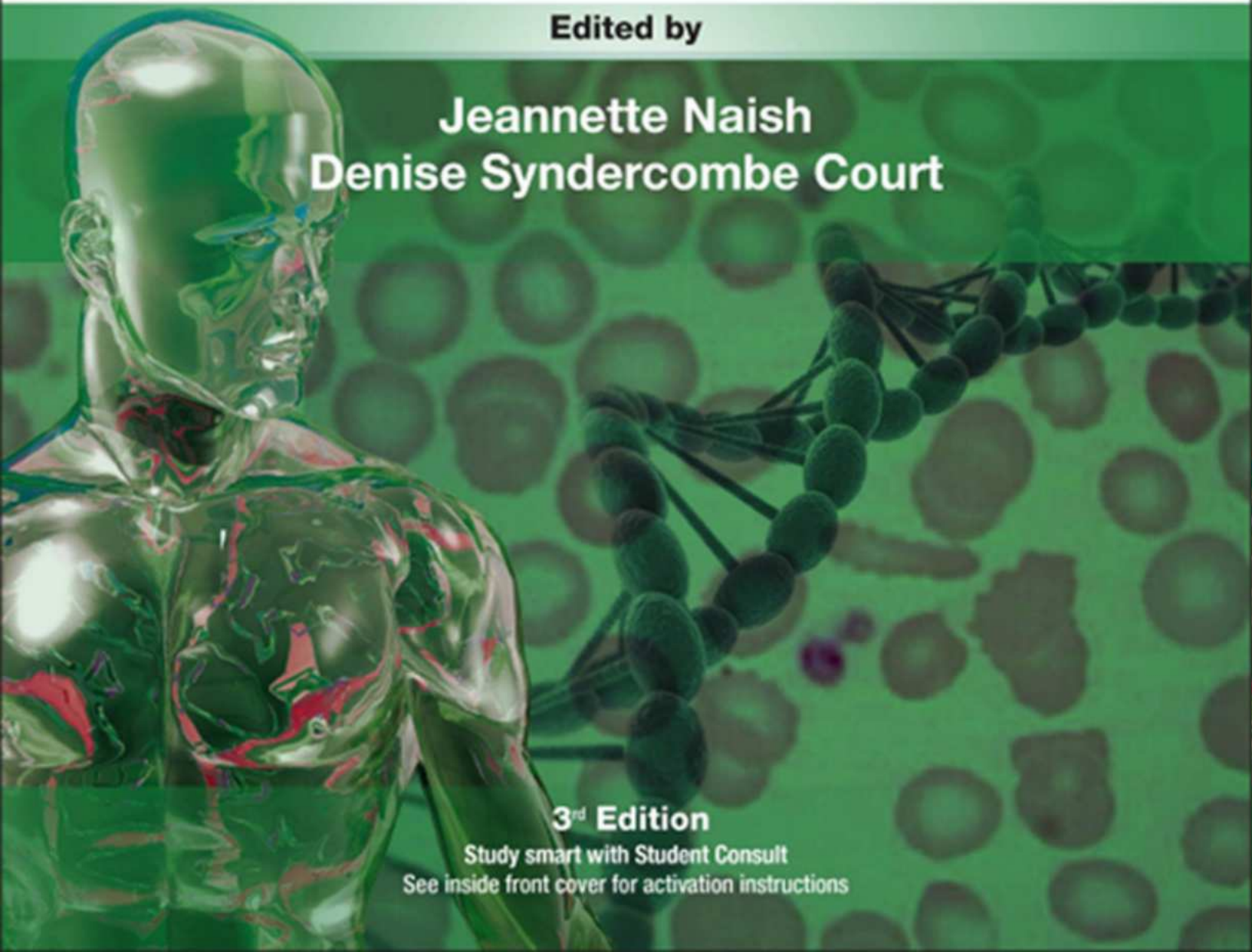
Enhanced DIGITAL VERSION Included. Details inside.

Medical Sciences

Edited by

Jeannette Naish
Denise Syndercombe Court



3rd Edition

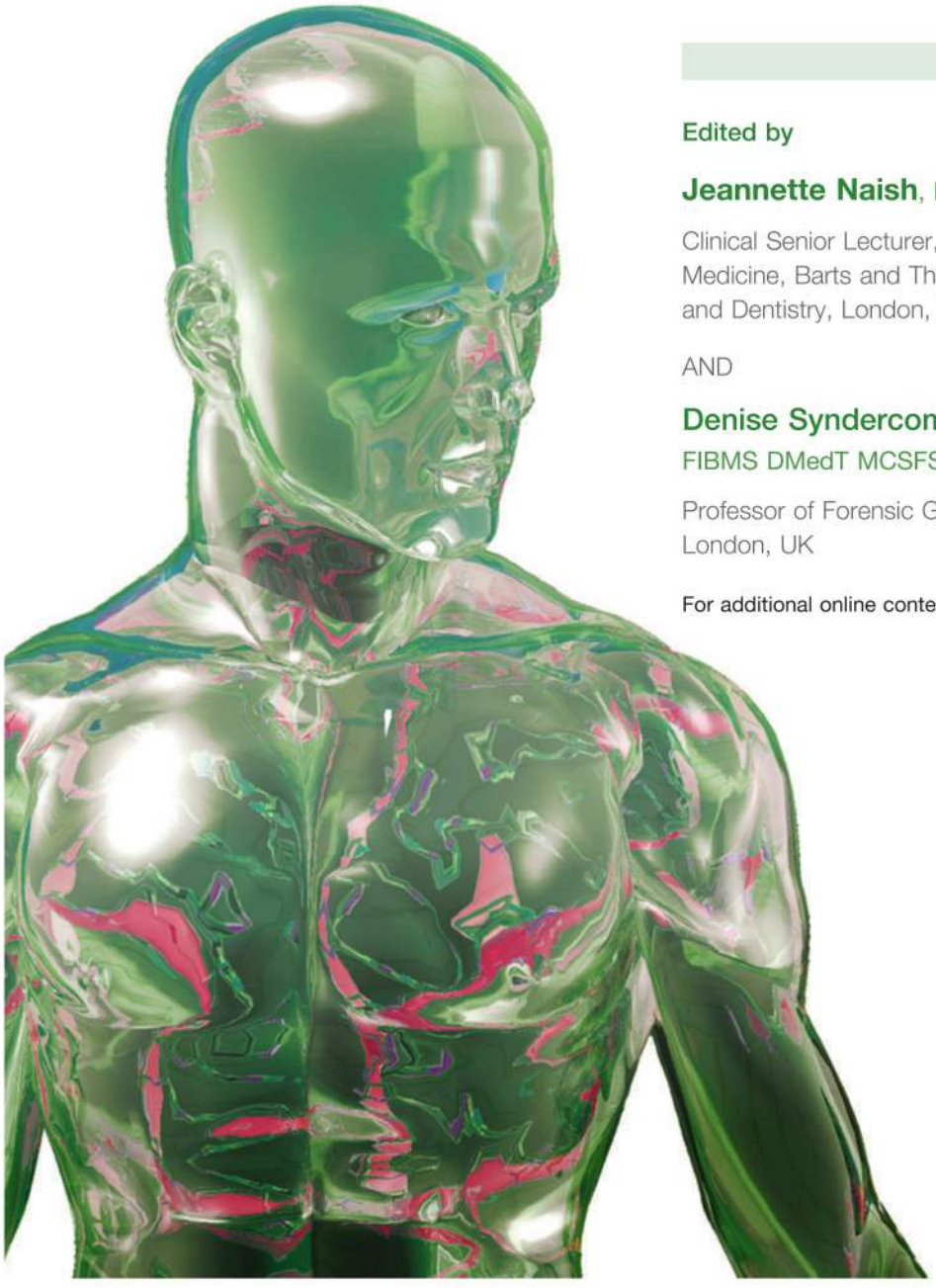
Study smart with Student Consult
See inside front cover for activation instructions

Medical Sciences

This page intentionally left blank

3rd Edition

Medical Sciences



Edited by

Jeannette Naish, MBBS MSc FRCGP

Clinical Senior Lecturer, Wolfson Institute of Preventive Medicine, Barts and The London School of Medicine and Dentistry, London, UK

AND

Denise Syndercombe Court, CBiol MRSB CSci FIBMS DMedT MCSFS PhD

Professor of Forensic Genetics, King's College London, London, UK

For additional online content visit StudentConsult.com

ELSEVIER

Edinburgh London New York Oxford Philadelphia St Louis Sydney 2019

ELSEVIER

First edition 2009
Second edition 2015
Third edition 2019

Notices
Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds or experiments described herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made. To the fullest extent of the law, no responsibility is assumed by Elsevier, authors, editors or contributors for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

ISBN: 978-0-7020-7337-3
Printed in Poland
Last digit is the print number: 9 8 7 6 5 4 3 2 1

Senior Content Strategist: Pauline Graham
Content Development Specialist: Carole McMurray
Project Manager: Anne Collett
Design: Christian Bilbow
Illustration Manager: Karen Giacomucci
Illustrator: Robert Britton, Antbits, Jade Myers
Marketing Manager: Deborah Watkins

Contents

Videos (www.studentconsult.com)

- **Nerve Impulse Transmission Made Easy**
  Shafiq Pradhan
- **Cardiac Physiology Made Easy**
  Shafiq Pradhan
- **Renal Filtration Made Easy**
  Shafiq Pradhan

Preface

We were delighted to be offered the opportunity to compile a third edition of Medical Sciences. This book was envisaged as a comprehensive introduction to medical studies, focussed on explaining the scientific foundation of core facts that are important to clinical medicine. It is unique in providing a text that integrates information across the diverse branches of medical science, focussing on body systems in health and linking to clinical phenomena. Accompanying the system chapters are more broadly ranging chapters that introduce the reader to concepts important to all students of medicine: homeostasis; biochemistry and cell biology; energy and metabolism; diet and nutrition; pharmacology; genetics; epidemiology and statistics.

Many aspects of medical science have developed or changed over the last few years and this new edition has provided us with the opportunity to update the material. Some chapters have been substantially rewritten. We have tried to avoid chemical formulae and mathematical equations that many students will not require, while maintaining an understanding of the processes that these relate to.

Some chapters include more clinical content than others, as clinical and information boxes. This is because these areas relate to more common, and therefore, important, clinical conditions. The student must, however, never forget that uncommon or rare conditions do exist and are, therefore, equally important.

It is never easy to get the balance right between basic and clinical sciences. We therefore welcome your feedback. We sincerely hope that you will enjoy reading this book and find it useful throughout your studies.

Jeannette Naish
Denise Syndercombe Court

Contributors

Andrew Archbold MD FRCP
Consultant Cardiologist, Barts Heart Centre, Barts Health NHS Trust, London, UK

Alistair Chesser MB BChir FRCP PhD
Consultant Nephrologist, Barts Health NHS Trust, The Royal London Hospital, Whitechapel, London, UK

Jeff Davies MA MRCP FRCPath PhD
Clinical Senior Lecturer, Department of Haematology, Barts and the London School of Medicine and Dentistry, London, UK

Marek H. Dominiczak dr hab med FRCPath FRCP (Glas)
Consultant Biochemist, NHS Greater Glasgow and Clyde, Department of Biochemistry, Gartnavel General Hospital, Glasgow, UK
Honorary Professor of Clinical Biochemistry and Medical Humanities, University of Glasgow

Gavin Donaldson BSc PhD
Reader in Respiratory Medicine, National Heart and Lung Institute, Faculty of Medicine, Imperial College London, London, UK

Joy Hinson BSc PhD DSc FHEA
Professor of Endocrine Science, Dean for Postgraduate Studies, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, UK

Peter MacCallum BMedSci MB ChB MD FRCP FRCPath
Senior Lecturer, Department of Haematology, Barts and the London School of Medicine and Dentistry, London, UK

Jeannette Naish MBBS MSc FRCGP
Clinical Senior Lecturer, Wolfson Institute of Preventive Medicine, Barts and The London School of Medicine and Dentistry, London, UK

Girish Namagondlu MBBS, MRCP
Consultant Nephrologist, Barts Health NHS Trust, Royal London Hospital, Whitechapel, London, UK

Penny Neild MD FRCP
Consultant Gastroenterologist and Honorary Senior Lecturer, St George's Hospital, St. George's University of London, London, UK

Adrian C. Newland BA MB BCh MA FRCP(UK) FRCPath
Department of Haematology, Royal London Hospital, Whitechapel, London, UK
Professor of Haematology, Institute of Cell and Molecular Science, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK

Clive Page BSc PhD OBE
Professor of Pharmacology, Head of Sackler Institute of Pulmonary Pharmacology, King's College London, London, UK

Despo Papachristodoulou MD
Reader in Biochemistry and Medical Education, Lead for the MBBS Graduate/ Professional Entry Programme, GPEP admissions tutor/MBBS senior tutor, GKT School of Medical Education, Faculty of Life Sciences and Medicine, King's College London, London, UK

Brian Pentland BSc MB ChB FRCP(Ed) FRCSLT
Head of Rehabilitation Studies (retired), University of Edinburgh; Professor (Honorary), Queen Margaret University, Edinburgh, UK
Consultant Neurologist (retired), Astley Ainslie Hospital, Edinburgh

Amrutha Ramu MBBS BSc MRCP MRCPGastro MSc
Consultant Gastroenterologist, Frimley Park Hospital, Portsmouth, UK

Peter Raven BSc PhD MBBS MRCP MRCPsych FHEA
Honorary Consultant Psychiatrist, Camden and Islington Mental Health Trust, London, UK
Faculty Tutor, Faculty of Medical Sciences, UCL

Lesley Robson BSc PhD
Reader, Institute of Health Sciences Education, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, UK

Armine M. Sefton MBBS MSc MD FRCP FRCPath FHEA
Emerita Professor of Clinical Microbiology, Centre of Immunology and Infectious Disease, Blizard Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, UK

Denise Syndercombe Court CBiol MRSB CSci FIBMS DMedT MCSFS PhD
Professor of Forensic Genetics, King's College London, UK

John Wilkinson BSc (Hons) PhD
Academic Manager (Retired), School of Life Sciences, University of Hertfordshire, Hatfield, UK

This page intentionally left blank

Acknowledgements

We thank all the contributors to this third edition of Medical Sciences; in particular our new contributors, as we recognise that joining an established writing team is often as difficult as to undertake an entirely new commission.

We would like to thank Elsevier for giving us the opportunity to update information in the previous edition as in some subjects especially, scientific development is fast moving. As Editors we have been supported through the project, in particular by Carole McMurray, Content Development Specialist. Pauline Graham, Senior Content Strategist, has been instrumental in commissioning this third edition and we would like to thank her for her encouragement and support through this process. We also thank the whole production team have been wonderfully efficient and thorough, providing the clarity necessary to communicate complex information through text and clear illustrations across the book pages to increase accessibility. We would also like to thank Shafiq Pradhan for creating the video animations which are a valuable addition to this third edition.

We would finally like to thank those contributors to the second edition who do not appear in the new addition, and acknowledge here, especially, the contribution of the late Patricia Revest, editor of the first edition. Without their contributions we would not be where we are today.

- Paola Domizio
- Mark Holness
- David Kelsell
- Drew Provan
- Mary Sugden
- Walter Wieczorek

Dedication

We would like to dedicate this book to all students of medicine and the medical sciences, past and future and hope that its contents will continue to provide knowledge for medical professionals in the future. We believe that you have to know the science in order to understand the practice of medicine.

Jeannette Naish
Denise Syndercombe Court

This page intentionally left blank

1 Introduction and homeostasis

Jeannette Naish

Disease ensues when normal physiological mechanisms and processes are disrupted. These processes take place in the basic unit of living organisms: the cell. It is therefore essential that all clinicians understand normal cellular and molecular mechanisms and processes in order to understand disease. The following chapters address specific mechanisms related to particular areas of human function and systems of the body.

The basic science concepts that attempt to explain disease processes cannot be undervalued. The best diagnostic and most effective therapeutic decisions made by clinicians have to be underpinned by sound scientific principles. The inclusion of all the relevant basic sciences in one book will, hopefully, be useful.

CHAPTER 2 BIOCHEMISTRY AND CELL BIOLOGY

This chapter gives an overview of the principles of mechanisms that enable the body to work as a biochemical system. The functional unit of the human organism is the cell (Ch. 2, ***Fig. 2.39***). All cells are surrounded by a **cell membrane**, also known as the **plasma membrane**. Other cell components are contained in the **cytoplasm** in which the cellular elements (organelles), including the nucleus, are suspended in the **cytosol** (**intracellular fluid** (**ICF**) or **cytoplasmic matrix**). Cells are suspended in fluid composed of water and a variety of biologically active molecules. Movement of these molecules into and out of cells, involving both active and passive transport, triggers the physiological mechanisms that enable the cells to perform their normal physiological functions. Examples include protein synthesis, regulation of cell function (signalling), cell movement, metabolism (glucose and respiration), cell division and death (apoptosis), skeletal and cardiac muscle contraction, the transmission of signals along nerve fibres, the digestion and absorption of nutrients in the alimentary system, the synthesis and secretion of hormones by the endocrine system, transport of oxygen and carbon dioxide by blood, the exchange of respiratory gases and the important functions performed by the renal system. These cells perform different functions, and therefore possess different properties, described in detail in the ensuing chapters on the systems of the body. Understanding cell and molecular biology – the similarities and differences between cell types, their components and functions – is essential to understanding the clinical sciences because disease results from the disruption of normal mechanisms. These principles underpin the development of disease, therapeutics and, in particular, the understanding of cancers and their treatment.

CHAPTER 3 ENERGY AND METABOLISM

Chapter 3 discusses the cellular mechanisms that enable human beings to produce the energy needed to survive, maintain body temperature and work. Most biological processes are driven by energy in the form of adenosine triphosphate (ATP), produced through metabolism of the food that we eat. The main metabolic fuels are carbohydrate, protein and fat. The most important source of energy is glucose, but the body has intricate and dynamic adaptive mechanisms for using alternate fuels under particular physiological conditions. Metabolism occurs in cells. It is tightly regulated by the actions of enzymes, gene expression and transcription in response to changing demands on the need for energy and by the action of hormones, which may take place rapidly or gradually. Energy metabolism is essential for life, and disturbances can lead to important diseases, such as diabetes mellitus.

CHAPTER 4 PHARMACOLOGY

This chapter describes how drugs work (pharmacodynamics) and how they are absorbed, distributed around the body (pharmacokinetics), metabolised and then eliminated. Knowledge of cell and molecular biology underpins the understanding of pharmacology and therapeutics. The pharmacokinetics and pharmacodynamics of synthetic drugs depend on their individual properties. Specific

classes of drugs share common properties, but there are variations between individual drugs. It is also important to remember that how a drug performs in the laboratory (in vitro) is not necessarily how it performs in the body (in vivo), which is important for the safety and effectiveness of drugs. Generally, pharmacokinetics follows the principles of cell biology. In pharmacodynamics, drugs work by targeting cellular processes to either enhance or inhibit the process. Examples include the targeting of enzymes, transport processes and receptors on cell surfaces. Here, understanding of the autonomic nervous system is essential because most drugs are designed to target elements of this system.

CHAPTER 5 HUMAN GENETICS

The understanding of genetics dates back to Charles Darwin's (1809–1882) *On the Origin of Species* (1859), later further explained by Gregor Mendel's (1822–1884) principles of inheritance and mutations. The most exciting modern development in genetics was the Human Genome Project, which mapped the complete set of genetic codes stored as DNA sequences in the whole 23 chromosomes of the human cell nucleus and took place between 1990 and 2003. The Human Genome Project published the working draft of the human genome in 2000; the complete genome was published in 2003.

Taking advantage of the multiplexing capabilities of new sequencing technologies, the 1000 Genomes Project ran from 2008 until 2015, targeting sequence variation in five continental regions. In 2015 *Nature* published this work, which mapped population genetic variation from more than 2500 human genomes, providing publically available data for research.

In the UK the 100,000 Genomes Project was launched by Genomics England at the end of 2012. Its aim was to provide sequence data from NHS patients diagnosed with cancer or a rare disease with the aim of stimulating the UK genomics industry. With its medicine-focussed approach the 100,000 Genomes Project is currently planned to run until the end of 2018 and has been expanded to include infectious disease. By mid-2017 the project had sequenced more than 36,000 genomes, putting the UK at the forefront of using genomic technology to transform patient care. The importance of partnering with industry is crucial to the project so that frontline clinicians of the future will be provided with the necessary infrastructure to benefit from this exciting future so that we can better understand disease processes and the development of preventive measures, diagnosis, prognosis and therapeutic strategies as genomic medicine moves into the mainstream.

CHAPTER 6 PATHOLOGY AND IMMUNOLOGY

Pathology and immunology are essential for understanding disease processes to enable the clinician to formulate sensible diagnostic and therapeutic decisions. Infectious diseases and the body's response to them, immunology, are discussed. Disorders of the immune system, including autoimmunity and hypersensitivity, are also discussed. In these conditions, it is thought that there is a defect in the genetic regulation of the immune response. The inflammatory response underpins the body's defence mechanisms and needs to be fully understood. This is followed by the pathology of neoplasia; cancers, which cause about 25% of all deaths in the UK. The pathology of common degenerative diseases is discussed in the chapters on systems of the body. Once again, molecular and cellular biology and medical genomics form the basis for understanding these processes.

CHAPTER 7 EPIDEMIOLOGY

Chapter 7 is about the epidemiological principles that underpin the discovery of patterns of diseases and their occurrence in populations, and how the effectiveness of therapeutic interventions is evaluated. It is, perhaps, unusual to consider this as a basic science. Epidemiology and the epidemiological approach, however, is the science that underpins the art of clinical medicine. Observational studies form the cornerstone of clinical medicine.

For example, how do we know how to diagnose disease from patient descriptions of symptoms? Our understanding of how disease presents and progresses clinically is based on repeated, multiple observations by many doctors and the sharing of their observations; e.g. whooping cough, which starts like a common cold, before the cough develops and continues for up to 100 days. The cough is characteristic in being spasmodic and prolonged, often ending in a sharp intake of breath – the 'whoop'.

In the example of John Snow and the Broad Street Pump, Snow found the association between the water from the Broad Street Pump and the cholera epidemic. The actual cause of cholera, the organism *Vibrio cholerae*, was not discovered until later by Filippo Pacini, an Italian anatomist, and was not widely known until published by Robert Koch some 30 years thereafter. Until the comma-shaped bacterium was identified, treatment and prevention could not be formulated. Careful and systematic observation thus formed the basis for further research into the cause of this disease.

Moreover, how do we select therapeutic interventions, whether pharmacological or surgical? How do we know that this intervention is effective, or more effective than another one? Here, the methodology for experimental studies, e.g. randomised controlled trials, and the statistical concepts that underpin the proof for the likelihood of a positive effect need to be understood. The mathematics might be daunting, but understanding the principles is essential. These principles also apply to diagnostic and screening tests.

SYSTEMS OF THE BODY

The next eight chapters are about all the systems of the body and discuss the cellular makeup of different organs, their functions, normal metabolic processes in health and the biological basis for disturbance leading to disease. Understanding these processes forms the rationale for diagnostic and therapeutic decisions. Despite their separation, the systems interconnect so that the body functions as a whole. Rather than describe each chapter in detail, it might be more helpful to think about the cellular mechanisms that ensure normal physiological function. These basic mechanisms are common to all living organisms, including *Homo sapiens*.

As mentioned previously, the basic unit of the human organism is the cell. Normal biological functioning is determined by molecular and cellular processes and controlled by human genomics and epigenetic modification, as outlined in Chapter 5. The cells in each system vary according to their physiological function. For example,

hepatic (liver) and muscle cells both store glycogen, but the primary function of the liver is to release glucose converted from glycogen (glycogenolysis) into the circulation when there is a shortage of glucose, whereas muscle cells (myocytes) are primarily need to break down the stored glycogen for generating ATP for muscle contraction. Skeletal muscle lacks the enzyme glucose-6-phosphatase (G6Pase); glucose-6-phosphate generated from muscle in glycogenolysis instead enters the glycolytic pathway after glucose, preserving one of the ATP molecules consumed at the start of glycolysis.

A more obvious example of the influence of genomics is sickle cell disease (SCD). This is a condition where there is a mutation in the haemoglobin gene (β-globin gene), leading to the red cells assuming a sickle shape and becoming rigid. Sickle cells confer a resistance to malarial infection, and the mutation arose historically among populations in tropical and subtropical regions where malaria is endemic. The disadvantage is that, under conditions of reduced oxygenation, infection, cold or dehydration, the sickle haemoglobin elongates and cannot flow smoothly through small blood vessels. It sticks to the vessel lining, leading to occlusion of the vessels and causing sickle cell crises, which may be life-threatening. An understanding of molecular and cell biology and human genomics for the cells in each system is therefore necessary for understanding disease processes.

CHAPTER 16 DIET AND NUTRITION

Chapter 16 is about the nutritional needs for humans to stay alive and, more importantly, the principles for assessing these needs in health and disease. What makes a human being eat or not eat is also addressed, with implications for dietary control of conditions such as obesity and some therapeutic diets for chronic conditions such as inflammatory bowel disease. The association between diet and disease is also discussed. Nutritional support during severe illness, artificial nutrition, and associated complications are discussed. Artificial nutrition includes enteral feeding, i.e. putting feeding liquid directly into the stomach or small intestine, and parenteral nutrition, which is intravenous feeding. The makeup of the feeding fluid will depend on the nutritional needs of the patient. These principles are important, especially during the foundation years. Inclusion of nutrition as a basic science in this book is perhaps unusual, but clinicians need to know about these principles for sustaining life.

HOMEOSTASIS

To maintain the normal physiological processes for sustaining life, all living organisms and cells have to maintain a stable internal environment in response to changes in external conditions. Physiologists have called this function **homeostasis**, from the Greek *homeo* meaning same or unchanging, and *stasis* meaning standing still. When an attribute of the organism or cell (such as pH or temperature) changes for any reason, this complex system of processes adjusts the attribute back to the set constant level needed for physiological functioning. Such an attribute is labelled a **variable**, something that is changeable.

Homeostatic systems are multiple, dynamic mechanisms that are **regulated** (or controlled) for making the adjustments necessary for a stable internal environment; this is unlike simple dynamic equilibrium or steady states that are not regulated. Many examples of human homeostasis are discussed in the following chapters. Disease ensues when homeostatic mechanisms break down and the body exhibits **symptoms** (what the patient experiences) and **signs** (what the clinician finds on clinical examination).

Many physiological parameters, such as blood glucose level (discussed in detail in Ch. 3, Energy and metabolism), water and electrolyte (sodium, potassium, calcium, etc.) balance and body temperature, are examples of precise control by homeostatic mechanisms. Of the homeostatic mechanisms that control body fluids, **fluid balance** (the control of fluid volumes) and **acid–base balance** (the control of acidity [H^+ ions]) are important to understand.

Homeostatic regulation mechanisms

Homeostatic control mechanisms have three (sometimes more) interdependent components for the variable being regulated.

1. A receptor that detects, monitors and responds to changes (sometimes wide variation) in a variable in the external environment; known as the **sensor**.
2. The sensor sends information to a **control centre** that sets the physiological range for the variable, and determines the necessary response for bringing the variable back to the set point. In humans, the control centre is usually in the brain. Many examples are discussed in Chapter 8.
3. The control centre sends signals to the tissues and organs, known as the **effectors**, that have to effect, i.e. make the adjustment, to changes in the relevant variable to bring it back to its set point.

A simplistic analogy would be ambient temperature control in air-conditioning systems, where the thermostat is the sensor responding to changes in environmental temperature, set at a comfortable level. It also acts as the control system that switches heating or cooling systems on and off. The effector would be the heating and cooling systems with their own, separate mechanisms.

Once the control centre receives the stimulus that a variable has changed from the set point, it sends signals to effectors to correct the change by:

- Negative feedback to depress the change if the variable level has increased beyond the narrow, set range. This is the commonest mechanism.
- Positive feedback to affect an increase or acceleration in the output variable that has already been triggered. The result is to push the level beyond the physiological range.
- Feedforward control to either depress or enhance the level of a variable before the change is needed, i.e. anticipatory (or open loop). Open-loop systems have no way to calibrate against the set point and so always need an accompanying closed-loop negative feedback to correct any over- or underanticipation.

Negative feedback

Negative feedback mechanisms can either increase or reduce the activity of tissues or organs back to normal, set levels, and the system is sometimes called a negative feedback loop ***(Fig. 1.1)***. Numerous examples of negative feedback exist in the metabolic processes of all physiological systems.

Homeostatic control of glucose metabolism

An example of a negative feedback system is the homeostatic control of blood glucose. Among the tissues of the body, red blood cells and the brain (under normal conditions) can only use glucose to

generate the energy needed to drive metabolic processes. Glucose is essential to ensure an adequate supply of energy for the vital functions performed by these and other tissues. Blood glucose concentration (measured as fasting blood glucose) is therefore tightly maintained within a narrow range (3.5–8.0 mmol/L) normally. Here, the sensor is specialised pancreatic cells that receive blood from the portal circulation. The control is the autonomic nervous system, and the effectors are the α (secreting glucagon) and β (secreting insulin) pancreatic cells in the islets of Langerhans. A simplified explanation of glucose homeostasis is shown in ***Fig. 1.2***.

- When blood glucose concentration is too high (hyperglycaemia), e.g. following a high-carbohydrate meal or the ingestion of excessive amounts of alcohol, increased secretion of the hormone insulin causes increased uptake of glucose into cells and inhibition of the glucagon-secreting α cells, thus reducing blood glucose concentration towards normal. These processes are described in detail in Chapters 3 and 16.

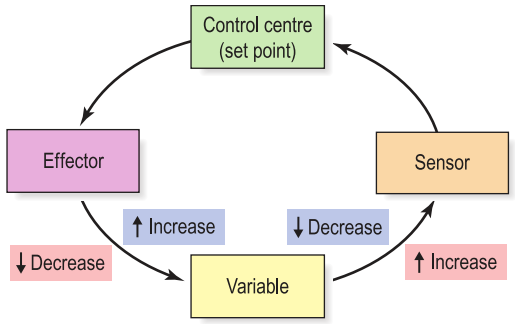


Fig. 1.1 Negative feedback loop. An increase in the variable produces an effector response to decrease it and a reduction in the variable leads to an effector response to increase it with the aim of returning to equilibrium.

- When blood glucose concentration is abnormally low (hypoglycaemia), as in prolonged fasting, glucagon is released from the pancreas to trigger alternative metabolic pathways to bring the level up: glycogenolysis, the process in which glucose stored in the form of glycogen is broken down to glucose; gluconeogenesis, in which, when glycogen stores are depleted, other metabolic fuels such as fat and protein are converted to glucose, and alternative fuels, such as fatty acids and ketone bodies, are used for generating energy.

Thermoregulation

Another example of a physiological homeostatic negative feedback system, addressed elsewhere in Chapter 8, is the control of body temperature: **thermoregulation.** Ambient environmental temperature can vary widely (–50°C to +50°C), but human body temperature has to be set at about 37°C (range 36°C to 38°C) for normal physiological functioning. For example, some mechanisms in glucose metabolism require energy, and if body temperature falls below a certain level, there will not be enough energy to drive the process. ***Fig. 1.3*** shows the mechanisms for controlling body temperature by negative feedback.

Body temperatures outside the normal range are defined as:

- Hyperthermia, when core temperature rises above 40°C
- Hypothermia, when core temperature falls below 35°C.

Prolonged and significant elevation (as in hyperthermia) or depression (as in hypothermia) in core body temperature (see below) can have fatal consequences.

Human body temperature (Clinical box 1.1)

As mentioned above, human body temperature is set at about 37°C. This can be measured through different anatomical orifices, such as the oral, rectal or vaginal orifice and the external auditory meatus. These are measurements of peripheral temperature and will vary between healthy subjects depending on where the measurement is taken. The **core** temperature (or core body temperature) is the temperature needed for normal physiological functions in deep organs such as the liver or brain, and is different from peripheral

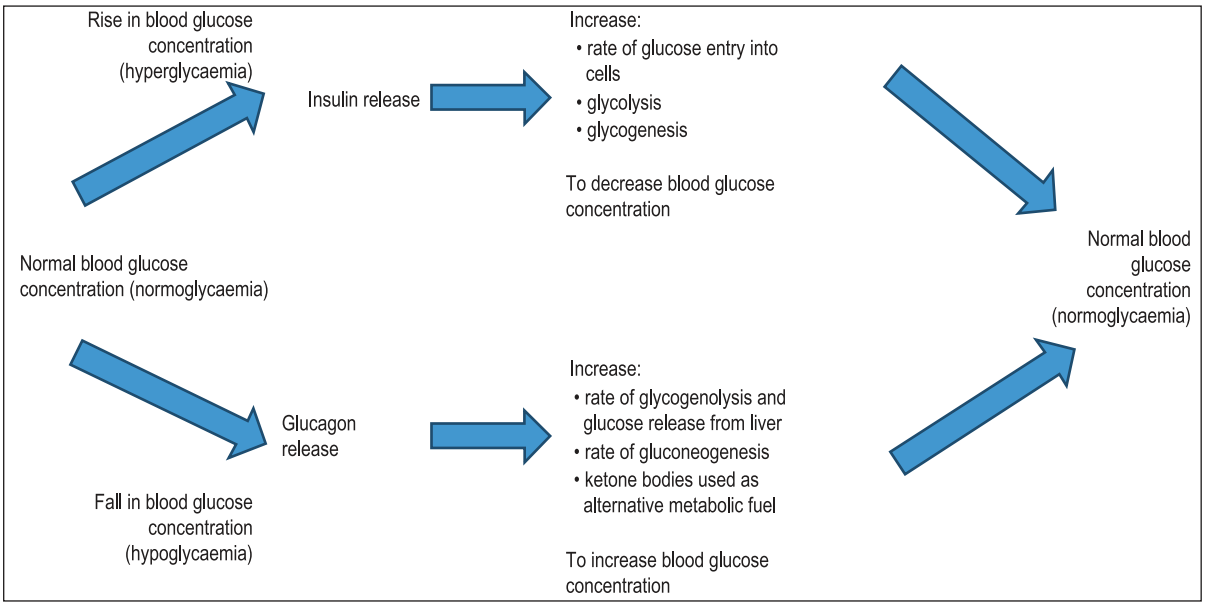


Fig. 1.2 Simplified scheme for glucose homeostasis. Increased blood glucose concentration leads to increased insulin secretion to lower blood glucose concentration back to normal, and a reduction in blood glucose concentration leads to the release of glucagon to raise blood glucose concentration to normal.

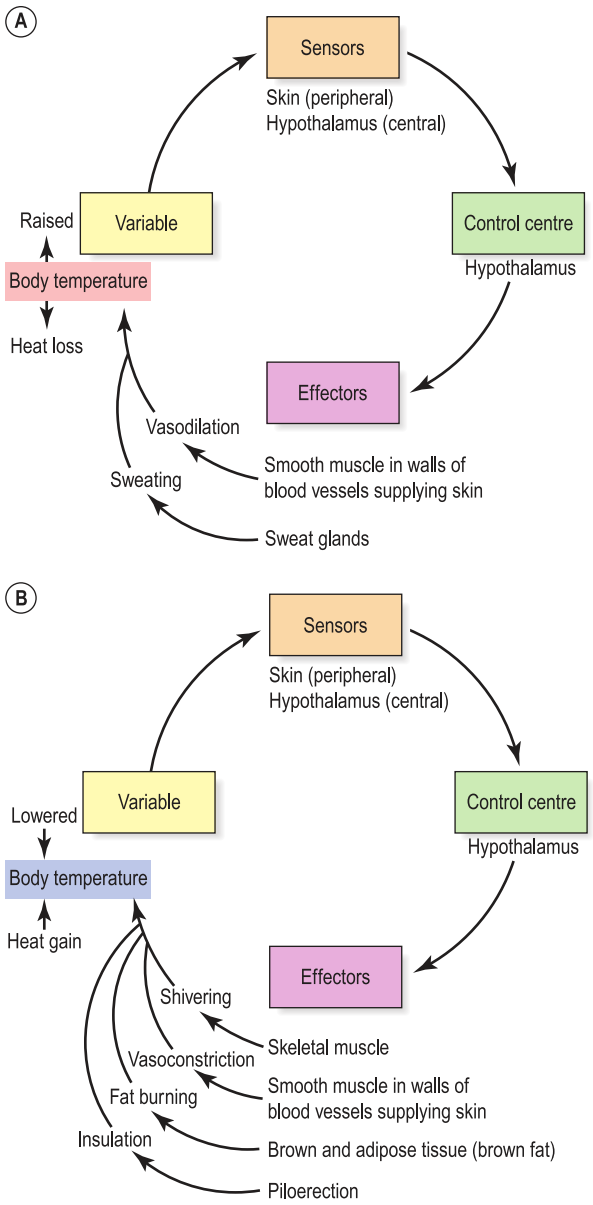


Fig. 1.3 Control of body temperature by negative feedback. (A) Responses to an increase in body temperature; (B) responses to a decrease in body temperature.

temperatures. Core temperatures have to be measured by inserting a deep probe, which is not always possible, so that rectal or vaginal temperatures are taken as an accurate reflection of core temperature.

Body temperature also varies according to the time of day, known as the **circadian rhythm**, an endogenous biological process driven by light and darkness in the external environment. Rhythmic physiological and behavioural patterns can be adjusted to follow the oscillations in circadian rhythm, a phenomenon known as **entrainment**. Sleep and wakefulness is an example (see Ch. 8). Core temperatures are higher in the evenings and lower in the morning, with the lowest temperature during the second half of sleep, about 2 hours before waking.

Behavioural and environmental factors can affect body temperature when homeostatic mechanisms will be called into play. For example, eating, drinking and exercise can raise body temperature; jet lag and shift work upsets circadian rhythm and thus the pattern of body temperature. Ambient temperatures will affect body temperature.

Using the negative feedback loop:

- The main **sensor** of temperature is the **skin**
- The **control centre** for body temperature is the hypothalamus (see Ch. 8 for details)
- The effectors for adjusting core temperature are the skin and muscles.

Heat-loss mechanisms

The skin is responsive to peripheral temperature, where capillary blood can be heated or cooled. When carried to the hypothalamus, the temperature of this blood is measured as core temperature. The hypothalamus then sends signals to the effector organs (see ***Fig. 1.3***).

Heat loss through the skin occurs when blood flow through the skin increases by vasodilation, and the skin becomes reddened and 'hot'. Heat is then lost through radiation to the atmosphere. Sweat glands in the skin also promote heat loss. Increased sweating increases water evaporation from the surface of the body and lowers temperature through the latent heat of evaporation. This mechanism is less effective in humid atmospheres. Also, excessive loss of water through sweating can lead to dehydration. For example, the combined effects of lack of drinking water and the hot, dry atmosphere in deserts are well documented and can be fatal.

Heat-gain mechanisms

Falls in core temperature trigger heat-gain mechanisms. Physiological adjustments include:

- In the skin, arteriolar vasoconstriction, leading to pallor and even cyanosis, reduces blood flow to prevent heat loss.
- Sweating almost ceases so that minimal heat loss occurs through water evaporation.
- Body hairs become erect through the action of erector pili muscles in the skin. Known as piloerection, warm air is trapped between the hairs to create an insulation effect.
- Increased muscular activity causes shivering to generate heat.
- The temperature in cells is raised by direct conversion of fat stores to heat energy by mitochondria (see also Ch. 3). Brown fat stores in infants are a form of fat specialised for conversion to heat energy; they are also abundant in hibernating animals. Their increased cellular fat droplets, mitochondria and associated increased capillary bed serve to supply oxygen and disperse heat faster.

Behavioural mechanisms can also operate; e.g. exercising to generate heat, putting on more clothing to conserve body heat and reduce heat loss, turning up the air-conditioning thermostat. In furry animals, piloerection is equivalent to putting on clothes. The

i Clinical box 1.1 Fever

The set point for temperature control is not always fixed. In infection, toxins released from bacteria and chemicals produced by cells of the immune system change the set point upwards (see Ch. 6). The normal mechanisms to generate heat, such as shivering, are triggered, leading to an increase in body temperature known as fever or **pyrexia**. The cause for this fever is thought to be a mechanism to activate certain immune cells and to limit bacterial growth. A higher rate of metabolism will also produce a faster rate of healing and more rapid induction of defence mechanisms. If the temperature becomes too high, however, the proteins inside the cells may be damaged.

erect hairs trap air, providing insulation, and the warmed air helps to conserve body heat.

Thermoneutral zones

In humans and other warm-blooded animals, the **thermoneutral zone** (**TNZ**) is the range of body temperatures in which the organism only needs to use a minimum amount of energy for maintaining normal body temperature. Within the TNZ, vasomotor response controls blood flow between the core and periphery to adjust for the amount of heat loss or gain from the body surface. Organisms use different mechanisms for adjusting body temperature within the TNZ; for example by changing posture or going into the shade to avoid heat, and into the sun for warmth. In humans, the TNZ is about 27°C at rest, and energy is expended at temperatures above and below this for maintaining body temperature.

Positive feedback (Information box 1.1)

Once triggered, some homeostatic mechanisms need to continue. This continuation is known as **positive feedback**. Unlike negative feedback, positive feedback has no set point, so the process can continue indefinitely if unchecked. As the process proceeds, small deviations from the original variable become amplified, and the process becomes a cascade. The 'brake' is usually the desired physiological outcome that ends the feedforward cascade, so this is normally a self-limiting mechanism. ***Fig. 1.4*** diagrammatically shows a positive feedback loop. Uncontrolled feedforward leads to disease.

i Information box 1.1 Some examples of positive feedback

- The coagulation cascade is described in Chapter 12. A positive feedback loop is triggered by tissue damage after injury, which initiates the coagulation cascade to stop bleeding. Coagulation (blood clotting) arrests the bleeding and then terminates the cascade.
- In the menstrual cycle (see Ch. 10), the rise in oestrogen levels during the follicular phase reaches a spike that triggers ovulation, after which oestrogen levels fall, terminating the follicular phase.
- During childbirth, the rhythmic uterine contractions for expelling the foetus are activated by the hormone oxytocin secreted by the pituitary gland (triggered by the hypothalamus). The pressure of the foetal head on the lower uterine segment continues to stimulate oxytocin release until the baby is delivered, when oxytocin secretion stops.
- In genetics, the production of gene transcription factors is accelerated by feedforward loops (see Ch. 5) in which one transcription factor regulates another, and they both regulate the target gene. The process is normally self-limiting, as the loop terminates when transcription is achieved.
- In cancer genetics (see Ch. 5), the proliferation of mutated cancer cells can occur through feedforward mechanisms that are unchecked. For example, in some forms of breast cancer the inflammation-associated enzyme nitric oxide synthase (NOS) is upregulated in response to factors such as hypoxia and exogenous NO production, consistent with a feedforward regulation of NO in the tumour microenvironment and associated with aggressive disease and poor prognosis.
- Production of the action potential is another example: the plasma membrane ion channels of the neuron plasma membrane are closed at the resting potential. Stimulated by chemical signals from the nerve ending, the sodium ion channels open, leading to an inflow of sodium that increases the membrane potential. This causes more ion channels to open, causing a rapid increase in the membrane potential as sodium ions pour in, resulting in a strong electrical signal. This continues until all the channels are open (the membrane reaches a threshold potential at which point the membrane polarity has reversed and the sodium channels become inactivated, reversing the process).

Feedforward (Information box 1.2)

Feedforward mechanisms trigger the change in a variable before the change is needed; it anticipates the need for the change and accelerates the change.

Water and electrolytes: homeostatic control of body fluids

The functional unit of the human organism is the cell; the human body is made up of some 10 trillion cells (10^{13}). The structure of the human cell is discussed in Chapter 2 (Biochemistry and cell biology). Each cell contains fluid, which is a solution of electrolytes and a variety of biochemical compounds and in which organelles, including the nucleus, are suspended. **Extracellular fluid** (**ECF**) compartments are separated from cells by the cell membrane. The cells and extracellular compartments of different organs have different components that perform biological processes that determine their functions. Depending on the requirements for these biological processes, the fluids, electrolytes and biochemical compounds are constantly being transferred from one compartment to another. The precise mechanisms for maintaining this balance are not yet fully understood, so laboratory measurements of serum electrolytes can only be a rough guide to the state of the internal environment.

The ensuing chapters describe the makeup of the various human organs, their composition and functions. In particular, Chapter 4 (Pharmacology) discusses drug distribution through the body, the mechanisms concerned with transferring compounds and metabolites into and out of cells; and Chapter 16 (Diet and nutrition) discusses the requirements for water and electrolytes, the balance of these substances in healthy people and the processes that could disrupt normal function.

Fluid compartments

Total body water in a healthy 70-kg adult male is between 40 and 45 litres. The amount of total body water is inversely related to body fat (adipose tissue); therefore a higher proportion of fat leads to a lower proportion of water.

Basically, there are two compartments of fluid in the human body (***Fig. 1.5***): the **ICF** compartment, which is fluid within all the cells and is the larger of the two, and the **ECF** compartment, which surrounds the cells. The ICF compartment occupies about two-thirds of total body water (40% body weight).

The ECF compartment is further divided into:

1. **Interstitial fluid** (ISF) between the cells, the larger of the ECF compartments, occupies two-thirds of the ECF.

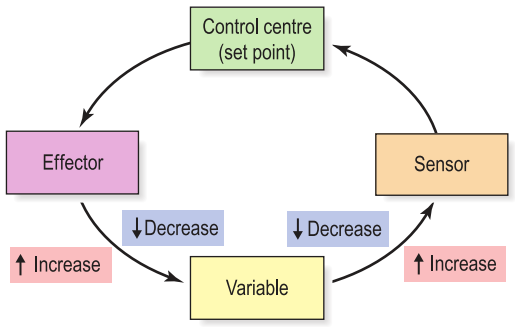


Fig. 1.4 Positive feedback loop. An increase in the variable produces an effector response to increase it, and vice versa, until the loop is terminated.

2. **Intravascular fluid** (IVF), which is mainly **plasma**, is contained in blood vessels and comprises about 25% of the ECF. All blood constituents, such as red and white blood cells, platelets, plasma proteins, various nutrients and electrolytes, are carried in plasma. Plasma makes up about 60% of blood volume (see Ch. 12). The lymphatic system contains the remainder of ECF.

Each compartment has a different ionic composition (see Ch. 2, Table 2.3). In healthy people, the distribution and constitution of the fluid compartments are homeostatically controlled to enable normal physiological function. Homeostatic imbalance leads to disease (Information box 1.3).

There are other small, discrete, collections of ISF, discussed in the systems of the body. Examples include the cerebrospinal fluid, which bathes the brain, the fluid inside the eye (Ch. 8), fluid in joints (Ch. 9) and fluid secreted into the intestines (Ch. 15).

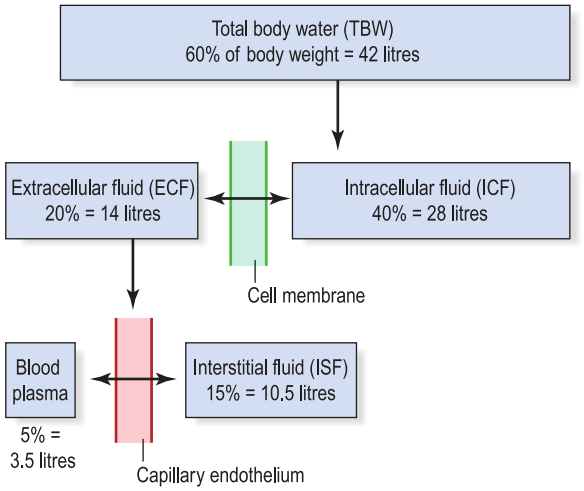


Fig. 1.5 **Distribution of body fluid in compartments.** Approximate values in an adult weighing 70 kg, showing percentage of total body water.

i Information box 1.2 Some examples of feedforward control

Many examples of physiological feedforward control exist. Some examples include:

- Salivation and increased stomach secretion in anticipation of food being eaten, as discussed in Chapter 15 (The alimentary system).
- In response to eating, a fast phase of insulin secretion by the β cells of the pancreas is triggered even before blood glucose levels rise in the portal circulation in anticipation of the requirement for insulin for glucose metabolism. When blood glucose levels rise, a negative feedback mechanism takes over for glucose homeostasis. This is discussed in Chapters 3 (Energy and metabolism) and 10 (Endocrinology: Endocrine control of glucose metabolism).
- When blood glucose is plentiful, as in the fed state (Ch. 3), muscle and fat cells express a glucose transporter (GLUT4, which is regulated by insulin) in anticipation of the need to transport glucose into the cells for storage.
- In gene regulation, the sequence of events that activate transcription factor genes may be seen as a feedforward control, where cell differentiation is controlled by a network of factors, each activating the next in sequence. This is discussed in Chapter 5 (Human genetics: Genes and development).
- In haematology (Ch. 12), the development of mature blood cells from stem cells follows a haemopoietic cell lineage, which is a feedforward mechanism.
- In the motor system, fast movements are controlled by feedforward mechanisms that anticipate what is required, based on learned, pre-existing motor programmes, e.g. playing the piano (Ch. 8, The nervous system: Motor control and pathways). Feedforward failure, where feedforward commands to alternating agonist/antagonist muscles cannot be properly timed, can lead to tremor, speech impairment and other rapidly alternating movements. Negative feedback, based on muscle stretch, is too slow. Training (practice) can accelerate feedforward.
- In neural signalling, the mechanism behind long-term potentiation (LTP) to continually strengthen synapses is still not well understood. It has been suggested that it may be associated with variation in dendritic spine length that can change over minutes or hours. This variation alters electrical resistance and increases synapse strength in a feedforward mechanism. LTP is associated with learning and memory (see Ch. 8).
- During exercise, a neurological feedforward mechanism can be triggered, where blood lactate levels rise in anticipation of the increased need for glucose. Lactate is a precursor for glucose in gluconeogenesis.

i Information box 1.3 Some clinical effects of homeostatic failure in fluid compartments: intravascular volume

Dehydration occurs when there is insufficient fluid (water) in all the compartments; some effects are described in Clinical box 1.5. Water overload leads to fluid accumulation in the tissues, particularly in the interstitial compartment, as described in Clinical box 1.6.

Homeostatic disturbance also occurs when there are abnormalities in the intravascular compartment, principally in blood plasma, known as the **intravascular volume status**:

- Fluid depletion in plasma is known as **hypovolaemia**. Hypovolaemia may be related to overall fluid depletion through severe diarrhoea and vomiting, or from renal or other extrarenal causes. Depending on whether the fluid loss is primarily water or solutes, hypovolaemia may be hyponatraemic, isonatraemic, or hypernatraemic, related to plasma sodium (Na^+) levels (see later). Plasma concentrations of Na^+, K^+, urea and proteins will rise, as will the volume of red cells: **packed cell volume** (PCV), also known as the haematocrit (see Ch. 12). The raised haematocrit (and to a lesser extent plasma protein concentration) increases blood viscosity so that blood flow through the vessels is much slower. Red cells and other blood constituents, such as platelets, tend to aggregate and stick together, increasing the risk of intravascular coagulation. The increased plasma protein concentration, particularly some proteins such as fibrinogen and immunoglobulins, also increases red cell aggregation. When chronic, increased blood viscosity is associated with the development of atheroma and coronary heart and peripheral vascular diseases. When there is acute and severe fluid or blood loss, as in dehydration or haemorrhage, hypovolaemic shock may occur. The heart is no longer effective as a pump to supply essential organs with blood, and multiple organ failure occurs. This is a medical emergency, as the consequences could be fatal.
- **Hypervolaemia** occurs when there is fluid overload and is usually associated with increased body sodium. The excess sodium causes an increase in extracellular fluid volume, which in turn leads to the entry of water into the intravascular compartment. The increase in sodium is related to homeostatic failure in sodium handling, as in congestive heart failure (CHF), renal failure or hepatic failure. Other causes include excessive sodium intake, intravenous infusions of saline or blood and some drugs.

Information box 1.4 Anion gap – a diagnostic aid

The body fluids, serum, plasma or urine, contain positively charged ions (**cations**), such as sodium and potassium, and negatively charged ions (**anions**), such as chloride and bicarbonate. The **anion gap** is a value calculated by subtracting the concentration of **anions** (Cl^- and HCO^-) from that of **cations** (Na^+ and K^+). Note that potassium is often omitted from the calculation, as the concentrations are very low. This value is clinically useful as an aid to diagnosis.

The anion gap may be high, normal or low. A high anion gap suggests metabolic acidosis. A low anion gap is rare and is due to the presence of abnormal cations, as in multiple myeloma, or a low serum albumin level. Some conditions leading to a high anion gap include renal failure (decreased HCO_3^- reabsorption and increased acid excretion), lactic acidosis, diabetic ketoacidosis, toxic effects of some drugs and poisons.

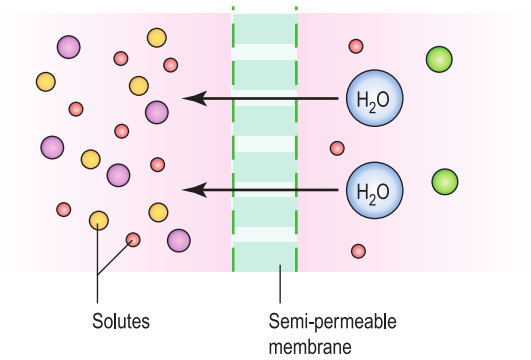


Fig. 1.6 Osmotic movement of water across a membrane.

The body fluid compartments are separated by semi-permeable barriers.

- The intracellular compartment is separated from the extracellular compartments by the cell membranes that allow water to move in and out of cells, but restrict the movement of the main extracellular ion, sodium, so that water can move freely between the compartments but sodium cannot move into cells except in disease conditions.
- In the extracellular compartment, ISF is separated from blood plasma by the endothelium in blood vessels (see Ch. 11, The cardiovascular system). In healthy people, the movement of blood cells and proteins between the interstitial and intravascular compartments is restricted. Water and ions can move freely between the two compartments. There are many ions in blood. The difference in the total number of positive and negative ions is calculated in the anion gap (Information box 1.4).

Movement of fluids between compartments

In healthy people, fluids constantly move between the different compartments of the human body. The driving forces consist of:

- Pressure generated by the pumping of the heart (hydrostatic pressure). Hydrostatic pressure in the circulation refers to the pressure exerted by the volume of blood in a blood vessel (see Ch. 11). Capillary hydrostatic pressure drives fluid out of the capillary bed (also known as filtration). It is highest at the arteriolar end and lowest at the venule end of the capillary bed. Interstitial hydrostatic pressure, determined by ISF volume and tissue compliance, opposes capillary hydrostatic pressure.
- Osmotic pressure, exerted by substances in solution, prevents the flow of water across a semi-permeable membrane, i.e. the cell membrane. Osmosis is the passage of a solvent through a semi-permeable membrane from a solution of higher concentration of solute to one of a lower concentration, and occurs when two solutions of different concentration are separated by a membrane which will selectively allow some solutes to move across. Thus, water osmotically moves from more dilute to more concentrated solutions, and osmotic pressure is pressure exerted by the solutes that must be applied to the solution from outside to prevent osmosis from occurring *(Fig. 1.6)*.

Disturbance of either the efficiency of the heart as a pump or the composition of body fluids as a consequence of disease would lead to manifestations of disease as symptoms and signs. These mechanisms are discussed in almost all the chapters on systems of the body, particularly in Chapters 11 (The cardiovascular system), 13 (The respiratory system), 14 (The renal system) and 15 (The alimentary system).

Properties of forces that drive fluid movement between compartments

The movement of water between compartments is driven by characteristics of the body fluids in the different compartments. These are complicated mechanisms and are perhaps best understood by considering the osmolarity and tonicity of the solutions.

Osmolarity

Osmolarity is a measure of the osmotic pressure exerted by a solution across a perfect semi-permeable membrane which allows free passage of water and completely prevents movement of solute. Osmolarity depends on the number of particles in solution, but not the nature of the particles. If two solutions contain the same number of particles they are **iso-osmotic** (**isosmotic**) with each other. If one solution has a greater osmolarity than another solution, it is **hyperosmotic** compared to the weaker solution. If one solution has a lower osmolarity than another solution then it is **hypo-osmotic** (**hyposmotic**) compared to the stronger solution.

Tonicity (Clinical box 1.2)

Tonicity is a measure of the osmotic pressure that a substance exerts across a semi-permeable membrane compared to blood plasma (as opposed to water for osmolarity), and is almost the same as osmolarity for substances that are impermeable to cell membranes. Tonicity depends on the number of particles in solution and also the nature of the solute. If a cell is suspended in a solution that exerts no osmotic pressure, the solution is **isotonic**; therefore there is no movement of water across the cell membrane. A solution containing more osmotically active particles than the cell is **hypertonic** and will draw water out of the cell, which shrinks. A solution with fewer particles is **hypotonic**, causing water to move into the cell, which swells and eventually bursts.

Clinical box 1.2 The effect of tonicity in disease

Uraemia (abnormally high levels of plasma urea) occurs in renal failure (see Ch. 14). Cell membranes are permeable to urea, so high levels in the extracellular compartment allow urea to enter the cell and the concentration of urea becomes higher in the intracellular compartment. Urea molecules are osmotically active, so the intracellular fluid becomes hypertonic. Water is drawn into the cell, which swells and then bursts, resulting in cell death. This can occur in any cell, but may be lethal in the brain.

Effect of solutes on body fluids

Some ions (solutes) in solution can penetrate the cell membrane while others cannot, so the three body fluid compartments contain different solutes. For example, cell membranes are impermeable to sodium (Na^+) but permeable to potassium (K^+); therefore Na^+ cannot move into cells by simple diffusion, whereas K^+ can diffuse out.

Within cells, the intracellular ions are mainly K^+ (together with phosphate and some large anions, e.g. proteins). Outside the cells, extracellular ions are mainly Na^+. These ions are moved in and out of cells by Na^+ and K^+ transporters, discussed in all the chapters about the systems of the body, and particularly in Chapter 4 (Pharmacology), where the action of drugs depends on their movement in and out of cells.

In the extracellular compartment, the solutes in the ISF differ from those in the IVF. While the ionic content of ISF and blood are the same, proteins in blood are barred from the interstitial compartment by the vascular endothelium. The higher hydrostatic pressure in blood tends to push water out, but the proteins exert a **colloid osmotic pressure** that opposes the blood hydrostatic pressure so that excess fluid does not enter the ISF and blood volume is maintained at a constant level.

Homeostatic control of fluid balance

The physiological processes that maintain life depend on the constant movement of fluids and solutes in and out of cells and between extracellular compartments. Fluids are also lost as urine and sweat and through respiration. In healthy people, fluids in the different body compartments have the same osmolarity (and tonicity). In disease states, when there are changes in either the water or solute content in the fluids, there will be a net movement of fluid between the compartments. The fluid and solutes lost have to be replaced by fluid intake and vice versa, so that the volume and composition of body fluids in each compartment are maintained in status quo. This is achieved through hormonal mechanisms. Human behaviour, such as drinking fluids in response to fluid loss, also helps to maintain fluid balance.

Hormonal control of fluid balance (Clinical box 1.3)

The renal system has a major role in the control of balancing fluid intake with fluid loss, as discussed in Chapter 14 (The renal system). The kidney forms urine in a process that filters fluid, retaining essential nutrients (e.g. protein and glucose) but excreting waste products of metabolism (e.g. urea), and some electrolytes (e.g. Na^+) are lost. These are complex processes that require hydrostatic pressure to drive filtration, and changes in osmolarity (or tonicity) in the different parts of the renal system for recovering water, small molecules, sugars, amino acids and electrolytes from the primary filtrate. An osmoreceptive complex in the hypothalamus that secretes vasopressin (antidiuretic hormone, ADH) controls the mechanisms that balance fluid intake with fluid loss is (see also Ch. 8, The nervous system). Vasopressin is the hormone that regulates urine volume. It is released from the posterior pituitary gland in response to small changes in osmotic pressure. Failure of any part of these mechanisms would lead to a failure in fluid balance.

Behavioural control of fluid balance (Clinical box 1.4)

In healthy people, lost body fluid is replaced through drinking. Daily fluid loss varies depending, in part, on environment and physical activity. For example, fluid loss increases with sweating in hot climates (see previously, Heat-loss mechanisms) and during exercise (Clinical box 1.5). The excretion of urea and salt (Na^+) is affected by dietary intake; a high-protein, high-salt diet would lead to increased urea and Na^+ excretion accompanied by increased urine output, whereas a low-protein, low-salt diet has the opposite effect (Clinical box 1.6). This effect is utilised in some therapeutic diets, as discussed in Chapter 16 (Diet and nutrition). Thirst is the sensation that drives the behaviour of drinking water.

i Clinical box 1.3 Some examples of renal causes for homeostatic failure in fluid balance

Disorders in the control of fluid balance related to renal mechanisms may be central, as in failure of control by the brain, or peripheral, as in kidney disease. Chapter 14 discusses these in detail.

Failure of central control

- Under-secretion or lack of vasopressin may occur due to hypothalamic or posterior pituitary damage. There are a variety of causes, discussed in Chapter 14, including tumours, trauma (head injury) and surgical damage, among others. This results in the daily excretion of a high volume of dilute urine (**polyuria**), accompanied by excessive drinking of water (**polydipsia**), which is characteristic of a condition known as diabetes insipidus. There is also a condition in which the receptors in renal tubules become unresponsive to vasopressin.
- Over-secretion of vasopressin, known as the syndrome of inappropriate ADH secretion, may be caused by low fluid intake or ectopic tumours and others. This can also occur post-operatively.

Failure due to kidney disease

This is more common than failure of central control.

- Impairment to filtration due to abnormal hydrostatic pressure may be caused by failure of the heart as a pump or abnormality of the blood vessel supplying the renal apparatus. Examples include heart failure, hypertension (high blood pressure), damage to renal vasculature from any cause, particularly atheroma/arteriosclerosis, and related conditions such as diabetes mellitus.
- Disease of the renal apparatus responsible for producing urine which could lead to disturbance in renal handing of either water, molecules (e.g. protein, glucose) or electrolytes (e.g. Na^+). This explains, in part, the appearance of glucose and protein in the urine in conditions such as diabetes mellitus. Significant conditions include acute or chronic renal failure from any cause.

i Clinical box 1.4 Some behavioural effects leading to disturbance of fluid balance

Excessive drinking (**polydipsia**) may be driven by thirst or be psychogenic. Polydipsia leads to a dilution of ECF, decreasing the osmolarity. The hypothalamus responds to the reduced osmotic pressure with decreased vasopressin secretion, leading to the passing of high volumes of dilute urine (polyuria). In extreme cases, water intoxication may occur, when the Na^+ concentration in ECF falls to the extent that water enters the cells, causing them to swell. In the brain, this can lead to coma and convulsion and may be lethal. Water overload may also occur through parenteral nutrition (see Ch. 16, Diet and nutrition).

- Polydipsia may be caused by thirst as the side effect of some drugs, e.g. phenothiazines, diuretics, anti-diabetic agents. It is also a symptom associated with diabetes mellitus (see Ch. 3, Energy and metabolism).
- Psychogenic polydipsia is associated with some mental illnesses, such as schizophrenia, or some form of intellectual disability, but may also be psychological, e.g. during a panic attack. This could lead to an inaccurate diagnosis of diabetes insipidus.

i Clinical box 1.5 Some clinical manifestations of fluid imbalance. 1. Dehydration

Urine overproduction, excessive loss of body fluids or inadequate fluid intake leads to **dehydration**. Dehydration occurs when water input is less than water loss. Physiologically, it also involves the loss of electrolytes, mainly Na^+.

- If electrolyte loss is the main problem, then it is known as hyponatraemic or hypotonic dehydration.
- If water is the primary loss, body fluids become hypertonic, i.e. hypernatremia or hypertonic dehydration.
- If water and Na^+ loss is balanced, then dehydration is isonatraemic or isotonic.

These issues are important when considering fluid replacement therapy (see Ch. 16, Diet and nutrition: Parenteral nutrition). Water moves from the intravascular to the extravascular compartment in hypotonic dehydration, which in turn affects intracellular osmolarity and, particularly if brain cells are affected, may lead to seizures. In hypertonic dehydration, the reverse occurs and may result in osmotic cerebral oedema if rehydration is too rapid.

To compensate for the reduced plasma volume, the heart and respiratory rates increase, leading to hypotension (low blood pressure). Further complications of hypotension include reduced renal perfusion (low hydrostatic pressure), which can cause renal failure. Body temperature rises due to the shutdown of heat-loss mechanisms that involve water loss.

- Symptoms of mild dehydration include dry mouth and thirst. Signs include decreased skin turgor (skin stays puckered if pinched gently) and low urine output. Infants may have a sunken anterior fontanel.
- Moderate to severe dehydration leads to anuria (reduced or no urine output). Symptoms of lethargy, delirium, seizures and orthostatic hypotension (fainting) may also occur. Death may ensue with increasing severity.

Dehydration may be caused by excessive fluid loss or inadequate intake.

Excess loss of fluids

- Increased urine output, as in diabetes mellitus (see Ch. 3, Energy and metabolism) and diabetes insipidus.
- Loss of other body fluids, as in severe vomiting and diarrhoea due to gastrointestinal disease (see Ch. 15, The alimentary system) or infection, such as cholera, and following surgical bowel resection (e.g. colostomy).
- Loss of plasma, as in haemorrhage or burns, and following some surgical procedures.

Inadequate fluid intake occurs

- In malnutrition and fasting (see Ch. 16, Diet and nutrition).
- In the elderly as a behavioural consequence and in infants due to inadequate feeding.

i Clinical box 1.6 Some clinical manifestations of fluid imbalance. 2. Oedema

Underproduction of urine leads to water overload, resulting in the accumulation of water in the tissues (**oedema**). Oedema may be a symptom or a sign and is clinically important. Patients may complain of facial puffiness or shoes feeling 'tight'. Clinical examination may demonstrate 'pitting' (a dent in the skin on gentle pressure that persists after pressure is released) over the ankles, shins or sacrum. There are many causes for oedema, including:

- Failure of renal mechanisms or due to abnormalities in hydrostatic pressure as a consequence of heart failure (see Ch. 11, The cardiovascular system)
- Failure of renal clearance of water and electrolytes (see Ch. 14, The renal system) or
- Failure of the lungs to clear CO_2 (see Ch. 13, The respiratory system).

To properly formulate diagnostic and therapeutic decisions, it is important to distinguish which physiological mechanisms are disordered. Oedema is commonly treated with diuretics (see Ch. 4, Pharmacology).

Thirst

Chapters 14 (The renal system) and 8 (The nervous system) describe the osmoreceptors in the hypothalamus. These also cause the sensation of thirst in response to stimulation by an increase in plasma osmolarity, and, less so, by a fall in plasma volume. The behavioural response to thirst is drinking. The presence of fluid in the mouth and pharynx abolishes thirst before the restoration of normal osmolarity and plasma volume.

Acid–base balance: homeostatic control of hydrogen ions (Clinical box 1.7)

Acid–base balance is the regulation of hydrogen ion (H^+) concentration in body fluids. The concentration of free H^+, i.e. not bound to other molecules such as proteins, determines the acidity of body fluids. This is measured as the partial pressure of free H^+ in solution in body fluids (pH). The partial pressure of a gas in solution is the pressure that the gas would exert if it were the sole occupant of that volume of fluid. Thus, the total pressure of a mixture of gases in solution is the sum of the partial pressures of each gas.

The pH of body fluids determines the rate of activity of the thousands of enzymes that control physiological processes. Enzymes are biological molecules that catalyse (accelerate) chemical reactions in cells at a rate sufficient to sustain life. Enzymes in general are discussed in detail in Chapter 2 (Biochemistry and cell biology) and also in almost all the systems. The various mechanisms for maintaining acid–base balance are discussed in detail under various headings in Chapters 3 (Energy and metabolism), 12 (Haematology), 13 (The respiratory system) and 14 (The renal system). Disturbances of acid–base balance could lead to life-threatening conditions.

Partial pressure of hydrogen ions

The pH of a neutral solution such as water is 7. Increased H^+ concentration lowers the pH, rendering the solution acid, whereas reduced H^+ concentration raises the pH, making the solution alkaline. In healthy people, the H^+ concentration in body fluids is regulated within a narrow physiological range (see below). Homeostatic failure, as the result of disease, leads to deviation of pH outside the

i Clinical box 1.7 Lactic acidosis

Cellular glucose metabolism is a two-stage process in which glucose is finally broken down to water (H_2O) and carbon dioxide (CO_2) (see Ch. 3). Put simply, glycolysis first breaks glucose down to pyruvate, which is then oxidised in mitochondria to CO_2 and H_2O. The second stage requires oxygen. If there is inadequate oxygen, either due to lack of tissue oxygen or a metabolic abnormality, the pyruvate is converted to lactate, which is released into the blood stream and may accumulate over time.

Lactic acidosis occurs when there is an accumulation of lactic acid, and blood and tissue pH is very low (acidosis). This can occur during conditions of extreme low blood pressure as in cardiogenic or septic shock, when there is a reflex vasoconstriction to abdominal organs, skin and other peripheral structures, leading to lack of oxygen. The acidosis compromises cardiac function, causing further vasoconstriction, oxygen lack and lactic acid production. This is potentially lethal if not corrected.

Some other causes of lactic acidosis include extreme, severe exercise, poisoning (e.g. ethylene glycol or anti-freeze), decreased lactate metabolism by the liver and some drugs (such as the accumulation of metformin in diabetics with chronic renal disease).

Table 1.1 pH values for some fluids

Solution	pH
Hydrochloric acid (0.1 moles/L)	1.0
Gastric juice	1.0–2.5
Lemon juice	2.1
Tomato juice	4.1
Urine (average)	6.0
Saliva	6.8
Milk	6.9
Pure water (25°C)	7.0
Blood (average)	7.4
Sea water	7.9–8.3
Ammonia (NH_3, 0.1 moles/L)	11.1

physiological range and gives rise to specific signs and symptoms. These mechanisms are discussed in detail in Chapters 13 (The respiratory system) and 14 (The renal system). Extreme deviations outside of this range are incompatible with life.

Physiological range of pH

The normal pH of blood is about pH 7.4. Arterial blood is slightly less acidic (at pH 7.45) than venous blood (pH 7.35). **Acidosis** occurs when blood pH falls below pH 7.35 and **alkalosis** is present at a pH value of over 7.45. Death would ensue if blood pH falls below pH 6.8 or rises above pH 8.0 for a significant period.

Some physiological process work best at pH values that are different from blood pH. Table 1.1 shows the pH values for some fluids. For example, muscle activity produces lactic acid, bringing the pH in myocytes to between pH 6.8 and pH 7.0, and sometimes below pH 6.4 (see Ch. 9).

Effect of pH on physiological processes

Many intracellular chemical reactions need to perform within a narrow range of pH, known as the **optimal pH**, at which they occur faster. Generally, optimal pH is similar to blood pH. Depending on the normal pH of the compartment in which the reaction takes place, however, optimal pH can vary.

Chapter 8 (The nervous system) discusses the effects of changes in pH on the brain and peripheral nerves. A reduction in pH (acidosis) causes a reduction in excitability, especially in the brain, which can lead to confusion and, in extreme cases, coma and death. Conversely, an increase in pH (alkalosis) will produce unwanted nervous activity in the peripheral and central nervous systems. Nerves become hypersensitive and transmit signals in the absence of normal stimuli, producing symptoms such as numbness and tingling, caused by overactivity of sensory nerves. Overactivity in the nerves that excite the muscles can cause muscle spasms, which in severe cases can lead to paralysis of the muscles required for breathing. Excessive nervous activity in the brain may lead to convulsions.

Chapter 13 (The respiratory system) discussed the mechanisms for the respiratory control of blood pH, and the effects of pH changes on respiration: known as respiratory acidosis or alkalosis. Chapter 14 (The renal system) discusses the renal control of acid–base balance and the effects of homeostatic imbalance: known as metabolic acidosis or alkalosis.

Table 1.2 Buffer capacity of the main buffer systems in the blood

Buffer system	Capacity (mmol H^+/L)
Bicarbonate/carbon dioxide	18
Protein	1.7
Haemoglobin	8
Phosphate	0.3
Total	28

Sources of acid and alkali

Acids are a by-product of metabolism. Carbon dioxide (CO_2), essential for the adjustment of pH, is a by-product of energy metabolism. Metabolism of dietary fats and proteins produces acid and exercising muscles release lactic acid. Chapters 3 (Energy and metabolism) and 16 (Diet and nutrition) discuss the breakdown of ingested food to provide energy for metabolic processes. Chapter 15 (The alimentary system) outlines the absorption and digestion of food, water and minerals. Chapter 14 (The renal system) describes how the kidneys produce acid.

Alkalis (bases) are not by-products of metabolic processes, but are ingested in food (vegetables). The excess meat (protein and fat) in Western diets may lead to an increased production of acid. In health, this excess acid is dealt with by homeostatic mechanisms to maintain acid–base balance:

- Chemical buffers act to limit the free H^+ concentration of body fluids in the short term.
- Renal and respiratory homeostatic mechanisms increase the excretion of CO_2 and H^+.

Buffer systems

A **buffer system** limits changes in free hydrogen ions [H^+] in body fluids. The H^+ (acid) combines with weak acids (bases) in free solution, thus limiting any large changes in [H^+]. These systems consist of buffer pairs, which are only partially dissociated at normal pH ranges found in the body, so both acid and base are present. If hydrogen ions are added to the system, they can bind to the base to prevent a fall in pH. If hydrogen ions are removed, then more acid dissociates to form H^+ so that the pH is unchanged. The amount of change that can be prevented depends on the amount of the buffer pair present. Three main buffer systems maintain a relatively constant pH in body fluids:

- Bicarbonate/carbon dioxide
- Proteins, particularly haemoglobin
- Phosphate.

The amount of H^+ that each buffer system is able to buffer is known as the **buffer capacity** (Table 1.2). Changes in pH are seen when the buffer capacity is exceeded.

Blood contains all three buffer systems. In ISF, buffering is done by the bicarbonate and phosphate systems because there is very little protein in ISF. The intracellular compartments also use all three systems, being rich in protein and phosphate, although the bicarbonate concentration is lower than in ECF (see Ch. 2, Table 2.3).

Bicarbonate/carbon dioxide

In blood, the bicarbonate/carbon dioxide buffer system performs an important role in acid–base balance. It is the main buffering

system in ECFs. It buffers H^+ from mechanisms that do not involve either bicarbonate (HCO_3^-) or CO_2, e.g. falls in H^+ in response to lactate production in exercise or the use of fat as fuel substrate in diabetes. These processes are discussed in Chapter 3 (Energy and metabolism).

In aqueous solutions, CO_2 forms carbonic acid (H_2CO_3), which dissociates into H^+ (acid) and HCO_3^- (base):

$$CO_2 + H_2O \rightleftharpoons H_2CO_3 \rightleftharpoons H^+ + HCO_3^-$$

The rate at which acids dissociate is known as the dissociation constant (*K*), referred to as the p*K*. A buffer is most effective when its p*K* is close to the desired pH.

Protein buffers

The acidic and basic protein side-chains accept or donate H^+ to limit changes in pH. Chapter 2 (Biochemistry and cell biology) discusses the structure of proteins. The carboxyl and amino groups found at the ends of each protein chain also accept H^+. Proteins, e.g. albumin, in plasma buffer significant amounts of H^+.

Haemoglobin buffer system

Haemoglobin, a protein in red cells, is discussed in detail in Chapter 12 (Haematology). Its main function is oxygen transport in the circulation. Carbon dioxide, a waste product of metabolism, has a complex transport system, but haemoglobin also has a role in CO_2 transport. When CO_2 from tissues is taken up by haemoglobin, it is converted by the enzyme carbonic anhydrase in red cells to carbonic acid ($H_2CO_3^-$) that takes part in the bicarbonate/carbon dioxide buffer system. Haemoglobin also absorbs H^+ when H^+ combines with deoxyhaemoglobin (formed by the release of O_2 from oxyhaemoglobin), which has a higher affinity for H^+ ions. The movement of HCO_3^- into plasma is counterbalanced by the chloride shift of chloride ions (Cl^-) into red cells ***(Fig. 1.7)***.

Bicarbonate diffuses back into the plasma to be transported by the venous circulation to the lungs, where haemoglobin takes up oxygen to form oxyhaemoglobin. Oxyhaemoglobin has a low affinity for CO_2, which is released and exhaled.

About 10% of CO_2 in haemoglobin is carried as **carbaminohaemoglobin**, where the combination of CO_2 to the haemoglobin molecule results in the release of further H^+ ions in addition to those generated from carbonic acid. Not all the H^+ ions are taken up by haemoglobin, so that venous blood is slightly more acid than arterial blood (see Ch. 12).

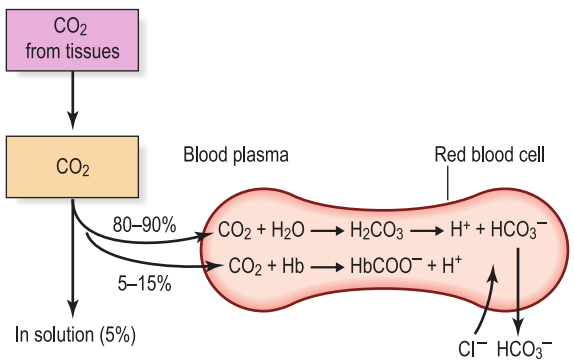


Fig. 1.7 Carriage of carbon dioxide in blood.

Chapter 13 (The respiratory system) discusses disturbances in acid–base balance when lung function is compromised (metabolic acidosis and alkalosis).

Phosphate buffer system

The phosphate buffer system is intracellular. It consists of dihydrogen phosphate, an acid, which dissociates to hydrogen phosphate and H^+.

$$H_2PO_4^- \rightleftharpoons HPO_4^{2-} + H^+$$

Phosphate buffering is an important mechanism for H^+ ion excretion by the kidneys. This is discussed in detail in Chapter 14 (The renal system). Whilst buffering systems control H^+ concentration, the excess H^+ ions have to be excreted. This is a renal function, where phosphate and ammonia are excreted and act as buffers for the H^+ secreted into urine.

Control of acid–base balance

The control of acid–base homeostasis is performed by the lungs and kidneys.

Respiratory control of pH (Clinical box 1.8)

The respiratory control of pH is dependent on the bicarbonate/carbon dioxide buffer system, which has the highest capacity of all the buffering systems. Chapter 13 (The respiratory system) discusses the mechanisms in detail.

In essence, the removal of CO_2 by the lungs restricts the amount of circulating free H^+ ions and thus the pH. If the pH falls, i.e. acidosis occurs, respiration increases either in the rate or depth of breathing to remove more CO_2. The removal of CO_2 allows the pH to rise. Conversely, if the pH rises (alkalosis), a compensatory decrease in respiration takes place, CO_2 is retained and pH falls. Breathing is therefore partly controlled by blood pH. Chronic lung diseases, with or without impairment of the respiratory centre, could lead to **respiratory acidosis or alkalosis** (Clinical box 1.8 and also see Ch. 13).

i Clinical box 1.8 Disturbance of acid–base balance related to lung ventilation: respiratory acidosis or alkalosis

Gas exchange in the lung relies on two mechanisms:

- The bellows, consisting of respiratory muscles, and the chest wall, which pump the gases in and out of the lungs; and
- The lung structures, i.e. the airways, alveoli and blood vessels, which are responsible for gas exchange.

If either or both are damaged, the take up of oxygen and the removal of carbon dioxide is affected. The CO_2 content in blood is the main determinant of acidity (pH) in respiratory acidosis or alkalosis.

Respiratory acidosis is caused by CO_2 retention, when the ability of the lungs to remove CO_2 is reduced. Failure of any part of the gas exchange mechanisms could do this. Some examples include:

- Disease of the airways: COPD, asthma, bronchial tumours
- Lung disease: emphysema
- Impairment of bellows: neuromuscular disease, chest wall and/or spinal deformities, interstitial fibrosis, disease of central respiratory control mechanisms (depression of respiratory centre, narcotics overdose, cardiopulmonary arrest).

Respiratory alkalosis is caused by abnormal, excessive removal of CO_2 through hyperventilation, which may be caused by anxiety or hysteria, brain injury or stroke, excessive mechanical ventilation and overdose of some drugs.

Renal control of pH (Clinical box 1.9)

The kidneys contribute to acid–base homeostasis by excreting H^+ and HCO_3^- in the urine. These are complex mechanisms and are discussed in detail in Chapter 14 (The renal system). The renal secretion of H^+ and reabsorption of bicarbonate are the functions that control blood pH.

Renal H^+ excretion

Carbon dioxide diffuses into the kidney tubule cells, where carbonic anhydrase converts CO_2 and H_2O to H_2CO_3. H_2CO_3 dissociates to H^+ and HCO_3^-; HCO_3^- is transported to blood and H^+ is excreted in urine ***(Fig. 1.8)***. Nearly all the H^+ found in urine is secreted by the kidneys. To prevent H^+ diffusing back into the tubules, urine is kept at a pH not less than 4.5.

Renal bicarbonate reabsorption

Bicarbonate is indirectly reabsorbed from the renal tubules because the cells lining the renal tubules are impermeable to HCO_3^-. The HCO_3^- in tubules reacts with secreted H^+ to form H_2CO_3 ***(Fig. 1.9)***. On

Clinical box 1.9 Disturbance of acid–base balance related to renal mechanisms: metabolic acidosis and alkalosis

Disturbance in acid–base balance from renal causes, known as metabolic acidosis or alkalosis, primarily affects the bicarbonate component of a system (see also Ch. 14).

Some causes of metabolic acidosis include:

- Overproduction of acid, as in lactic acidosis (see Clinical box 1.7)
- Chronic renal failure when renal H^+ excretion is reduced
- Diabetic ketoacidosis, where alternative metabolic fuels other than glucose are used
- Bicarbonate loss from gastrointestinal disease
- Some drugs.

Some causes of metabolic alkalosis, although uncommon, include:

- Excessive acid loss, as in severe, prolonged vomiting of acid gastric contents
- Renal disease
- Some drugs, such as diuretics.

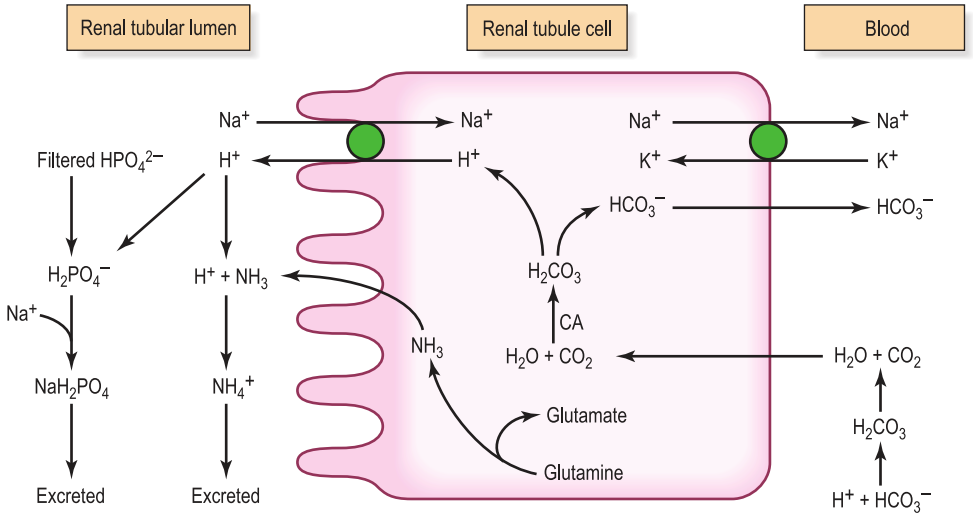


Fig. 1.8 Renal excretion of H^+ ions. H^+ combines with phosphate or ammonia before being excreted. CA, carbonic anhydrase.

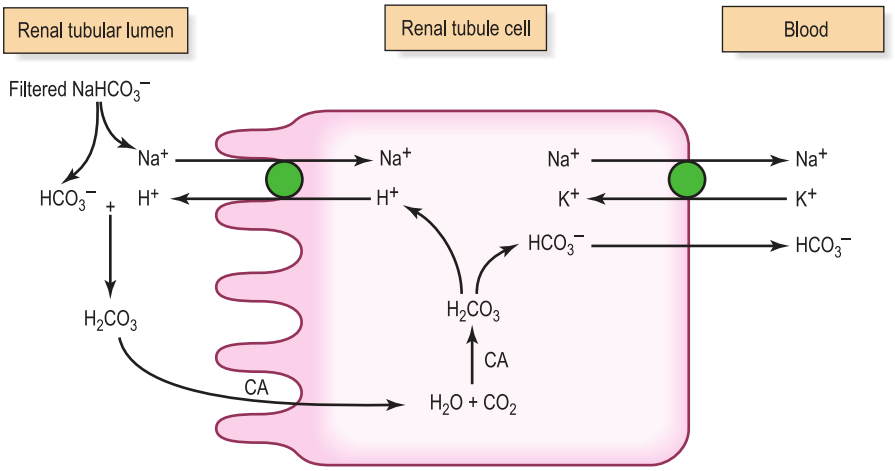


Fig. 1.9 Renal reabsorption of bicarbonate. One molecule of HCO_3^- is transferred from the tubular lumen to blood. CA, carbonic anhydrase.

i Clinical box 1.10 An example of acid–base disturbance

Table 1.3 was constructed using data from patients with different respiratory and metabolic acid–base disorders. This shows the range of changes in H^+ and HCO_3^- levels and pH values (***Fig. 1.10***). The table is used identify the type of acid–base disorder. Serial measurements for individual patients in graphic form are used to monitor the progress and effectiveness of treatment.

The arterial blood gases of an elderly man with emphysema showed the following:

- pH = 7.30
- pCO_2 = 50 mmHg
- Standard $[HCO_3^-]$ = 32 mM

The pH value is below 7.35, so he has acidosis. The pCO_2 is greater than 45 mmHg, so the primary cause of the acidosis is respiratory. The standard $[HCO_3^-]$ is greater than 28 mM, so he has respiratory acidosis with renal compensation.

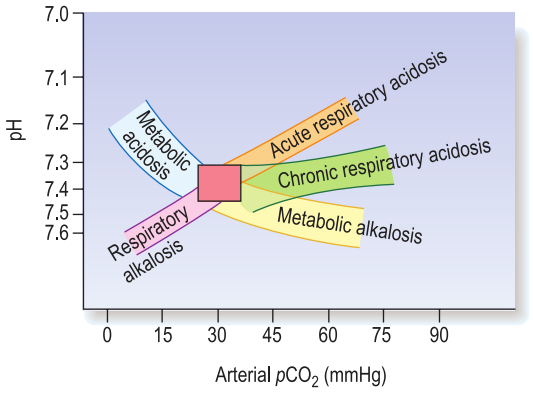


Fig. 1.10 Acid base graphs showing H^+ and pCO_2 ranges for acid–base disturbance. Normal range in red.

Table 1.3 Investigation of acid–base disturbance

1. Measure arterial pH (normal range pH 7.35–7.45)			
pH < 7.35: acidosis		pH >7.45: alkalosis	
2. Measure arterial pCO_2 and $[HCO_3^-]$ (standard bicarbonate). Normal values: pCO_2 35–15 mmHg (4.8–6.1 kPa), $[HCO_3^-]$ 22–28 mM. The following shows the values for the different types of acid–base disturbance			
pCO_2 > 45 mmHg (6.1 kPa): respiratory acidosis	$[HCO_3^-]$ < 22 mM: metabolic acidosis	pCO_2 < 35 mmHg (4.8 kPa): respiratory alkalosis	$[HCO_3^-]$ > 28 mM: metabolic alkalosis
3. Interpret the two measurements together		**3. Interpret the two measurements together**	
$[HCO_3^-]$ > 28 mM: respiratory acidosis with renal compensation	pCO_2 < 35 mmHg (4.8 kPa): metabolic acidosis with respiratory compensation	$[HCO_3^-]$ < 22 mM: respiratory alkalosis with renal compensation	pCO_2 > 45 mmHg (6.1 kPa): metabolic acidosis with respiratory compensation

the surface of the tubule cells, carbonic anhydrase converts H_2CO_3 to H_2O and CO_2. The CO_2 diffuses freely into the tubule cells where intracellular carbonic anhydrase catalyses the reverse reaction to produce H_2CO_3. This then dissociates into HCO_3^- and H^+. The H^+ is secreted into the urine and the HCO_3^- diffuses into the blood. The net result is the transfer of one molecule of HCO_3^- from the urine to the blood.

When a mechanism fails, for example in respiratory acidosis, a compensatory renal mechanism may operate to retain bicarbonate and $[H^+]$ could return to normal. In respiratory alkalosis, when CO_2 levels are persistently low a compensatory metabolic acidosis may occur, although the response is usually slight (Clinical box 1.10 and Table 1.3).

Respiratory compensation of metabolic acidosis also occurs, when respiration increases to 'blow off' CO_2 and allows $[H^+]$ to rise in a respiratory alkalosis. There is usually a delay in this respiratory compensatory mechanism. Similarly, respiratory compensation for metabolic alkalosis, although slight, can occur.

2 Biochemistry and cell biology

Marek H. Dominiczak

INTRODUCTION

Cell biology tells us about the structure and functions of the cell and its organelles, whereas biochemistry addresses the chemical basis of the composition of the human body, its structure and its functions, and also the preservation and continuity of the structure and function through generations. The body is an open system and, thus, there are the interfaces with surroundings that are both sensory and metabolic, the latter through nutrition and excretion of metabolic products. The function of the organism may become disrupted during both the lack of nutrients (starvation, malnutrition) and their excessive intake (the diseases of excess: obesity, cardiovascular disease and diabetes).

Central to metabolism and homeostasis is the energy flow: metabolism includes reactions that are energy requiring (endergonic) and the ones that yield energy (exergonic). Some chains of reactions (pathways) accumulate energy in a range of biosynthetic products (anabolic pathways), and others release energy (catabolic pathways). Reactions take place in an aqueous environment, and electrolytes and ions are both bystanders and participants in chemical reactions. Most reactions happen at physiological temperature within a narrow range of pH, thanks to the action of biocatalysts (enzymes).

There has been an enormous acceleration in our understanding of the chemical aspects of body structure and function. Perhaps the most significant development in biochemistry in the last few years was the expansion of knowledge concerning cellular signalling systems and their links to the control of gene expression ***(Fig. 2.1)***. It is so important because the aim of a large number of medical therapies is to control these processes when the normal regulatory mechanisms fail.

In this chapter, we will first consider the fundamentals of the atomic structure, chemical bonds and chemical reactions, highlighting the energy flow in biological systems. We will go on to discuss the chemical composition of the human body and the most important classes of biological compounds: carbohydrates, fats, proteins, nucleic acids and a range of 'hybrid' molecules, such as glycolipids or proteoglycans.

We will then discuss the cell and its organelles, highlighting the function of cell structures that enable cells to communicate, transport nutrients and interact with each other.

PRINCIPLES OF MOLECULAR INTERACTIONS

Atoms

All atoms have a nucleus surrounded by shells of electrons. Each of these shells is characterised by a different energy level. Subshells (orbitals) exist within each shell. Atomic orbitals are described by quantum numbers. The principal quantum number (X) corresponds to the energy level, and the angular quantum number l (type, denoted by a small letter) describes the shape of the subshell. The superscript (y) in the convention Xtype ^{y} describes the number of electrons in an orbital.

The orbitals are designated as 1s, 2s, 2p, 3s, 3p and 4s. The 1s orbital is closest to the nucleus. Each orbital can be occupied by a maximum of two electrons, each of them having a different spin. The order in which the atomic orbitals are filled goes from the lowest energy level (closest to the nucleus) to the higher levels. The order (from first to last) is 1s, 2s, 2p, 3s, 3p, 4s, 3d, 4p, 5s, 4d, 5p, 6s, 4f, 5d, 6p, 7s, 5f, 6d and 7p. When orbitals of the same energy are available, the electrons fill them singly first (this is known as the Hund rule).

Each of the electron shells contains a defined number of electrons in their orbitals:

- The first (innermost) shell can have one orbital (1s) and a maximum of 2 electrons.
- The second shell 2s2p has 4 orbitals and a maximum of 8 electrons.
- The third shell can have 3 orbitals (3s3p3d), which together makes 9 orbitals and up to 18 electrons.
- The fourth shell has 16 orbitals and up to 32 electrons.

Electrons always occupy orbitals of the lower energy first. Note the anomaly that occurs between the third and the fourth shell. The 4s orbital is filled before 3d (contravening the general rule, it has a lower than 3d energy level).

A fully occupied shell makes an atom chemically inert (examples are the noble gases such as helium and neon). Atoms that have incompletely filled outer shells can react until their shells become fully occupied. The outermost orbital of the atom contains the so-called valence electrons that participate in forming chemical bonds. The configuration of eight electrons (octet) in the outer shell is the most stable one (*Fig. 2.2*).

Atoms that possess an unpaired electron that is not shared with other atoms are known as **free radicals** and are highly reactive.

The nucleus contains **protons** and **neutrons** (the hydrogen atom has only a single proton as a nucleus). The number of protons is the **atomic number** of an element. The sum of the protons and neutrons is the **atomic mass**. Different **isotopes** of a given element differ with respect to the atomic mass (Information box 2.1).

i Information box 2.1 The carbon atom

Carbon is the element present in all organic molecules. The carbon atom is assigned an atomic number of 6 because it has 6 protons. However, it can have 6, 7 or 8 neutrons, forming different isotopes with different atomic masses. Carbon isotopes show no differences in chemical reactivity (Table 2.1).

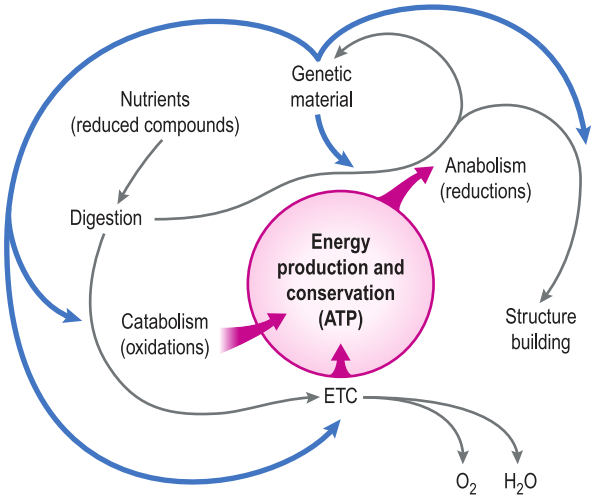


Fig. 2.1 Metabolism and its genetic control. Supply of nutrients (reduced compounds) and their oxidation yields energy that is conserved in the form of adenosine triphosphate (ATP). ATP is subsequently used in energy-requiring biosyntheses (predominantly reductions again). The work of the entire system is controlled by information contained in the genetic code, which also utilizes the metabolic machinery to preserve its continuity. ETC, electron transport chain.

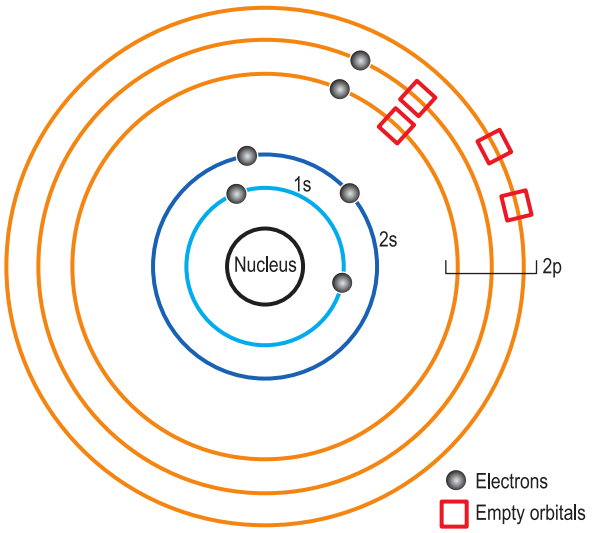


Fig. 2.2 The carbon atom. The carbon atom contains 2 shells and 5 subshells (orbitals). Note that the 2 orbitals in shell 2 are filled by 1 electron each. Sharing of 4 electrons in shell 2 with another atom would form an octet – a stable configuration.

Table 2.1 Common functional groups and classes of compounds

Group	Formula	Class
Hydroxyl	R-OH	Alcohols
Aldehyde	R-COH	Aldehydes
Ketone	R-COR′	Ketones
Carboxy	R-COOH	Carboxylic acids
Ester	R-COO-R′	Esters
Amino	R-NH_2	Amines
Imino	R-NH	Imines
Sulphydryl	R-SH	Thiols

Ions

In an atom, there is normally a balance between the positive charge of the protons and the negative charge of the electrons, rendering the atom electrically neutral. Consequently, when it gains or loses electrons, it acquires an electrical charge: such an atom is called **an ion**. The charge can also be associated with groups of atoms, forming ionised (functional) groups. Thus,

- Anions are negatively charged ions and are generated by the gain of electrons.
- Cations are positively charged ions and are generated by the loss of electrons.
- Metals tend to form cations, and non-metals form anions. Formation of ionic bonds between atoms is one of the principal mechanisms of chemical reactions.

Acids and bases

The concept of an acid and a base is associated with the movement of protons and electrons in aqueous solutions. According to the **Brönsted-Lowry** definition, an acid is a molecule that can donate protons and a base is a molecule that accepts protons. Therefore, an acid in a sense 'contains' a base: when an acid loses a proton, the remaining species (now negatively charged) is its conjugate base. Another definition of an acid, the Lewis definition, defines it as a molecule that accepts a pair of electrons and a base as a molecule that donates a pair of electrons (Information box 2.2)

Chemical bonds

Chemical bonds determine how molecules join together. Atoms can form chemical bonds with other atoms of the same or different kind (*Fig. 2.3*). Bonds differ in their strength and stability, and are also determine the spatial conformation of molecules. The bonds most relevant to biomolecules are the following:

- Ionic bonds
- Covalent bonds
- Hydrogen bonds.

Ionic bonds

Ionic bonds form when ions are attracted to each other by their opposite electrical charge. An electron(s) from one atom move(s) closer to the nucleus of another, forming a molecule. Importantly, such molecules dissociate into their component ions in an aqueous solution. For instance, **sodium chloride** (table salt) is formed when sodium (Na) and chlorine (Cl) atoms attract each other. In this reaction, Na loses an electron, becoming a cation, Na^+, whereas Cl acquires the electron, forming an anion, Cl^-. A strong ionic bond forms sodium chloride (NaCl). When the water is removed, salt crystals consisting of Na^+ and Cl^- held together by ionic bonds are formed (see *Fig. 2.3A*).

Covalent bonds

The principle behind covalent bonds is **electron sharing**. A pair of electrons is shared between two atoms, making both atomic shells

Information box 2.2 Strength of acids

The strength of an acid is the ease with which it donates proton or accepts electrons. Strong acids dissociate completely, whereas weak acids dissociate to a limited extent. This tendency is described by the acid's dissociation constant (K_a), which is a ratio between its undissociated and dissociated forms:

$$K_a = [H^+] + [A^-]/[HA]$$

The derivative of the dissociation constant is its negative logarithm, the pK_a. The lower the pK_a is, the more active a molecule is as a proton donor (a stronger acid). Examples of strong acids are inorganic acids such as hydrochloric or sulphuric acid. The weak acids are carbonic acid (an important blood buffer) and the carboxylic acids.

Conversely, increasing pK_a means an increase in the alkalinity- of the conjugate base.

The acidity or alkalinity of a solution also relates to the concentration of hydrogen ions (H^+) measured as its negative logarithm, the pH. The neutral pH of pure water is close to 7.0. Solutions with a pH of less than 7.0 are acidic, and solutions with a pH of greater than 7.0 are alkaline. The normal range of human blood pH is 7.35–7.45. The maintenance of stable pH of the body fluids is necessary for survival (Ch. 1).

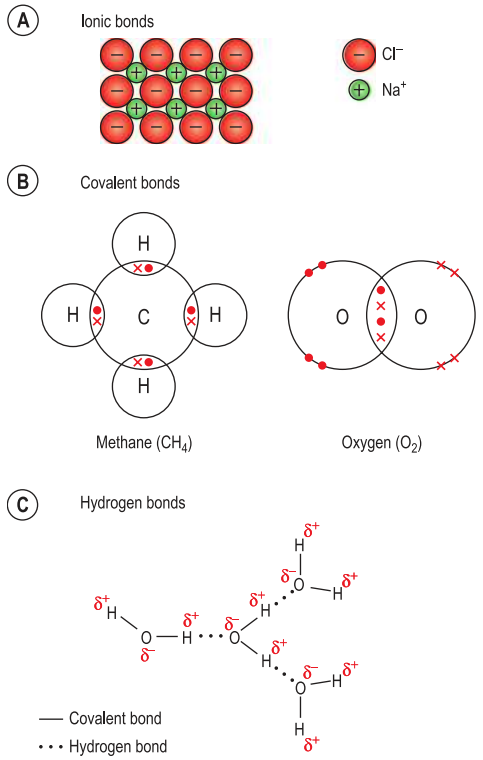


Fig. 2.3 Different types of chemical bond. (A) Ionic bonds. (B) covalent bonds, (C) hydrogen bonds. In (C), δ^+ and δ^- denote partial positive and partial negative charges, respectively.

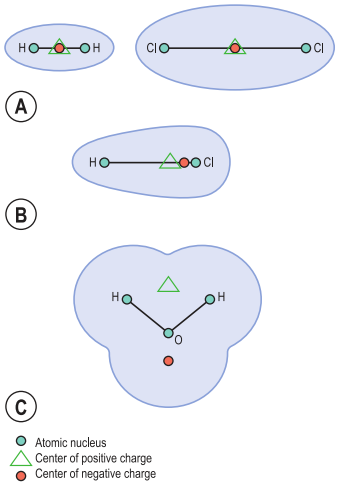


Fig. 2.4 Polar and non-polar covalent bonds. (A) Covalent bonds formed between atoms of the same kind are usually non-polar, and the shared electrons are equidistant from both atomic nuclei. (B) A polar covalent bond may form between atoms of different elements, where the electrons are closer to the nucleus of one element than to that of the other. Such a molecule acquires a partial electric charge. (C) Water is a dipolar molecule. Oxygen has a higher electronegativity than hydrogen and results in a charged molecule. Dipole–dipole interactions are the basis of attraction between many molecules known as hydrogen bonding (see *Fig. 2.3C*).

complete. This may occur between atoms of the same or different elements. Examples include oxygen (O_2), where two oxygen atoms form a double bond (O=O); hydrogen atoms sharing electrons with a single bond (H–H); and carbon and hydrogen atoms sharing electrons to form methane (CH_4) (see *Fig. 2.3B*).

Polar covalent bonds

A covalent bond is electrically neutral when the electrons remain equidistant from the two participating atoms *(Fig. 2.4A)*. However, when the nuclei of two atoms differ in their positive charge, this distance becomes unequal, creating partial electric charges across the covalent bond. Such a bond becomes a **polar covalent bond** (see *Fig. 2.4B*) and the resulting molecule becomes a **dipole**. The water molecule (H_2O) is an example of a dipole (see *Fig. 2.4C*). The strongly negative oxygen atom attracts electrons away from the two hydrogen atoms, which become positive. Because of this, they keep at a distance from each other. The water dipole can form hydrogen bonds with non-water molecules: this is the principle behind the solubilisation of substances by water. The hydrogen ions can also associate with other water molecules. This results in **water dissociation** into the hydronium ion (H_3O^+) and the hydroxide ion (OH^-):

$$H_2O + H_2O \rightleftharpoons H_3O^+ + OH^-$$

We normally simplify this in our notation and show water dissociation as

$$H_2O \rightleftharpoons H^+ + OH^-$$

Hydrogen bonds

The hydrogen atom has a single proton as the nucleus, and only one electron shell, occupied by a single electron. When this electron is lost, a cation (H^+) forms. Such a cation can attract the negative pole of a dipolar molecule, forming a hydrogen bond (see *Fig. 2.3C*). These bonds are weaker than covalent and ionic bonds, and are easily disrupted by pH and temperature changes. They are important in stabilising the spatial structures of proteins and nucleic acids. Many large molecules contain numerous hydrogen bonds.

Non-polar molecular interactions

Non-polar molecules are water-insoluble (**hydrophobic**). Such molecules tend to aggregate in polar solvents; this is known as **hydrophobic interaction**. An example is the behaviour of lipid molecules in the plasma membrane. Another type of weak interaction, the **Van der Waals force**, makes molecules align in an energetically optimal conformation. Van der Waals forces are effective at a relatively long range (up to 50 nm) and are easily reversible. For example, they contribute to the binding of substrates to enzyme molecules and to the binding of antibodies to antigens.

Organic compounds

An organic compound is a compound containing carbon atoms linked by covalent bonds. This definition usually excludes some small molecules, such as carbon dioxide. Organic compounds may also contain oxygen, nitrogen and sulphur.

The carbon atoms commonly bond with each other and with hydrogen, forming **hydrocarbon chains**. The hydrogen atoms in hydrocarbons may be replaced by other atoms or functional groups. In addition, carbon atoms can share more than one electron with another atom, forming double or triple covalent bonds. Table 2.1 shows examples of the chemical groups that occur in different classes of organic compounds.

The carbon atoms may share four electrons with other atoms, including carbon (see *Fig 2.2*). For example, in methane (CH_4), the carbon atom links covalently with four hydrogen atoms. Importantly, the carbon can form long chains (e.g. in some fatty acids), which can also branch, or form rings (e.g. the steroids), containing either carbon only or carbon linked to other atoms, such as nitrogen.

Spatial arrangement of organic molecules

Within an organic molecule, a carbon atom that shares its four available electrons with four different atoms (or groups) is known as the **chirality centre** or **stereocentre**. The chiral compound cannot be superimposed on its mirror image. Such a molecule can exist as variants that, while having the same formula, have different spatial orientations (known as **stereoisomers**). Stereoisomers are identified by the way they rotate the plane of polarised light. Those that rotate the compound anticlockwise (to the left) are called L-isomers, whereas those that rotate it clockwise (to the right) are called D-isomers.

Another type of isomerism is **cis–trans isomerism**, which relates to the arrangement of atoms across the carbon–carbon double bonds. The *cis* configuration is when two linked atoms reside on the same side of the double bond, whereas the *trans* one is when they reside on the opposite sides. *Cis–trans* isomerism is particularly important in lipid chemistry.

Chemical reactions

Chemical reactions are exchanges of protons and electrons. They result in the formation, or breakage, of chemical bonds between atoms and molecules. Chemical reactions are associated with energy transfer – some require energy to proceed, whereas others release energy. The strength (potential energy) of different chemical bonds is shown in Table 2.2.

Table 2.2 Potential energy of some chemical bonds in biological systems

Type of bond	Energy (kJ/mol)
Ionic	12.6–29.3
Covalent (single)	210–160
Covalent (double)	500–710
Covalent (triple)	815
Hydrogen	4.2–8.4
Van der Waals interactions	4.2

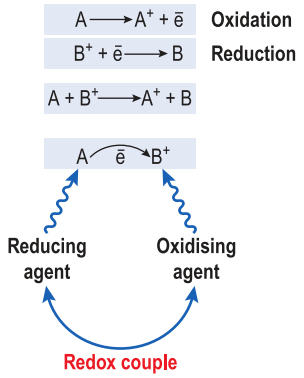


Fig. 2.5 **Oxidation–reduction reaction.** Note: The reducing agent becomes oxidised by donating electrons. The oxidising agent becomes reduced by accepting electrons.

The main types of chemical reactions are the following:

- **Synthesis**, when a larger molecule is formed from smaller substrates
- **Lysis**, when a molecule is broken down into smaller compounds
- **Exchange reactions**, where atoms, or groups of atoms, are exchanged between molecules (e.g. transamination reactions, where the amino group is transferred between molecules).

Electrophiles and nucleophiles

During a chemical reaction, atomic structures of reacting molecules are modified by the transfer of electrons (or protons). The excess electrons from one atom may 'invade' the orbitals of another, forming new shared orbitals. The relevant terminology is as follows:

- **Nucleophiles** are negatively charged atoms that are electron-rich. They tend to lose pairs of electrons.
- **Electrophiles** are positively charged atoms that are electron-poor. They accept electrons.
- A **nucleophilic attack** is a situation where the electrons from a nucleophile move into an electrophilic atom. Nucleophilic attack underpins, for instance, very common reactions of hydrolysis.

Oxidation–reduction (redox) reactions

Oxidation–reduction (redox) reactions are paired reactions in which electrons pass from one molecule to another (*Fig. 2.5*). In the process, the energy trapped in the chemical bonds in the molecule being oxidised is transferred to the molecule being reduced.

Oxidation is a net *loss* of electrons with an *increase* in oxidation state by a molecule (the complete oxidation state is when an atom is maximally charged after all its electrons have participated in ionic bonds).

Reduction is a net *gain* of electrons (or protons – hydrogen ions) with a *decrease* in oxidation state by a molecule.

Oxidation has also been defined as a loss of a hydrogen atom from a compound, and reduction as an addition of a hydrogen ion to a compound.

The oxidation–reduction (redox) reactions are key to the cellular flow of energy.

In redox reactions, electrons will flow from carrier A to carrier B, which has the higher redox potential (a tendency to be reduced or to accept electrons).

An electron added to an atom (a reductant) traps the energy. Reduced molecules possessing such trapped energy are more stable. The energy can then be released during oxidations.

Specific compounds in the cell transfer protons and electrons. The coenzyme nicotinamide adenine dinucleotide (NAD^+) is the most important electron acceptor. Another one is flavin adenine dinucleotide (FAD), which is an integral part (the prosthetic group) of several enzymes. Both are also hydrogen ion carriers. The structurally almost identical nicotinamide adenine dinucleotide phosphate (NADP) is a reductant that participates in many biosynthetic reactions.

Energy in biological systems

Energy is the capacity to do work. There is potential energy and kinetic energy. Potential energy is the capacity of an object to do work (e.g. because of its position in space), whereas kinetic energy is the capacity to do work associated with the movement of an object. Each chemical reaction is associated with a change in free energy (ΔG). Exergonic reactions are reactions that release energy and are characterised by a negative change in free energy ($-\Delta G$). Endergonic reactions are reactions that require energy to proceed (positive change in free energy ($+\Delta G$).

Energy cycle in biology

The energy required by living organisms is trapped in foodstuffs by plant photosynthesis, which consumes carbon dioxide and generates oxygen. Organic molecules thus produced then become metabolic fuel for animals. Their absorption and digestion by living organisms provide the energy necessary for their survival. This energy is released in a highly controlled, stepwise manner during biological oxidations, resulting finally in the production of carbon dioxide and water. The released energy is used for the building up of sophisticated biological structures. It is also used to support cellular transport, neural transmission and mobility, as well as growth, reproduction, defence and repair.

Potential energy of chemical bonds

Formation of a chemical bond requires energy input, and some potential energy accumulates in the formed bond (see Table 2.2). This energy can be recovered when the bond is broken. An analogy is putting a bucket of water on a high shelf (inputting energy) and recovering the energy when the water is poured down to drive a wheel. The biologically important bonds that have a particularly high potential energy are the phosphoanhydride bonds formed between phosphate groups in molecules such as adenosine triphosphate (ATP).

Another mechanism of energy trapping is building up an ion gradient across a biological membrane. Energy input is required to create such a gradient; it is released when the ions 'return' across the barrier.

Energy flow in chemical reactions

The content of free energy (G) changes during chemical reactions. We describe the energetics of reactions in relative terms, indicating free energy change (ΔG), which is negative when the energy content decreases or positive when it increases. If a reaction releases energy, it is **exergonic**, and when it requires energy input to proceed, it is **endergonic**. The products of exergonic reactions have a lower potential energy than their substrates, and the products of endergonic reactions have a higher potential energy.

Reactions where products have less potential energy than the substrates tend to occur spontaneously, whereas those where the potential energy of substrates increases require energy input. The key concept in biochemistry is that the energetically favourable reactions are used to drive the unfavourable ones. Let us imagine two reactions:

1. $A + B \rightarrow X$ and (endergonic reaction)
2. $X + Y \rightarrow Z$ (exergonic reaction).

Reaction 1 is energetically unfavourable. It occurs only very slowly, yielding small amounts of the product X, according to its equilibrium. Reaction 2, on the other hand, proceeds spontaneously. As its substrate is X, it depletes it, shifting the equilibrium of reaction 1 and 'forcing' it to proceed. This is a common pattern in various metabolic pathways. Note that the hydrolysis of ATP has a highly negative G. Thus, ATP can drive the otherwise energetically unfavourable reactions.

Anabolic and catabolic pathways

Metabolic pathways are chains of chemical reactions, classed according to their purpose. The **anabolic pathways** use energy in order to build up large molecules and are usually associated with reductions; the **catabolic pathways**, on the other hand, break down large reduced precursor molecules. They are associated with oxidations.

Generation of metabolic energy

Metabolism generates energy to support body functions, growth and regeneration. The required energy is acquired from nutrients. Cellular energy flow depends on biological oxidations. The overall oxidation process in living organisms is the multistep conversion of highly reduced organic compounds, in the presence of oxygen, to carbon dioxide and water. It is analogous to the combustion reaction, but in biology this 'combustion' proceeds by minute, carefully controlled, stages and under physiological conditions. The substrates for oxidation are **highly reduced** compounds, mainly carbohydrates and fats. We call them metabolic fuels (Information box 2.3). A substantial part of the released energy is chemically trapped for subsequent use. The major stages in this process are as follows:

1. **Ingestion of nutrients** and their digestion in the gut, which results in the liberation of the molecules of metabolic fuels from more complex compounds. These are absorbed into plasma.
2. **Conversion of the fuel molecules.** Metabolic fuels are then oxidised. The electrons (and protons) are first transferred to intracellular coenzymes, such as NAD and NADP, forming the reduced NADH and NADPH, respectively (the NADPH is used in biosynthetic reactions). The NADH then carries the electrons, accompanied by protons, to the mitochondria and transfers them to the electron transport chain (ETC; Ch. 3).
3. In the ETC, electrons and protons are transferred along the chain, participating in the sequence of redox reactions involving ferric iron complexes, cytochromes and other electron carrier molecules. The final recipient of electrons is the oxygen atom, and a molecule of water is formed.

As the electrons proceed through the ETC, the 'remaining behind' protons are pumped out, creating a concentration gradient across the inner mitochondrial membrane. This an electrogenic pump system generates an electrochemical potential because it pumps ions across the membrane without the ion movement in the opposite direction. The formed gradient acquires potential energy. These protons are then returned to the mitochondrial matrix through a membrane channel, which constitutes part of the enzyme ATP synthase; as a result the released energy is trapped in high-energy bonds of the ATP, formed from ADP. The process is known as oxidative phosphorylation.

Information box 2.3 Metabolic fuels

Metabolic fuels are highly reduced compounds that can be oxidised in the body to release energy. The main metabolic fuels are carbohydrates and fats. Among the carbohydrates, the most important fuel is glucose with the caloric value 16.7 kJ (4 kcal)/g. Under normal circumstances, glucose is the only fuel used by the brain. Red blood cells are also completely dependent on glucose for energy. Muscle preferentially uses glucose at the start of exercise, switching subsequently to the use of fatty acids. A limited amount of glucose, sufficient for approximately 12 hours, is stored in the body in the form of glycogen.

Fatty acids, which yield 37.7 kJ (9 kcal)/g, are the most efficient metabolic fuel. The amount of fat stored in the adipose tissue of a lean average human is sufficient for survival for more than 2 months without food.

Proteins can be converted to glucose when its supply is insufficient. Their caloric value is also 16.7 kJ (4 kcal)/g. The pathway of conversion of non-carbohydrate compounds to glucose is known as gluconeogenesis. The primary substrates for gluconeogenesis are the amino acid alanine, glycerol and lactate.

CHEMICAL COMPOSITION OF THE HUMAN BODY

Chemical elements

The body contains a vide variety of organic compounds. The skeleton contains substantial inorganic, mineral components. Carbon, oxygen, hydrogen and nitrogen make up 88.5% of the dry body mass. The most abundant mineral is calcium (4%), followed by phosphorus (2.5%), potassium, sulphur, sodium, chlorine and magnesium (0.1%). Other elements are present in much lesser amounts: iron, for instance, constitutes only 0.01%. The elements that are present in less than 0.01% are known as the trace elements (e.g. manganese or iodine). Despite their minute amounts, many perform important metabolic roles, particularly as prosthetic groups and cofactors of enzymes.

Sodium and potassium are fundamental for the maintenance of membrane potential and impulse transmission. Sodium and chloride also maintain osmolality and, therefore, fluid volume in the body compartments, including cell volume. Magnesium serves as a cofactor for multiple enzymes, including the ones participating in DNA replication and ATP synthesis. Calcium and phosphate are the main components of bone. Calcium is also fundamental for the transmission of impulses and cell signalling. Fluoride is present in bone and teeth.

Iron is a component of haem and, thus, of haemoglobin and myoglobin. Iodine is a component of thyroid hormones. Copper scavenges for free radicals, is associated with oxygenases and is important for collagen cross-linking. Zinc, like magnesium, is a cofactor for multiple enzymes and influences wound healing, tissue proliferation and growth, and immune function. Selenium is a cofactor for enzymes, including the ones participating in the metabolism of thyroid hormones (deiodinases). It also affects the immune function.

Water content and the main fluid compartments

Water makes up approximately 60% of the lean body mass (see Chs 1 and 16). A 70-kg adult male has approximately 42 L total body water. Two-thirds of it is in the **intracellular fluid** (ICF) and one-third is in the **extracellular fluid** (ECF). ECF is further compartmentalised into the interstitial fluid, plasma and transcellular fluids, such as lymph and cerebrospinal fluid (CSF) (Table 2.3). The main cations, i.e. sodium and potassium, are differentially distributed across the plasma membrane, with higher concentrations of sodium in the ECF and higher concentration of potassium in the ICF. ICF also has a higher concentration of proteins and phosphates than ECF. The transfer of water, ions and metabolites across cell membranes is fundamental to maintaining body functions.

The concentration gradient of sodium and potassium **ions across the cell membrane** is the key driver of cellular transport systems. The distribution of sodium and potassium ions across the cell membrane creates an electrical potential known as the **resting membrane potential.** In most cells, it is of the order of –60 mV, the inside being negative (Ch. 8). Changes in the electrical potential underlie the electrical signals generated by excitable cells and **nerve impulses** (Clinical box 2.1, see also Information box 2.4).

Another ionic concentration gradient fundamental for cell function is the gradient of calcium ion (Ca^{2+}) concentration. Calcium concentration outside the cell is approximately 10,000 higher than that inside. Therefore, a cell can be flooded with calcium very quickly. After an increase in concentration, calcium can be compartmentalised within the cell, for instance, in the endoplasmic reticulum. This decreases its concentration in the cytoplasm. Sudden calcium ion concentration changes are very important for intracellular signalling and are essential, for muscle contraction (see Ch. 9), neurotransmission and most secretory processes. For details of laboratory assessment of water and electrolyte balance see Information box 2.4.

Table 2.3 Composition of the extracellular and intracellular fluid

Ion	Approximate extracellular concentration (mmol/L)	Approximate intracellular concentration (mmol/L)
Na^+	140	12
K^+	4	140
Cl^-	110	4
Bicarbonate, HCO_3^-	25	12
Phosphates	2	13
Protein anions	9	138
Ca^{2+}	2.4	<0.0002
Mg^{2+}	1.0	0.8

Clinical box 2.1 Cardiac muscle is very sensitive to changes in K^+ concentration

Changes in the extracellular K^+ concentration alter the resting membrane potential of cardiac myocytes. This changes their excitability. Normal extracellular potassium concentration is 3.5–5.0 mmol/L, but if levels fall below this, the myocytes become hyperpolarised and cardiac excitation is reduced. On the other hand, if the K^+ concentration rises above 5.5 mmol/L, cardiac excitation increases. I both situations there is a potential risk of arrhythmias and cardiac arrest (Ch. 11).

The role of vitamins

Vitamins are essential organic nutrients. They are structurally diverse and are different from the main nutrient groups. They are only required in very small amounts. Vitamins act as coenzymes, and some as signalling molecules. Most vitamin deficiencies lead to specific diseases. Vitamin A and vitamin D are also toxic in excess.

Vitamins are classified into fat soluble (A, D, E and K) and water soluble (vitamin B, folate, biotin and vitamin C).

Vitamin D and vitamin A (retinol) act in a steroid hormone-like way through intracellular receptors. Vitamin D (calciol) plays a crucial role in bone metabolism. Vitamin D can be synthesised in the skin, and vitamin A is key to the process of vision and also affects cell growth and proliferation. Vitamin E is an important antioxidant, and vitamin K (phylloquinone) participates in the clotting of blood and is required for the synthesis of prothrombin and several other coagulation factors.

The B vitamins serve predominantly as coenzymes. Vitamin B_1 (thiamine) is important for carbohydrate metabolism, and B_2 (riboflavin) participates in FAD synthesis. Vitamin B_3 (niacin) is required for the synthesis of NAD. Vitamins B_6 (pyridoxine) and B_7 (biotin) participate as coenzymes in carbohydrate, lipid and amino acid metabolism. Biotin is also required for the synthesis of some neurotransmitters.

Vitamin B_9 (folic acid) participates in the so-called single-carbon transfer reactions and is necessary for the synthesis of purines and pyrimidines, and, thus, of nucleic acids. Vitamin B_{12} is part of the haem structure. Panthotenic acid is important for the synthesis of coenzyme A (CoA). Vitamin C is a reducing agent and is necessary for the synthesis of collagen.

Information box 2.4 Laboratory assessment of water and electrolyte balance

The water and electrolyte balance of a patient is assessed by recording the patient's intake and output of fluids (the record includes the amount drunk, the volume of given intravenous solutions, the urine volume, the volumes of fluids obtained during surgical drainage, etc.).

This is complemented by measurements of the ionic balance of the plasma. The set of common measurements, known in hospital jargon 'urea and electrolytes', includes measurements of sodium and potassium as the main cations, and chloride and bicarbonate as the main anions. The measurements of plasma urea and creatinine are normally included because they reflect kidney function – the most important determinant of the ionic balance.

Normally, the sum of sodium and potassium is greater than the sum of bicarbonate and chloride. The difference, which is normally approximately 10 mmol/L, is known as the anion gap (AG):

$$AG = (Na^+ + K^+) - (Cl^- + HCO^{2-}_3)$$

The anion gap includes anions that are not routinely measured, such as lactate or ketones. The gap can increase substantially when these anions accumulate (e.g. ketones in poorly controlled diabetes) and, thus, is of diagnostic significance for the physician.

Organic biomolecules

The main classes of organic biomolecules are carbohydrates, fats and proteins, and also nucleotides and nucleic acids. They may combine in forming 'hybrid' molecules such as glycoproteins or glycolipids.

Carbohydrates and fats are the main energy substrates. Lipids are also a major structural component of cell membranes.

Proteins are required to maintain body structure and functions, but can be transformed into fuel when there is a large demand for energy and/or when the supply of carbohydrates is poor. Proteins in the form of enzymes, signalling molecules and antibodies underpin most body functions. They are synthesised, maintained and regulated according to the genetic information contained in the DNA, which itself is a molecule containing amino acid–derived structures (the purine and pyrimidine bases), carbohydrates and phosphate.

Amino acids, the basic units of proteins, and lipid molecules such as cholesterol, also serve as precursors of a wide range of biologically important molecules, including hormones and neurotransmitters.

CARBOHYDRATES

Carbohydrates form only approximately 2% of the body mass. They consist of carbon, oxygen and hydrogen, with a ratio of hydrogen to oxygen of approximately 2:1. Carbohydrates have the general formula $(CH_2O)_n$, where n is the number of carbon atoms. They are single molecules (simple sugars, **monosaccharides**, or their polymers, **polysaccharides**), which may contain thousands of units.

Monosaccharides

Monosaccharides are present in the diet **or** may be obtained by digestion of more complex carbohydrates. Because they can also be synthesised in the body from non-carbohydrate sources, their external supply is not required for survival. The simplest monosaccharides are the three-carbon sugars (trioses; *Fig 2.6*). Those that contain an aldehyde group are known as **aldoses**, and those with a ketone group in their structure are called **ketoses** (see *Fig. 2.6*).

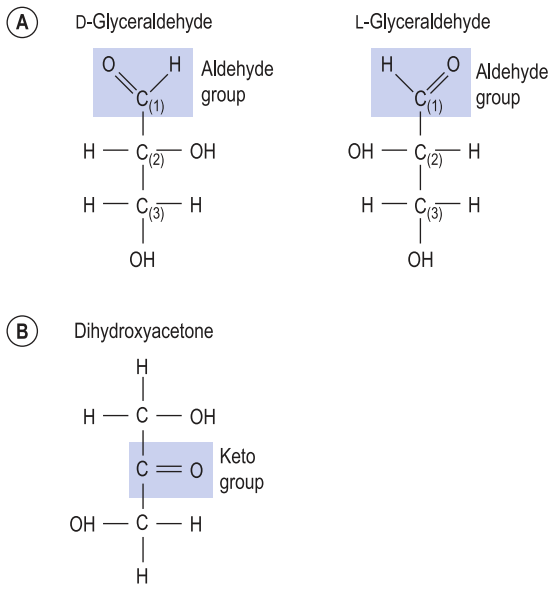


Fig. 2.6 Structure of trioses. (A) Aldoses (showing stereoisomerism). (B) A ketose.

Like the other types of organic molecules, carbohydrates exhibit **isomerism**, i.e. variability of structure with the same atomic content. For instance, in glyceraldehyde, which is a triose, the central carbon atom is chiral and therefore it can exist as two (D- and L-) stereoisomers. All sugars that have chiral groups are divided into the D- and L-series, according to the isomer of glyceraldehyde from which their asymmetrical carbon farthest from C1 derives, Virtually all sugars found in biological systems are D-isomers. There is also optical isomerism, which makes the molecules rotate polarised light to the right (dextrorotatory; d/D) or to the left (laevorotatory; l/L).

The five-carbon sugars (pentoses)

Two important pentoses are **ribose** and **deoxyribose**, the components of the nucleic acids RNA and DNA, respectively.

The six-carbon sugars (hexoses)

The most important hexose is glucose, $C_6H_{12}O_6$. Glucose is the central molecule in human energy metabolism. In an aqueous solution, it can form *(Fig. 2.7)* both straight chain (less than 1% of all glucose

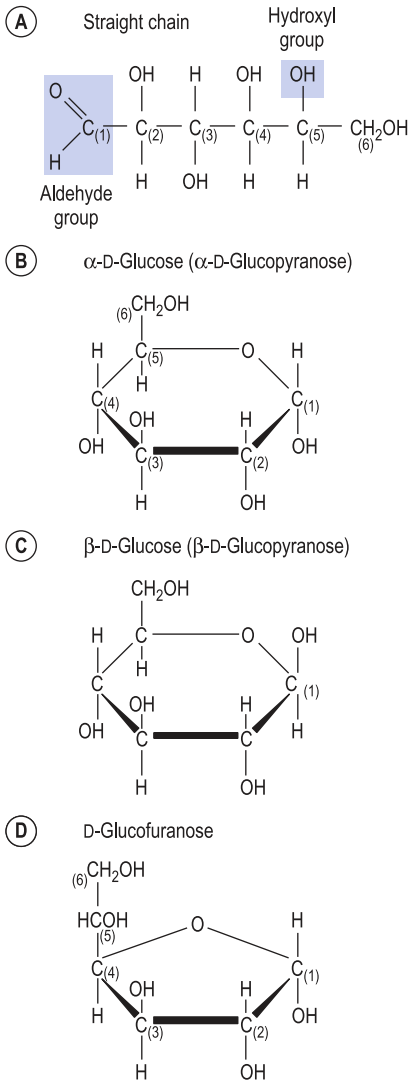


Fig. 2.7 Different forms of glucose. Each of these molecules has an identical chemical composition. (A) Open chain molecule. (B) and (C) A reaction between the hydroxyl group on C5 and the aldehyde on C1 produces the pyranose ring structure. (D) A reaction with the hydroxyl group on C4, produces a furanose ring.

molecules) and ring structures. Furthermore, the so-called anomeric carbon in ring structures forms α- or β-glucose anomers that differ by the arrangement of the hydrogen atoms and hydroxyl (–OH) groups on C1.

Other common hexoses are **mannose, galactose** and **fructose** *(Fig. 2.8)*. Mannose and galactose only differ from glucose in the configuration of one of the carbons (mannose around C2 and galactose around C4) and are called **epimers**.

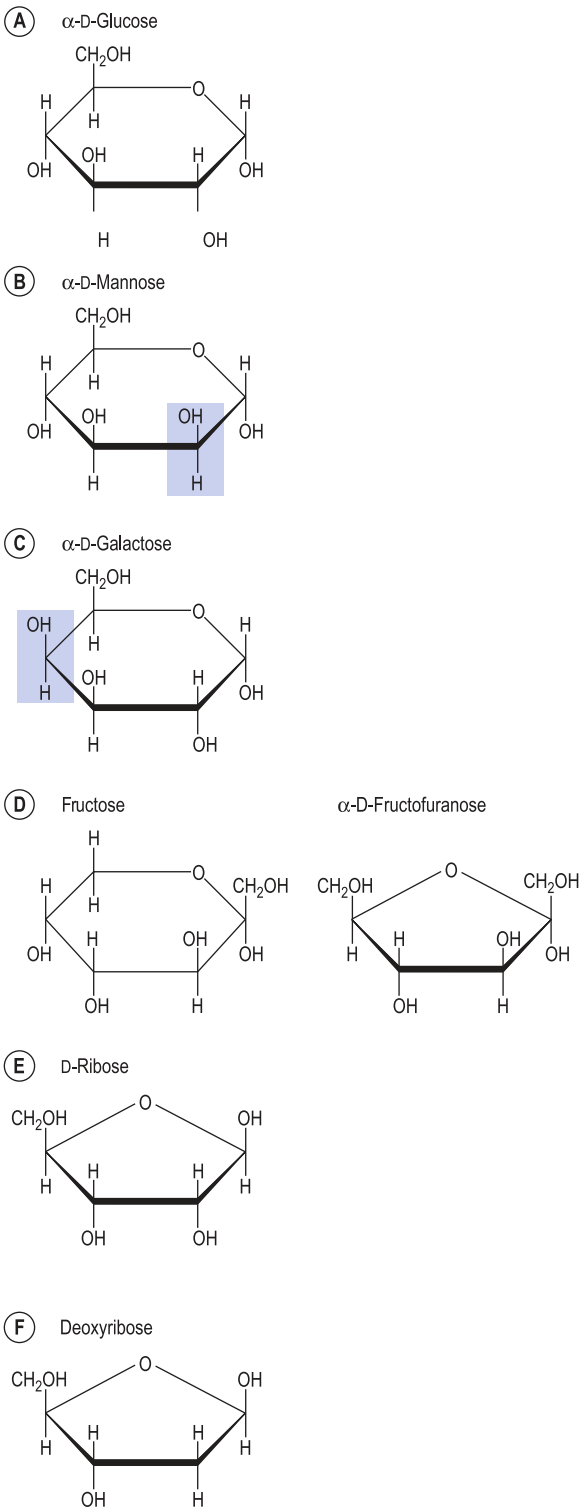


Fig. 2.8 Common monosaccharides. In (B) and (C), the part of the molecule that differs from glucose is highlighted.

Modified monosaccharides: amino sugars and sugar-derived acids

Monosaccharide molecules can be modified by other chemical groups. An introduction of the amino group on C2 of glucose and galactose yields the amino sugars **glucosamine** and **galactosamine**, respectively. The insertion of carboxyl groups creates glucuronic and sialic acids. The sulphate group is a component of the proteoglycans, which are key substances present in the extracellular matrix (ECM).

Disaccharides

Disaccharides form from two either identical or different monosaccharide units linked together by a **glycosidic bond**. These bonds link C1 of one sugar and the hydroxyl group of another. They are either α or β linkages, depending on the orientation around C1 (Table 2.4). Sucrose, the ordinary sugar, is a disaccharide made up of α-glucose and β-fructose linked by an α1→2 linkage. Two other dietary disaccharides are **maltose**, a dimer of glucose molecules, and **lactose** (milk sugar), made from glucose and galactose.

Polysaccharides

Starches, the polymers of glucose, serve as storage carbohydrates in plants. The simplest, **amylose**, has long linear glucose chains linked by α1→4 bonds. **Amylopectin**, which makes up approximately 80% of the starch found in foods, has 1→4 linked glucose chains, which branch by forming 1→6 bonds. The starches found in grains such as wheat, rice and potatoes are the major source of dietary carbohydrates. Amylases secreted by the salivary glands and the pancreas digest them, releasing glucose, maltose and isomaltose.

Cellulose is an unbranched glucose polymer linked by β1→4 bonds. It is the principal component of plant cell walls and is the most common organic compound on the planet. Humans have no enzymes capable of hydrolysing these bonds. Cellulose and other non-metabolised polysaccharides form **dietary fibre**.

In human tissues, the carbohydrate storage molecule is **glycogen**. It is a large, branched polymer of glucose. The high level of glycogen branching means that the molecule has a large number of free ends. This facilitates its rapid degradation to glucose.

Glycogen stores glucose in a highly concentrated form, without the osmotic problems (attraction of large amounts of water, and thus increased volume) associated with a large number of free glucose molecules. However, the size of glycogen stores is modest in humans, with approximately 75 g present in the liver and 250 g in skeletal muscle. This can supply glucose for 12–18 hours, after which time the glucose needs to be synthesised from non-carbohydrate sources, such as amino acids, through **gluconeogenesis**.

Table 2.4 Principal disaccharides

Disaccharide	Carbon-1 sugar	Other sugar	Linkage	Digesting enzyme
Sucrose	α-Glucose	β-Fructose	1→2	Sucrase-isomaltase
Lactose	β-Galactose	β-Glucose	1→4	Lactase
Maltose	α-Glucose	β-Glucose	1→4	Maltase
Isomaltose	α-Glucose	β-Glucose	1→6	Sucrase-isomaltase

COMPLEX CARBOHYDRATES

Complex carbohydrates are carbohydrates covalently linked to proteins or lipids.

Thus, **glycoproteins** contain carbohydrate chains linked to serine and threonine residues of proteins. Their carbohydrate content is 10%–15%. The carbohydrates involved are most commonly mannose, galactose, fucose, xylose, amino sugars such as *N*-acetylglucosamine and *N*-acetylgalactosamine, and glucuronic and sialic acids. Carbohydrate chains in the glycoproteins are relatively short, containing 10–15 molecules, and often look like two- or three-pronged forks. They form highly variable structures important for cell recognition and immune recognition: such as, a range of membrane receptors. Together with glycolipids (see below), they constitute **blood group antigens** on the surface of red cells, the most important being the **ABO blood group system**.

The formation of a bond between the hydroxyl group of the amino acid side chain and C-1 of *N*-acetylgalactosamine is known as **O-glycosylation**, and the linkage between the amide group of the amino acid asparagine and C-1 of *N*-acetylglucosamine is known as **N-glycosylation**.

Complex carbohydrates linked to proteins, the **proteoglycans**, are the key components of the ECM. They typically contain a large proportion (95%) of carbohydrate. The carbohydrates are attached to a polypeptide chain, forming long, linear chains. These molecules are known as glycosaminoglycans (GAGs) or **mucopolysaccharides**. Many GAGs are negatively charged and thus attract cations and large amounts of water, forming a gel-like substance. Thus, they impart a degree of flexibility to the tissue and act as shock absorbers.

LIPIDS

Lipids make up approximately 20% of body mass in adults of normal weight, more in women. They play a number of essential roles:

- They are the key components of cell membranes.
- They are the major form of energy storage.
- They play important roles in cell signalling.

Fatty acids

Fatty acids are hydrocarbon chains ***(Fig. 2.9A, B)***, with a carboxy group at one end (the α-carbon) and a methyl group (CH_3) at the other (the ω-carbon):

- S**aturated** fatty acids have all the carbon atoms in the chain linked by single bonds.
- **Monounsaturated fatty acids** possess one double bond.
- **Polyunsaturated** fatty acids have more than one double bond.

Many of the fatty acids found in cells have 16, 18 or 20 carbon atoms and up to 3 double bonds. Most of the double bonds in biological molecules are in the *cis* configuration. Such a configuration puts a kink in the hydrocarbon chain. The most common fatty acid is the C18 **palmitic acid** (Information box 2.5). Some fatty acids have one of the double bonds placed three carbons from the ω-carbon (they are known as the ω-3 fatty acids). An example is the **eicosapentaenoic** acid, the precursor of some prostaglandins.

The ω-3 fatty acids are found in fish oil. Their ingestion confers some cardiac benefits; for instance, it prevents some disturbances of the heart rate (arrythmias).

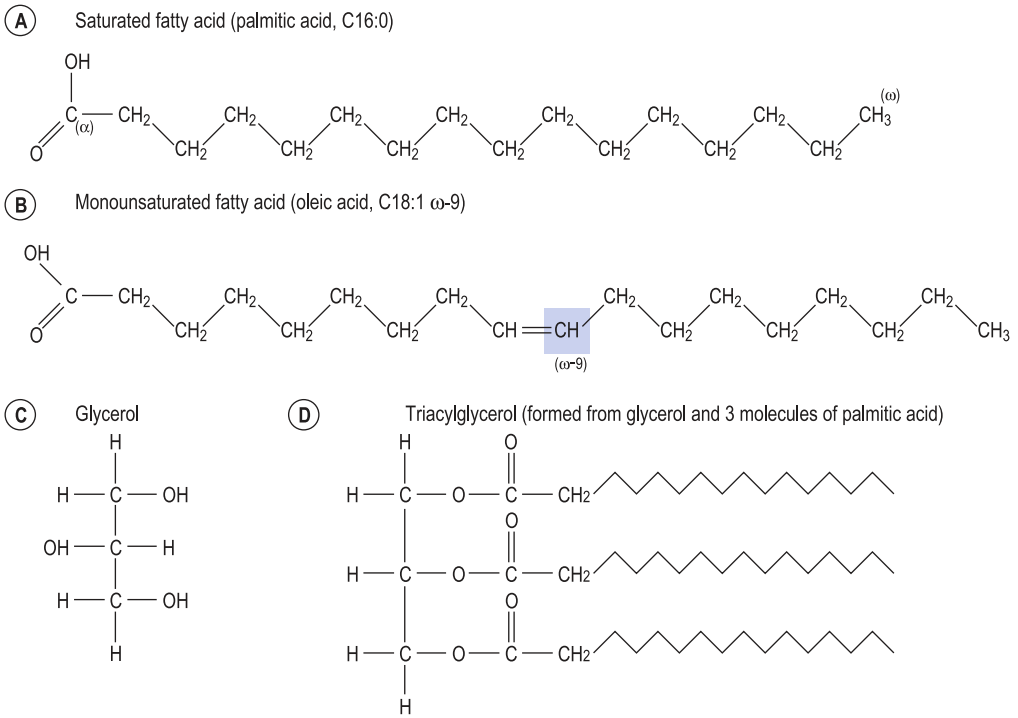


Fig. 2.9 Fatty acids, glycerol and triacylglycerol.

Information box 2.5 Fatty acids can be named in several ways

Fatty acids can be classified according to the number of carbon atoms in the chain and the number and position of the double bonds.

A notation that is widely used gives the number of carbon atoms, the number of double bonds and the position of the first double bond, counting from the ω-carbon. For example, the saturated fatty acid found in palm oil is called palmitic acid. The formula for palmitic acid is $CH_3(CH_2)_{14}COOH$, its chemical name is *n*-hexadecanoic acid and the shorthand notation is C16:0, indicating that it has 16 carbon atoms and no double bonds.

For arachidonic acid, the *cis*-5,8,11,14-eicosatetraenoic acid, the shorthand notation is C20:4 ω-6. The positions of all of the double bonds are noted as C20:4 all *cis*-$\Delta^5,\Delta^8,\Delta^{11},\Delta^{14}$ (see ***Fig. 2.7***).

Information box 2.6 Prostanoids facilitate blood flow through the vessels and aspirin acts by inhibiting the synthesis of prostaglandin

Prostacyclin released from the blood vessel wall is a vasodilator and thus encourages blood flow. It also inhibits the aggregation of blood platelets, preventing clot formation.

Conversely, thromboxane A_2 acts as a vasoconstrictor, thus reducing flow (and thus blood loss from a damaged vessel), and stimulates platelets to aggregate, facilitating coagulation.

The enzyme that acts on arachidonic acid in the first step in the synthesis of prostanoids is cyclo-oxygenase (COX). Aspirin (acetylsalicylic acid) belongs to a group of drugs called non-steroidal anti-inflammatory drugs (NSAIDs), which inhibit COX and reduce inflammation and pain. Aspirin is also used in cardiovascular prevention, to inhibit blood clotting.

Triacylglycerols

Triacylglycerols (also called **triglycerides**) are esters of glycerol and fatty acids and are the storage form of lipids. Glycerol is a C3 sugar alcohol, formed by reduction of the aldehyde group of a triose to the hydroxyl group ***(Fig. 2.9C)***. Triacylglycerols are formed by dehydration reactions between the carboxyl group (COOH) of a fatty acid at each of the three hydroxyl (OH) groups of the glycerol molecule (see ***Fig. 2.9D***). Triacylglycerols are stored in the adipose tissue as subcutaneous fat and as visceral fat, which surrounds the abdominal organs.

Dietary fats

Triacylglycerols are the main components of solid and liquid dietary fats. Fatty acids that make up the triacylglycerol molecules determine the physical properties of a particular fat. Triacylglycerols composed of short-chain fatty acids, or of unsaturated fatty acids, are liquid at room temperature. Examples of these are olive oil (containing oleic acid) and sunflower oil (containing polyunsaturated fatty acids). Triacylglycerols that contain more saturated fats and longer fatty acid chains are fats solid at room temperature (e.g. butter).

Essential fatty acids

Most of the fatty acids used in the body are supplied in the diet. Saturated fatty acids can be synthesised from carbohydrates (Ch. 3). They can be converted to unsaturated fatty acids. However, **linoleic acid** (ω-6 fatty acid C18:2; *cis*, *cis* 9,12) and **linolenic acid** (ω-3 C18:3; *cis*, *cis*, *cis* 9,12,15) are **essential fatty acids**. They cannot be manufactured in the body because none of the human enzymes can insert double bonds beyond C-10 in a fatty acid molecule.

Eicosanoids

Eicosanoids are derived from C20 fatty acids with between three and five double bonds. **Arachidonic acid**, synthesised from linoleic acid, is the precursor of a large number of eicosanoids. It can be released from the membrane by the action of the enzyme phospholipase A_2 (PLA_2) on phosphatidylcholine, one of the membrane phospholipids. Different groups of eicosanoids include prostaglandins, thromboxanes and leukotrienes. They are locally acting hormones with a very short half-life. They are important in the response to inflammation and in the control of vascular smooth muscle contraction (Information box 2.6).

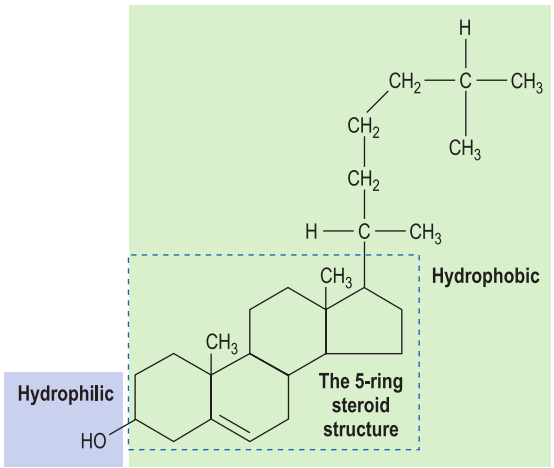


Fig. 2.10 Structure of cholesterol. The figure inside the *dashed rectangle* is the basic steroid structure.

Cholesterol and steroids

The **cholesterol** molecule has a four-ring structure known as the cholestane structure ***(Fig. 2.10)***. There is a hydroxyl group attached to ring number 1, and it can interact with water. Cholesterol is the major sterol in animals, whereas plants produce related sterols such as sitosterol and campesterol. Cholesterol is an essential structural component of the cell membranes. It decreases the fluidity of cell membranes. It is also an important precursor of a range of biologically important substances: bile salts, vitamin D and steroid hormones.

Most cells are capable of synthesising cholesterol. It can also be absorbed from the diet. It is distributed to tissues by lipid-transporting particles known as lipoproteins.

Cells regulate their cholesterol supply, balancing the rate of synthesis against the external provision. The key regulatory points are the enzyme **HMG-CoA** reductase (a key enzyme in cholesterol synthesis and a target of statin drugs) and the **LDL** receptor (which controls LDL, and thus cholesterol, uptake from plasma). Dysregulation of cholesterol balance leads to its high concentrations in plasma and contributes to the formation of atherosclerotic plaques (Clinical box 2.2).

Bile acids

The human body cannot metabolise cholesterol. It is excreted by the liver in bile.

Clinical box 2.2 Atherosclerosis

There is a link between raised plasma cholesterol or low-density lipoprotein (LDL-cholesterol; Ch. 16) and the increased risk of coronary events such as heart attacks. It is beneficial to reduce plasma cholesterol levels, particularly in persons who already suffer from heart disease. Diets rich in polyunsaturated fats reduce serum cholesterol to some extent. To obtain greater reductions, one needs to use the statins, drugs that inhibit the regulatory enzyme in cholesterol synthesis pathway, the 3-hydroxy-3-methylglutaryl-CoA reductase (HMG-CoA reductase).

Table 2.5 Selected transport proteins in plasma

Protein	Ligand
Albumin	Non-specific transport protein: metal cations, free fatty acids, steroids, bilirubin, haem, therapeutic drugs
Transferrin	Iron
Thyroid-binding globulin	Thyroxine (T_4), tri-iodothyronine (T_3)
Cortisol-binding globulin	Cortisol
Sex-hormone-binding globulin	Androgens and oestrogens

Bile acids emulsify fats in the intestine and thus are essential for their absorption and digestion. They are synthesised in the liver from cholesterol and are secreted as conjugates with the amino acid taurine or glycine. They are stored in the gallbladder and are eventually secreted into the duodenum (Ch. 15). The secreted bile acids (known as the **primary bile acids**) are further modified by the intestinal bacteria, yielding the **secondary bile acids.** Bile acids are conserved through reabsorption in the large intestine. This recirculation process is known as the **enterohepatic circulation.**

Steroid hormones

Steroid hormones are all derived from cholesterol after its conversion to another key metabolite, pregnenolone. The subsequent reactions in the pathway yield **glucocorticoids** such as cortisol, **mineralocorticoids**, such as aldosterone, and **oestrogens** and **androgens**, including progesterone, the hormone that is important in maintaining pregnancy (see also Ch. 10).

Steroids are lipid soluble; they cross the plasma membranes and bind to cytoplasmic receptors. They are usually carried in the plasma bound to specific binding proteins (Table 2.5).

Vitamin D

Vitamin D is a hormone with action mediated through the intracellular receptor. It is essential for calcium and bone metabolism. Its precursor (cholecalciferol; calciol) is synthesised from 7-dehydrocholesterol in the skin under the influence of ultraviolet light. It then undergoes two hydroxylation reactions. The first one, in the liver, yields 25-hydroxycalciferol (calcidiol), and the second one, in the kidney, generates the active form, 1,25-dihydroxycholecalciferol (calcitriol). Vitamin D facilitates the absorption of calcium in the intestine and its reabsorption in the kidney, and stimulates bone resorption by increasing the number of osteoclasts (Ch. 16).

Complex lipids

Complex lipids are lipid molecules with incorporated non-lipid (particularly carbohydrate) components. Their lipid component is hydrophobic, whereas the non-lipid part is often hydrophilic. Therefore, these molecules are **amphipathic** (both hydrophilic and hydrophobic) and orient themselves at the lipid/water interfaces. They are important components of cell membranes. Their hydrophilic part faces the 'outside' of the membrane, and the hydrophobic part orients towards the membrane core. Complex lipids also play a major role as components of the ECM.

Phospholipids

Glycerophospholipids are formed from phosphatidic acid, a molecule where glycerol phosphate is esterified with two acyl residues.

Choline linked to the phosphate group of the phosphatidic acid produces **lecithin**, another cell membrane phospholipid. Instead of choline, serine or ethanolamine or inositol can also be linked to the phosphate group. The structure of phosphatidylserine is shown in *Fig. 2.11A*. The hydrophilic groups of phospholipids are negatively charged and, thus, the overall surface charge of the membrane is negative.

Sphingolipids

Sphingolipids are structurally similar to phospholipids, but instead of glycerol they contain the C18 **sphingosine** (an alcohol formed from palmitic acid and serine) (see *Fig. 2.11B*). A molecule consisting of sphingosine linked to a fatty acid (acylsphingosine) is known as a **ceramide.** In turn, when ceramide is linked to a choline, it forms **sphingomyelin.** When a ceramide binds a sugar, it forms a **cerebroside**, and when it binds sialic acid, it forms a **ganglioside.**

Sphingomyelin, cerebrosides and gangliosides are components of cell membranes abundant in nerve tissues and the brain. Defects in the degradation of cerebrosides and gangliosides result in several rare clinical disorders known as **lysosomal storage diseases.**

Sphingolipids also contribute to the structure of blood group substances. Phosphorylated sphingosine is a signalling molecule.

PURINES AND PYRIMIDINES

Purines and pyrimidines contain nitrogen in their carbon rings. They also have basic amino groups (hence the alternative name 'nitrogenous bases'). They are components of nucleotides and nucleic acids (*Fig. 2.12*).

Purines are two-ring structures. The major purines are **adenine** and **guanine.** Pyrimidines consist of a single ring, and the major pyrimidines are cytosine, thymine and **uracil.**

Nucleosides are molecules in which a sugar (**ribose** or **deoxyribose**) phosphate is linked to a purine or a pyrimidine. **Nucleotides** are phosphorylated nucleosides (*Fig. 2.13*, *Fig. 2.14* and Table 2.6). There could be one, two or three phosphates attached to C5 of the sugar, each via phosphoanhydride bonds, forming nucleotide mono-, bi- or tri-phosphates, respectively. Nucleotides are the building blocks that form nucleic acids. In addition, free nucleotides play important roles in energy transfer and cell signalling. Purines are components of nucleotide coenzymes such as ATP and GTP (*Fig. 2.14* and Information box 2.7). Cyclic nucleotides derived from them, also containing purines, are cyclic adenine monophosphate (**cAMP**) and cyclic guanosine monophosphate (**cGMP**), both being fundamental for signal transduction (Table 2.6). Other key coenzymes, such as **NAD**, **NADP**, **FAD** and **coenzyme A**, are all derived from adenine nucleotide.

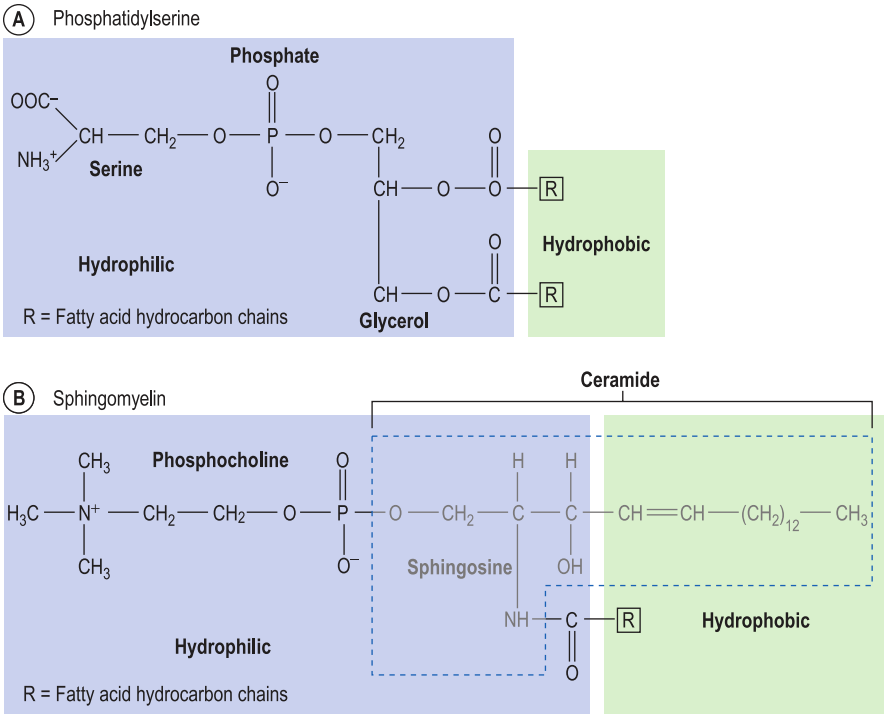


Fig. 2.11 Structure of (A) phosphatidylserine and (B) sphingomyelin.

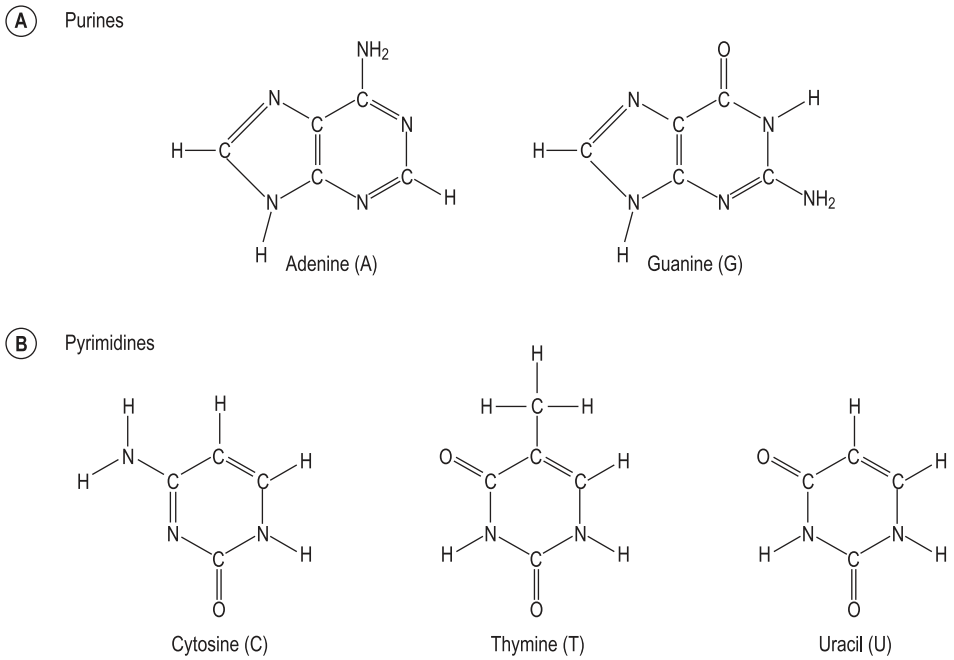


Fig. 2.12 Structure of (A) purine and (B) pyrimidine bases.

Synthesis and degradation of nucleotides

The key intermediate in purine synthesis is 5-phosphoribosyl pyrophosphate (PRPP), a derivative of ribose supplied by the pentose phosphate pathway, a pathway that branches off from glycolysis. Reactions in the synthetic pathway require the amino acids glutamine, aspartate and glycine; CO_2; and the coenzyme tetrahydrofolate, which transfers single-carbon residues between molecules. The product is the nucleotide inosine monophosphate (IMP), the precursor of purines.

The precursor of pyrimidines is a 1-carbon molecule, carbamoyl phosphate. The synthetic pathway involves the amino acids glutamine and aspartate, and also the bicarbonate anion. The pathway yields uridine monophosphate (UMP), from which other pyrimidines are formed.

Ribonucleotides are reduced to deoxyribonucleotides by nucleotide reductase, producing purine and pyrimidine nucleotides that are used in DNA synthesis. Folic acid (a B vitamin) is

required for the synthesis of the deoxyribonucleotide TMP (a thymidine nucleotide) from deoxyuridine monophosphate (dUMP) (Information box 2.8).

Purines are degraded to uric acid, and pyrimidines to a Krebs cycle metabolite, succinyl-CoA.

i Information box 2.7 Adenosine triphosphate (ATP)

Adenosine triphosphate or ATP (***Fig. 2.14***) is a highly mobile cellular energy store. It also participates in many metabolic regulatory loops, controlling the activity of key enzymes.

The phosphoanhydride bonds between the second (β) and the third (γ) phosphate group of the ATP release energy on hydrolysis, yielding about 7.3 kcal/mol each. These bonds are usually depicted by ~, indicating that they are high-energy bonds.

Most of the ATP in the cell is formed during oxidative phosphorylation in the mitochondrial electron transfer chain (ETC). A reaction that results in ATP synthesis by the transfer of a phosphate group from another phosphorylated compound outside the ETC is known as substrate-level phosphorylation.

i Information box 2.8 Anti-cancer drugs may act by inhibiting nucleotide synthesis

One of the ways of treating cancer is by blocking the rapid proliferation of the cancer cells. Many of the side-effects of anti-cancer drugs can be accounted for by their effect on cells that are normally rapidly dividing, such as hair and intestinal cells.

A commonly used anti-cancer drug, methotrexate, acts by inhibiting an enzyme involved in the production of thymidine nucleotides. Methotrexate is an analogue of dihydrofolate, which acts as a competitive inhibitor of the enzyme dihydrofolate reductase. Other cytotoxic agents act by inhibiting other points in the production of nucleic acids.

Nucleotide recycling: the salvage pathways

De novo synthesis of nucleotides requires large amount of energy. There are, however, pathways that recycle the purines and pyrimidines before they are completely degraded and re-incorporate them into the nucleotides. They are known as salvage pathways. Both excess uric acid production and enzyme deficiencies in the salvage pathways can manifest clinically as gout (Clinical box 2.3).

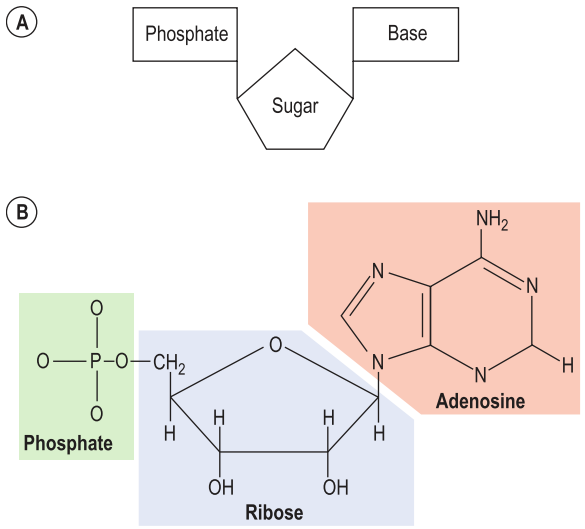


Fig. 2.13 Nucleotide structure. (A) Components of a nucleotide. (B) Structure of ribo-adenosine monophosphate (rAMP), usually referred to as adenosine monophosphate.

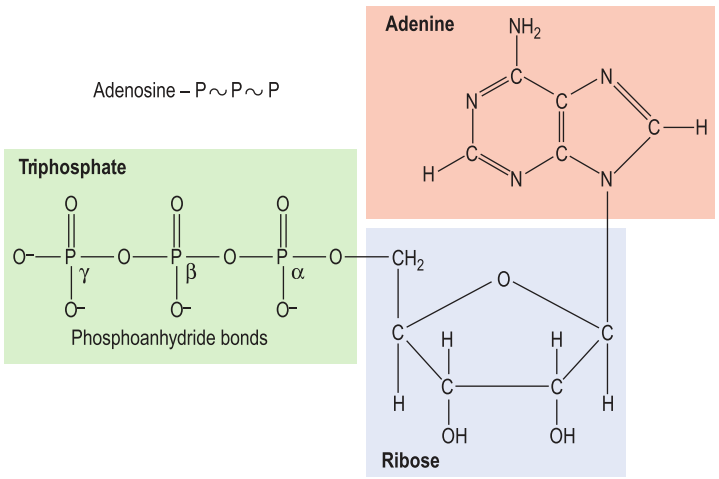


Fig. 2.14 Structure of adenosine triphosphate (ATP). The symbol "~" denotes high-energy bonds.

Table 2.6 Nucleosides and nucleotides (nucleoside phosphates)

Base	Nucleoside	Nucleoside monophosphates	Nucleoside diphosphates	Nucleoside triphosphates	Cyclic nucleotides
Adenine	Adenosine	AMP	ADP	ATP	cAMP
Guanine	Guanosine	GMP	GDP	GTP	cGMP
Cytosine	Cytidine	CMP	CDP	CTP	
Uracil	Uridine	UMP	UDP	UTP	
Thymine	Thymidine	TMP	TDP	TTP	

Clinical box 2.3 Excess purines can cause gout

In humans, excess purines are broken down to uric acid. If large amounts of uric acid are produced and not removed by the kidneys, uric acid crystals are deposited in joints and soft tissues, resulting in gout. One treatment of gout blocks the enzyme xanthine oxidase, which catalyses the final steps in the production of uric acid. The intermediate breakdown products are more water soluble than uric acid and can be excreted.

Deficiencies in one of the enzymes involved in the purine recycling pathways lead to increased purine synthesis and a concomitant increase in uric acid, leading to a very rare condition known as Lesch–Nyhan syndrome. One of the features of this syndrome is gout. Patients may also go on to develop arthritis and severe mental disorder.

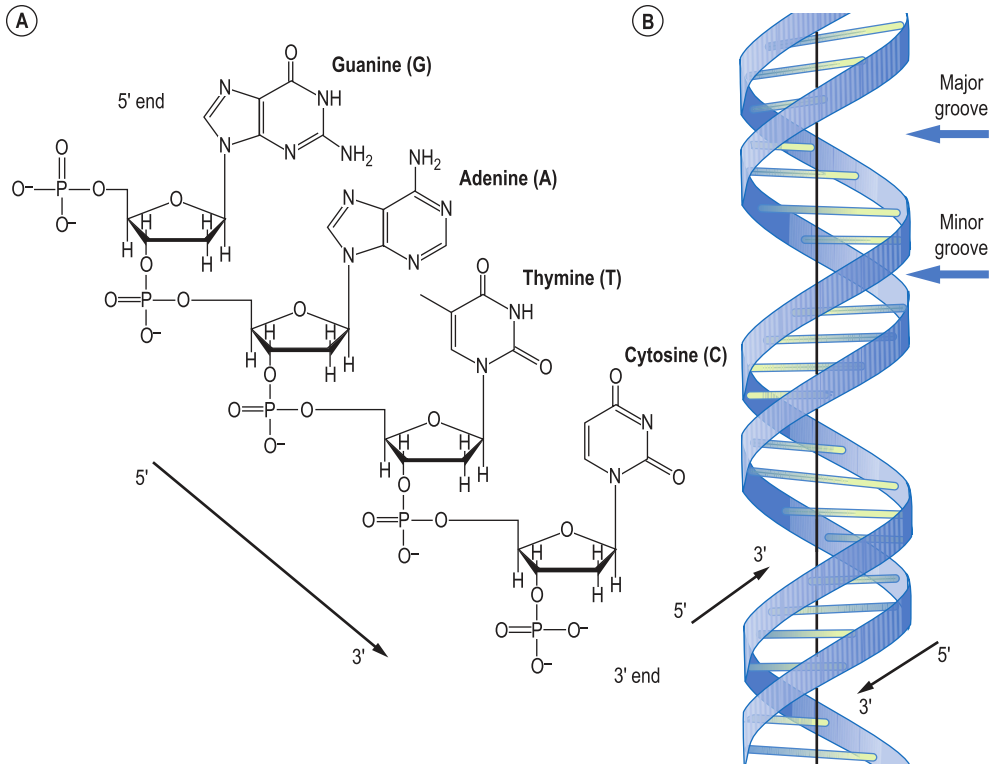


Fig. 2.15 Structure of DNA. (A) Details of the nucleotide sequence in a fragment of the DNA strand. (B) The double helix showing the major and the minor groove. *(Reproduced with permission from Dominiczak MH 2012 Flash cards in biochemistry. Elsevier, London, with permission.)*

NUCLEIC ACIDS

Deoxyribonucleic acid (DNA) and the **ribonucleic acids** (RNAs) contain and transmit genetic information. They safeguard the structural and functional continuity of an organism through generations. Both DNA and RNA are polymers of nucleotides. DNA forms very long chains, whereas the sizes of different types of RNAs vary widely. DNA and RNA are made from only four different nucleotides each.

Deoxyribonucleic acid

DNA is found in all cells capable of dividing. Most of the DNA is in the cell nucleus. A small amount is present in the mitochondria. In the prokaryotic cells (bacteria and some viruses), DNA is found in the cytoplasm.

DNA is a polymer of deoxyribonucleotides and forms a double-stranded molecule, with the characteristic double-helical structure. The building blocks in the DNA and RNA are four nucleotides: two are purine and two are pyrimidine nucleotides. The purines adenine (A) and guanine (G) are found in both DNA and RNA. With regard to pyrimidines, DNA contains cytosine (C) and thymine (T), and RNA contains cytosine (C) and uracil (U) (see *Fig. 2.12*).

Primary structure of the nucleic acids

The primary structure of the nucleic acids is their sequence of nucleotides (Information box 2.9). Nucleoside phosphates are linked by the phosphodiester bonds between the sugar and the phosphate groups (*Fig. 2.15*): the hydroxyl group at C3 of the sugar links with the phosphate on C5 on the next nucleotide. The bases are attached to C1 of each sugar.

The substrates for nucleic acid synthesis are nucleotide triphosphates. Pyrophosphate (PPi) is released during their incorporation into the nascent nucleic acid chain. The nucleotide chain is assembled in the 5′→3′ direction.

Secondary structure of DNA

The secondary structure of DNA is the spatial arrangement of the nucleotide chains. The two strands of the DNA molecule are held together by hydrogen bonds. Its secondary structure is dependent

on the **pairing** between the bases that project from the sides of the sugar–phosphate backbone of each of the two strands.

The pairs of nitrogenous bases, each consisting of a purine and a pyrimidine, are linked by hydrogen bonds. The bases can only form specific pairs (**AT and GC**) (*Fig. 2.16*). The AT pairs form two hydrogen bonds, and the GC pairs have three hydrogen bonds. The two types of pairs are of the same width; therefore, the double strands are held at the same distance apart throughout the length of the molecule. The two complementary strands of the DNA molecule run in opposite directions from one another (they are anti-parallel). One is the 3′→5′ chain and the other is the 5′→3′ chain.

The base pairs can be 'stacked' one on top of another in the core of the DNA molecule. In addition to the hydrogen bonds between the bases, hydrophobic interactions occur 'vertically' between the stacked base pairs.

The double helix

The DNA molecule forms a right-handed helix with 10 base pairs (bp) and 3.4 nm of distance for each turn. As there are no 'vertical' hydrogen bonds between nucleotides, the DNA molecule is not rigid. There are two grooves of different width, namely, the major groove and the minor groove, formed along the outside of the molecule. Most DNA exists in this form, called the B form (see *Fig. 2.15*).

Complementarity of the DNA strands

The two strands of DNA are **complementary**. This means that a particular sequence of bases in one strand is reflected by a specific sequence on the other. For example, the sequence TGCT in one strand would be reflected by the sequence ACGA in the other. Complementarity is fundamental for the generation of identical copies of DNA during its replication.

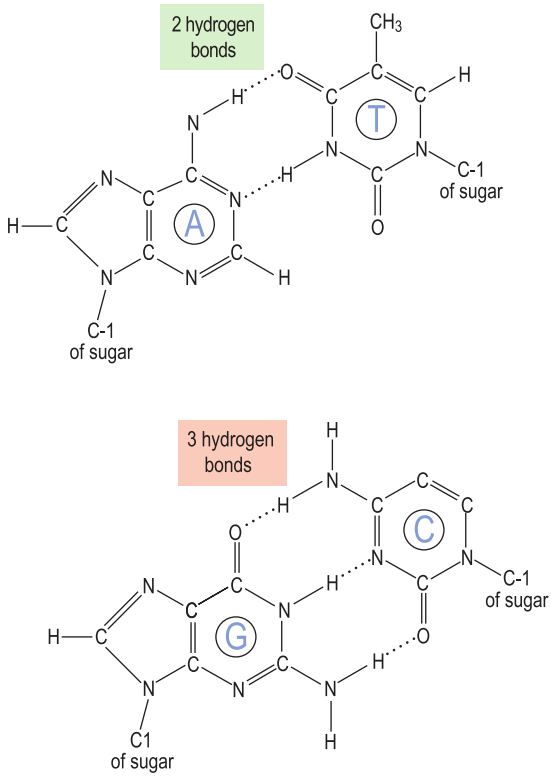


Fig. 2.16 Pairing between adenine (A) and thymine (T) (or uracil (U)) and between guanine (G) and cytosine (C).

Tertiary structure of DNA: chromatin and the nucleosomes

The largest human DNA molecules are nearly 10 cm long when fully extended and have 2–3 × 10^8 nucleotide pairs. These large molecules must be packed into the cell nucleus and must uncoil easily for copying segments.

In the nucleus, DNA is combined with proteins, forming **chromatin**. Chromatin consists of a series of structures called **nucleosomes** (*Fig. 2.17*). Each nucleosome contains a protein core consisting of histones, around which loops of DNA are wrapped, with a piece of linker DNA between each nucleosome (see *Fig. 2.17A*). The nucleosome string is then folded again, forming a supercoiled structure (see *Fig. 2.17B, C*).

The genetic material in eukaryotes is organised into chromosomes, each of which contains a single DNA molecule. There are 23 pairs of chromosomes in human cells, each characterised by a specific size and shape. When cells are about to divide, the chromatin concentrates still further (see *Fig. 2.17D*) and the chromosomes become visible under a light microscope.

Mitochondrial DNA

Mitochondria contain their own mitochondrial DNA (mtDNA), as well as the machinery to make proteins. mtDNA is present in the mitochondrial matrix and codes for some of the molecules required by the mitochondrion, including its own rRNA and tRNA. The mtDNA is a circular molecule similar to the bacterial DNA.

DNA replication

DNA replication ensures the continuity of the genetic material through generations. The DNA molecule is copied during the interphase, which precedes the cell division (Ch. 5). During replication, the DNA unwinds (the hydrogen bonds holding the double helix together break) and a new complementary strand is synthesised against each of the two parent strands. Thus, each new DNA molecule contains one new and one old strand: the replication is **semi-conservative**.

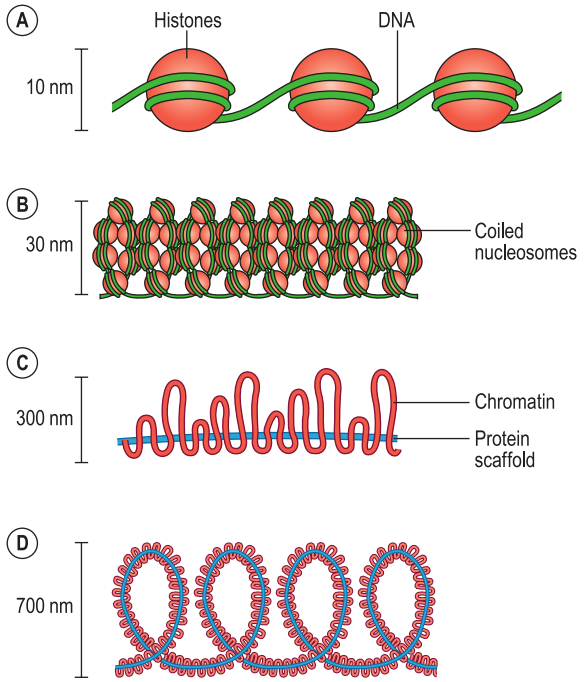


Fig. 2.17 (A–D) Tertiary structure of DNA.

The double-stranded DNA is unwound by the action of a number of enzymes, including helicases, which separate the strands at a location called the **origin of replication.** In eukaryotes, there are many such replication sites. The separation is maintained by proteins that bind to the separate strands, known as the single-strand binding proteins *(Fig. 2.18)*. The two DNA strands are called the sense and the antisense strand.

The replication process is 'primed' by the binding of the priming RNA polymerase called primase. The new antisense strand is then synthesised by **DNA polymerase III**; the RNA primer is subsequently removed. Replication can proceed in both directions away from the primer, forming two so-called **replication forks.** Another enzyme, topoisomerase, travels on the DNA ahead of the replication fork and maintains the strand in the non-coiled state.

The DNA polymerase travels along the leading strand in the 3′ to 5′ direction, adding new nucleotides to the 3′ end of the new strand. The other strand, known as the lagging strand, is replicated in fragments known as Okazaki fragments. At the end of the process, these fragments are joined together by DNA ligase. The DNA polymerase III is kept in place on the replication fork by a protein called the **sliding clamp.**

DNA proofing and repair

DNA polymerase III controls the accuracy of the replication by checking that the bases on the original strand and the new strand are complementary. If they are not, then the incorrect nucleotide is excised and replaced (Clinical box 2.4). The error rate in the replication is very low. It is estimated that the error rate before proofreading is approximately 1 in 10,000, but with the replacement of wrong nucleotides, this falls to 1 in 10^8–10^{12}. However, it does not eliminate all errors.

There are two other major mechanisms of continuous repair of the DNA. They have been studied extensively in bacteria. The repair process is initiated by **DNA glycosylases**, which are enzymes that remove nitrogenous bases but leave the sugar-phosphate chain intact. Larger segments are repaired by removal of the damaged section by **DNA helicase** and its replacement by DNA polymerase. The new fragment is attached to the DNA strand by **DNA ligase.**

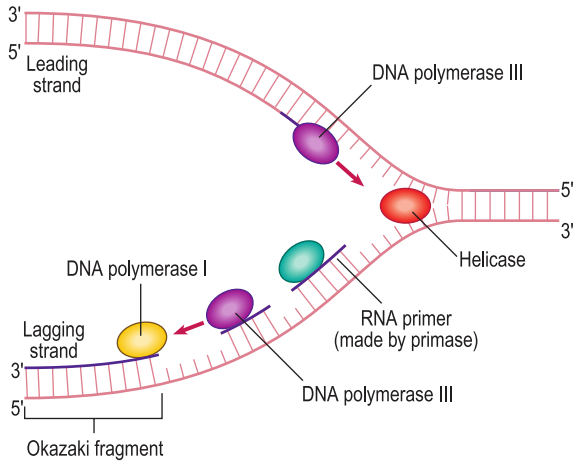


Fig. 2.18 DNA replication. See text for details.

Damage to DNA

The DNA of a cell is subjected to damage by high-energy radiation, the free radicals, and many mutagenic chemicals. These processes have been estimated to produce up to 60,000 base modifications per day (see Information box 2.9). Many of the modifications produce no discernible effect, possibly because they occur in non-essential regions of DNA or do not change the activity of the gene product. Some, however, are detrimental to the cell. Mutations in the somatic cells can lead to uncontrolled proliferation of cells, and thus to the development of cancers. Mutations in gametes can lead to inherited metabolic errors.

Ribonucleic acids

RNA is found in all cells (except red blood cells), in eukaryotes, prokaryotes and some viruses. Several types of RNA are present in eukaryotic cells:

- **Messenger RNA** (mRNA) serves as a template for protein synthesis and is synthesised during transcription. It is a copy of the transcribed gene. It subsequently moves from the nucleus to the cytoplasm, where it associates with protein-synthesising ribosomes.
- **Transfer RNAs** (tRNAs) are a range of relative molecules that bind specific amino acids in the cytoplasm and move them to the ribosome for protein synthesis. Approximately 50 different tRNA molecules are produced in animal cells, each between 70 and 90 nucleotides long. Each tRNA possesses the anticodon sequence for a particular amino acid, which reads the codon on the mRNA molecule.
- **Ribosomal RNA** (rRNA) together with proteins contributes to the structure of the ribosome. The ribosome in eukaryotes has two subunits, 60S and 40S, formed by roughly equal proportion of RNA and protein.
- **Small nuclear RNAs** (snRNAs) and small nucleolar RNAs (snoRNAs) are involved in mRNA and rRNA processing and maturation within the nucleus, respectively. Small cytoplasmic RNAs (scRNAs) are components of the spliceosome, a complex containing small ribonucleoprotein particles (snRNPs)

i Clinical box 2.4 AZT prevents the replication of HIV because HIV does not proofread its DNA

The human immunodeficiency virus (HIV) is an RNA virus that uses the enzyme reverse transcriptase, packaged in the viral particles, to copy its RNA into the host genome. Reverse transcriptase, however, lacks the proofreading functions of DNA polymerase. This characteristic is exploited in the use of the drug azidothymidine (AZT). AZT is a nucleoside and is metabolised to an analogue of thymidine triphosphate. The latter is incorporated into the elongating DNA chain by reverse transcriptase and is not corrected. Because the azido group cannot form a phosphodiester bond with the next nucleotide, this prevents further elongation of the chain, blocking replication of the virus.

i Information box 2.9 Deletion of a single base from DNA causes frameshift mutation

If a single nucleotide is removed from a gene, the entire subsequent triplet sequence of the genetic code will change (this is known as a frameshift mutation). If the deletion is near the end of the coding sequence, then the protein may be nearly normal, but if it is close to the beginning, then the changes can be major. Normal haemoglobin is composed of two polypeptide chains, α- and β-globin. β-Thalassaemia is an inherited anaemia caused by a point mutation in the β-globin chain. The resulting frameshift produces a protein that is unstable. The globin chains in red cells can precipitate, causing haemolysis, and a severe anaemia, in individuals who have inherited the abnormality from both parents.

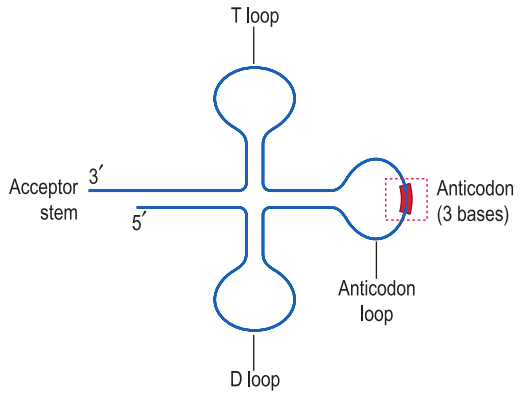


Fig. 2.19 General structure of tRNA.

Clinical box 2.5 Errors in amino acid metabolism

When the metabolism pathway of an amino acid is disrupted, its precursors may accumulate in the blood or appear in the urine (normally, most amino acids are reabsorbed in the kidney proximal tubules). Thus, measurements of serum and urine levels have diagnostic significance in the detection of inborn errors of metabolism.

For example, in phenylketonuria (PKU), where there is a deficiency of phenylalanine hydroxylase (or two other enzymes in the same pathway), and there is accumulation of the intermediate metabolites phenylacetate, phenylacetylglutamine and phenylpyruvate. Untreated PKU leads to brain damage (Chs 3 and 5).

that removes intron sequences from precursor mRNA. The splicing apparatus contains RNA and proteins, and it assembles on mRNA.

- **Micro RNAs (miRNAs)** including small interfering RNA (siRNA) participate in the regulation of gene expression.

miRNAs are single-stranded non-coding RNAs 20–25 nucleotides long. They downregulate gene expression through mRNA cleavage and deadenylation, as well as through repression of the translation which involves a multiprotein RNA-induced silencing complex (RISC). siRNA is double stranded.

Secondary structure of RNA

Most RNA molecules are single-stranded. Parts of the molecule can form loops or hairpins between complementary regions of its single strand. The tRNAs have a cloverleaf shape with four stem-loops, each of which forms a short double helix ***(Fig. 2.19)***.

AMINO ACIDS

Amino acids are the structural units of proteins. They contain short hydrocarbon chains of variable length, oxygen atoms and nitrogen. Apart from being incorporated into proteins, amino acids are precursors of nucleosides, neurotransmitters and haem, among many other molecules (Ch. 3) (Clinical box 2.5).

Neurotransmitters are released from nerve endings into the synaptic space. They are small molecules, such as acetylcholine, epinephrine (adrenaline) or norepinephrine (noradrenaline).

Amino acids can also be converted into carbohydrates through gluconeogenesis and, therefore, are also a potential energy source. In contrast to carbohydrates and fats, amino acids are not stored in the body. The amino acid 'reserve' is the muscle mass. Thus, when the fuel supply is short, muscle proteins are used up to synthesise glucose. This is why starvation and chronic disease are associated with muscle wasting.

Table 2.7 Amino acids found in human proteins

Amino acid	Abbreviation	
Glycine	Gly	G
Alanine	Ala	A
Valine	Val	V
Leucine	Leu	L
Isoleucine	Ile	I
Proline	Pro	P
Phenylalanine	Phe	F
Tyrosine	Tyr	Y
Tryptophan	Trp	W
Serine	Ser	S
Threonine	Thr	T
Asparagine	Asn	N
Glutamine	Gln	Q
Glutamic acid	Glu	E
Aspartic acid	Asp	D
Lysine	Lys	K
Arginine	Arg	R
Histidine	His	H
Cysteine	Cys	C
Methionine	Met	M

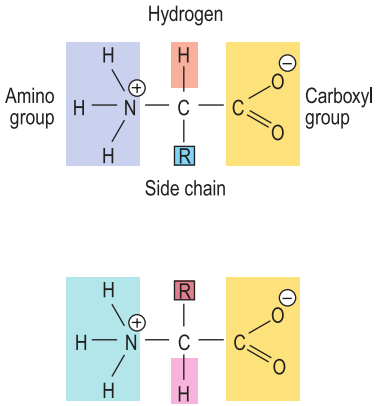


Fig. 2.20 **Two stereoisomers of an amino acid with ionised amino ($-NH_3^+$) and carboxyl ($-COO^-$) groups.** The two molecules are mirror images of each other and cannot be superimposed.

In humans, 20 amino acids are used in protein synthesis (Table 2.7). Of these, 8 (9 in infants) are known as **essential amino acids** because they must be supplied by the diet (Ch. 16). By convention, amino acids are designated by two sets of letter codes: a set of three-letter abbreviations and a set of single-letter abbreviations.

Structure of the amino acids

The structure of an amino acid includes a chiral carbon atom (the α-carbon) attached to four different chemical groups ***(Fig. 2.20)***: the hydrogen atom (–H), the amino group ($-NH_2$), the carboxyl group (–COOH) and a side chain (–R). The chirality means that amino

acids (except glycine) can exist as two stereoisomers (enantiomers) named D and L (*Fig. 2.20*). With very few exceptions, only the L-amino acids are incorporated into proteins. D-amino acids are present in bacterial walls.

Classification of amino acids

The side chains of amino acids vary in length and complexity, and determine the characteristics of these molecules. They range from a single hydrogen atom (in glycine) to the relatively complex cyclic side chains of tryptophan and the branched molecules of lysine and valine (*Fig. 2.21*). Amino acids are classified into the following groups:

- **Non-polar aliphatic**: amino acids with linear hydrocarbon side chains.
- **Non-polar aromatic**: amino acids with side chains containing a ring structure.
- **Polar neutral**: amino acids with side chains containing polar hydroxyl or amide groups. The hydroxyl groups of serine,

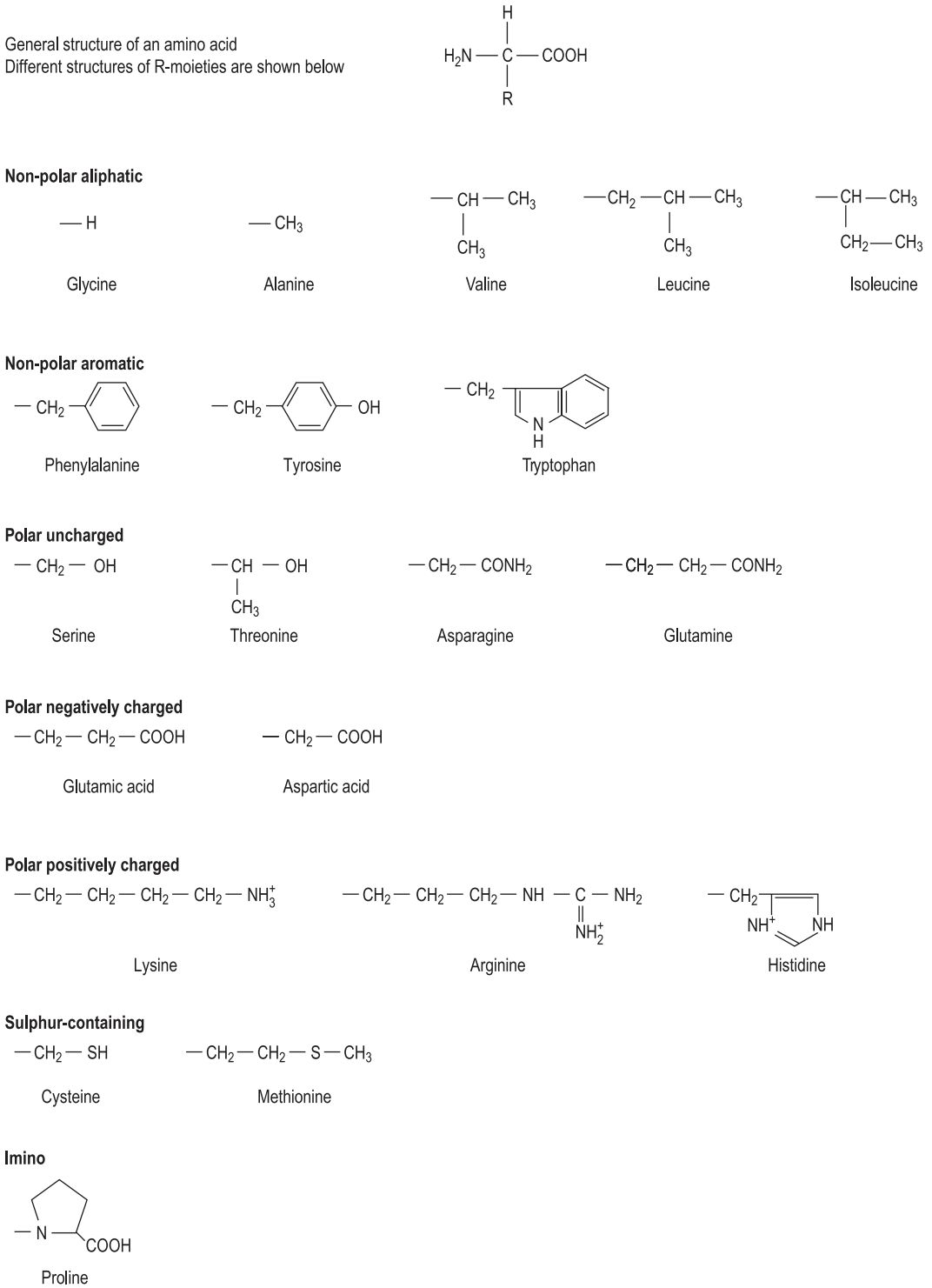


Fig. 2.21 Side chains of the common amino acids.