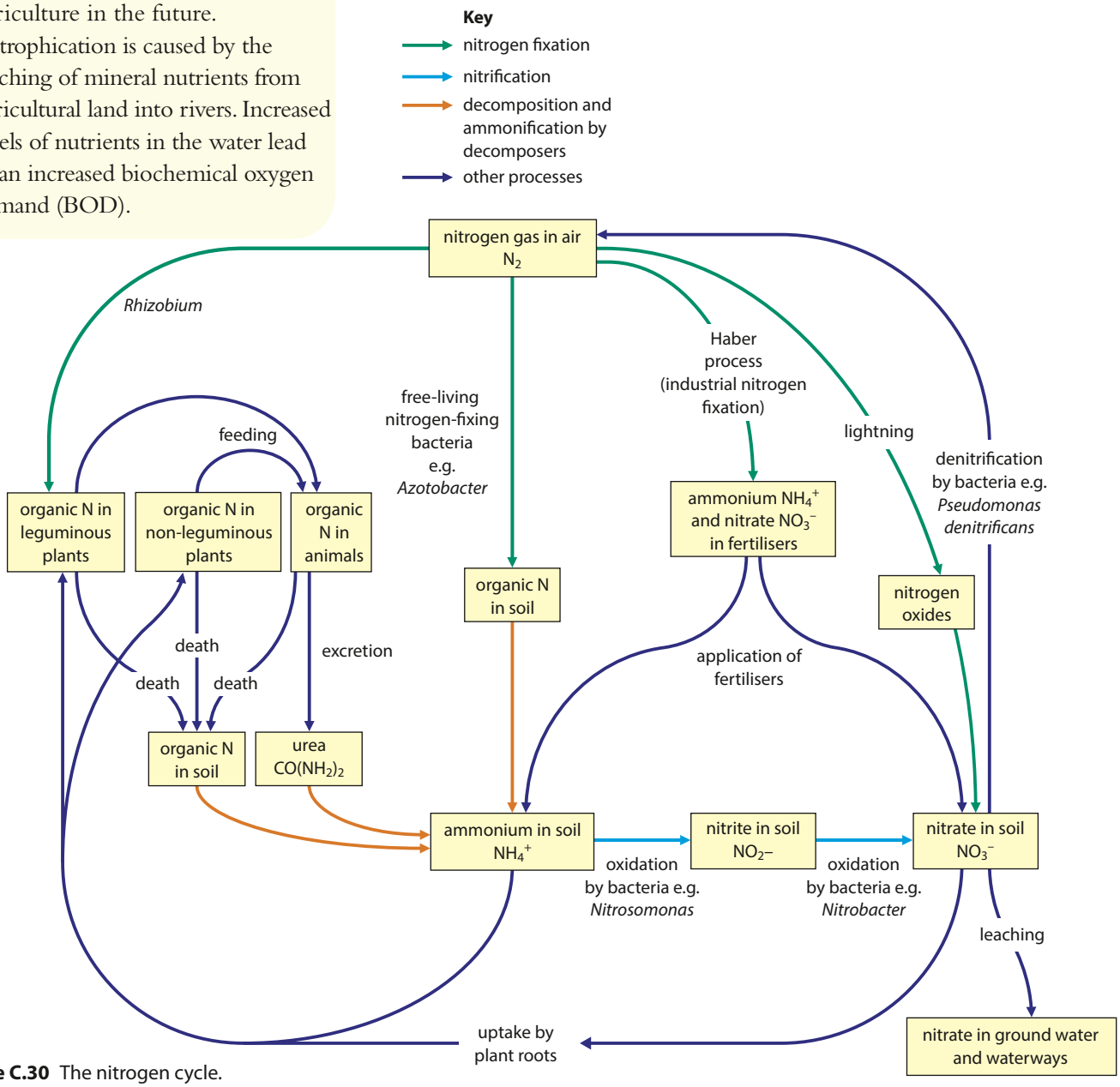


Figure C.30 The nitrogen cycle.

For example, pyruvate (pyruvic acid) reacts with ammonia to form the amino acids alanine, valine and leucine. *Rhizobium* invades the roots of leguminous plants such as peas, beans and clover to form nodules on the roots (Figure **C.31**). The bacteria and plants form a **mutualistic** relationship in which the bacteria receive sugars from the plant, and the plant in turn receives nitrates from the bacteria. The only other natural method of **nitrogen fixation** is the effect of lightning, which combines nitrogen gas in the air with oxygen, forming nitrates that enter the soil, where they are useful to living things. Humans also fix nitrogen in the Haber process, which is used to manufacture fertilisers.

Figure C.31 *Rhizobium* nodules on the roots of a bean plant.

Nitrification

Other important groups of bacteria in the nitrogen cycle are the nitrifying bacteria *Nitrosomas* spp., which use oxygen to convert ammonia from excretory material into nitrites, and *Nitrobacter* spp., which use oxygen to convert nitrites into nitrates. **Nitrification** is an important part of the nitrogen cycle because both ammonia and nitrite are toxic to plants. Nitrates are all soluble compounds that can be absorbed by plants through their roots and assimilated into their biomass. Nitrification is favoured by a neutral pH, warmth and well-aerated soil, as it is an oxidative process. Ammonium compounds and nitrites cannot be taken in directly by plants.

Denitrification

Denitrification reduces the fertility of the soil, depleting it of nitrates so that it may not be useful for cultivation. Denitrifying bacteria (for example, *Pseudomonas denitrificans*) – which are found mainly in anaerobic conditions in compacted or waterlogged soils – convert nitrates to nitrogen gas. Waterlogging therefore has an impact on the cycling of nitrogen – it causes a lack of nitrifying and nitrogen-fixing bacteria, and favours the return of nitrogen to the atmosphere by denitrification.

Insectivorous plants

A few plants are able to survive in nitrogen-poor and phosphorous-poor environments such as bogs and acidic moorlands where waterlogging encourages the growth of denitrifying bacteria. Conditions like these limit the amount of nutrients that plants can extract from the soil.

Insectivorous plants, including the Venus fly trap (*Dionaea muscipula*) and the sundews (*Drosera* spp.), supplement their nutrition by trapping and digesting insects and other small arthropods, which provide the nitrogen they need to form proteins.

The Venus fly trap, which is endemic in the wetlands of the East Coast of the USA, has spines on the edges of its leaves (Figure **C.32**). When they are triggered, the leaves can fold to form a cage, which traps an insect inside. The closing of the trap is triggered when an insect comes into contact with hairs on the surface of the leaves. The leaves squeeze tightly together and release digestive enzymes on to the prey inside.

Sundews, which can be found growing in bogs and moorland on every continent except Antarctica, use a sticky mucilage to trap their prey. Insects are attracted to the plant by sweet, sticky secretions from their 'tentacles' (Figure **C.2**). If an insect touches the tentacles it becomes trapped in the sticky fluid and is unable to escape – it will either die of

Although nitrogen gas (N_2) is inert and unreactive, many important compounds contain nitrogen. Ions of some of these compounds are:

NO^{2-} nitrite
NO^{3-} nitrate
NH^{4+} ammonium

The nitrogen cycle and soil fertility

Plants obtain the nitrogen they need to grow in the form of nitrates, which they absorb from fertile soil through their roots.

Good for plant growth are:

- nitrogen fixation, which converts nitrogen gas to useful nitrates
- nitrification, which converts ammonia to useful nitrates.

Bad for plant growth is:

- denitrification, which converts useful nitrates into nitrogen gas that plants cannot use.

Figure C.32 Venus fly trap.

exhaustion or of suffocation if the mucilage fills its spiracles. Tentacles bend towards the centre of the leaf and secrete digestive enzymes, which dissolve the prey. Nutrients are absorbed through the surface of the leaves.

The phosphorus cycle

Phosphorus is an essential nutrient for life and is vital to many physiological and biochemical processes. It is an essential raw material needed for the formation of DNA. Most of the phosphorus on Earth is found stored in soil and rocks as inorganic phosphate (PO_4^{3-}). Phosphates are released naturally through the weathering and dissolution of rocks and minerals. Phosphorus is usually the limiting element for animal and plant production and throughout history phosphate has often been in short supply in agriculture.

The phosphate that dissolves in the soil as rocks are weathered and eroded is absorbed by plants through their roots and flows through food chains and webs, eventually being returned to the ground when organisms die and decompose (Figure **C.33**).

The rate of turnover in the phosphorus cycle is much lower than in the nitrogen cycle. Weathering and erosion are slow, long-term processes so that phosphates remain locked up in the abiotic part of the cycle for long periods of time.

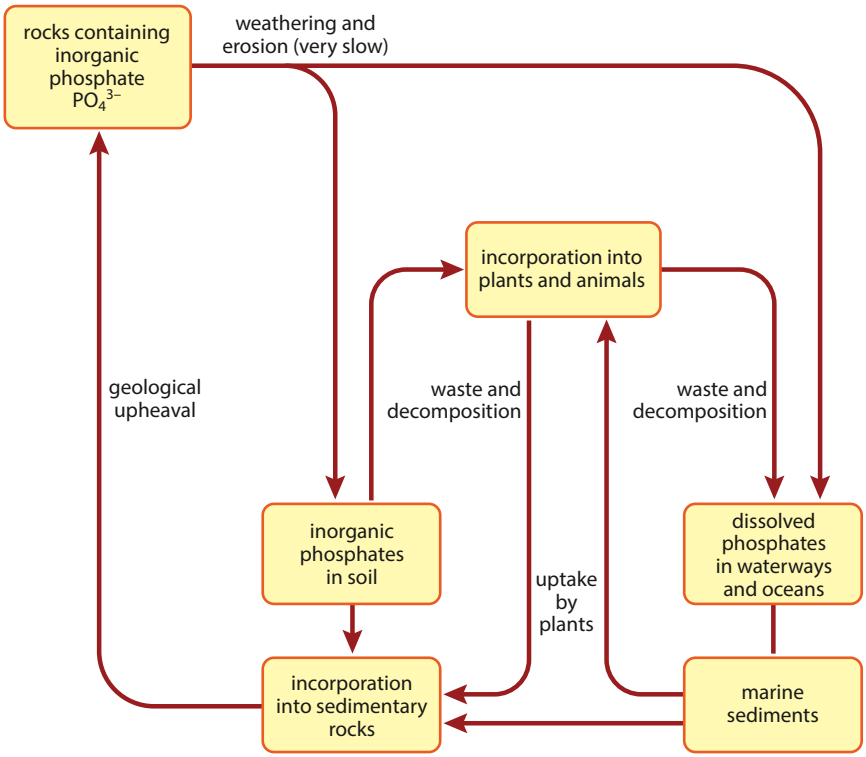


Figure C.33 The phosphorus cycle.

Agricultural demands for phosphate

Farmers have known about the importance of organic manure, which contains phosphate, for thousands of years and animal manure, compost and sewage sludge are still important sources of both phosphate and nitrates. But with the increasing world population and demand for food, the natural recycling of phosphorus has been interrupted and phosphate fertilisers have become more important. Where fertilisers are available the quantity and quality of agricultural output has been increased.

Phosphate is mined on a large scale to make fertilisers. Today plant and animal health problems caused by lack of phosphate have been eliminated in developed countries and more than 30 countries extract phosphate rock for use in fertilisers. The three major producing countries are the USA, China and Morocco, who together produce approximately two thirds of the world's phosphate. Moroccan reserves account for around 50% of the world total. But phosphate remains in short supply in many countries due to economic or political limitations.

Increased life expectancy, lower child mortality and improved farming methods allowing increased food production have led to exponential global human population growth over the last 150 years. Rising populations and increased wealth have also increased the demand for higher-grade foods – for example, the proportion of meat and dairy products in the world diet is rising. Increasing meat consumption leads to an increase in the need for cereals to feed farmed animals and thus a greater demand for phosphate fertilisers. It is possible that availability of phosphate may one day limit the expansion of agriculture worldwide.

Eutrophication

Eutrophication is defined as the natural or artificial addition of nutrients (especially nitrates and phosphates) to water, which leads to a reduction or depletion of the oxygen content of the water.

Phosphorus is present in agricultural fertilisers, animal manure and sewage, and household detergents have also been a major source. Nitrates from fertilisers and animal manure also contribute to eutrophication. Excess fertiliser can run off the land, particularly in areas where large numbers of livestock are kept or where slurry is used as a fertiliser. Soil erosion also deposits both manure and artificial fertilisers in waterways, especially where forests have been cut down and **leaching** of minerals from the soil by rainwater is therefore increased. Nitrate and phosphate flowing into rivers and streams can cause ecological problems and eventually lead to eutrophication.

If manure or sewage enters a river, the following processes occur:

1. Saprotrophic bacteria and fungi feed on the organic material in the raw sewage as a source of nutrients, and multiply. These aerobic organisms use up a large amount of oxygen and reduce its concentration in the water. They are said to cause an increased **biochemical oxygen demand** (BOD).
2. When the oxygen level drops, river organisms, including fish and many invertebrates that are highly dependent on high oxygen levels, die or move to other unpolluted areas if they can.

Crop rotation

Crop rotation is a traditional method of farming which involves changing the type of crop grown in a field on an annual or regular basis. Farmers alternate different types of plants to increase soil fertility and prevent pests and diseases becoming established. A typical rotation might be a root crop, followed by cereals, then brassicas (cabbages) and finally legumes (Figure **C.34**). Legumes, such as clover, add nitrogen to the soil, while wheat and potatoes use up nutrients. A crop rotation will often include a 'rest' period for an individual field. In a resting field, grass or clover can be planted for a season or longer, and then grazed or ploughed into the soil to add fertility.

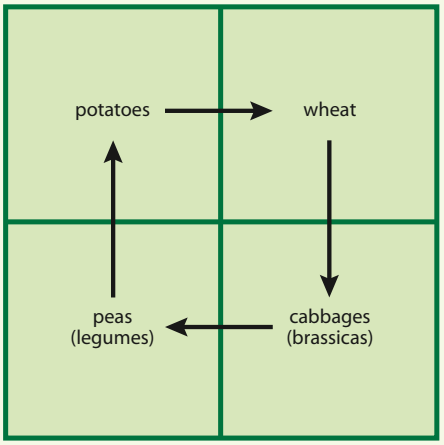


Figure C.34 Crop rotation.

3 Death and decay of the sensitive organisms leads to a build-up of ammonia, phosphate and minerals.
4 Ammonia is converted to nitrate and with this increased concentration of nutrients, algae reproduce rapidly. This is known as **eutrophication** (Figure **C.35**).
5 In time, the increased photosynthesis by the large amounts of algae that use the nitrate to grow can restore the levels of oxygen, so the river returns to normal.
6 But if the algae produce an **algal bloom** and then die and decay, this may cause a cycle of events that reduce oxygen concentration again and lead to the death of other organisms. In this case river, the takes longer to recover. If the algae do not die, then the river can recover from sewage pollution, although this may be several kilometres downstream from where the pollution entered the river.

Biochemical oxygen demand (BOD) the amount of dissolved oxygen that aerobic organisms need to break down organic material present in a body of water, at a certain temperature, over a certain period of time.

In many countries, fertiliser use is controlled, and in modern farming the requirements of crop plants are closely monitored. Farmers are sometimes blamed for causing eutrophication through inappropriate use of fertilisers that are high in nitrates and phosphates. However, farmers cannot always be held responsible, for example, if it rains heavily after fertiliser is applied, much of it will pass through the soil to ground water, before crop plants have had a chance to absorb it.

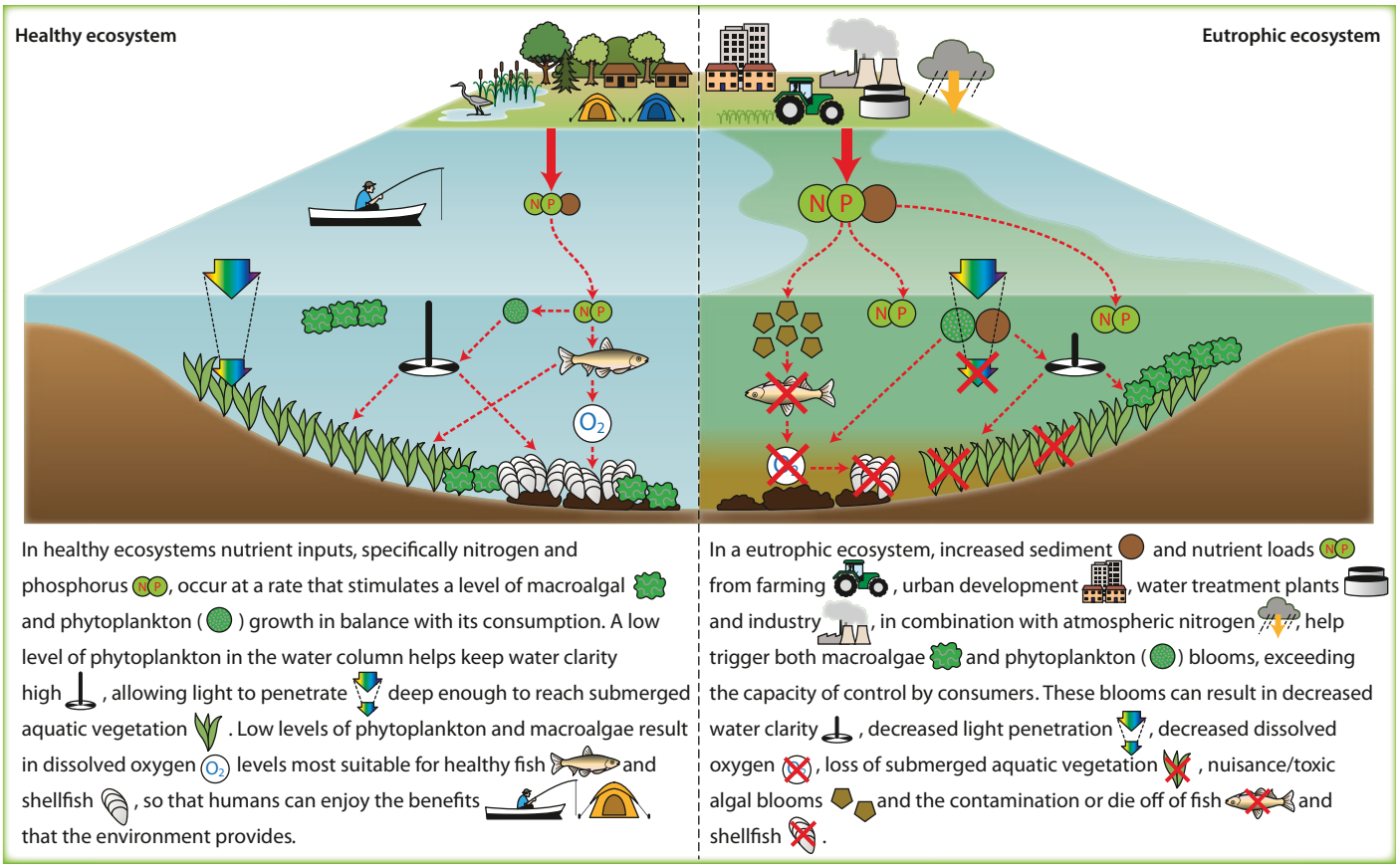


Figure C.35 Comparing healthy and eutrophic systems.

Nature of science

Assessing risk in science – use of phosphate fertilisers

Research has revealed that 'phosphate efficiency' is low in agricultural systems. One year after it has been applied, only about 15–25% of the phosphate in a fertiliser is taken up by crops. Much of it remains in the soil, bound to soil particles or other elements, and unavailable to crops for a long period of time. This may increase phosphate reserves in the soil – indeed, in many western countries the soil fertility has been improved by year-on-year application of phosphate fertilisers so that it may now be possible to reduce the level of phosphate in fertilisers and still maintain yields – but up to 25% will never be available to crops because it is in form they cannot use. Increased phosphate reserves in the soil are at risk of being leached out by rain and polluting waterways and thus causing an imbalance in the natural cycling of phosphorus.

The benefits of using phosphate and nitrate fertilisers must therefore be balanced against the potential damage that can be caused to ecosystems such as rivers and streams. Monitoring of these systems by environmental agencies produces data about risks to their structure, which must be weighed against the value of the sustained increases in food production that the burgeoning human population requires.

Test yourself

17 Name **one** species of nitrogen-fixing bacterium.

18 State what is meant by the term 'eutrophication'.

19 Outline the reasons why some plant species have evolved to be insectivorous.

Exam-style questions

1 Explain **three** factors that affect the distribution of plant animal species. **[4]**

2 **a** Discuss the strengths and weaknesses of an approach to conservation that favours a high-profile mammal. **[3]**

b Explain the importance of shape and size in the design of a nature reserves. **[4]**

c Outline **two** advantages of wildlife corridors. **[2]**

3 The Simpson's diversity index is used to calculate the diversity of an ecosystem.

a Calculate the diversity index (*D*) of this area of sand dune (using the formula in Subtopic **C4**). **[2]**

Species	Number (*n*)
sea holly	20
sedge	80
sea bindweed	10
Portland spurge	30
sea spurge	10
Total	150

b A similar area of sand dune at a different location was sampled and found to have a higher diversity index. What does this tell you about the second sand dune? **[1]**

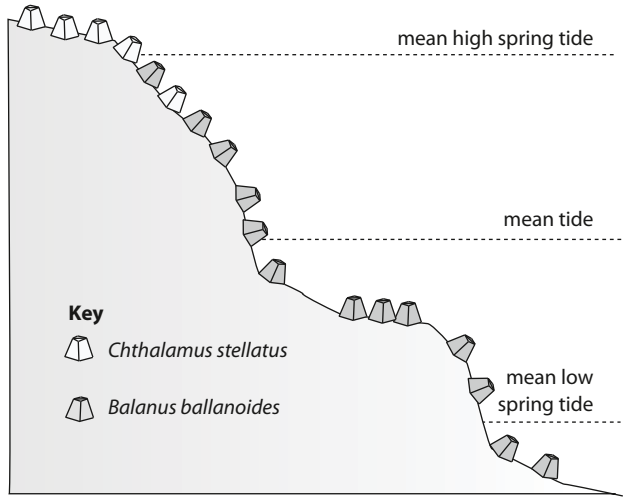
4 **a** Bacteria in a culture vessel divide so that the population doubles every 20 minutes. Complete this table to show the number of bacteria present over a period of 160 minutes. **[4]**

Time / mins	Population / 1000s
0	1
20	2
40	4
60	8
80	16
100	
120	
140	
160	

b Construct a graph of these data and describe the shape of the curve. **[2]**

c If no additional nutrients were added to the culture of bacteria, predict what would happen to the shape of the curve and explain your answer. **[3]**

5 Two species of barnacle *C. stellatus* and *B. balanoides* live on rocky shores of Europe. The distributions of the two species are shown in the diagram. Below the tide line there is abundant growth of seaweed on the rocks and large numbers of a predatory whelk that feeds on barnacles.



a Suggest **two** factors that may be limiting the distributions of each of the two species at the upper and lower limits of their ranges. **[4]**

b If a rock on which *C. stellatus* is found is placed lower down the shore the barnacles can survive very well. Suggest a reason why they are not normally found on the lower shore at this location. **[1]**

6 Calculate the species frequency of an organism if 200 quadrats are used to sample an area and the species is present in 86 of them. **[2]**

7 Outline **four** ways in which macroplastic debris accumulating in the oceans has harmed marine species. For each suggestion outline how the animal is harmed. **[4]**

8 The cane toad is an example of an alien species that was introduced into Australia and has become invasive. Outline **three** reasons why this species has caused the reduction of the number of endemic and other species. **[3]**

9 Explain the use of indicator species and biotic indices in monitoring pollution in a stream environment. **[4]**

10 Discuss the advantages of *ex situ* conservation of endangered species. **[5]**

11 **a** Outline the technique of capture–mark–release–recapture to estimate a population. **[3]**

b List **three** limitations of this method. **[3]**

Option D Human physiology

Introduction

Human physiology involves examining not only the structures of the human body but also how they function together in harmony. Nerves and hormones work together under the control of the brain to ensure that the heart beats at the correct rate, that digestion is completed by all the right enzymes produced at the right times, and that our breathing rate matches our activity level. After digestion, nutrients are processed, stored or disposed of so that our cells have the resources they need to live, grow and repair themselves. If an emergency arrives, the body is prepared; and if we should move to a high altitude, our physiology is modified so that we adapt to the different conditions.

D1 Human nutrition

The food we eat keeps us alive and provides the nourishment we need to grow, repair our bodies and stay active. A balanced diet gives us all the essential substances that we need in just the right quantities. Our needs differ depending on our age, activities and lifestyle. Some people who do not get the balance right may become overweight, underweight or even seriously ill. The choices people make about what to eat depend on where they live but are also influenced by social issues.

Nutrients

Nutrients are chemical substances, found in foods that are used in the human body. We need a number of nutrients to build our bodies and to stay healthy. We obtain these from the foods we eat. **Essential nutrients** are those that cannot be made in the body and must therefore be included in the diet. Essential nutrients are:

- essential amino acids
- essential fatty acids
- vitamins
- minerals
- water.

Minerals and vitamins are both needed in very small quantities. Their chemical structures are quite different – vitamins are organic compounds, whereas minerals are usually derived from inorganic ions. For example, sodium in the diet is available as Na^+ ions.

Although carbohydrates are a very important part of the human diet, there are no specific carbohydrates that are essential so they are not included in this list.

Nutrient a chemical substance taken in by a living organism and used for growth or metabolism – for humans, nutrients are found in the food we eat

Learning objectives

You should understand that:

- Essential nutrients must be included in the diet because they cannot be synthesised by the body.
- Dietary minerals are chemical elements that are essential nutrients.
- Vitamins are chemically diverse carbon compounds that are essential nutrients.
- Some amino acids and some fatty acids are essential nutrients.
- A lack of essential amino acids affects protein production.
- Malnutrition can be caused by an imbalance, deficiency or excess of nutrients in the diet.
- A centre in the hypothalamus controls appetite.
- Overweight people are more likely to suffer with type II diabetes and hypertension (high blood pressure).
- Starvation can lead to the breakdown of tissue in the body.

Exam tip
Remember that the term 'essential nutrients' does not refer to all substances that the body needs, but only to necessary substances that the body can't synthesise for itself.

Figure D.1 Kwashiorkor is a protein deficiency disease seen in young children, resulting from a diet low in protein, energy and other nutrients. The children do not grow properly and suffer from edema, which causes the swollen appearance of their abdomen.

Malnutrition the insufficient, excessive or imbalanced consumption of nutrients, which leads to health problems

Aspartame

The artificial sweetener aspartame contains phenylalanine, so children with PKU must avoid it. You might have noticed labels on things like chewing gum and diet drinks, warning that the product contains phenylalanine – this warning is aimed at people with PKU, since phenylalanine is not harmful to most people.

Amino acids and proteins

To synthesise all the proteins in a human body, 20 different amino acids are needed. We can make some of these in our cells by converting certain nutrients into amino acids, but there are nine that cannot be synthesised and must be taken in as part of a healthy diet. These are known as **essential amino acids**.

Protein deficiency malnutrition can occur if an individual does not have enough of one or more of these essential amino acids. Protein deficiency can lead to poor growth and lack of energy, as well as loss of body mass. One of the most common conditions associated with protein deficiency is swelling of the abdomen (Figure **D.1**). A lack of protein in the diet prevents blood plasma proteins being produced properly. Blood plasma protein assists with the reabsorption of tissue fluid into blood capillaries and without it fluid remains in the tissues causing **edema** (swelling).

Phenylketonuria

Phenylketonuria (PKU) is a rare genetic metabolic disorder. In the USA, PKU occurs in only 1 in 15 000 births. It is caused by a mutation to a gene on chromosome 12. People who suffer from PKU lack an enzyme that is needed to process the amino acid phenylalanine. They are unable to make the liver enzyme tyrosine hydroxylase, which converts phenylalanine into another non-essential amino acid called tyrosine. Phenylalanine is essential for normal growth but if too much builds up in the blood, brain damage can result. The condition is treatable if it is diagnosed soon after birth. Babies and children with untreated PKU develop serious physical and mental health problems as levels of phenylalanine in their blood rise. In many parts of the world, a simple blood test at birth is used to identify babies with PKU. Children who are identified as having PKU must be given a special diet that is low in protein and especially low in the amino acid phenylalanine. They must avoid many common, high protein foods such as milk and dairy products, nuts, fish and meat. PKU only affects children until puberty. After this, they can have a normal diet.

Malnutrition, starvation or deficiency?

Malnutrition occurs when a person does not eat a balanced diet. The diet may mean the person is deficient in one or more nutrients, or suffers an imbalance from eating an excess of a particular nutrient. A person can eat a lot of food but still be malnourished.

Starvation is different from malnutrition – it occurs when an individual simply does not have enough to eat. Starvation can lead to the breakdown of body tissues as the individual first uses up stored carbohydrate and then protein from body structures as a source of energy for respiration.

A **deficiency** occurs when a person does not have enough of one particular nutrient and suffers health problems as a result.

Vitamins and minerals

Vitamins and minerals are usually listed together in diet information because, although they are both vital for good health, they are both needed only in very small quantities. Vitamins are chemically quite different from minerals and the two nutrient groups come from many different sources. Some key differences are shown in Table **D.1**.

The ability to synthesise vitamins varies in different animals. Most animals can synthesise vitamin C but there are a few notable exceptions including bats, guinea pigs, monkeys, apes and humans, which all need a dietary supply of vitamin C. Humans can synthesise vitamin D in the skin (see below) but cannot synthesise any other vitamins. A range of vitamins must be taken in as part of a healthy diet.

Vitamins are:	Minerals are:
made in plants and animals	substances derived from rocks or found dissolved in water
compounds	elements in ionic form e.g. phosphate (PO_4^-)
organic e.g. vitamin C ($C_6H_8O_6$)	inorganic e.g. iron (Fe^{2+}), calcium (Ca^{2+}), iodine (I^-)

Table D.1 Comparing vitamins and minerals.

Vitamin C

Nutritional labels on food products show the quantities of different nutrients that they contain, together with a recommended daily amount (RDA) for each one. The recommended level for vitamin C is about 50 mg per day. Vitamin C helps to protect the body from infection and is important in keeping bones, teeth and gums healthy and for synthesis of the protein collagen. A shortage of the vitamin leads to the deficiency disease called **scurvy**.

Vitamin C is found in citrus fruits such as oranges and lemons, and strawberries and kiwi fruit are also rich sources. Fresh vegetables also contain vitamin C but the amount is reduced by cooking.

Two main techniques have been used to work out how much vitamin C a person needs each day. The first involves the use of animal tests and the second uses human test subjects.

During tests involving animals, small mammals such as guinea pigs, which cannot manufacture vitamin C, are fed diets containing different levels of the vitamin, while all other nutrients are controlled. Levels of vitamin C in the blood can be measured and the health of the animals is monitored. After a time, animals receiving insufficient vitamin C show signs of deficiency, such as poor collagen in bones and increased rates of infection. The data collected can be used to calculate the amount of vitamin C required by a human.

Humans were directly monitored during a number of medical investigations carried out in Sheffield, UK, during World War II. The subjects were conscientious objectors – pacifists, who were allowed to volunteer for experiments as an alternative to military service. The young men and women were fed diets lacking in vitamin C for 6 weeks but were given supplements of 70 mg of L-ascorbic acid each day. (L-ascorbic acid is the chemical term for vitamin C.) The subjects were then divided into three groups. The first group continued to receive 70 mg of L-ascorbic acid per day, the second was given 10 mg per day and the third group received no L-ascorbic acid at all. After 6–8 months on this regime,

the volunteers deprived of vitamin C developed signs of scurvy while the other two groups did not.

The results indicated that 10 mg of vitamin C per day would be sufficient for good health but it is generally agreed that the recommended level should be higher to account for variation between people and provide a suitable level to protect people from scurvy and infection.

Test yourself

1 Explain why all the human volunteers were deprived of vitamin C in their diet but given L-ascorbic acid for 6 weeks before the experiments.

Recommended daily amounts

The Vitamin and Mineral Nutrition Information System (VMNIS) was set up in 1991 to strengthen the surveillance of micronutrient deficiencies in the world. Different countries establish different values for RDAs of various nutrients. The RDAs for vitamin C for an adult from various authorities are shown below.

World Health Organization	45 mg day^{-1}
Canada	75 mg day^{-1} for women; 90 mg day^{-1} for men
UK	75 mg day^{-1}
USA	60–95 mg day^{-1}

Human experimentation

Between 1942 and 1946, Professor John Pemberton was a Medical Officer to the Research Team that carried out the medical experiments in Sheffield, including those on vitamin C, and he has written about that research. In his paper, published in the *International Journal of Epidemiology* (2006), he says:

'The Sheffield conscientious objectors demonstrated, once again, how valuable medical knowledge can be obtained by human experimentation, and sometimes in no other way, if volunteers can be found who are willing to undergo considerable discomfort, pain and even serious risks to the health. The contribution of the volunteers to medical knowledge during 1939–45 should not be forgotten.'

Questions to consider

Today, experiments like these would not be permitted under the Helsinki Agreement of 1975, which promotes human rights.

- What are the ethical issues involved in such trials?
- Can experiments in which subjects may be put at risk ever be justified?

A reputation on the line

Linus Pauling (1901–94) was an American biochemist who won two Nobel Prizes for his work. In 1986, in his book *How to Live Longer and Feel Better*, he suggested that very large doses of vitamin C, as high as 1000 mg per day, would provide protection against colds and other minor respiratory tract infections. There was no conclusive experimental evidence to back up his claims but he is believed by many people because of his reputation and fame.

Question to consider

- How important is reputation in deciding whether or not a new theory or proposal is a good one?

Vitamin D

Vitamin D (calciferol) is needed to ensure that sufficient calcium is absorbed in the digestive system to build healthy bones. Vitamin D deficiency can lead to softening or malformation of the bones, a condition known as **osteomalacia** or **rickets** (Figure **D.2**). This condition can be a problem in growing children and breastfeeding mothers whose vitamin D intake is low.

Vitamin D is obtained from foods such as oily fish, particularly salmon and tuna, egg yolk, liver and dairy products including milk, cheese and butter. In some countries, milk is fortified by adding supplements of vitamin D.

Vitamin D is one of the few vitamins that can be made in the body. It is synthesised in the skin when it is exposed to ultraviolet (UV) rays from the sun. Only a short exposure is needed but in some countries at extreme latitudes there is insufficient sunlight in winter months for vitamin D to be made. Fortunately, the liver can store vitamin D that is produced during the summer.

Osteomalacia and rickets

Osteomalacia causes bones to become softened because they do not contain sufficient of two vital minerals – calcium and phosphorus. There are several causes of osteomalacia but most occur as a result of faulty metabolism of vitamin D or phosphorus. **Rickets** is the term used for the condition when it occurs in children, for whom the consequences can be particularly severe because the bones are still growing.

Balancing risks

For vitamin D to be produced, sufficient light must reach the skin. If an individual stays out of the sunshine, or protects the skin with a sun-blocking cream or clothing, they may not receive enough UV rays. On the other hand, excessive exposure to UV light leads to an increased risk of malignant melanoma, a form of skin cancer. It is important to balance the need for sunlight with the risk of too much exposure.

Fair-skinned people should always protect their skin and avoid intense sunlight. To minimise the risk of skin cancer, they should only expose their skin during early morning or late afternoon when the sunlight is less intense. Individuals with darker skin, which is protected by increased amounts of melanin pigment, require more time in sunlight than fair-skinned people to produce the vitamin D they need. The amount of time depends on the intensity of the sunlight, and therefore varies with distance from the equator and with the seasons of the year.

Both exposure to the sun and complete protection from it carry some risks. It is important to consider these risks but impossible to avoid them completely.

Question to consider

- How can conflicting knowledge claims be balanced?

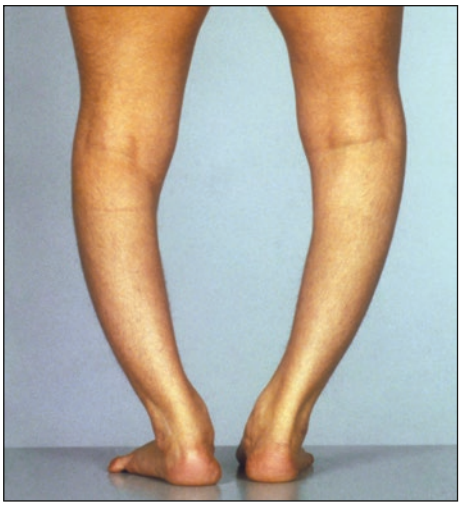


Figure D.2 People who suffer from rickets have bone deformities because vitamin D is vital for the incorporation of calcium and phosphorus into the bone matrix.

Appetite control

The **appetite control centre** is found in the **hypothalamus** at the base of the brain. Its job is to signal to the body when sufficient food has been eaten and the body has reached satiation. The hypothalamus receives information in four important ways:

- When the stomach is full, receptors in the stomach wall send messages to the brain via the vagus nerves.
- When food enters the small intestine, it secretes a peptide hormone called PYY 3-36, which suppresses appetite.
- Another hormone, leptin, is released by adipose (fat) tissue and this also suppresses appetite – a person who has more adipose tissue has more leptin-secreting cells (Subtopic **6.6**).
- Insulin from the pancreas, released as food is absorbed, also suppresses appetite.

Evidence for the importance of the hypothalamus in controlling the feeling of hunger has come from people who have damage in that area of the brain. In many cases, they have severe appetite-control problems being either unable to eat or unable to stop eating. Despite the controls that the hypothalamus provides for us, almost everyone can be encouraged to eat when they are not hungry by the appearance or smells of delicious food, or if we are tempted by enticing advertisements.

Risks associated with excessive body weight

Carbohydrates in the form of sugars and starch are the body's main source of energy, but if the energy is not used for day-to-day activities, the excess is stored either as glycogen in the liver and muscles, or as fat. The body's capacity to store glycogen is limited so eating an excess of carbohydrate means that fat reserves build up and over a long period of time can lead to weight gain and obesity.

Fat contains twice as much energy as carbohydrate, so eating an excess of fatty foods is very likely to lead to obesity. There is also a substantial risk of heart disease and other cardiovascular problems associated with an excess of saturated and *trans* fatty acids in the diet. (You can review the structures of these fats in Subtopic **2.3**.)

Type II diabetes

Type II diabetes is the most common form of diabetes, accounting for nine out of ten cases worldwide. It is also known as late-onset diabetes or non-insulin-dependent diabetes mellitus. Individuals who have the condition develop insulin resistance, which means that the receptor cells that normally respond to insulin fail to be stimulated by it, even though the beta cells in the pancreas still produce insulin (Subtopic **6.6**).

Causes and symptoms

The causes of type II diabetes are not fully understood but there is a strong correlation of risk with weight and diet. High levels of fatty acids in the blood may be a factor causing the condition and people whose diets are high in fat but low in fibre seem to be most at risk. Obesity, associated with a lack of exercise or a genetic makeup that influences fat metabolism, is a key risk factor. The condition is more common in older people but there are an increasing number of cases in overweight children.

Some ethnic groups are more likely to develop type II diabetes and this provides evidence for a genetic link to a predisposition to the condition. Aboriginal Australians, people of Asian and Afro-Caribbean origin, Native Americans and Maori peoples are all at a higher risk.

The symptoms of type II diabetes tend to develop slowly but include:

- high glucose levels in the blood
- glucose in the urine
- frequent need to urinate, which leads to dehydration and increased thirst
- tiredness and fatigue
- some loss of weight.

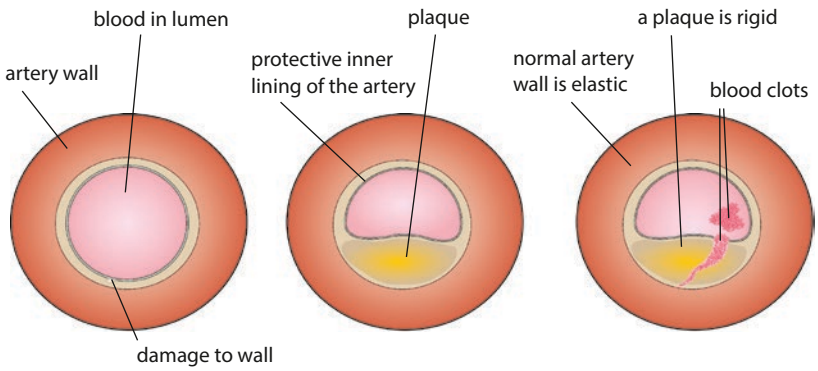
Hypertension and CHD

The relative amounts of different types of fatty acid in a person's diet can, in many cases, be correlated with health issues. Diets in societies around the world are very different and so is the incidence of coronary heart disease (CHD) and other diet-related illnesses. Diets in Mediterranean and Asian populations, such as those in Greece and Japan, tend to include a high intake of vegetables, fruit and wholegrain products, and lower amounts of fish and red meat, high-fat dairy products and other animal products. These populations have relatively low rates of CHD.

Eating a diet that is high in saturated fatty acids has been shown to have a positive correlation with an increased risk of CHD and other diseases of the circulatory system, including hypertension (high blood pressure). Saturated fatty acids can be deposited inside the arteries, and if the deposits combine with cholesterol they may lead to atherosclerosis, which reduces the diameter of the lumen and leads to hypertension (Figure **D.3**). Reliable evidence suggests that in countries where the typical diet is high in saturated fatty acids, and many high-fat foods, animal products and processed foods are eaten, there is likely to be a high incidence of CHD. Since all fatty acids are high in energy, an excess of these foods in the diet can also lead to obesity, which places a further strain on the heart.

Correlation and cause

Differences in CHD rates between countries may also be due to differences in other CHD risk factors, including physical activity and obesity, and not only be due to diet. It is important to remember that correlation between two variables does not necessarily indicate cause. In addition, dietary patterns vary with gender, socioeconomic status, culture and ethnic group.



1 Part of the lining of the artery is damaged.

2 Gradually, over time, cells divide in the artery wall and there is a build-up of lipids.

3 Exercise or stress can make the plaque break. Blood enters the crack. Platelets in the blood are activated and a clot forms. Part of the clot may break off.

Figure D.3 The development of atherosclerosis in an artery.

Another indicator of the risk of CHD is the level of cholesterol in an individual's blood. **Cholesterol** is a steroid that is synthesised in the liver and found almost exclusively in foods of animal origin. It forms part of the cell membrane and it helps in the transport of substances in and out of cells, communication between cells and the conduction of impulses along nerve cells. Cholesterol is carried in the blood to the liver, where it can either be broken down or excreted in the bile (Subtopic **D3**). Cholesterol is transported around the body in the form of two types of **lipoprotein**.

- **Low-density lipoproteins** (**LDLs**) are often referred to as 'bad cholesterol' because they do not travel easily in the bloodstream and can clog up arteries, causing atherosclerosis, CHD or stroke. Raised LDL levels in the blood frequently occur in people who have high levels of saturated or *trans* fatty acids in their diet.
- **High-density lipoproteins** (**HDLs**) are sometimes known as 'good cholesterol'. HDLs are carried easily in the blood and do not contribute to blockages in the arteries. Evidence also suggests that HDL cholesterol can help to remove LDLs from arteries.

A definite link has been found between the intake of saturated fats in the diet and the level of cholesterol in the blood. High levels of saturated fats in food increase both LDL and total blood cholesterol levels. There is also a clear correlation between saturated fats in the diet and CHD. Medical professionals recommend reducing the intake of saturated fats to reduce blood cholesterol and the risk of a heart attack. However, it may be that reducing the levels of cholesterol in the form of LDLs is more important than reducing the total cholesterol level in the blood.

Starvation

Starvation occurs when an individual simply does not have enough to eat. Starvation can lead to the breakdown of body tissues as the individual first uses stored carbohydrate and then protein from muscles and other tissues as a source of energy.

One extreme case of starvation is **anorexia nervosa**, an eating disorder in which individuals stop eating a balanced diet and become seriously underweight because they have an obsessive fear of gaining body mass. The condition has psychological causes and people with the disorder are unable to appreciate that they have a problem, often perceiving themselves as overweight despite being normal or underweight. The condition affects mostly young women – approximately 90% of cases of anorexia nervosa are in females, with few men or boys being affected.

Among the many consequences of anorexia nervosa are serious disturbances of the endocrine system, anemia, and loss of hair and muscle mass. The lack of food causes dehydration and low blood pressure, which in turn can cause fainting and, in the longer term, damage to the kidneys and liver. In extreme cases, not only voluntary muscle but also heart muscle may be broken down to supply amino acids that can be used for respiration. A shortage of calcium can permanently damage the teeth and weaken bones, while a shortage of other ions can lead to irregular heart beat or even a heart attack.

Nature of science

Falsification of theories – replacement of one theory with another

Early investigations into vitamin C deficiency and scurvy proved unsuccessful because the animals used in the studies were laboratory rats and mice. If these animals were deprived of dietary vitamin C they showed no signs of scurvy. This led to the suggestion that scurvy was exclusively a human condition. It was not until laboratory guinea pigs were studied that the ability of animals to synthesise their own vitamin C was considered. Researchers concluded that scurvy was not exclusive to humans but was limited to those mammals that are unable to synthesise the vitamin in their own bodies.

Test yourself

2 Define what is meant by 'essential nutrient'.

3 State where in the body the appetite control centre is located.

D2 Digestion

Digestive juices

Saliva, gastric juice and pancreatic juice

Digestive juices contain many different enzymes and other substances that assist with digestion (Subtopic **6.1**). The contents of secretions from exocrine glands in the mouth, stomach and pancreas are summarised in Table **D.2**.

Digestive juice	Site of production	Contents
saliva	salivary glands in the mouth	• water • mucus • salivary amylase
gastric juice	gastric glands in the stomach wall	• water • mucus • pepsin secreted as pepsinogen • hydrochloric acid
pancreatic juice	exocrine cells in the pancreas	• water • pancreatic amylase • trypsin secreted as trypsinogen • pancreatic lipase • carboxypeptidase chymotrypsin • hydrogencarbonate (HCO_3^-) ions

Table D.2 The contents of saliva, gastric juice and pancreatic juice.

Pepsin and trypsin are protease enzymes that digest protein and are therefore potentially harmful to the cells that produce them and structures that they come into contact with. In order to prevent damage

Learning objectives

You should understand that:

- The secretion of digestive juices is controlled by nervous and hormonal stimulation.
- Exocrine glands secrete into the lumen of the intestine or onto the surface of the body.
- The volume and content of gastric secretions released into the stomach and small intestine is carefully regulated by nervous and hormonal mechanisms.
- The acidic nature of the stomach provides a favourable environment for some types of hydrolysis and also helps to kill pathogens in ingested food.
- The cells of the epithelium of the villi are adapted to absorb food.
- There is a positive correlation between the rate of transit of material through the large intestine and the fibre content of the material.
- Materials that are not absorbed in the gut are egested.

to the proteins in body cells, both these enzymes are secreted as inactive **precursors**: pepsinogen and trypsinogen, respectively. These are converted to their active form where they are needed, after secretion.

- **Pepsinogen** is converted to **pepsin** in the stomach when in the presence of hydrochloric acid, which is secreted by different cells in the stomach lining. The inner lining of the stomach is protected from both hydrochloric acid and pepsin by a thick layer of mucus.
- **Trypsinogen** is activated by the enzyme enterokinase (or enteropeptidase), which converts trypsinogen to **trypsin**. The activating enzyme is secreted by the walls of the small intestine when food enters from the stomach.

Exocrine glands

Digestive juices are produced in **exocrine glands** in the mouth (salivary glands), in the stomach wall (gastric glands) and in the exocrine tissue of the pancreas. All exocrine glands secrete their products via ducts to where they are needed. Cells of an exocrine gland that produces digestive juices are arranged in a single layer around small ducts that are connected to the intestine. One group of exocrine cells arranged around a duct is called an **acinus** (Figure **D.4**) and one exocrine gland contains many acini. The small ducts join together to form one larger duct, which carries the secretions to their destination.

Exocrine gland cells of the digestive system produce enzymes that are proteins and so the cells contain an extensive rough endoplasmic reticulum, which is the site of protein synthesis. Also visible inside the cells are numerous ribosomes, Golgi apparatus for packaging and processing the enzymes and large numbers of vesicles, which store enzymes before they are secreted by exocytosis into the ducts of the gland. Exocrine cells also contain a clearly visible nucleolus in the nucleus for the production of ribosome subunits, and numerous mitochondria to produce ATP for protein synthesis (Figure **D.5**).

Exam tip

Remember the differences between exocrine and endocrine glands:

- **Exocrine** glands, such as salivary glands and sweat glands, produce secretions that pass along ducts to their site of action. They are not part of the endocrine (hormonal) system.
- **Endocrine** glands secrete hormones directly into the bloodstream and they travel in the circulation to their target organs.

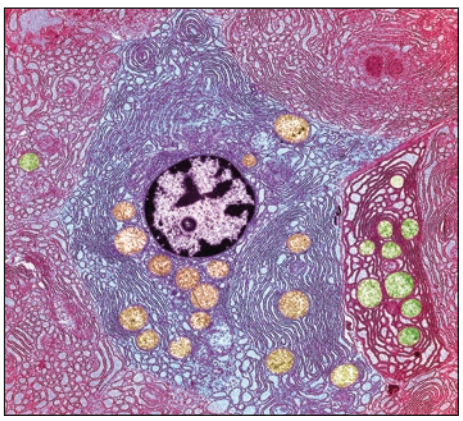


Figure D.5 Coloured electron micrograph of a pancreatic cell showing rough endoplasmic reticulum and many mitochondria.

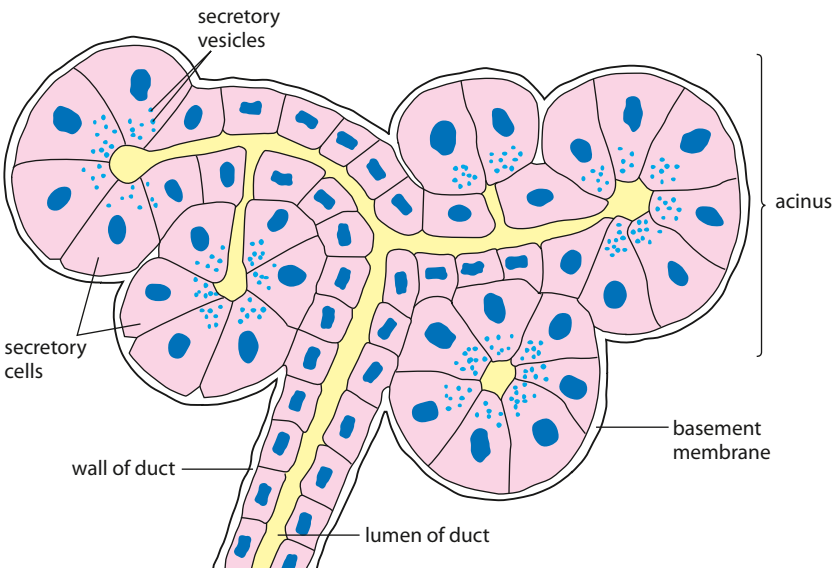


Figure D.4 A group of acini in exocrine tissue of the pancreas.

Stomach acid

Hydrochloric acid released into the stomach maintains the contents at pH 2, which is the optimum pH for the hydrolysis of protein molecules by protease enzymes. The acid conditions denature proteins so that peptide bonds in protein molecules are exposed to protease enzymes. Acidic conditions also kill bacteria that may be present on ingested food.

Gastric acid is produced by cells known as **parietal cells**, which line the stomach. In the plasma membranes of these cells are hydrogen–potassium pumps with ATPase enzymes, which pump H^+ ions from the cell into the stomach lumen. Parietal cells are stimulated to increase acid production when it is needed. Other cells in the stomach release hydrogen carbonate ions, which prevent the contents becoming too acidic, and also produce mucus forming a barrier against the stomach contents so that the lining is not damaged by the acid.

Excess acid in the stomach can cause acid indigestion, which, in mild cases, can be treated with proprietary antacid preparations. But if a person suffers from excess acid for long periods of time, the amount of acid that their stomach produces may need to be controlled. Acid production can be reduced using proton pump inhibitors (PPIs), a group of drugs that are very effective inhibitors of acid secretion, and are used to treat indigestion, acid reflux and stomach ulcers. PPIs work by blocking the hydrogen–potassium pumps and ATPase enzymes in the plasma membranes of the parietal cells, the last step in the H^+ secretion process.

Control of secretion of digestive juices

The production of digestive juices requires energy, and large numbers of mitochondria can be seen in the exocrine cells that produce them (Figure **D.5**). The body controls the secretion of digestive juices so that they are released at the correct time when food is present in the alimentary canal and so that neither enzymes nor energy are wasted. Both the nervous and hormonal systems are involved in controlling the volume and content of digestive juices that are produced.

Experience tells us that the sight or smell of food stimulates the production of saliva, and the nerve impulses that cause this response also stimulate the release of gastric juice in the stomach. Both responses are **reflex actions**. When food enters the stomach, more gastric juice is released as touch and stretch receptors in the stomach wall send impulses to the brain. Chemoreceptors in the stomach lining also send impulses to the brain, which controls the continued stimulation of the gastric glands.

Nerve impulses also pass to **endocrine glands** in the stomach wall that produce the hormone gastrin. Gastrin is released into the bloodstream and stimulates gastric glands to produce more hydrochloric acid and continue the production of gastric juices for periods of several hours. If lipids are present in the stomach, the hormone enterogasterone is also released. This hormone decreases the flow of gastric juice and delays the exit of fat-containing food from the stomach. As partly digested food leaves the stomach, two further hormones – secretin and CCK-PZ (cholecystokinin-pancreozymin) – are released from the small intestine. The effects of these hormones, which are all produced by endocrine glands and released into the bloodstream, are summarised in Table **D.3**.

Exam tip

Sketch a flow diagram to show the interaction of hormones, glands and nerve impulses involved in controlling digestive juices.

Hormone	Site of production	Effect of hormone
gastrin	stomach wall	• stimulates production of hydrochloric acid (HCl) and gastric juices
enterogasterone	stomach wall	• slows flow of gastric juice • slows exit of fats from stomach
secretin	small intestine	• stimulates the pancreas to release hydrogen carbonate (HCO_3^-) ions to neutralise acidic chyme (partly digested food) from the stomach
CCK-PZ (cholecystokinin)	small intestine	• stimulates the release of bile from the gall bladder (Subtopic **D3**) • and the release of pancreatic enzymes into the small intestine

Table D.3 Hormones involved in the control of intestinal secretions.

Important digestive disorders

Helicobacter pylori infection and stomach ulcers

Helicobacter pylori is a spiral-shaped bacterium that is able to grow in the human stomach. Unlike other bacteria, it can tolerate acidic conditions and survives well at the pH values found in the stomach. The organism was brought to the attention of the medical profession in the 1980s by two Australians – Dr Barry Marshall and Dr Robin Warren – who isolated it from the stomach linings of patients who were suffering from stomach ulcers and inflammation of the stomach lining (gastritis). Marshall and Warren proposed that *H. pylori* caused these symptoms. Until this time, stomach ulcers were thought to be caused by the excess secretion of acid, which caused damage to the stomach lining. Ulcers were associated with a stressful lifestyle, which was said to cause excess production of gastric juice. Since the 1980s, plenty of evidence has been gathered to support the hypothesis proposed by Marshall and Warren.

- Ulcers used to be treated with antacid treatments, which relieved the symptoms for relatively short periods of time. Today, antimicrobial drugs that kill bacteria and remove *H. pylori* infection provide long-term relief of symptoms and cure ulcers.
- *H. pylori* is regularly found in patients with both gastritis and ulcers. Many strains of the bacterium produce toxins that cause inflammation of the stomach lining.

Stomach cancer occurs far more frequently in patients who are infected for many years with strains of *H. pylori* than in non-infected people, so the bacterium seems to increase the risk of stomach cancer. But millions of people are infected with these bacteria and most of them do not get stomach cancer so there must be other factors at work. *H. pylori* has not been established as the sole cause of stomach cancer.

Cholera and the effect of cholera toxin

Cholera toxin is produced by virulent strains of the cholera bacterium *Vibrio cholerae*, which can infect the intestine if ingested in contaminated water or food. The toxin binds to surface receptors on mucosal cells of the intestine known as enterocytes. The cells take in the toxin, which disrupts proteins inside the cell leading to over-production of cAMP. cAMP modifies a chloride channel protein in the cell membrane so that chloride ions are pumped out of the cell in a process involving ATP. This efflux of chloride ions is followed by the secretion of water, sodium ions, potassium

ions and hydrogencarbonate ions into the intestine. The cells replace lost water and ions from the blood in a continuous process, so that the overall effect of the toxin can be a loss of up to 2 litres of fluid per hour from the intestine. A person with cholera quickly becomes severely dehydrated, suffering from devastating diarrhea (so-called 'rice water' feces), which can contaminate water and spread the cholera bacteria to others.

Cholera in the world today

The worldwide incidence of cholera is a cause for concern but it is less prevalent today in areas where public sanitation is good and clean water is available. International agreements mean that people must be vaccinated before they travel to infected areas. In 2011, 32% of all cholera cases reported to the WHO were from Africa, which was a considerable reduction from 2001–09, when about 95% of cases were from that continent. Cholera outbreaks can be caused by damage to sanitation systems during times of war, natural disasters such as earthquakes, or when people are displaced from their homes. Following the 2010 earthquake in Haiti, there has been a serious cholera outbreak which began in October that year and has continued. It is thought to be the worst epidemic of cholera for two decades. By August 2013, it had killed over 8000 people and the situation was made worse by the fact that many people were living and being treated in tented camps, without proper sanitation systems.

A paradigm shift

'No one believed it', Staffan Normark, a member of the Nobel Assembly at the Karolinska institute, said at a news conference.

The discovery of *Helicobacter pylori* is an example of a paradigm shift. This term was first used by an American philosopher Thomas Kuhn in his book *The Structure of Scientific Revolutions* (1962). It describes a change in assumptions within a ruling scientific theory. A paradigm shift occurs when there are a significant number of anomalies that counter the accepted paradigm – in this case, the belief that bacteria could not survive in the stomach and that ulcers were caused by excessive acid production. The accepted theory is thrown into a state of crisis until a new paradigm is formed and gains its own followers. For a time, an intellectual 'battle' will occur between the followers of the old and new paradigms.

For some time after Marshall and Warren's initial discovery of *H. pylori* in their patients, the well-established idea that bacteria could not survive in acid conditions persisted, despite evidence to the contrary. The men's proposed hypothesis was outside the mainstream view of the time (even though there had been some anecdotal and published evidence regarding antibiotic treatment of ulcers), and Marshall and Warren had to persevere in the face of considerable scepticism. They tested their theory, gathered evidence to support it, published their results and eventually they overturned the prevailing notion that ulcers were caused by stress and diet, based on the evidence of their experiments. Marshall even decided to deliberately infect himself with the bacterium in 1985 in order to show from his own experience that it caused stomach inflammation, a potential precursor of an ulcer. Their persistence paid off and the two men were awarded Nobel Prize for medicine in 2005 for showing that bacterial infection was to blame for painful ulcers in the stomach and intestine.

Questions to consider

1 Marshall and Warren used themselves as experimental subjects to provide evidence to support their hypothesis. Do you think there is a case for carrying out research on human subjects?
2 Discuss the ethical implications of doing research on humans.

Absorption of food

Structure of the small intestine

Absorption of digested food occurs in the ileum, which is part of the small intestine (Subtopic **6.1**). The structure of this region is related to its function. The surface area provided for absorbing food is increased enormously by folding of its inner lining into structures known as villi (Figure **D.6**) and each villus contains capillaries and a lacteal to transport absorbed molecules.

The inner lining of the small intestine is known as the **intestinal mucosa**. It is this layer that is responsible for absorbing food. The longitudinal and circular muscles in the intestine wall contract to move food along.

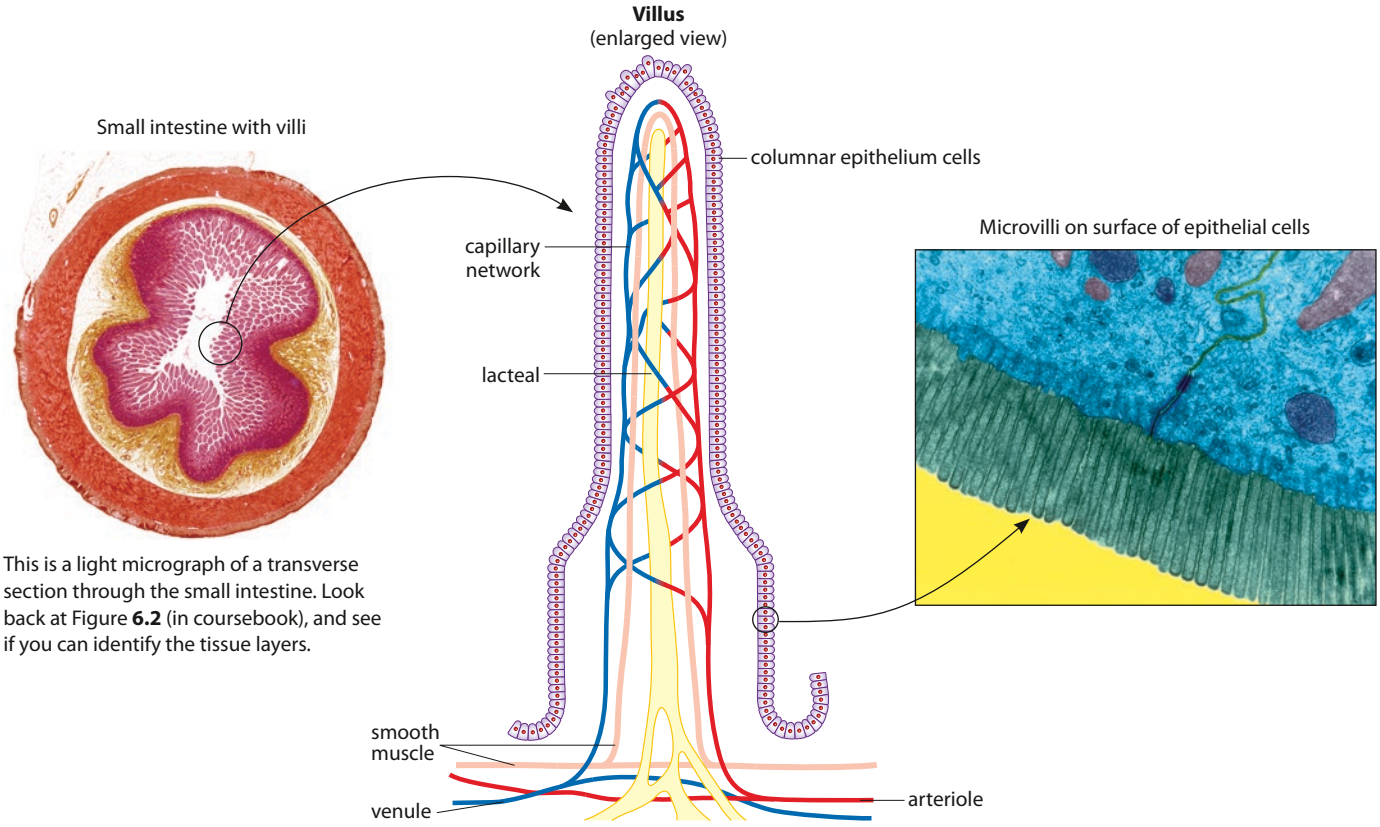


Figure D.6 The structure and microstructure of the small intestine.

Structure of a villus

Each fold of the intestinal mucosa is known as a **villus** and each villus also has many tiny projections known as **microvilli**. These structures produce a very large surface area for the absorption of digested materials.

Digested material must pass through the microvilli of the epithelial cells that make up the villi in order to reach a capillary or lacteal vessel. These epithelial cells contain structures that are vital to the processes of absorption. They have many mitochondria indicating that some absorption occurs using active transport and requires energy. In addition, many vesicles are present and these structures show that some materials

are taken in from the intestine by the process of **pinocytosis**. The epithelial cells are linked together by **tight junctions**, which seal each cell from the adjacent cell. The two adjacent membranes share some proteins and the tight junction between them prevents materials passing in between the cells. Most molecules are forced to pass straight through the cells from the lumen of the small intestine and into capillaries on the other side.

Digested molecules are small enough to pass through the epithelial cells and into the bloodstream. Movement can occur by a number of means (Subtopic **1.4**).

- **Simple diffusion** can occur if molecules are small and can pass through the hydrophobic part of the plasma membrane. For example, short-chain fatty acids and vitamins A, D, E and K are absorbed by simple diffusion.
- **Facilitated diffusion** occurs in the case of molecules such as fructose, which are hydrophilic. Channel proteins in the epithelial cell membrane enable these molecules to move, provided they are small enough and there is a concentration gradient which permits diffusion.
- **Active transport** is used to transport molecules that do not have a sufficiently high concentration gradient to pass by diffusion. Glucose, amino acids and mineral ions are all absorbed by this method. Mitochondria produce the ATP needed for active transport by the membrane pumps.
- **Pinocytosis** also draws in small drops of liquid from the ileum. Each droplet is surrounded by small sections of membrane that invaginate to form a vesicle. The vesicles are taken into the cytoplasm where their contents can be released.

Undigested material

Materials such as cellulose, which cannot be digested or absorbed, pass right through the intestine and are **egested** as part of the solid waste, or feces. Feces contain not only cellulose but also lignin from plant cell walls and bacteria that live in the digestive system and are carried through it. Cells of the intestine wall that are worn away as food travels past them also form part of the feces, as well as bile pigments containing material from the breakdown of red blood cells, which give the feces their familiar colour. The rate at which these materials pass through the large intestine is influenced by the amount of fibre in the feces.

Dietary fibre is found in cereals, fruits and vegetables. Fibre is made up of the indigestible parts or compounds of plants, which pass relatively unchanged through the stomach and intestines. There are two categories of fibre and both are beneficial to the body and important in a healthy diet. Most plant foods contain a mixture of both types.

- **Soluble fibre** includes pectins, gums and mucilage, which are found mainly in fruits, vegetables, oat bran, beans and lentils. Soluble fibre soaks up water like a sponge, which helps to slow down the rate of digestion and helps the contraction of intestinal muscles. This type of fibre lowers LDL ('bad') cholesterol levels (Subtopic **D1**).
- **Insoluble fibre** includes cellulose, hemicelluloses and lignin, which make up the structural parts of plant cell walls and are found in bran, fruit skins, vegetables and wholegrain foods. The most important role of insoluble fibre is to add bulk to feces, which therefore pass through

Rates of passage

The '**transit time**' is the time food takes to pass through the whole intestine, from ingestion to egestion, and it varies depending on many factors including stress levels and the type of food that has been eaten. On average, a meal spends about 2 hours being digested in the stomach – approximately one hour after a meal, half the stomach contents will have passed into the small intestine and total emptying of the stomach takes about 2 hours. The small intestine then takes 1–2 hours to process half this food and pass it into the large intestine. Finally, food must pass through the colon, a process that can take between 12 and 50 hours. The wide variation in the time that material spends in the colon is related to an individual's diet and fibre intake.

the colon more easily, preventing constipation and associated problems. Insoluble fibre, which does not absorb water, speeds up the time that food takes to pass through the gut – there is a positive correlation between the amount of insoluble fibre in a person's diet and the rate of transit of materials through the large intestine. This type of fibre helps to reduce the risk of colon cancer.

Nature of science

Serendipity in science – the role of gastric acid in digestion

William Beaumont (1785–1853) was a United States Army doctor. As a result of a chance meeting with a Canadian fur trapper who had been shot in the stomach, he became famous for his studies of digestion in the stomach and is sometimes known as the 'father of gastric physiology'.

The Canadian fur trapper, Alexis St Martin, was wounded in an accident in 1822 and Beaumont was called to treat him. The wound was reported to be 'more than the size of the palm of a man's hand' and, although it healed, a year later small holes were still present in St Martin's skin. Over a period of years, Beaumont was able to lift a flap of skin and observe St Martin's stomach beneath. He not only watched the digestive process, he also extracted gastric juices and studied how they worked. He used extracted gastric juice to digest meat in his laboratory and then compared this with the digestion of a similar piece of meat that he inserted, attached to a string, into St Martin's stomach. Beaumont studied the digestion of a range of foods and also investigated the effect of emotion on digestion. Beaumont and St Martin worked together over a period of years and in 1833, Beaumont published his research as *Experiments and Observations on the Gastric Juices and the Physiology of Digestion*. He described many aspects of the digestive process and even suggested that alcohol could cause damage to the stomach.

The serendipitous meeting of Beaumont and St Martin had an important impact on the understanding of digestion and Beaumont became the first person to observe digestion in the stomach directly. Interestingly, St Martin did not suffer any ill effects from his wound or Beaumont's experiments. He died at the age of 86, outliving Beaumont by almost 30 years.

Test yourself

4 Outline the importance of the acid conditions in the stomach.
5 Explain how the structure of intestinal epithelial cells is related to their function.
6 State the importance of fibre in the human diet.

D3 Functions of the liver

Learning objectives

You should understand that:

- Toxins are removed from the blood and detoxified by the liver.
- The liver recycles components of red blood cells (erythrocytes).
- The breakdown of red blood cells (erythrocytes) begins with phagocytosis by Kupffer cells.
- Recycled iron is taken to the bone marrow and used to make hemoglobin in new red blood cells.
- Excess cholesterol is converted to bile salts.
- Plasma proteins are produced in hepatocytes, by the endoplasmic reticulum and Golgi apparatus.
- Blood from the small intestine flows to the liver, which regulates nutrient levels in the blood.
- The liver can store certain nutrients that are present in excess.

Structure of the liver

The liver is the largest internal organ in the body and makes up 3–5% of our body weight. It is situated just below the diaphragm and has many important roles including detoxification of poisons, recycling the constituents of worn out erythrocytes, storage of nutrients and the production of bile and plasma proteins.

The liver is an unusual organ because it is supplied by two large blood vessels, rather than one (Figure **D.7**). The first is the hepatic artery, a branch of the aorta, which carries oxygenated blood to the liver, and the second is the hepatic portal vein, which carries blood from the intestine. Blood in the hepatic portal vein is rich in nutrients that have been absorbed by capillaries in the small intestine. Of the total volume of blood in the body, 20% flows through the liver at any time.

The liver is divided into **lobules**, which are rows of **hepatocytes** (liver cells) arranged in a circular pattern around a central vein (Figure **D.8**). Between the rows of cells are **sinusoids**, which are a type of blood capillary that is much larger than the capillaries of other tissues. Blood from branches of both the hepatic portal vein and the hepatic artery flows along the sinusoids. The endothelial cells that line the sinusoids are very thin and well spaced. These structural features help with absorption of substances into the surrounding hepatocytes (Figure **D.8**).

Attached to the walls of the sinusoids are numerous phagocytes called Kupffer cells, which remove and break down bacteria and damaged red blood cells (erythrocytes). Blood from the sinusoids flows into the central veins, which unite to form the hepatic vein. The hepatic vein leaves the liver and joins the vena cava.

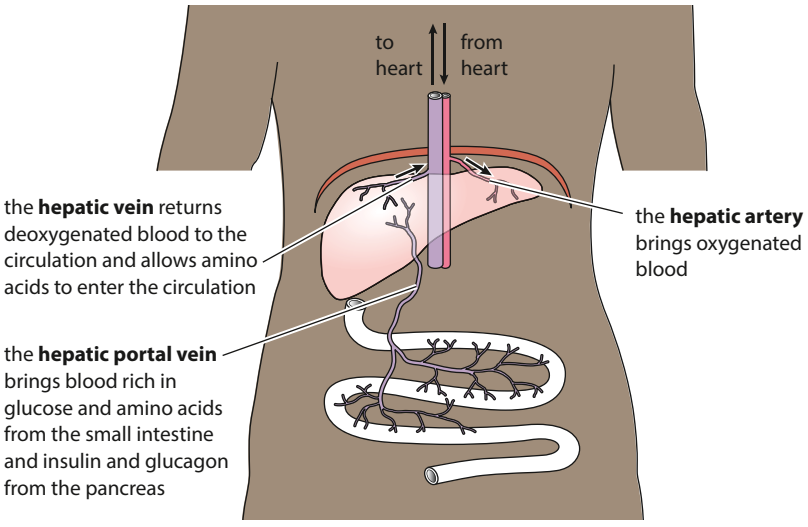


Figure D.7 Blood supply of the liver.

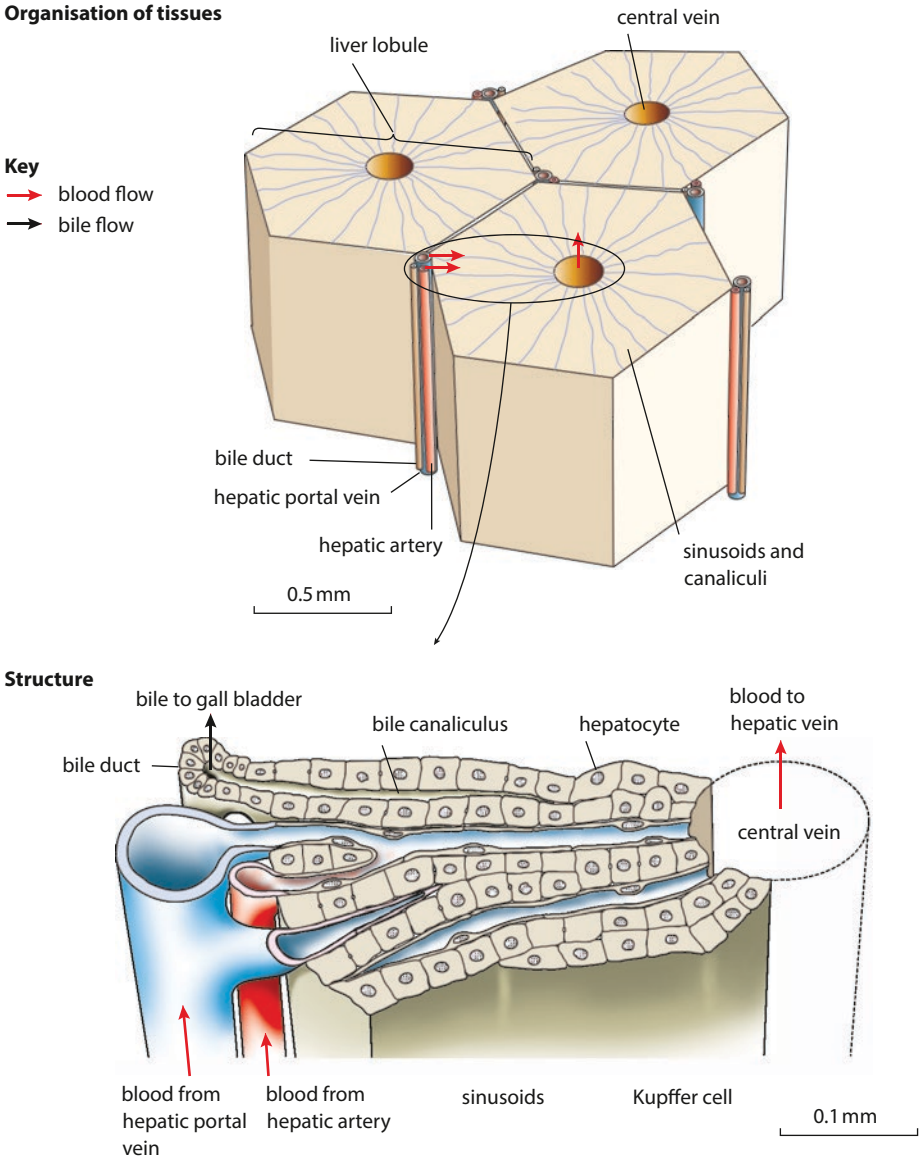


Figure D.8 Structure of a liver lobule.

Liver functions

Detoxification

The liver has an essential role in removing toxins from the blood and **detoxifying** them. Hepatocytes absorb toxins and convert them into non-toxic or less toxic products. Some of the toxins are by-products of metabolic reactions, such as lactate from anaerobically respiring muscles or hydrogen peroxide produced by processes including fatty acid metabolism. These toxins are broken down by the enzyme catalase. Other toxins that the liver processes are ingested substances such as alcohol, food additives and pesticides.

Erythrocyte and hemoglobin breakdown

Red blood cells survive in the bloodstream for about 120 days before they must be replaced with new cells from the bone marrow. At the end of their lives, red blood cells may break into fragments as their membranes become weakened and thus release free hemoglobin into the bloodstream. Cell fragments and hemoglobin are taken in by phagocytosis by the Kupffer cells in the sinusoids of the liver and the component parts of hemoglobin are broken down for recycling or excretion. This is shown in Figure **D.9**.

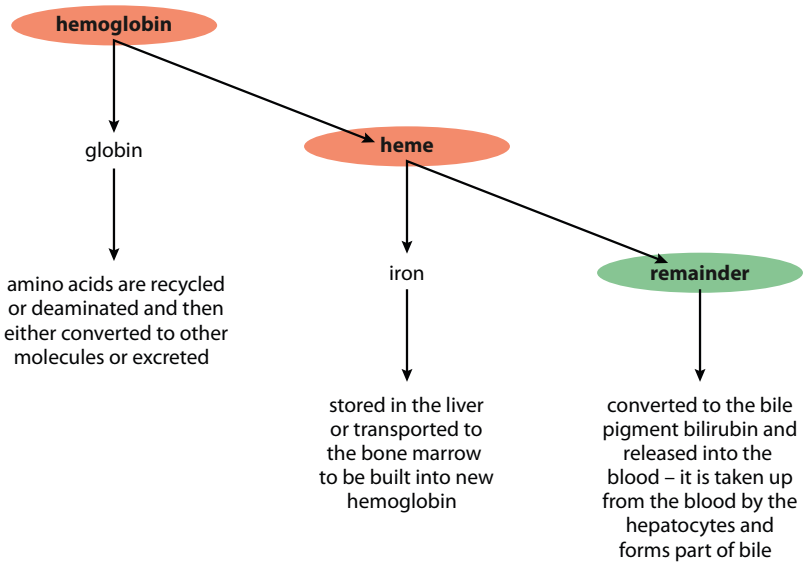


Figure D.9 The breakdown of hemoglobin. Hemoglobin is split into heme groups and globin. Globin is hydrolysed to amino acids, which can be re-used. Iron is removed from the heme group and either stored or taken to the bone marrow where it is used to produce hemoglobin for new red blood cells. The remaining part of the molecule becomes part of bile.

Synthesis

The liver synthesises both plasma proteins and cholesterol. Plasma proteins, found in the blood, play an important part in blood **homeostasis**. They are key to regulating the osmotic balance of body fluids and regulate the movement of water between plasma and tissue fluid, as well as affecting ultrafiltration in the kidney. The plasma proteins synthesised by the endoplasmic reticulum and assembled in the Golgi apparatus of hepatocytes include globulins and albumen and the blood-clotting protein, fibrinogen.

Hepatocytes also synthesise cholesterol, which is essential in membrane structure and is the precursor for several other molecules including the steroid hormones testosterone, estrogen and progesterone. Cholesterol is found in many of the foods of animal origin that we eat but all the cholesterol required by the body is made by the liver.

Cholesterol

The liver can regulate the amount of cholesterol in the blood. If there is sufficient, the liver may stop its synthesis of cholesterol, in a **negative feedback** process. Excess cholesterol is excreted in the bile, but high levels in bile can cause deposits in the gall bladder called gallstones, which obstruct the bile duct. Excess cholesterol in the blood can contribute to blockages in the walls of certain arteries (Figure **D.3**) and lead to cardiovascular disease.

Exam tip

Check you understand the role of cholesterol in the body as well as the link it has to CHD.

Bile

Small channels called bile canaliculi pass between the rows of hepatocytes. Bile is produced in the hepatocytes and secreted into these channels, which connect to form the bile duct. Bile contains bile salts, which are important in the digestion of lipids in the intestine, and bile pigments, which are derived from the breakdown of hemoglobin. Bile is stored in the gall bladder and released when food enters the small intestine.

Vitamin D

Vitamin D is made in the skin when exposed to sunlight (Subtopic **D1**). People who live in areas close to the Poles, where there is little sunlight during winter and the temperature is very cold, cover most of their skin during this time. In this situation, the liver is able to release vitamin D that has been stored during the warmer, summer months, when more skin is exposed to sunlight. However, sunlight contains ultraviolet rays that can cause skin cancer, and so a balance is needed between exposure, to synthesise vitamin D, and protection from the damaging UV.

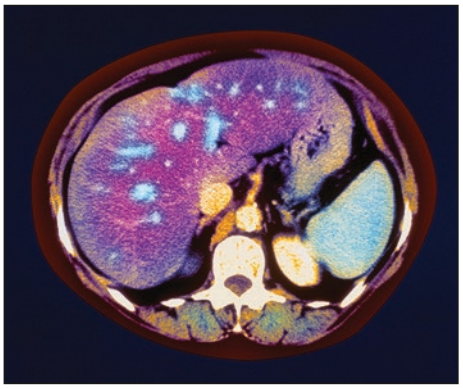


Figure D.10 This is a coloured MRI scan of a person suffering from a fatty liver. The fat deposits can be seen as blue patches in the liver (purple).

Storage and regulation of nutrients

The liver plays a key role in homeostasis. It regulates blood sugar levels and is able to store lipids, iron (from the breakdown of hemoglobin) and the fat-soluble vitamins A and D. Absorbed nutrients are carried in the blood directly from the intestine to the liver via the hepatic portal vein. The blood that leaves the liver contains regulated amounts of nutrients.

One of the most important roles of the liver is to maintain the correct level of glucose in the blood. After a meal containing carbohydrate, the glucose level in the blood will rise, but during exercise it will fall as glucose is respired by the muscles. The liver helps to balance out these fluctuations by storing glucose as **glycogen** in the hepatocytes when levels are high and breaking down and reconverting the glycogen to glucose when the level falls. Two pancreatic hormones insulin and glucagon control this process.

- **Insulin** is released when glucose levels are high and stimulates the hepatocytes to take up glucose and convert it to glycogen.
- **Glucagon** is released when glucose levels are low and stimulates hepatocytes to convert glycogen back to glucose.

Liver damage

Alcohol and the liver

Alcohol, absorbed from the gut, passes straight to the liver in the hepatic portal vein and is absorbed by the hepatocytes. Hepatocytes remove and detoxify the alcohol, but if it is present in large amounts, blood may have to flow through the liver many times before all the alcohol can be absorbed. This makes liver cells very susceptible to damage by alcohol. If large quantities are consumed, fatty deposits begin to build up in the liver lobules, replacing damaged cells, and reducing liver function (Figure **D.10**). The liver can become inflamed, a condition known as alcoholic hepatitis, whose symptoms include nausea and **jaundice**. In the longer term, the liver may become permanently damaged as scar tissue develops in place of damaged blood vessels and hepatocytes. This is called **cirrhosis** of the liver. These areas of the liver are no longer able to function efficiently and if cirrhosis is extensive, liver failure can be the result. Damage like this is fatal unless the liver tissue can be replaced by a transplant.

Causes and consequences of jaundice

Jaundice is caused by a build-up of the pigment bilirubin in the blood and tissues of the body. People with jaundice have yellow skin and whites of the eyes as a result of the liver's inability to metabolise and excrete the bilirubin. They may also have yellowing of mucous membranes in the nose and mouth, dark urine and pale feces.

Bilirubin is a waste product of the breakdown of red blood cells by the liver (Figure **D.9**) and in healthy individuals it is removed in bile. Bile is stored in the gall bladder and is released into the digestive system through the bile duct. Jaundice can occur as a result of any condition that disrupts the movement of bilirubin from the blood to the liver and then out of the body. There are three slightly different types of jaundice that are caused in different ways.

- **Pre-hepatic jaundice** occurs when a condition or infection speeds up the breakdown of red blood cells and increases bilirubin levels in the blood. This type of jaundice can be caused by malaria, sickle cell anemia and thalassemia, as well as some rare genetic conditions.
- **Intra-hepatic jaundice** occurs when the liver is damaged and its cells cannot process bilirubin. This may be due to an infection such as hepatitis, leptospirosis or glandular fever. Cells may also be damaged by exposure to harmful substances, such as alcohol and drugs, including ecstasy and overdoses of paracetamol (also called acetaminophen). Certain chemicals including phenol can also harm hepatocytes.
- **Post-hepatic jaundice** occurs when the bile duct is damaged, inflamed or blocked, so that the gall bladder is unable to send bile to the digestive system. Gallstones can cause this type of jaundice.

Exam tip
The liver is the body's largest organ. Make a checklist of all the roles it has and problems that occur if it isn't working properly.

Attitudes to knowledge

Alcohol is widely consumed in many parts of the world and in some cultures there are major health and social concerns about the rise of 'binge drinking' in young people. Binge drinking is the consumption of large amounts of alcohol in a short period of time.

Attitudes to both alcohol consumption and the use of certain drugs vary considerably in different cultures, despite widespread knowledge of the well-established correlation between excessive alcohol consumption and liver disease, as well as the longer-term problems of alcohol addiction.

Question to consider

- How is the interpretation and application of knowledge affected by culture?

Nature of science

Public understanding of science – the cholesterol story

Our understanding of health issues changes as new discoveries are made. The link between cholesterol and heart disease is one very good example. As new knowledge about different types of cholesterol – LDL and HDL – has emerged, scientists have been forced to reconsider previously held views and to try and communicate the new information to others.

The fact that heart disease is only partially related to cholesterol levels in the blood has been known since the 1950s. In addition, there have long been clues that dietary cholesterol is not directly linked to developing heart disease. Nevertheless, heart disease and its causes are often misunderstood and many people are confused about the effect of cholesterol and try to avoid all foods that contain it.

Chemically, cholesterol is a fat, but it does not provide the body with energy. Instead, it's an essential building block for molecules, cells and tissues. It forms part of all cellular membranes and is particularly important in nerve cells. It is a component of hormones and skin cells can also convert it to vitamin D in the presence of sunlight, making it a most useful and vital substance.

In metabolic terms, cholesterol in the diet is not directly correlated to cholesterol in the blood but fat in the diet is. Cholesterol that is present in foods such as eggs or shellfish is broken down during digestion and the body manufactures its own cholesterol from fats that we eat.

Our knowledge of cholesterol has changed. New scientific advances have shown that there are different types of cholesterol, some of which have beneficial effects. The public perception of the issue is confused, with some people still believing the original view that 'all cholesterol is bad', while others now think that all the cholesterol in their diet is doing them good. Scientists must be aware that good communication is vital and time is needed for new ideas to be disseminated. Communication of new research must be clear so that people who would benefit from scientific advances do not take the view that 'science keeps changing its mind' and decide not believe any of it.

Test yourself

7 Outline **three** important roles of the liver.

8 Outline the stages in the breakdown of red blood cells and the fate of the products.

9 State the names of the main blood vessels that enter and leave the liver.

Learning objectives

You should understand that:

- The structure of cells in cardiac muscle enables electrical impulses to be propagated through the heart wall.
- Impulses from the sinoatrial node cannot pass directly from atria to ventricles.
- Impulses are delayed at the atrioventricular node.
- The delay allows time for atrial systole to occur before the atrioventricular valves close.
- Conducting fibres enable the contraction of the entire ventricular wall to be coordinated.
- Normal heart sounds are caused by the closing of the atrioventricular and semilunar valves.

D4 The heart

The structure of cardiac muscle

Heart muscle works throughout our lives and never rests. It requires constant supplies of blood, which carries oxygen and nutrients to it. Three large coronary arteries branch from the aorta and supply heart muscle with oxygen-rich blood (Subtopic **6.2**, Figure **6.8**).

Heart muscle has a unique composition that adapts it for the conduction of waves of excitation from fibre to fibre. It is made up of short, striped muscles fibres, which branch and are also joined together at their ends by linking structures known as intercalated discs. (Figure **D.11**). This arrangement of linkages between cells allows action potentials to spread rapidly and enables the heart muscle fibres to act together and produce a more powerful effort as they contract simultaneously. Blood vessels, found in the spaces between the fibres, are branches of the two coronary arteries, which come from the aorta. In this way the heart is provided with a good blood supply carrying oxygen and glucose for its activity.

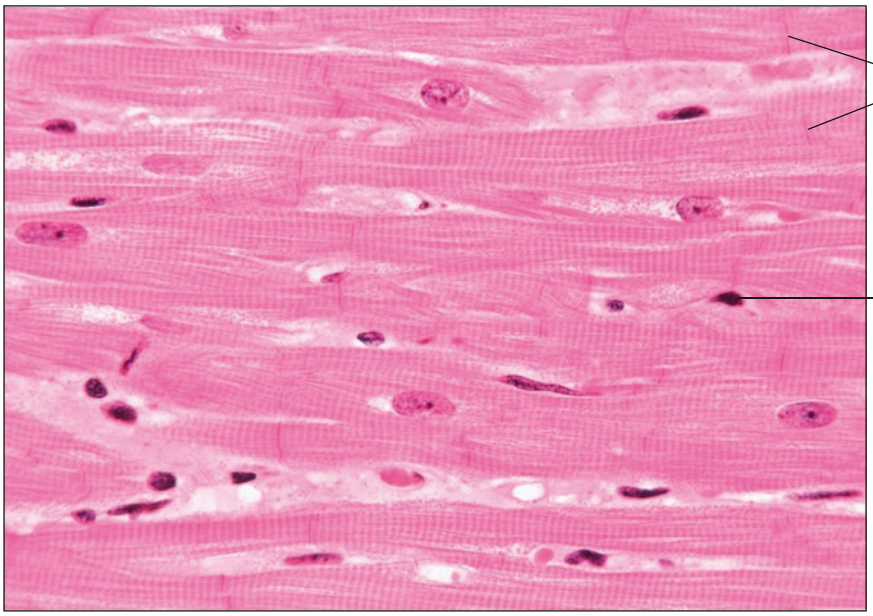


Figure D.11 Stained light micrograph of structure of cardiac muscle showing striations formed by actin and myosin filaments, branching cells and intercalated discs. (×300).

Control of the heart beat

An individual's heart rate changes with the level of their activity, emotions or stress. Heart muscle is unique in that it can contract without stimulation – it is said to be **myogenic**. However, under normal circumstances, heart rate is controlled by nervous or hormonal stimulation. Impulses pass to the pacemaker, the **sinoatrial node** (SAN), in the left atrium via nerves from the medulla oblongata in the brain. The SAN is also stimulated by the hormone epinephrine (adrenalin).

The SAN initiates contraction of the heart. Cells in the SAN produce action potentials that spread through the muscle cells in the walls of the atria and cause atrial **systole** (contraction). The impulses are prevented from passing directly to the ventricles but they do stimulate a group of cells known as the **atrioventricular node** (AVN). This node is situated in the lower part of the atrium, close to the ventricles. The AVN sends out impulses down two bands of conducting fibres that run down the centre of the heart, between the two ventricles, to the base of the heart (Figure **D.12**). From here, fibres branch out between the cells of the thick ventricular walls. As impulses arrive, coordinated contraction occurs across the muscle tissue of the ventricle walls. This sequence of events is known as the **cardiac cycle**.

Heart rate is speeded up during exercise as a rise in carbon dioxide and a fall in the pH of the blood cause impulses to be sent from the medulla in the brain via the sympathetic nerve to the SAN. When exercise stops and blood pH returns to normal, impulses pass via the vagus nerve, which slows the heart rate down. Increasing levels of epinephrine (adrenalin), the 'fight or flight' hormone, are produced at times of stress or anxiety – epinephrine stimulates the SAN to increase the heart rate.

The sympathetic nerves are part of the nervous system that deals with functions such as heart rate, which are automatically controlled by the brain.

Artificial pacemakers

Artificial heart pacemakers are fitted to patients who have a problem with electrical conduction through the heart. This can mean that the heart beats too slowly or too fast. They are also used for patients who have suffered heart failure so that their heart beat is uncoordinated. A typical pacemaker is approximately 4 cm long and weighs about 30 g. Pacemakers are fitted below the collar bone (Figure **D.13**) and have leads that connect to the heart via a vein. Pacemakers are battery powered and generate pulses that stimulate the heart at an appropriate, adequate rate.

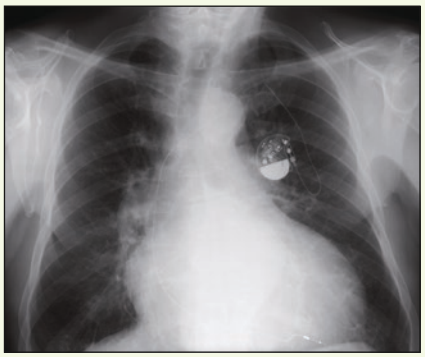


Figure D.13 An X-ray photograph showing a pacemaker in position in a patient.

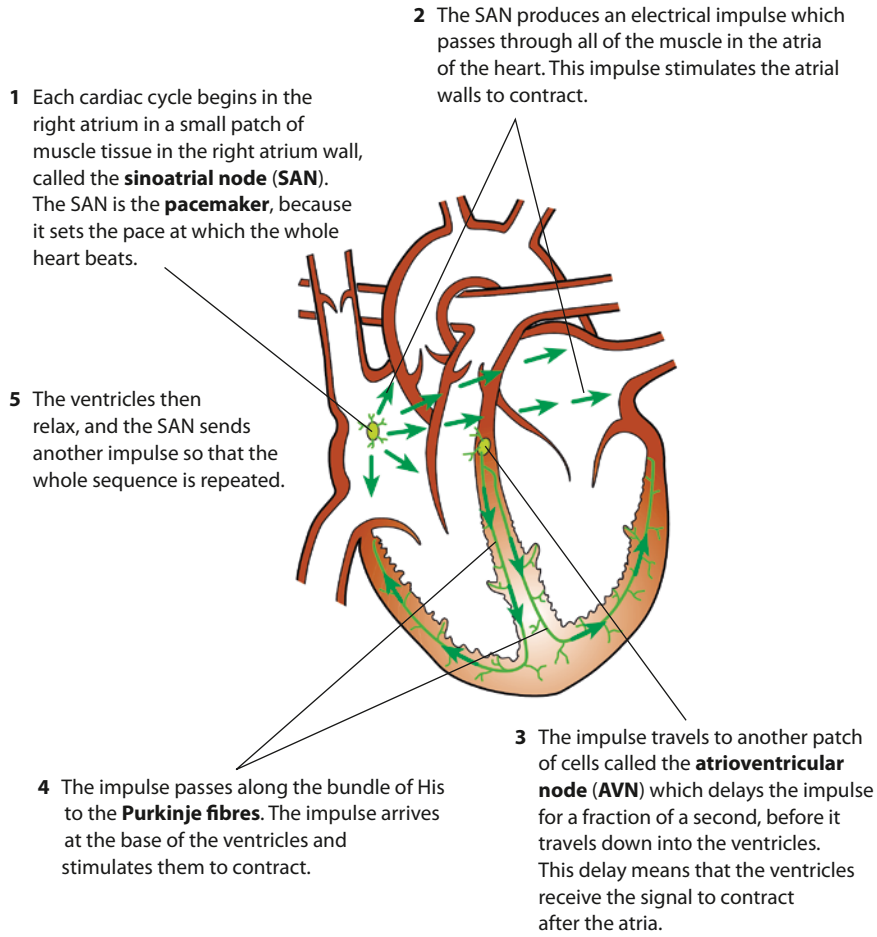


Figure D.12 How electrical impulses move through the heart.

Defibrillators

If a person's heart stops beating properly they may to go into **cardiac arrest**. In such cases, a **defibrillator** can be used to deliver an electric shock to the heart, which will be stimulated into re-establishing its proper rhythm. Defibrillators are not usually used to restart a heart that has stopped completely but are vital in cases when uncoordinated contraction of the ventricular muscle (ventricular fibrillation) makes the ventricles 'quiver' rather than contract fully. Defibrillation can also treat **tachycardia**, a fast heart rhythm that originates in one of the ventricles. A defibrillator consists of a pair of electrodes, which are placed on the patient's chest, and a battery that delivers an electrical impulse between them.

The cardiac cycle

The cardiac cycle describes the events that go to make up one heart beat. The heart rate, normally about 70 beats per minute, is a measure of the frequency of the cardiac cycle. The key structures that are important in the cardiac cycle are the muscles of the walls of the atria and ventricles, the atrioventricular and semilunar valves (Figure **6.6** in Subtopic **6.2**), the sinoatrial node (SAN) or pacemaker, and the atrioventricular node (AVN).

Contraction of the heart is called **systole** and relaxation is known as **diastole** (Figure **D.14**). Atria and ventricles always contract separately, with contraction of the two atria being followed – after a short pause due to the delay of the impulse at the AVN – by contraction of the two ventricles. The four valves, two in the heart and two in the main arteries, keep blood flowing in one direction. As these valves close, they produce the characteristic 'lub-dub' or heart sounds that can be heard through a stethoscope. The two sides of the heart work together so that the 'lub' sound is made as the two atrioventricular valves flap shut after atrial systole, and the 'dub' sound is the closing of the two semilunar valves once ventricular systole is complete.

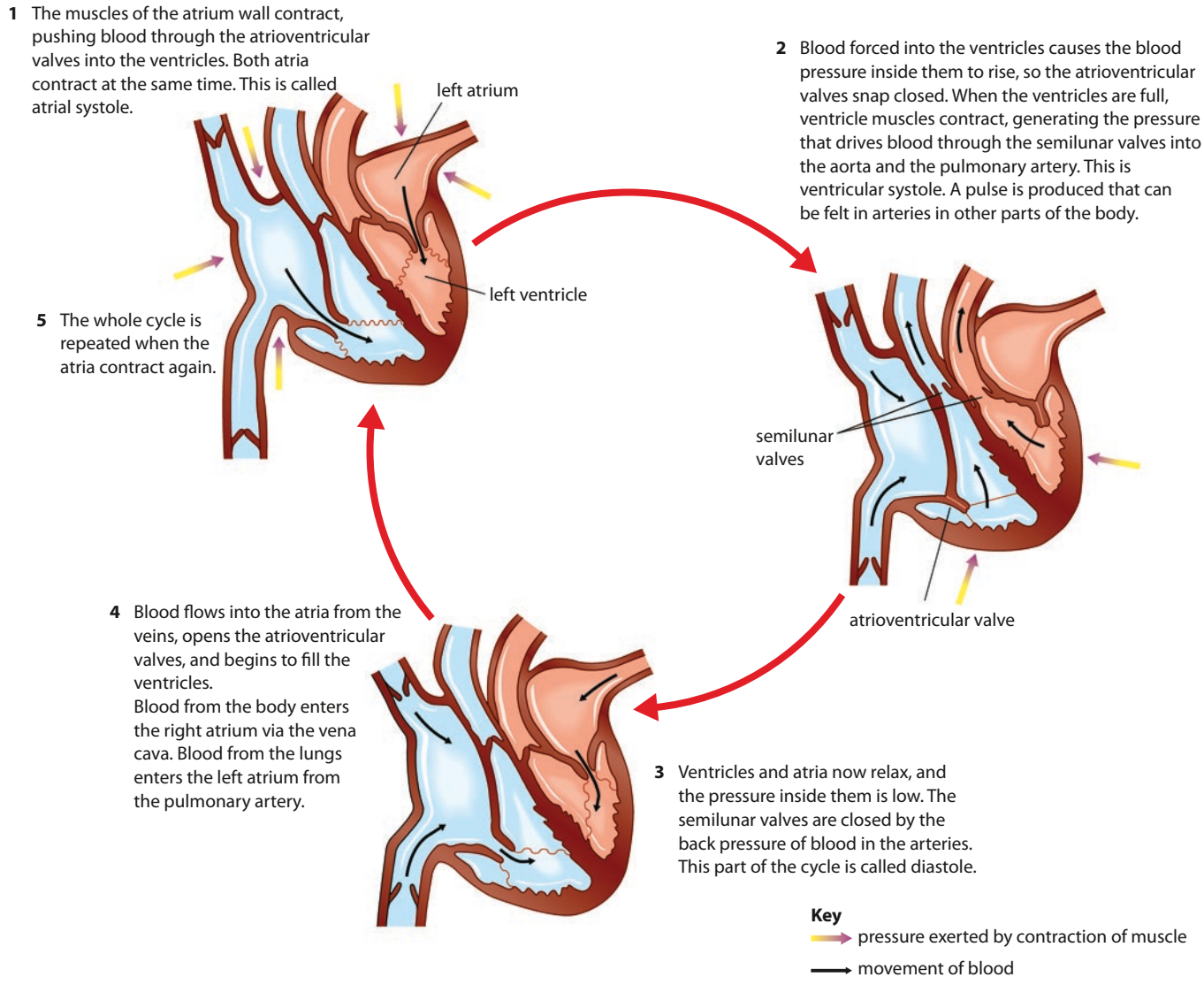


Figure D.14 The cardiac cycle.

As the heart beats, the pressure and volume in each of its four chambers change and these changes can been shown on graphs such as the one shown in Figure **D.15**.

At the end of the cardiac cycle, both atria and ventricles are in diastole (relaxed). Blood has been pumped out of the ventricles and blood is re-entering the atria from the pulmonary veins and vena cava. The pressure in the atria is slightly greater than in the ventricles so blood flows through the atria via the atrioventricular valves and into the ventricles. Blood pressure in the arteries is higher than that in the ventricles so the semilunar valves remain closed.

When the ventricles are approximately 70% full a new cardiac cycle begins. It starts with contraction of the walls of the atria, **atrial systole**. Blood is pumped through the atrioventricular valves, filling the ventricles to capacity. The thin walls of the atria do not generate much pressure, but after the contraction is complete, most of the blood from the atria has entered the ventricles.

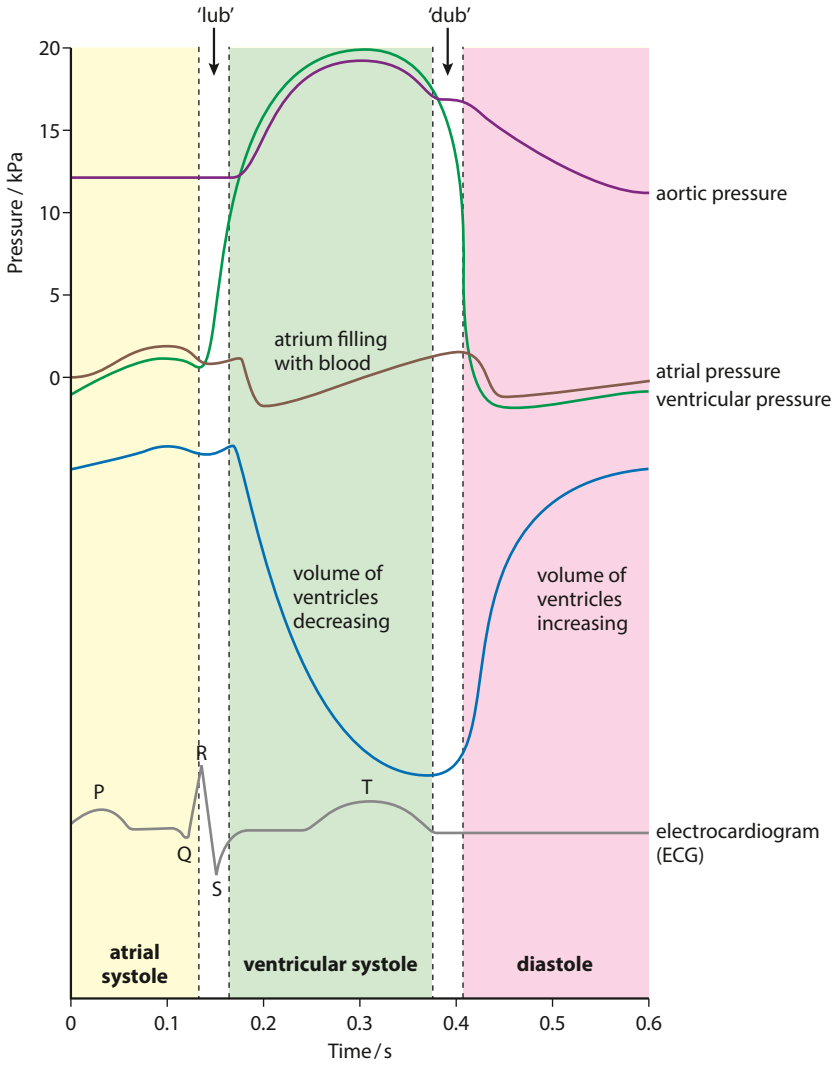


Figure D.15 Pressure and volume changes in the heart during the cardiac cycle, with an electrocardiogram (ECG) trace. An ECG trace records the rhythm and electrical activity of the heart via electrodes attached to the chest.

As **ventricular systole** begins, it produces sufficient pressure to snap the atrioventricular valves closed and produce the first heart sound. As the ventricles contract, pressure inside the chambers rises so that it becomes greater than the pressure in the arteries (the pulmonary artery and aorta) that leave the heart. This pressure forces the semilunar valves open and blood is pumped out of the ventricles into the arteries. At the end of ventricular systole, blood pressure in the ventricles is lower than that in the arteries and the back pressure forces the semilunar valves shut, causing the second heart sound.

When pressure in the ventricles falls below that in the atria, the atrioventricular valves re-open. Blood from the veins flows passively through the atria and into the ventricles. All four chambers of the heart return to diastole and the cycle begins all over again.

Exam tip
Check that you can interpret what is happening during the cardiac cycle from graphs of pressure in the chambers and an ECG trace.

Coronary heart disease

Coronary heart disease (CHD) is often caused by damage to the arteries – for example, by **atherosclerosis**. Atherosclerosis is a slow degeneration of the arteries caused by a build-up of material known as **plaque** inside them. Plaque becomes attached to the smooth endothelium lining an artery and can accumulate over many years. Few people suffer from any symptoms before middle age. Fibrous tissue in the lining may become damaged and thickened so that lipids, cholesterol (released from low-density lipoproteins) and cell debris accumulate. Calcium may also be present, causing the artery to loses elasticity and become hard and inflexible. Over time, the diameter of the artery becomes restricted so that blood cannot flow along it properly, leading to **hypertension** (Figure **D.3**). As the rate of flow slows down, blood may clot in the artery, further restricting the movement of blood along it. Clots may also break free and travel to block another smaller artery elsewhere in the body – this is called a **thrombosis**. If any of the three coronary arteries is blocked (Figure **6.8**), an area of the heart muscle will receive less oxygen and cells in that region may stop contracting or even die. A blockage in a coronary artery or one of its branches is known as a **coronary thrombosis** or heart attack. If an artery is blocked in the brain, the clot may cause a **stroke**.

Risk factors associated with CHD

The incidence of heart disease varies from country to country (Table **D.4**) and between individuals. Some factors associated with the likelihood of developing CHD are related to a person's environment. By making personal choices about lifestyle, it is possible to lower the risk of developing CHD. Other factors that increase the risk cannot be controlled. For example, the risk of CHD is affected by:

- **genetic factors** – CHD tends to occur more frequently in some families than others with similar lifestyles
- **the person's sex** – men are more likely to have CHD than women
- **the person's age** – CHD is more prevalent in older people.

Country	Death rate / 100 000
Kiribati	11
France	29
Switzerland	52
Australia	60
Canada	66
UK	68
Germany	75
China	79
USA	80
Brazil	81
Colombia	85
Ghana	120
Poland	122
Rep. of Congo	125
Malaysia	138
Turkey	157
Iran	194
Pakistan	222
Lithuania	233
Afghanistan	328
Ukraine	399

Table D.4 Death rates from CHD per 100 000 people in different countries (from UNESCO and WHO).

Lifestyle factors that increase the risk of CHD include:

- **smoking** – smokers are significantly more likely to suffer from CHD than non-smokers
- **lack of exercise** – a lifestyle involving little physical activity may contribute to obesity and high blood pressure
- **hypertension (high blood pressure)** – causes strain on the heart, which has to work harder to pump blood
- **obesity** – increases the work of the heart; people who are overweight are also more likely to have high blood pressure and high cholesterol levels in their blood
- **diet** – there have been many claims that diet can increase the risk of CHD; for example, there is a positive correlation between intake of saturated fat and CHD, but cause and effect have not been proven. Reliable evidence suggests that in countries where many high-fat foods, animal products and processed foods are eaten, there is likely to be a high incidence of CHD. Since all fatty acids are high in energy, an excess of these foods in the diet can also lead to obesity, which places a further strain on the heart.

There is some correlation between CHD and blood cholesterol levels. Reducing the amount of cholesterol in the diet can reduce blood cholesterol levels to a certain extent. But while LDL cholesterol (low-density lipoprotein) is associated with an increased risk of CHD, HDL cholesterol (high-density lipoprotein) is correlated with a reduced risk. So it is difficult to predict the effect of reducing dietary cholesterol on risk with any certainty.

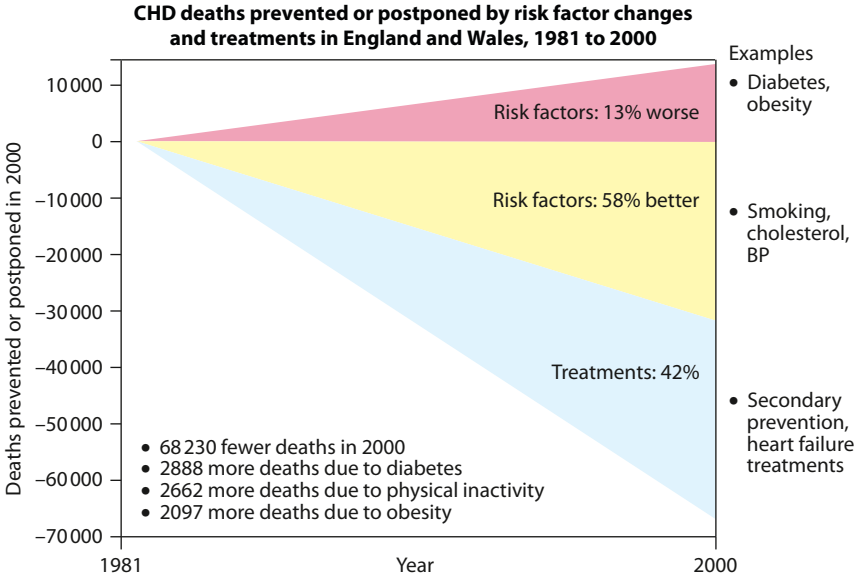
Many aspects of lifestyle are interrelated and it is very difficult to isolate a single factor that can be said to cause CHD. Research focusing on just one aspect of risk may underestimate the contribution of other important factors. If a person changes one aspect of their lifestyle, other risk factors may become important.

Analysis of data

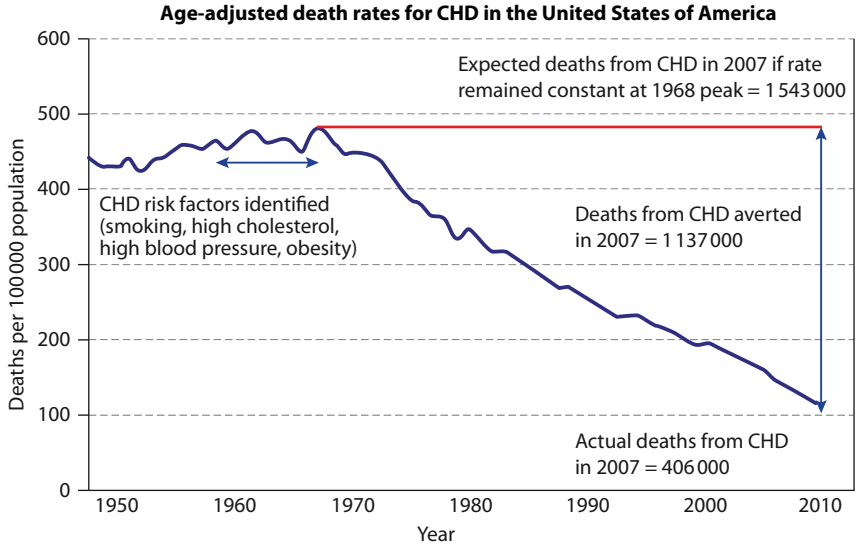
Medical researchers collect and collate large amounts of data related to heart disease and risk factors. For example, the two graphs shown in Figure **D.16** present data on the prevention and reduction of heart disease in the USA and UK over a period of 30 and 20 years, respectively.

The upper graph shows the reductions in CHD deaths as a result of campaigns to reduce smoking and the effect of health programmes which screen and treat high risk patients. The lower graph shows the results of similar programmes in the USA.

Consider how important data like these are. Why do you think long term studies are vital to medical research? Why is it important to include data from large numbers of people? How can this information be used to help in planning public health strategies and education?



Source: B. Unal *et al* in *Circulation*, (2004).



Source: *Vital Statistics of the United States,* CDC/National Centre for Health Statistics.

Figure D.16 Graphs showing age-adjusted death rates from CHD in the UK and USA in recent decades.

Nature of science

Scientific advance follows technical innovation – the stethoscope

Our understanding of the way the heart works took a big step forward following the invention of the stethoscope by René Laennec (1781–1826) in France in 1816. His device was a simple wooden tube, similar to an ear trumpet, which had only one earpiece that he used to listen to a patient's breathing. It was not until the mid 1800s that a flexible stethoscope with two earpieces first appeared, and the first instrument for commercial sale was designed by the American physician George Cammann in 1852. Since that time, the stethoscope has become one of a doctor's most important tools.

The invention of the stethoscope gave doctors a chance to listen to a person's heart and lungs, and for the first time they were able to detect faults in internal anatomy that could cause symptoms of disease. Today, stethoscopes are not only used to listen to the chest but also other parts of the body such as the bowels, and the heart of a fetus during pregnancy. Learning to listen for and diagnose the sounds is an important part of any doctor's training.

Symbols of knowledge

Symbols are often used as a form of non-verbal communication. The heart is used as a symbol in many different cultures. In the Egyptian Book of the Dead a heart was shown being weighed as a measure of worthiness to enter into paradise. In the Bible, the heart has been used as a devotional image for many centuries. It has also appeared on playing cards since the 1400s. In Japan, the oldest heart symbol known is one on a helmet from the 1550s where it was used to represent the goddess of archers. In Chinese, the symbol for a heart is the same one used to mean feelings and the mind. In the contemporary world, a heart is best known as a symbol of love.

Questions to consider

- Why is the heart a symbol of love?
- Do symbols make communication easier or do they exclude those who do not understand them?
- How are symbols used in other areas of knowledge?

Test yourself

10 Outline the risk factors for coronary thrombosis.

11 Describe what happens to the chambers of the heart during atrial systole, ventricular systole and diastole.

12 Define the term 'myogenic'.

13 Using the information in Figure **D.16** and your knowledge of heart disease, state the most important factors contributing to the fall in deaths from CHD in the USA and UK in recent decades.

D5 Hormones and metabolism (HL)

The body is under the control of two systems: the nervous system and the endocrine system. These are mostly independent of one another but there are situations in which the two work together to control activities such as heart rate (Subtopics **6.2** and **D.4**).

The chemical structures of hormones

Hormones are chemical substances that are secreted directly into the bloodstream from endocrine glands found throughout the body (Figure **D.17**).

Since hormones circulate in the bloodstream, they come into contact with all cells in the body but only cells that have specific, genetically determined receptors will respond. These **target cells** have receptors on the plasma membrane that recognise and bind to the hormone.

Different hormones have different chemical structures and can be divided into three categories as shown in Table **D.5**.

Chemical category of hormone	Examples
steroids derived from cholesterol	testosterone, progesterone
proteins	insulin, FSH, LH
tyrosine derivative	thyroxin – each thyroxin molecule has four iodine atoms

Table D.5 The different chemical forms of hormones.

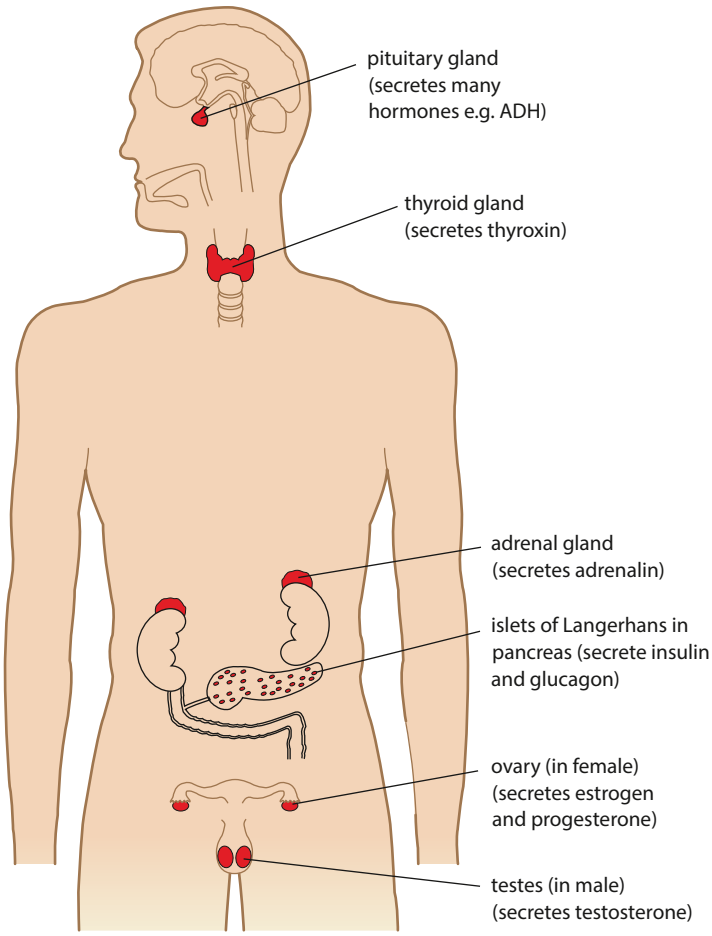


Figure D.17 Endocrine glands of the body.

Learning objectives

You should understand that:

- Endocrine glands secrete the hormones that they produce directly into the bloodstream.
- Steroid hormones bind to receptor proteins in the cytoplasm of target cells and form a receptor–hormone complex.
- The receptor–hormone complex promotes the transcription of particular genes.
- Peptide hormones bind to receptors in the plasma membranes of target cells.
- A cascade is activated as hormones bind to membrane receptors, mediated by a second messenger inside the cell.
- Hormone secretion by the anterior and posterior lobes of the pituitary gland is controlled by the hypothalamus.
- Pituitary hormones control growth, developmental changes, reproduction and homeostasis.

How hormones control cells

Protein hormones and steroid hormones control their target cells in different ways (Figure **D.18**).

Protein hormones bind to a surface receptor, very often a glycoprotein, but do not enter the cell. Instead, the binding process triggers the release of a second messenger chemical which cascades from the cytoplasmic side of the plasma membrane and this messenger controls the activity of the cell. This may be achieved by regulating the activity of a specific enzyme in the cell, either activating it or inhibiting it.

Steroid hormones do enter target cells as they can easily pass through the plasma membrane. They bind to a specific receptor in the cytoplasm forming a hormone–receptor complex, which is transported through a nuclear pore into the nucleus. Here, the hormone regulates the process of transcription of one or more specific genes.

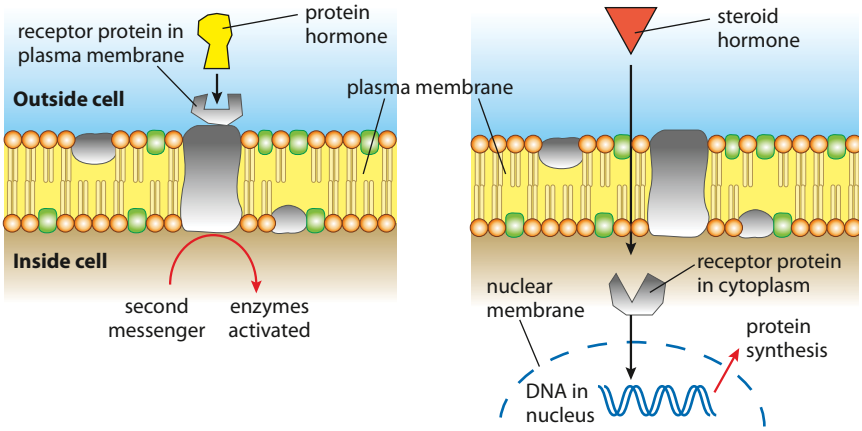


Figure D.18 The modes of action of protein and steroid hormones.

The hypothalamus and the pituitary gland

The **hypothalamus** is a small area of the brain that monitors hormone levels and indirectly controls functions including body temperature, hunger and sleep. It links the hormonal and nervous systems and secretes releasing hormones that regulate the hormones of the anterior pituitary gland. It has a range of receptors of its own, which allow it to act independently, but it also receives information from other parts of the brain.

The **pituitary gland**, situated just below the hypothalamus (Figure **D.19**), is made up of two different parts – the anterior and posterior lobes. The posterior lobe develops from the brain and has neurons connecting it directly to the brain. The anterior lobe develops separately and has no direct neural connection with the brain.

The hypothalamus has to communicate with each lobe of the pituitary gland in a different way.

- The hypothalamus contains the cell bodies of many **neurosecretory cells**, which have their terminal ends in the posterior lobe of the pituitary. A neurosecretory cell is simply a neuron that has been modified to secrete and store a large quantity of hormone at the terminal end of the cell body. Surrounding the terminal ends of the neurosecretory cells is a capillary network so that when the cells receive

Portal veins always connect two capillary networks. The hepatic portal vein has branches at both ends, in the liver and in the digestive system. The portal vein that connects the hypothalamus and the pituitary gland also has branches in both structures.

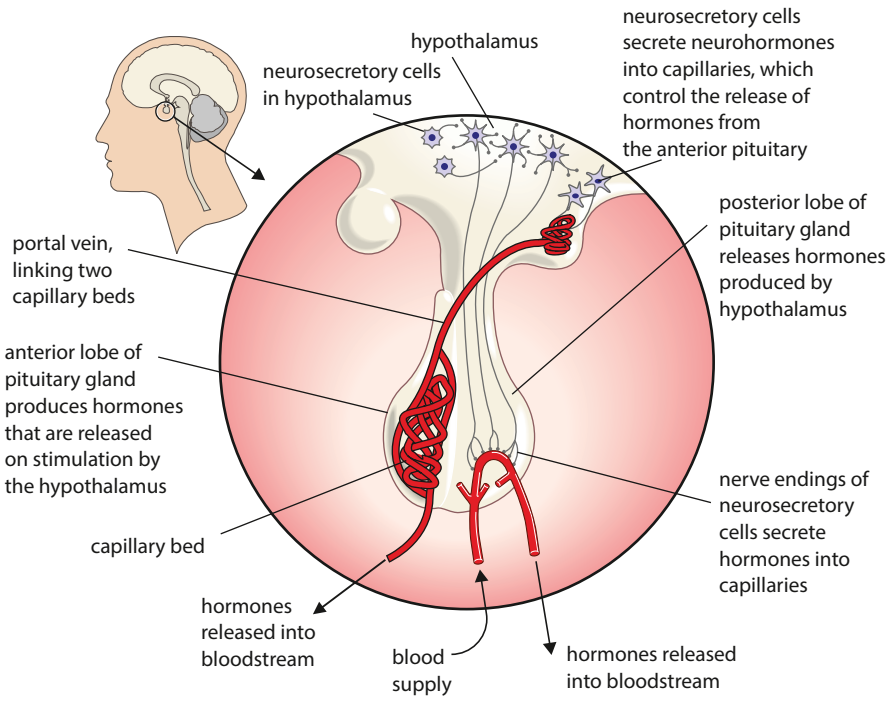


Figure D.19 The hypothalamus and pituitary gland.

the appropriate information they can release the hormone directly into the blood. Two examples of posterior lobe hormones released in this way are antidiuretic hormone (ADH) and oxytocin.

- Control of the anterior lobe of the pituitary is regulated by another set of neurosecretory cells in the hypothalamus. These cells end in a different capillary bed just above the pituitary gland. The blood from these capillaries flows into a portal vein, which passes into capillaries within the anterior lobe of the pituitary. These neurosecretory cells secrete releasing hormones (RH), which control the release of the hormones from the cells of the anterior lobe. One example is gonadotrophin releasing hormone (GnRH), which controls the release of follicle stimulating hormone (FSH) and luteinising hormone (LH).

The secretions from neurosecretory cells are often called hormones. They are more correctly termed **neurohormones** because they are not produced by endocrine glands. These neurohormones are usually peptides that travel as droplets along the axons of the cell.

Pituitary hormones

Hormones from the pituitary gland control reproduction, growth and changes that take place as the body develops. They are also important in homeostasis. The hormones produced and secreted by the anterior and posterior pituitary glands are summarised in Table **D.6**.

Hormone	Secreted by	Effect on the body
human growth hormone (HGH or somatotrophin)	anterior pituitary	controls growth and metabolism
thyroid stimulating hormone (TSH)	anterior pituitary	stimulates the thyroid gland to produce thyroxin, which increases metabolic rate
adreno-corticotropic hormone (ACTH)	anterior pituitary	produced in response to stress – targets the adrenal gland to produce corticosteroid hormones
follicle stimulating hormone (FSH)	anterior pituitary	targets the reproductive system – stimulates follicular development in the ovary and meiotic division of primary spermatocytes in the testes (Subtopic **11.4**)
luteinising hormone (LH)	anterior pituitary	targets ovaries and testes to produce sex hormones (Subtopic **11.4**)
prolactin	anterior pituitary	stimulates milk production by the mammary glands and estrogen production by ovaries
oxytocin	posterior pituitary	stimulates contraction of the uterus wall during childbirth and milk release during suckling
antidiuretic hormone (ADH)	posterior pituitary	important in the homeostatic regulation of blood plasma – causes reabsorption of water from the collecting ducts of the kidneys (Subtopic **11.3**)

Table D.6 Hormones produced and secreted by the anterior and posterior pituitary glands.

Hormonal changes and control of milk secretion

During the second and third trimesters of pregnancy (3–9 months), the placenta produces progesterone and estrogen, which suppress the menstrual cycle and promote the growth of breast tissue for **lactation** (milk production).

As the end of pregnancy approaches, the level of progesterone produced by the placenta falls (Figure **D.20**) and this signals the onset of the uterine contractions known as labour. At this time, the hormone **oxytocin** is secreted by the posterior lobe of the pituitary gland. Oxytocin stimulates the uterus muscles to contract and is controlled by **positive feedback**. A small contraction of the uterus muscle stimulates the release of further oxytocin, which in turn stimulates more and stronger contractions until the baby is born. Oxytocin is also important in the control of milk release from the mammary glands.

After birth, blood levels of the hormone **prolactin**, from the anterior pituitary gland, increase. This hormone stimulates milk production by the mammary glands. As a baby suckles, prolactin secretion is maintained and oxytocin is also released from the posterior pituitary gland. Oxytocin is necessary for the milk-ejection reflex, or 'let-down', to occur – it causes contraction of the smooth muscle cells that squeeze milk into the duct system of the breasts. The two hormones continue to be released in proportion to the amount of milk the baby consumes as it suckles so that the supply is matched to the demand. When the baby is weaned and no longer takes milk, the levels of these hormones decrease.

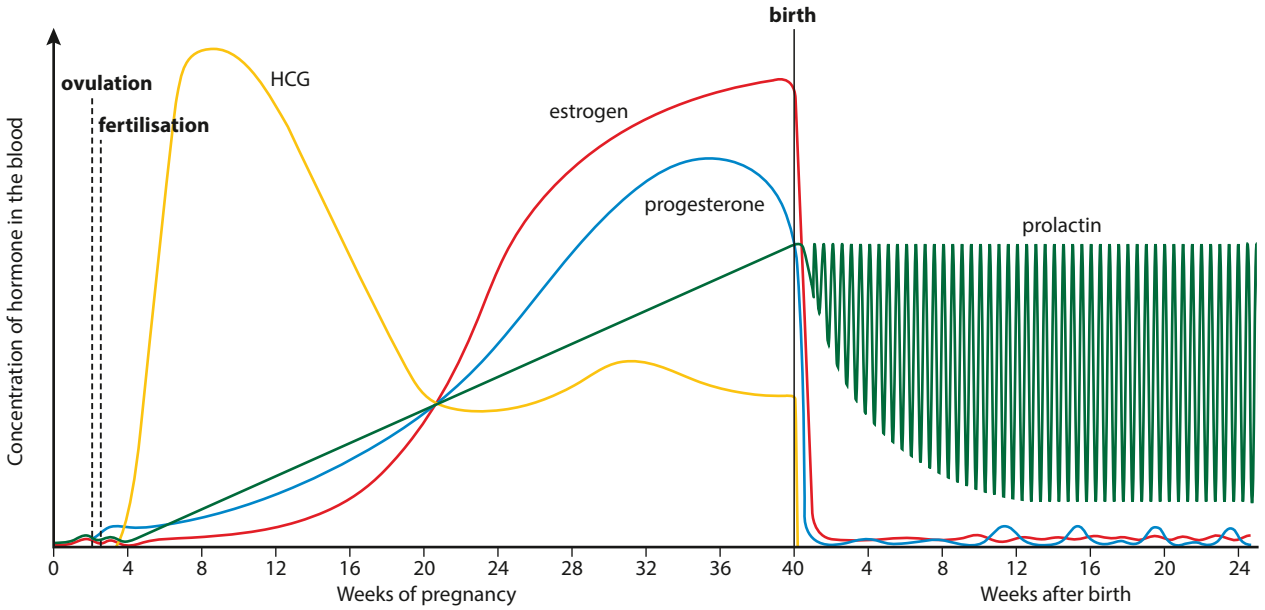


Figure D.20 Changes in the levels of hormones during pregnancy and birth. Oxytocin is not shown, but peaks during birth and then drops, continuing to rise and fall less dramatically each time the baby feeds. Prolactin levels also fluctuate regularly with the baby's feeding pattern.

Human growth hormone use by athletes

Human growth hormone (HGH) has been used by some athletes to promote muscle growth – particularly in power sports such as bodybuilding, swimming and weight lifting – in an attempt to improve performance. The hormone was declared a banned substance by the Olympic Committee in 1989 but there is evidence that it is still used. In the USA, HGH is available with a doctor's prescription.

Scientific evidence about the effect of HGH on muscles is mixed. The hormone appears to reduce body fat and increase lean body mass, but does not seem to increase the strength of muscles. It has been shown to build up connective tissue in muscles, making them appear larger and also increasing the ability of the muscle to resist injury and repair itself quickly. But none of these effects make the muscle physically stronger. Researchers in a 2010 study funded by the World Anti-Doping Agency in Sydney, Australia, reported that HGH could provide an energy boost in short-duration events such as running or swimming – but the study failed to find any evidence of HGH use producing an increase in power, strength or endurance.

Despite the results of such studies, and HGH's side-effects – which can include muscle and joint pain and swelling – some athletes still use the hormone. Part of its attraction may lie in the fact that, because it is a naturally occurring substance, it is hard to detect in the body.

Nature of science

Cooperation and collaboration – the international effort to eliminate iodine deficiency

People living far from the coast or those who have little seafood in their diet may suffer from a shortage of iodine. Iodine is needed to synthesise the hormone **thyroxine**, which regulates growth and controls metabolism. Without iodine, the thyroid gland can become enlarged, producing a swelling known as **goitre**.

Iodine deficiency disorder (IDD) is a serious problem for babies whose mothers were iodine deficient during pregnancy. Unborn babies with IDD can suffer brain damage and have poor mental development after birth. The most severe cases lead to **cretinism**, which UNICEF estimated affected more than 11 million people worldwide in 2000.

Goitre and cretinism are rare in Europe and the Americas because in these regions sodium iodide (NaI) has been added to table salt since the early part of the twentieth century. In the Mid-Western states of the USA, iodine added to salt in the 1920s reduced the incidence of goitre in children from 40% to 10% in just 4 years.

UNICEF has been involved in efforts to eliminate iodine deficiency since the 1950s. The organisation has persuaded and assisted many governments to iodise salt in an effort to eliminate IDD. It has provided salt iodisation equipment and iodine supplements to many countries. At the time of the World Summit for Children in 1990, only about 20% of households in the world used iodised salt and the campaign was stepped up. By the end of 2000, this figure had risen to around 70%. This global progress has meant that by the 21st century more than 91 million newborn children were protected against significant losses in learning ability caused by IDD.

Test yourself

14 List the **three** chemical categories of hormones.

15 State which hormone type:

a enters the cell directly

b binds to a cell surface receptor.

16 Describe the effect of oxytocin in the control of milk secretion.

D6 Transport of respiratory gases (HL)

Oxygen dissociation curves

The oxygen content of air is measured as a **partial pressure**. In a mixture of gases, each component gas exerts a pressure (the partial pressure) in proportion to its percentage in the mixture. It is calculated as follows.

For normal dry air at sea level, atmospheric pressure is 101.3 kPa. The partial pressure of oxygen, which makes up 21% of the air, is:

$$\frac{21}{100} \times 101.3\,\text{kPa} = 21.3\,\text{kPa}$$

The partial pressures of other gases in dry air at sea level are shown in Table **D.7**.

Gas	Approximate percentage composition / %	Partial pressure / kPa
oxygen	21	21.3
carbon dioxide	0.0035	negligible
nitrogen	79	80.0

Table D.7 The partial pressures of gases in dry air at sea level. At high altitude, the pressure of air falls but the percentage of oxygen in the air remains approximately the same. At 5000 m, the partial pressure of oxygen is 11.5 kPa; at 10 000 m, it falls to just 5.5 kPa.

Oxygen is transported from the lungs to respiring tissues bound to the hemoglobin molecules that fill red blood cells. Each hemoglobin molecule can bind four oxygen molecules via the iron in the heme groups it contains (Figure **D.21**). When hemoglobin comes into contact with normal air, containing approximately 21% oxygen, it binds readily with oxygen molecules and becomes almost 100% saturated.

It follows that in an area of the body where there is a lot of oxygen (a high partial pressure), such as the lungs, most hemoglobin molecules will be carrying the maximum amount of oxygen and will be fully saturated. However, in areas where the oxygen level is lower, fewer hemoglobin molecules carry their maximum complement of oxygen and the hemoglobin may be only 50% saturated. As blood travels from the lungs to actively respiring tissues, the amount of oxygen bound to hemoglobin changes as the partial pressure of oxygen decreases. Hemoglobin readily releases oxygen where the partial pressure is lower, so it acts as an oxygen delivery service for respiring cells. Figure **D.22** shows the percentage saturation of hemoglobin at different oxygen concentrations and is known as an **oxygen dissociation curve**.

The steep S-shape of the dissociation curve shows how the affinity of hemoglobin changes at different partial pressures of oxygen. At a partial pressure of 10 kPa, which might be found in the lungs, hemoglobin is 95% saturated. At a partial pressure of 4 kPa, found in the tissues, hemoglobin does not bind with oxygen and will release it, so saturation falls to only about 50%. About half of the oxygen collected by hemoglobin in the lungs is released at this low partial pressure to supply the needs of actively respiring cells.

Learning objectives

You should understand that:

- The affinity of hemoglobin for oxygen at different partial pressures of oxygen can be shown in a graph called a dissociation curve.
- Carbon dioxide is carried in the blood both in solution and bound to hemoglobin.
- Carbon dioxide is converted to hydrogen carbonate ions in the red blood cells.
- The increased release of oxygen by hemoglobin in respiring tissues can be explained by the Bohr shift.
- Chemoreceptors are sensitive to pH changes in the blood.
- The respiratory centre in the medulla oblongata controls the ventilation rate.
- Ventilation rate increases in response to the amount of carbon dioxide in the blood during exercise.
- Fetal hemoglobin differs from adult hemoglobin, which means that oxygen can be transferred across the placenta to the fetus.

Partial pressure the proportion of the total pressure that is due to one component of a mixture of gases

The **pascal** (Pa) is the SI unit of pressure, named after the French physicist and philosopher Blaise Pascal. It is a measure of force per unit area, defined as 1 newton per square metre.
1000 Pa = 1 kilopascal (kPa).

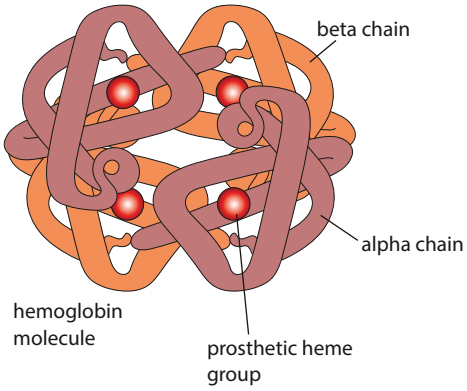


Figure D.21 Hemoglobin is a protein that has quaternary structure. It consists of four subunits bound together – two α chains and two β chains – each of which contains an iron-containing heme group.

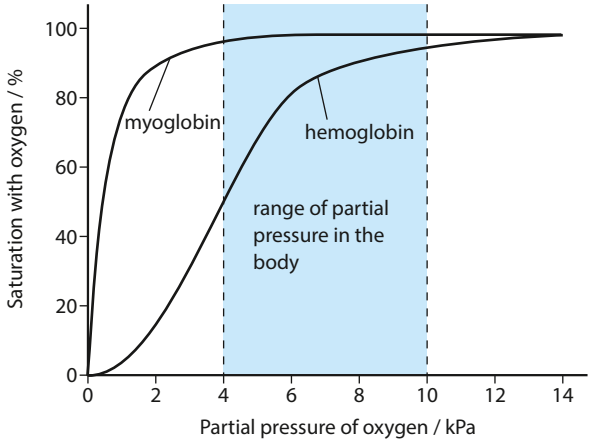


Figure D.22 Dissociation curves for hemoglobin and myoglobin. The curves are constructed using the normal range (at sea level) of partial pressure of oxygen in the body. The partial pressure of oxygen in alveolar air is about 14 kPa due to the presence of water vapour, which forms about 6% of alveolar air.

Sigmoid curve

The sigmoid shape (S-shape) of a dissociation curve is due to the allosteric effect that occurs when the first molecule of oxygen binds to hemoglobin. Once the first oxygen molecule has bound to one of the heme groups in hemoglobin, the shape of the molecule changes making it much easier for other oxygen molecules to bind. The steep slope of the curve shows that as the partial pressure of oxygen rises, hemoglobin readily becomes saturated.

Fetal hemoglobin

Whereas adult hemoglobin is composed of two α and two β chains, fetal hemoglobin is composed of two α and two γ (gamma) chains. Fetal hemoglobin must be able to collect oxygen from the mother's hemoglobin across the placenta. Fetal hemoglobin has a higher oxygen affinity than adult hemoglobin because the γ chain has fewer positive charges than the adult hemoglobin β chain. At birth, a baby's blood contains between 50% and 90% fetal hemoglobin. During the first 6 months of life, adult hemoglobin synthesis begins and production of fetal hemoglobin stops.

Myoglobin is another oxygen-binding protein found in muscle cells. Each myoglobin molecule has only one heme group and can bind to just one oxygen molecule. It is used to store oxygen, which is released as oxygen supply falls and the muscles begin to respire anaerobically. The dissociation curve for myoglobin (Figure **D.22**) is to the left of the curve for hemoglobin. At almost all partial pressures of oxygen, myoglobin remains saturated. It is still fully saturated with oxygen at partial pressures well below those that cause hemoglobin to release oxygen, and myoglobin can hold onto oxygen until the partial pressure falls extremely low. It provides a 'reserve supply' of oxygen during vigorous activity because oxygen is only released when the partial pressure falls to 1 or 2 kPa, and this reserve means that muscles can continue to respire aerobically for longer.

Fetal hemoglobin

The molecular structure of hemoglobin in the blood of a fetus is different from that of an adult. The dissociation curve for fetal hemoglobin lies to the left of the adult curve for all partial pressures of oxygen (Figure **D.23**). This tells us that fetal hemoglobin has a higher affinity for oxygen than maternal (adult) hemoglobin, whatever the concentration of oxygen. In the capillaries of the placenta, the partial pressure of oxygen is low. Here the mother's adult hemoglobin releases oxygen, which is easily picked up and bound to fetal hemoglobin. At a partial pressure of 4 kPa, the mother's hemoglobin is only 50% saturated, but fetal hemoglobin becomes approximately 70% saturated. The fetal hemoglobin carries the oxygen to the baby's body and releases it into the respiring fetal tissues.

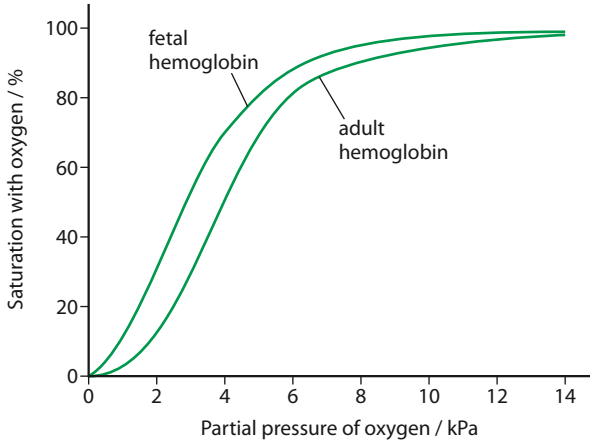


Figure D.23 Graph showing dissociation curves for adult and fetal hemoglobin.

Exam tip
Double check the dissociation curves for different respiratory pigments and practise explaining the differences.

Transport of carbon dioxide in the blood

Carbon dioxide produced during aerobic respiration is carried back to the lungs by the blood. It diffuses into capillaries close to respiring cells and is transported in one of three ways.

- About 70% of carbon dioxide enters red blood cells and is converted to HCO_3^- (hydrogencarbonate) ions.
- About 7% remains in the blood and is transported dissolved in plasma.
- The remainder is bound to hemoglobin.

Carbon dioxide reacts with water to form carbonic acid, which dissociates to form hydrogencarbonate ions and hydrogen ions:

$$CO_2 + H_2O \rightarrow H_2CO_3 \rightleftharpoons H^+ + HCO_3^-$$

The hydrogen ions bind to plasma proteins and this process has a buffering effect, preventing an excessive fall in pH in the blood. This reaction in the plasma is slow, but most of the carbon dioxide (about 70%) diffuses into the red blood cells where the reaction is catalysed by the enzyme carbonic anhydrase.

The hydrogen carbonate ions that are formed move out of the red blood cells via special channel proteins by facilitated diffusion. Hydrogencarbonate ions are exchanged for chloride ions, so the balance of charges on each side of the membrane is maintained. This is known as the **chloride shift** (Figure **D.24**).

The hydrogen ions remaining in the red blood cells bind reversibly to hemoglobin and prevent the pH of the cell falling, a process known as pH **buffering**.

Some carbon dioxide binds to hemoglobin. Each hemoglobin molecule can combine with one carbon dioxide molecule to form carbaminohemoglobin but, as it does so, it must release its oxygen. As carbaminohemoglobin is formed, oxygen is released into actively respiring tissues, just where it is needed. When red blood cells reach the lungs, the carbaminohemoglobin releases carbon dioxide, which is exhaled and hemoglobin is available again to collect oxygen. About 20% of carbon dioxide produced by respiration is carried in this way. As hydrogen ions and carbon dioxide bind to hemoglobin they cause the Bohr shift, which is described below.

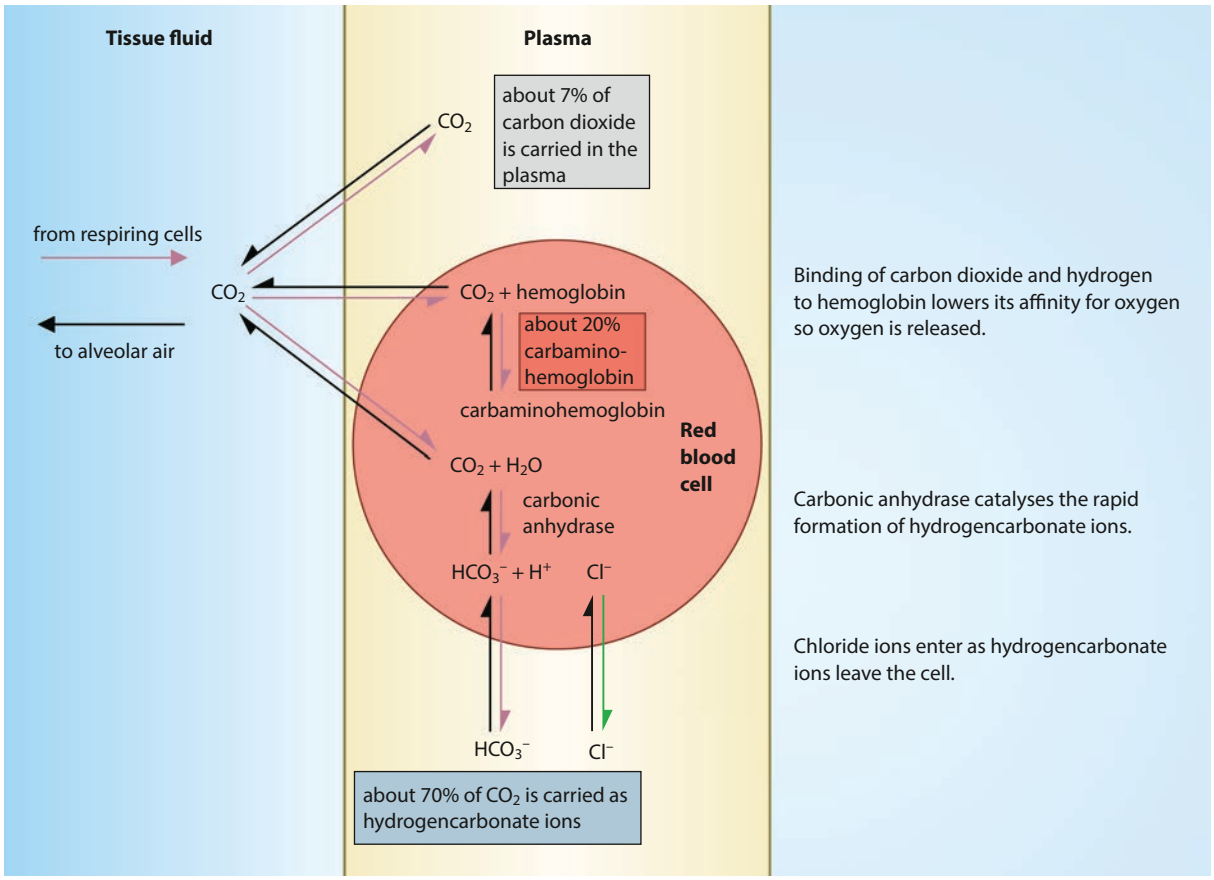


Figure D.24 Carbon dioxide transport in the blood and the chloride shift.

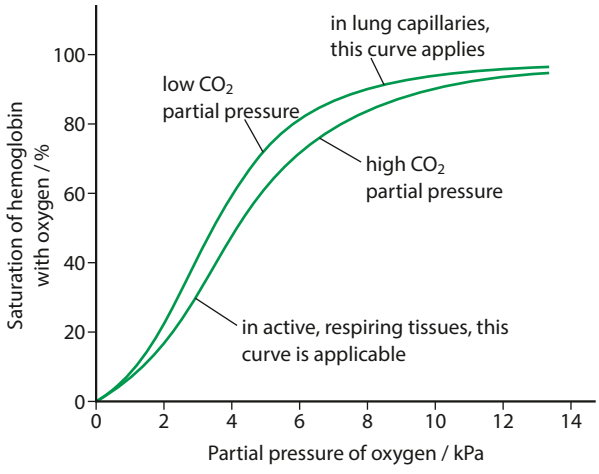


Figure D.25 The effect of carbon dioxide concentration on hemoglobin saturation: the Bohr shift.

The Bohr shift

The affinity of hemoglobin for oxygen is not only affected by the partial pressure of oxygen, but also reduced in the presence of high carbon dioxide concentrations. As the partial pressure of carbon dioxide rises, the ability of hemoglobin to combine with oxygen falls and so the dissociation curve moves to the right. This effect is known as the **Bohr shift**. It is caused when hydrogen ions produced from carbonic acid combine with hemoglobin. Figure **D.25** shows the effect of two different partial pressures of carbon dioxide on the oxygen dissociation curve. In an environment where the partial pressure of carbon dioxide is high, such as in actively respiring tissue, the curve moves to the right – which means that, at any given oxygen partial pressure, oxygen is more likely to dissociate from hemoglobin if the partial pressure of carbon dioxide is high. This effect promotes the release of oxygen in active tissues where respiration is producing high levels of carbon dioxide, so cells receive the oxygen they need.

Ventilation rate and exercise

When a person exercises, their ventilation rate and tidal volume (depth of breathing) increase. Muscles need oxygen for aerobic respiration and as the rate of exercise increases, so does the rate of oxygen consumption. Blood returning to the lungs also has a higher level of carbon dioxide, produced as a result of the increased activity. An increase in ventilation rate and tidal volume draws in more fresh air to maintain the concentration gradient between the alveolar air and the blood. Thus oxygen can be absorbed at a faster rate and the body can get rid of the additional carbon dioxide produced. These changes in ventilation are adjusted to match the body's metabolic needs.

Ventilation rate is controlled by the breathing centre of the **medulla oblongata** in the brain stem, which receives nerve impulses from sensory cells in different parts of the body. The breathing centre responds to match ventilation rate to activity levels (Figure **D.26**). **Chemoreceptors** in the inner wall of the aorta and carotid arteries respond to an increase in carbon dioxide in the blood. This excess carbon dioxide forms carbonic acid that cannot be buffered and so the pH of the blood falls. Impulses are passed to the medulla, which increases the ventilation rate by sending motor impulses to the intercostal muscles and diaphragm to increase their

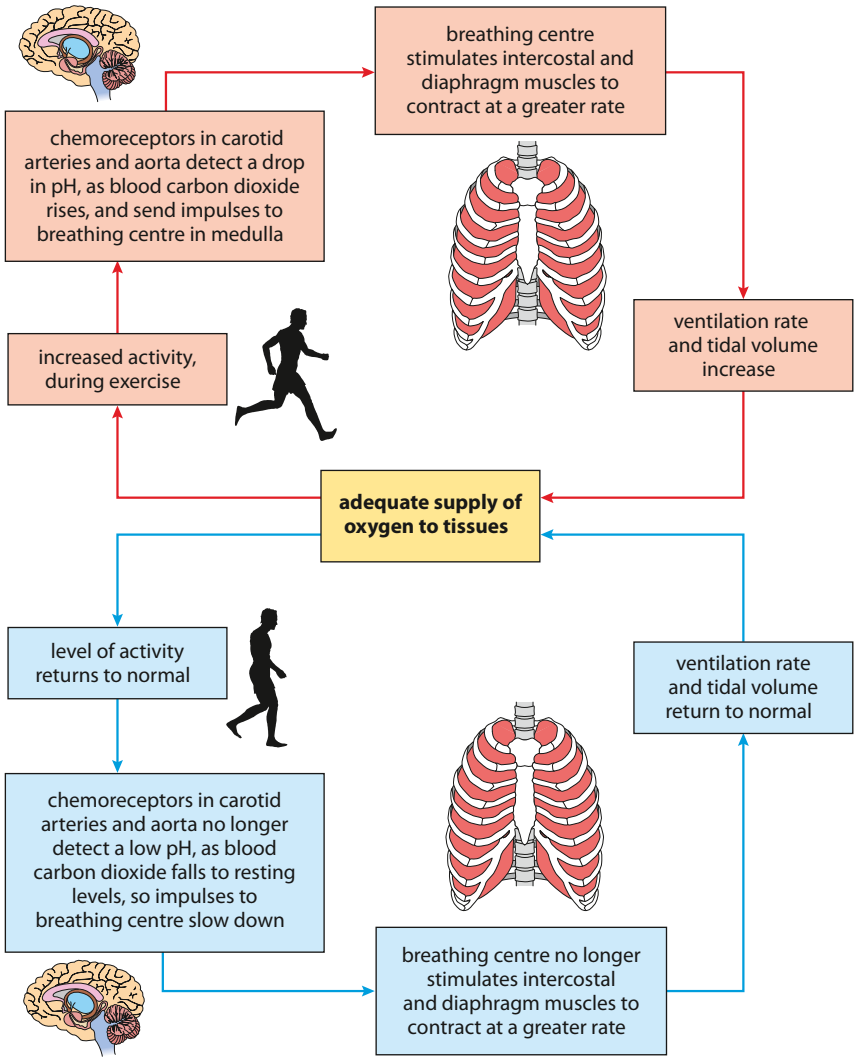


Figure D.26 The control of breathing.

rate of contraction (Subtopic **6.4**). The breathing centre contains similar chemoreceptors, which also respond to deviations of blood pH from the normal level. An increase in ventilation rate causes carbon dioxide to be removed from the body at a faster rate and blood pH returns to its normal level of between 7.35 and 7.45.

After exercise, as the level of carbon dioxide in the blood falls, ventilation rate decreases.

Gas exchange at high altitude

At high altitude, the percentage of oxygen in the air is the same as it is as sea level, but because air pressure is lower, the partial pressure of oxygen is reduced. At these lower partial pressures, hemoglobin does not become fully saturated with oxygen so that when a person moves suddenly from low to high altitude they may experience 'altitude sickness'. Symptoms include headache, nausea, dizziness and breathlessness as well as an increased heart rate. Altitude sickness can be avoided by travelling to high altitude gradually over a period of days so that the body has the opportunity to acclimatise.

During acclimatisation, the body adjusts both the ventilation and circulatory systems to cope with the lower oxygen availability. Ventilation rate increases temporarily and adjustments to the circulatory system ensure that the rate of oxygen delivery to the tissues increases. Over a period of weeks, the number of red blood cells and thus the concentration of hemoglobin steadily increase and the density of capillaries in the lungs and muscles also rises. After long periods at high altitude, the size of the lungs and the tidal volume increase so that the volume of air breathed can be 25% greater than at sea level. Once these adjustments have been made, heart and ventilation rates can return to their previous levels. People who live permanently at high altitude have a larger lung capacity and surface area for gas exchange, as well as a greater concentration of myoglobin in their muscles. Animals from high-altitude environments, such as llamas and vicunas, are also adapted to cope with the lower oxygen availability.

Some athletes choose to train at high altitude to develop increased levels of red blood cells and myoglobin. These adaptations are retained for a short time when they return to sea level and can improve athletic performance.

Causes and treatment of emphysema

Emphysema is a chronic obstructive pulmonary disease (COPD), which slowly destroys the alveoli in the lungs (Figure **D.27**) so that their bunch-like structure is lost and they eventually burst, becoming large, irregular sacs, which trap air in the lungs. The area available for gas exchange is reduced, so that insufficient oxygen reaches the blood, eventually leading to low oxygen levels (hypoxemia) and high levels of carbon dioxide (hypercapnia) building up in the blood. Emphysema also damages the elastic fibres in the lungs so that airways collapse during exhalation and air cannot be expelled properly. The main symptom of emphysema is shortness of breath, which usually begins gradually, so a person may not realise they have the disease at first, but in more advanced stages of emphysema the lungs become so damaged that sufferers become short of breath even while resting.

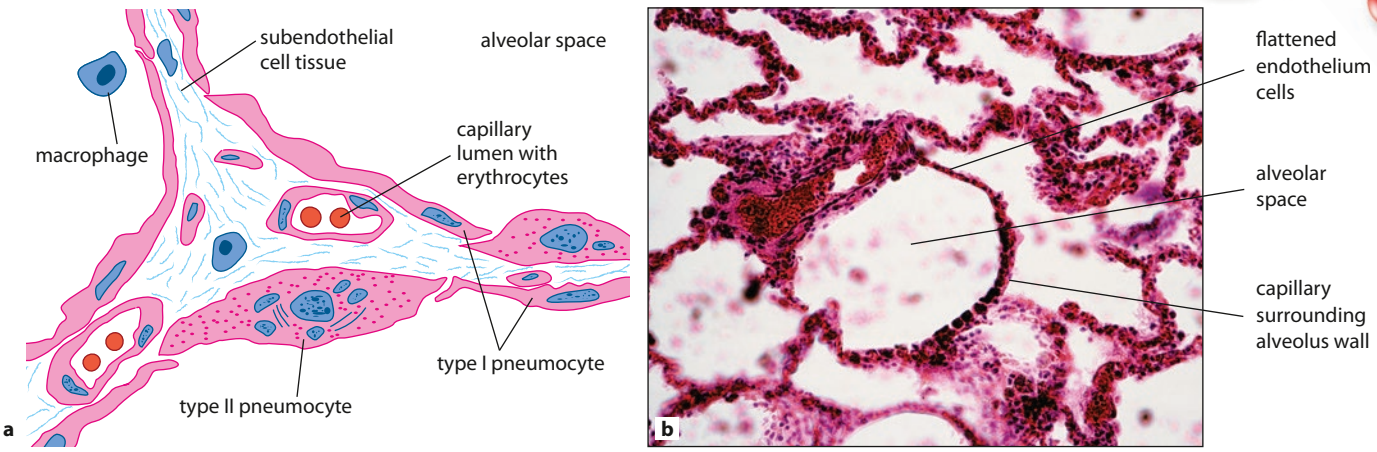


Figure D.27 **a** Enlarged diagram showing part of an alveolus. Pneumocytes line the alveoli where their secretions assist with gas exchange. They are easily damaged by smoking. **b** In this stained micrograph of healthy lung tissue, the structure of the alveoli, with capillaries running adjacent to the alveolar epithelium, is clear. The proximity of the red blood cells to the alveolar walls makes gas exchange highly efficient.

The most common cause of emphysema is smoking, although long-term exposure to air pollution, industrial pollutants and coal or silica dust can also be causes. Treatments can be used to reduce some of the problems of emphysema, but the damage it causes is permanent. At present, there are no drug treatments that can alter the rate of decline in lung function. The most important single factor in treatment is for the person to give up smoking so that lung damage is slowed down. Other treatments include bronchial dilators, delivered via an inhaler. These medicines relax the muscles in the bronchi and ease a person's breathing, in a similar way to treatments for asthma. In more severe cases patients are treated with oxygen therapy, and inhale oxygen from a cylinder either continuously or during periods of exercise.

Nature of science

Scientific understanding drives technical advance – athletics training

Athletes have always understood that regular training, with repetition of specific exercises to increase fitness, has an effect on the muscles and the cardiac and ventilation systems. The maximum ventilation rate during exercise can increase by 10–15% as a result of training, which will also lead to an increase in the vital capacity and development of new capillaries in the alveoli. These factors lead to a higher rate of oxygen absorption and of carbon dioxide removal. After regular training the heart becomes stronger and can pump the necessary blood to working muscles. If an athlete trains at high altitude, additional red blood cells form to carry the oxygen that is needed for improved performance.

Today elite athletes are more likely to use specific training regimes designed for them by sports scientists and physiologists who study individual athletes to maximise their performances. Precise measurement of metabolic products such as creatine in the blood, as well as the number of red blood cells a sample contains may make a difference of vital seconds or minutes to an athlete's performance.

But knowledge of science can be used in other ways, which many regard as undesirable. Erythropoietin (EPO) is a hormone that is naturally produced by the kidneys to stimulate red blood cell production in bone marrow. It is also used as a performance-enhancing substance by some athletes. Artificially increasing the proportion of red blood cells in the blood is known as **blood doping**. It can be achieved by transfusion – an athlete stores a blood sample in advance, and then some time later, just before a competition or event, receives a transfusion of his or her own red blood cells. Alternatively, an athlete may receive an injection of EPO (which can be produced by genetic engineering) to stimulate additional red blood cell production.

Is EPO OK?

Both professional and amateur athletes have been known to use EPO and there has been much debate about the acceptability of this practice. EPO has been used for blood doping in competitive endurance sports such as cycling, rowing and marathons. Until recently, there was no way to test for it directly, but in 2000 French chemists developed a method of distinguishing pharmaceutical EPO from the natural hormone normally present in an athlete's urine. This was accepted by the World Anti-Doping Agency and now EPO tests are conducted on both blood and urine samples at international events.

Some argue that, since it is possible to increase red blood cell counts by training at high attitude or sleeping in low-oxygen environments, it should be acceptable to achieve the same results by taking an injection of EPO.

EPO is also used therapeutically to treat patients who have kidney failure and cannot produce it naturally.

Questions to consider

- Can the use of EPO be regarded as cheating?
- What constitutes an acceptable level of risk for professional athletes when using substances such as EPO?
- Is the case of EPO any different from other substances used to enhance athletic performance?
- How might the views of medical practitioners and those of spectators differ in relation to the effects of EPO?

Test yourself

17 Define 'partial pressure'.

18 Explain what is meant by the term 'Bohr shift' and why it is important in supplying oxygen to respiring tissues.

19 Describe how fetal hemoglobin differs from adult hemoglobin in its affinity for oxygen.

Exam-style questions

1 Outline the cause and treatment of phenylketonuria (PKU). **[5]**

2 Distinguish between the mode of action of steroid and protein hormones. **[2]**

3 Explain why pepsin is initially synthesised as an inactive precursor and how it is subsequently activated. **[3]**

4 Outline the importance of acidic conditions in the stomach. **[3]**

5 Explain how the structural features of the epithelium cells of a villus are related to their function. **[5]**

6 Outline how blood is brought to the liver, circulates through the liver, and then leaves the liver. **[4]**

7 **a** Describe the process of erythrocyte and hemoglobin breakdown in the liver. **[5]**

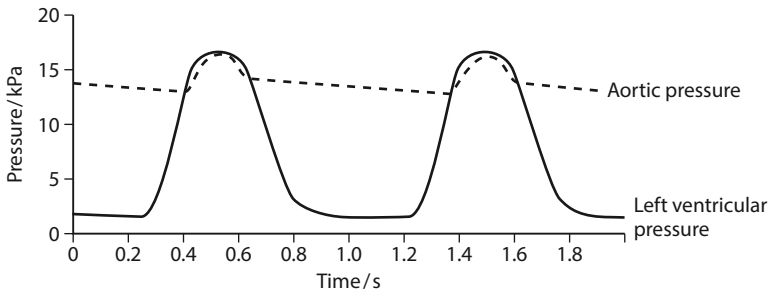
b Outline the role of the liver in the storage of nutrients. **[2]**

8 Explain the events of the cardiac cycle including systole, diastole and heart sounds. **[6]**

9 **a** Define the term 'partial pressure'. **[1]**

b Outline how the body acclimatises to high altitudes. **[3]**

10 The information in the graph shows blood pressure changes in the left ventricle and the aorta during two cardiac cycles.



a Draw an arrow on the graph to show when the atrioventricular valves close. **[1]**

b Calculate the heart rate shown by the graph. **[2]**

c What is the maximum pressure in the left ventricle during the cardiac cycle? **[1]**

d The pressure in the right ventricle rises to a maximum of about 3.2 kPa during the cardiac cycle. Suggest reasons why this pressure is different from that in the left ventricle. **[3]**

e State **two** reasons why the pressure in the left ventricle falls to a lower level than the pressure in the aorta. **[2]**