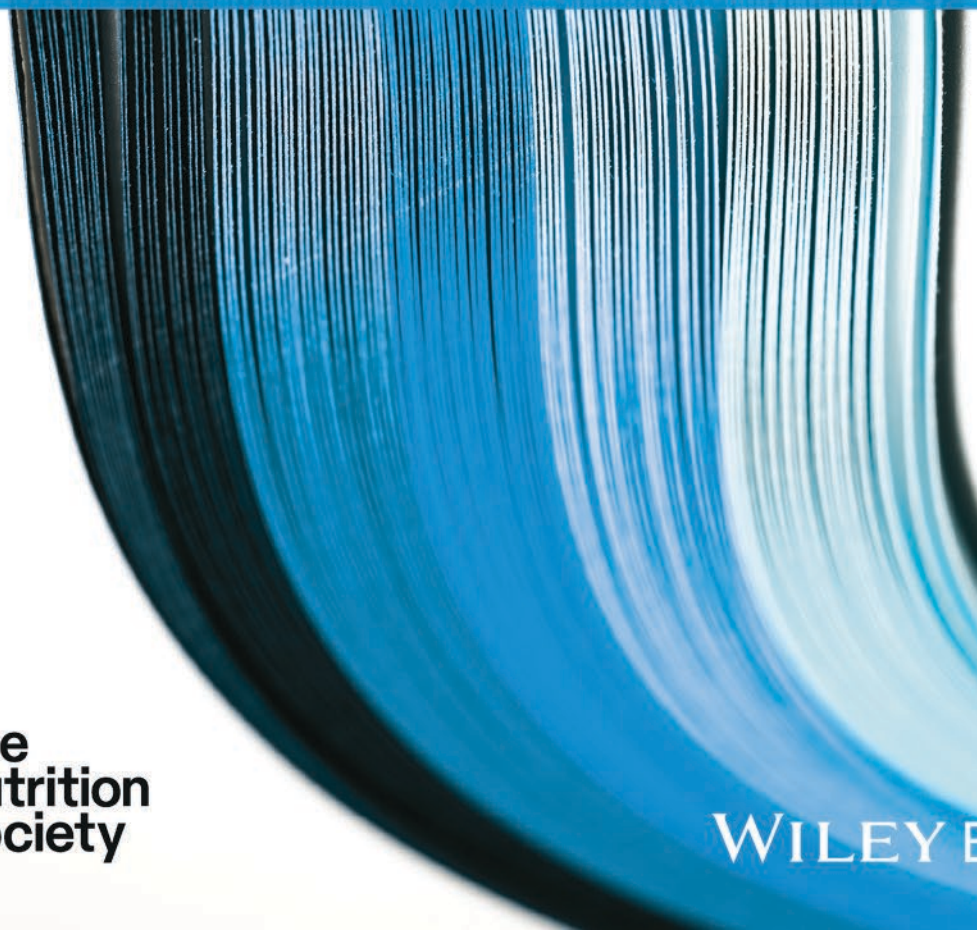
THE NUTRITION SOCIETY TEXTBOOK SERIES

Nutrition and Metabolism

Third Edition

Edited by **Helen M. Roche, Ian A. Macdonald, Annemie M.W.J. Schols, Susan A. Lanham-New**



The Nutrition Society

WILEY Blackwell

Nutrition and Metabolism

The Nutrition Society Textbook Series

Sport and Exercise Nutrition, 2nd Edition
Susan A. Lanham-New, Samantha Stear, Susan Shirreffs, Adam Collins, Richard Budgett

Clinical Nutrition, 2nd Edition
Marinos Elia, Olle Ljungqvist, Rebecca J. Stratton, Susan A. Lanham-New

Nutrition Research Methodologies, 1st Edition
Julie A. Lovegrove, Leanne Hodson, Sangita Sharma, Susan A. Lanham-New, Lord John Krebs

Public Health Nutrition, 2nd Edition
Judith L. Buttriss, Ailsa A. Welch, John M. Kearney, Susan A. Lanham-New

Introduction to Human Nutrition, 3rd Edition
Susan A. Lanham-New, Thomas R. Hill, Alison M. Gallagher, Hester H. Vorster

Nutrition and Metabolism, 3rd Edition
Helen M. Roche, Ian A. MacDonald, Annemie M.W.J. Schols, Susan A. Lanham-New

Nutrition and Metabolism

Third Edition

Edited on behalf of The Nutrition Society by

Helen M. Roche
Full Professor of Nutrigenomics
Director, UCD Conway Institute
University College Dublin
Dublin, Ireland

Ian A. Macdonald
Emeritus Professor of Metabolic Physiology
School of Life Sciences, Faculty of Medicine and Health Sciences
University of Nottingham
Nottingham, UK

Annemie M.W.J. Schols
Professor of Nutrition and Metabolism in Chronic Diseases
Dean Faculty of Health, Medicine and Lifesciences, Maastricht University
Vice-president Executive Board, Maastricht University Medical Centre
Maastricht University
Maastricht, The Netherlands

Susan A. Lanham-New
Professor of Human Nutrition & Head, Nutritional Sciences Department
School of Biosciences & Medicine, Faculty of Health & Medical Sciences
University of Surrey
Guildford, UK

WILEY Blackwell

The Nutrition Society

This edition first published 2024

Edition History
1e, 2002; 2e, 2009 The Nutrition Society

Registered Offices
John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial Office
9600 Garsington Road, Oxford, OX4 2DQ, UK

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Wiley also publishes its books in a variety of electronic formats and by print-on-demand. Some content that appears in standard print versions of this book may not be available in other formats.

Library of Congress Cataloging-in-Publication Data
Names: Roche, Helen M., editor. | MacDonald, Ian A., editor. | Schols, Annemie M.W.J., editor. | Lanham-New, Susan A., editor.
Title: Nutrition and metabolism / edited on behalf of The Nutrition Society by Helen M. Roche, Full Professor of Nutrigenomics, Director, UCD Conway Institute, University College Dublin, Dublin, Ireland, Ian A. Macdonald, Emeritus Professor of Metabolic Physiology, School of Life Sciences, Faculty of Medicine and Health Sciences, University of Nottingham, Nottingham, UK, Annemie M.W.J. Schols, Professor of Nutrition and Metabolism in chronic diseases, Dean, Faculty of Health, Medicine and Lifesciences, Maastricht University, Vice-president Executive Board, Maastricht University Medical Centre, Maastricht University, Maastricht, The Netherlands, Susan A. Lanham-New, Professor of Human Nutrition & Head, Nutritional Sciences Department, School of Biosciences & Medicine, Faculty of Health & Medical Sciences, University of Surrey, Guildford, UK.
Description: Third edition. | Hoboken, NJ, USA : John Wiley & Sons Ltd, 2024. | Series: The nutrition society textbook series | Revised edition of: Nutrition and metabolism / edited on behalf of the Nutrition Society by Susan A. Lanham-New, Helen M. Roche, Ian MacDonald. 2nd ed. 2011. | Includes bibliographical references and index.
Identifiers: LCCN 2022054800 (print) | LCCN 2022054801 (ebook) | ISBN 9781119840237 (Paperback) | ISBN 9781119840244 (ePDF) | ISBN 9781119840251 (ePUB)
Subjects: LCSH: Nutrition. | Metabolism.
Classification: LCC QP141 .N7768 2023 (print) | LCC QP141 (ebook) | DDC 613.2--dc23/eng/20221215
LC record available at https://lccn.loc.gov/2022054800
LC ebook record available at https://lccn.loc.gov/2022054801

Cover Design: Wiley
Cover Images: © jxfzsy/Getty Images

Set in 10/12pt Minion by 9.5/12.5pt STIXTwoText by Integra Software Services Pvt. Ltd, Pondicherry, India

Contents

Contributors

Abayomi O. Akanji
Quinnipiac University, USA

Ellen E. Blaak
Maastricht University, the Netherlands

Nicholas A. Burd
University of Illinois at Urbana-Champaign, USA

Louise M. Burke
Australian Catholic University, Australia

Philip C. Calder
University of Southampton, UK

Tyler A. Churchward-Venne
McGill University, Quebec, Canada

Emilio Martínez de Victoria
Universidad de Granada, Spain

Connor M. Delahunty
CSIRO, Canberra, Austrailia

Emily J. Ferguson
Queen's University, Canada

Graham Finlayson
University of Leeds, UK

Albert Flynn
University College Cork, Ireland

Keith N. Frayn
University of Oxford, UK

Ángel Gil
Universidad de Granada, Spain

Gijs H. Goossens
Maastricht University, the Netherlands

Wendy L. Hall
King's College London, UK

Remco C. Havermans
Maastricht University, the Netherlands

Phoebe Hodges
Queen Mary University of London, UK

Mark Hopkins
University of Leeds, UK

Asker E. Jeukendrup
Loughborough University, UK

Paul Kelly
Queen Mary University of London, UK

Mara Ioana Iesanu
Carol Davila University of Medicine and Pharmacy, Romania

J.A. Lovegrove
University of Reading, UK

Ansuyah Magan
University of the Witwatersrand, Johannesburg, South Africa

Mariano Mañas
Universidad de Granada, Spain

John C. Mathers
University of Newcastle, UK

Chris McGlory
Queen's University, Canada

Michelle K. McGuire
University of Idaho, USA

Ronald P. Mensink
Maastricht University, the Netherlands

Jennifer L. Miles-Chan
University of Auckland – Waipapa Taumata Rau, New Zealand

Siobhain M. O'Mahony
University College Cork, Ireland

Sharleen L. O'Reilly
University College Dublin, Ireland

John M. Pettifor
University of the Witwatersrand, New Zealand

Herman E. Popeijus
Maastricht University, the Netherlands

Ann Prentice
University of Cambridge, UK

Harriët Schellekens
University College Cork, Ireland

Jeremy P.E. Spencer
University of Reading, UK

Mark L. Wahlqvist
Monash University, Australia

Kate A. Ward
University of Southampton, UK

Gary Williamson
Monash University, Australia

Jayne V. Woodside
Queen's University Belfast, UK

María D. Yago
Universidad de Granada, Spain

Parveen Yaqoob
University of Reading, UK

Series Foreword – Nutrition and Metabolism 3rd Edition

In 1941, a group of leading physiologists, biochemists and medical scientists recognised that the emerging discipline of nutrition needed its own Learned Society and The Nutrition Society was established. The original mission remains "*to advance the scientific study of nutrition and its application to the maintenance of human and animal health*". The Society is now one of the world's largest Learned Societies for nutrition, with over 2,500 members. You can learn more about it and how to become a member by visiting www.nutritionsociety.org. The Society's first journal, *The Proceedings of the Nutrition Society*, published in 1944, records the scientific presentations made at the time. Shortly afterwards, 1947, scientists worldwide established the *British Journal of Nutrition* to publish primary research on human and animal nutrition. Recognising the needs of students and their teachers for authoritative reviews on topical issues in nutrition, the Society began publishing *Nutrition Research Reviews* in 1988. *Public Health Nutrition*, the first international journal dedicated to this critical and growing area, was launched in 1998 and made open access in January 2022. The Society's first open-access journal, the *Journal of Nutritional Science*, was launched in 2012. The Society is constantly evolving in response to emerging areas of nutritional science. It has most recently launched the journal *Gut Microbiome*, an open-access journal published in partnership with Cambridge University Press.

Now 25 years old, the Nutrition Society Textbook Series was first established by Professor Michael Gibney (University College Dublin) in 1998. Under the direction of the second Editor-in-Chief, Professor Susan Lanham-New (University of Surrey), the Series continues to be extraordinarily successful for the Society. This series of Nutrition textbooks is designed for use worldwide. This has been achieved by translating the Textbook Series into many different languages, including Spanish, Greek, Portuguese, Italian and Indonesian. The success of the Textbook Series is a tribute to the quality of the authorship and the value placed on them in the UK and Worldwide as a core educational tool and a resource for practitioners. *Nutrition and Metabolism* is an important textbook, with more than 10,000 copies sold to date and with it now transcending into its 3rd Edition.

Therefore, writing the Foreword for this book gives me great pleasure. As President of the Royal College of General Practitioners, I know how important it is to understand the scientific basics of nutrition in the context of a systems and health approach. This Textbook brings together science and the practical application of methodologies in nutrition and is a most valuable resource to all those working in the field.

Professor Dame Clare Gerada, DBE, PRCGP FRCPsych FRCP (Hons) FRCGP

President of the Royal College of General Practitioners

Series Editor's Preface

I am absolutely delighted in my capacity as Editor-in-Chief (E-i-C) of the Nutrition Society Textbook Series to introduce the 3rd Edition of *Nutrition and Metabolism (N&M3e)*. The production of this Third Edition represents a significant milestone for the Textbook Series, given that it is now exactly 25 years on since the production of the first textbook, *Introduction to Human Nutrition*, and 22 years since the production of the 1st Edition of N&M.

The Team of *Nutrition and Metabolism* 3rd Edition, namely Professor Helen Roche (University College Dublin), Professor Ian Macdonald (University of Nottingham) and Professor Annemie Schols (Maastricht University) have been meticulous in ensuring that each chapter is updated & accurate, and to ensuring that new aspects of N&M3e are also brought into the book. N&M3e comprises of a total of 19 chapters, each with their own unique summary of the take home messages. How indebted we are to have so many experts in the field who have written chapters to make N&M3e a complete and thorough review of the area of Metabolic Science - a must read!

N&M3e is intended for those with an interest in nutrition and metabolic science whether they are nutritionists, food scientists, dietitians, medics, nursing staff or other allied health professionals. We hope that both undergraduate and postgraduate students will find the book of great help with their respective studies and that the book will really put nutrition and metabolic science as a *discipline* into context.

It is a great honour for our 3rd Edition of N&M to have the Foreword written by Professor Dame Clare Gerada, DBE, PRCGP FRCPsych FRCP (Hons) FRCGP, and we are most grateful for her support of our work at the Society, particularly in her role as President of the Royal College of General Practitioners. We are also most grateful to the following individuals for their support and most generous Forewords in *Introduction to Human Nutrition3e, Public Health Nutrition2e, Sport and Exercise Nutrition1e, Clinical Nutrition2e* and *Nutrition Research Methodologies1e*; namely – the late The Earl of Selbourne, Her Royal Highness The Princess Royal; Professor Richard Budgett OBE, Chief Medical Officer for the London 2012 Olympic and Paralympic Games and now Medical and Scientific Director at the International Olympic Committee (IOC); Dame Sally Davies, former Chief Medical Officer (CMO) for England, and the UK Government's Principal Medical Adviser; Professor Lord John Krebs, Principal, Jesus College, University of Oxford and our first Chairman of the UK Food Standards Agency. We are now planning ahead with respect to the production of the 3rd Edition of *Clinical Nutrition* and 2nd Edition of *Sport and Exercise Nutrition* as well as bringing a *seventh* book to the Textbook Series, *Animal Nutrition*, published in collaboration with the British Society of Animal Sciences.

The Society is most grateful to the Textbook publishers, Wiley-Blackwell, for their continued help with the production of the textbook and in particular: Tom Marriott - Commissioning Editor; Charlie Hamlyn and Vallikannu Narayanan - Managing Editors; Durgadevi Shanmugasundaram - Content Refinement Specialist. In addition, I would like to acknowledge formally my great personal appreciation to: Professor G.Q. Max Lu AO, DL, FREng, FAA, FTSE, FIChemE, FRSC, FCAS, FNAI, Vice-Chancellor & President of the University of Surrey; Professor Tim Dunne, Provost & Executive Vice-President of the University of Surrey; Professor Paul Townsend, Pro-Vice-Chancellor and Executive Dean of the Faculty of Health and Medical Sciences and Professor Roberto La Ragione BSc (Hons) MSc PhD FRSB CBiol FIBMS CSci AECVM FRCPath HonAssocRCVS, Head of the School of Biosciences, University of Surrey, for their respective great encouragement of the nutritional sciences field in general, especially in light of Surrey's success in the 2017/2018 Queen's Anniversary Prize for our work in *Food and Nutrition for Health*, and for their support of the Textbook Series production in particular.

Sincerest appreciation indeed to the Nutrition Society President, Professor Mary Ward (Ulster

University), and our past President, Professor Julie Lovegrove (University of Reading) for their great support and belief in the Textbook Series. With special thanks to Honorary Publications Officer Professor Jayne Woodside (Queen's University Belfast) for her support of the Textbook Series and for being such a great sounding board, and to past-Honorary Publications Officer, Professor Paul Trayhurn for his wise counsel to me during the six years we worked together on the Textbooks. And finally an enormous thank you indeed to: Mark Hollingsworth MBA, FInstLM, Chief Executive Officer of the Nutrition Society for his unstinting support of the Textbook Series, to Cassandra Ellis, Science Director and Deputy Editor, for her pivotal continued contribution to the development of the Textbook Series, and to Caroline Roberts, (formerly Science Communications Manager), whose excellent support made this latest edition possible.

Finally, as I always write and mean with absolute sincerity, how the Textbook Series is indebted to the forward thinking focus that Professor Michael Gibney (University College Dublin) had at that time of the Textbook Series development. It shall remain for always, such a tremendous privilege for me to continue to follow in his footsteps as the second E-i-C.

I really hope that you will find the textbook a great resource of information and inspiration ……please enjoy, and with so many grateful thanks to all those who made it happen!

With my warmest of wishes indeed

Professor Susan A. Lanham-New
RNutr, FAfN
E-i-C, Nutrition Society Textbook Series
Professor of Human Nutrition and Head,
Department of Nutrition, Food &
Exercise Sciences
School of Biosciences,
Faculty of Health and Medical Sciences
University of Surrey

About the Companion Website

www.wiley.com/go/nutrition/metabolism3e

- Multiple choice questions
- Short answer questions
- Essay questions

1
Core Concepts of Nutrition

Ian A. Macdonald and Annemie M.W.J. Schols

Key messages

- The change in body reserves or stores of a nutrient is the difference between the intake of that nutrient and the body's utilisation of that nutrient. The time-frame necessary to assess the body's balance of a particular nutrient varies from one nutrient to another.
- The concept of turnover can be applied at various levels within the body (molecular, cellular, tissue/organs, whole body).
- The flux of a nutrient through a metabolic pathway is a measure of the rate of activity of the pathway. Flux is not necessarily related to the size of the pool or pathway through which the nutrient or metabolite flows.
- Nutrients and metabolites are present in several pools in the body. The size of these metabolic pools varies substantially for different nutrients/metabolites, and a knowledge of how these pools are interconnected greatly helps us to understand nutrition and metabolism in health and disease.
- The Darwinian theory of evolution implies a capacity to adapt to adverse conditions, including adverse dietary conditions. Many such examples can be cited. Some allow for long-term adaptation and others buy time until better conditions arrive.

1.1 Introduction

Nutrition and metabolism are addressed in this textbook in an integrated fashion. Thus, rather than considering nutrients separately, this book brings together information on macronutrients, energy and substrate metabolism in relation to specific nutritional or disease states or topics (e.g. undernutrition, overnutrition, cardiovascular disease). Before considering these topics in detail, it is necessary to outline the core concepts of nutritional metabolism. The core concepts covered in this chapter are nutrient balance, turnover and flux, metabolic pools and adaptation to altered nutrient supply.

1.2 Balance

As discussed in Chapters 3 and 4, nutrient balance must be considered separately from the concepts of metabolic equilibrium or steady state. In this chapter, the concept of balance is considered in the context of the classic meaning of that term, i.e. the long-term sum of all the forces of metabolic equilibrium for a given nutrient.

The concept of nutrient balance essentially restates the law of conservation of mass in terms of nutrient exchange in the body. It has become common practice to refer to the content of the nutrient within the body as a 'store' but in many cases this is not appropriate and the term 'reserves' is better. Thus, the idea of nutrient balance is summarised by the equation:

$$\begin{bmatrix}\text{nutrient}\\ \text{intake}\end{bmatrix} - \begin{bmatrix}\text{nutrient}\\ \text{utilisation}\end{bmatrix} = \begin{bmatrix}\text{change in body}\\ \text{nutrient reserves}\end{bmatrix}$$

The above equation can have three outcomes:

- *Zero balance* (or nutrient balance): Intake matches utilisation and reserves remain constant
- *Positive balance* (or positive imbalance): Intake exceeds utilisation and reserves expand

Nutrition and Metabolism, Third Edition. Edited on behalf of The Nutrition Society by Helen M. Roche, Ian A. Macdonald, Annemie M.W.J. Schols and Susan A. Lanham-New.

Companion Website: www.wiley.com/go/nutrition/metabolism3e

- *Negative balance* (or negative imbalance): Utilisation exceeds intake and reserves become depleted.

In relation to macronutrient metabolism, the concept of balance is most often applied to protein (nitrogen) and to energy. However, many research studies now subdivide energy into the three macronutrients and consider fat, carbohydrate and protein balance separately. This separation of the macronutrients is valuable in conditions of altered dietary composition (e.g. low-carbohydrate diets) where a state of energy balance might exist over a few days but be the result of negative carbohydrate balance (using the body's glycogen reserves to satisfy the brain's requirement for glucose) matched in energy terms by positive fat balance.

Balance is a function not only of nutrient intake but also of metabolic requirements and metabolically elevated requirements. Fat balance is generally driven by periods where energy intake exceeds energy expenditure (positive energy balance) and by periods when intakes are maintained below energy expenditure, such as in dieting, acute and chronic disease, hypoxia (negative energy balance). However, nutrient balance can also be driven by metabolic regulators through hormones or cytokines. For example, the dominance of growth hormone during childhood ensures positive energy and nutrient balance. In pregnancy, a wide range of hormones lead to a positive balance of all nutrients in the overall placental, fetal and maternal tissues, although this may be associated with a redistribution of some nutrient reserves from the mother to the fetus (see Chapter 6). By contrast, severe trauma or illness will dramatically induce energy and protein losses, an event due to elevated metabolic requirements unrelated to eating patterns.

Balance is not something to be thought of in the short term. Following each meal, there is either storage of absorbed nutrients (triacylglycerol [TAG] in adipose tissue or glucose in glycogen) or a cessation of nutrient losses (breakdown of stored TAG to non-esterified fatty acids or amino acid conversion to glucose via gluconeogenesis). As the period of postprandial metabolism is extended, the recently stored nutrients are drawn upon and the catabolic state commences again. This is best reflected in the high glucagon to insulin ratio in the fasted state before the meal and the opposite high insulin to glucagon ratio during the meal and immediate postprandial period. However, when balance is measured over a sufficient period, which varies from nutrient to nutrient, a stable pattern can be seen: zero, positive or negative (Figure 1.1).

It is critically important with respect to both obesity and malnutrition that the concept of balance is correctly considered. While at some stage energy balance must have been positive to reach an overweight or obese stage, once attained most people sustain a stable weight over quite long periods.

In the context of the present chapter, it is worth reflecting on the reasons why the time taken to assess energy balance correctly varies for different nutrients. Because of differences in capacity and mobilisation as summarised below and explained in Chapters 4 and 12, calcium

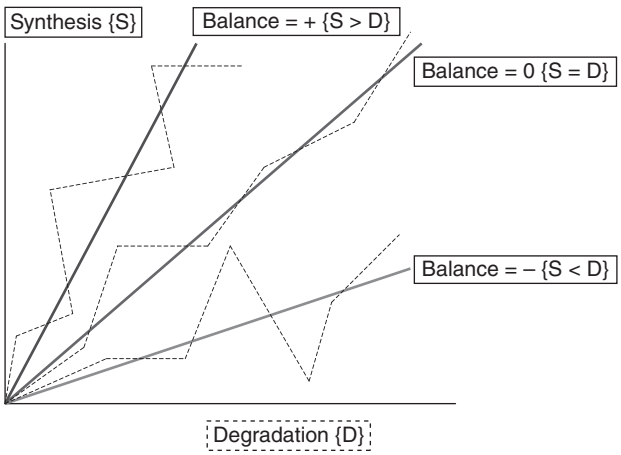


Figure 1.1 Positive, zero and negative nutrient balance over time with fluctuations upwards and downwards within that time.

balance, for example, will require months of equilibrium while fat balance could be equilibrated in days or at most a few weeks.

Fat and adipose tissue (Chapter 4)

- There is a very large capacity to vary the body's pool of adipose tissue. One can double or halve the level of the fat reserves in the body.
- The capacity to vary the level of TAG in blood en route to adipose tissue can vary.
- Almost all of the TAG reserves in adipose tissue are exchangeable.

Calcium and bone (Chapter 12)

- The human being must maintain a large skeleton as the scaffold on which the musculature and organs are held.
- There is a very strict limit to the level of calcium that can be transported in blood. Excess or insufficient plasma calcium levels influence neural function and muscle function, since calcium is a major component of muscle and nerve function.
- Only a small fraction (the miscible pool) of bone is available for movement into plasma.

1.3 Turnover

Although the composition of the body and of the constituents of the blood may appear constant, the component parts are not static. In fact, most metabolic substrates are continually being utilised and replaced (i.e. they turn over). This process of turnover is well illustrated by considering protein metabolism in the body. Daily adult dietary protein intakes are in the region of 50–100 g and the rates of urinary excretion of nitrogen match the protein intake. However, isotopically derived rates of protein degradation indicate that approximately 350 g is broken down per day. This is matched by an equivalent amount of protein synthesis per day, with most of this synthesis representing turnover of substrate (i.e. degradation and resynthesis) rather than being derived de novo from dietary protein (see Chapter 5).

Similar metabolic turnover occurs with other nutrients; glucose is a good example, with a relatively constant blood glucose concentration arising from a match between production by the liver and utilisation by the glucose-dependent metabolic tissues (see Chapter 4).

The concept of turnover can be applied at various levels within the body (molecular, cellular, tissue/organs, whole body). Thus, within a cell, the concentration of adenosine triphosphate (ATP) remains relatively constant, with utilisation being matched by synthesis. Within most tissues and organs, there is a continuous turnover of cells, with degradation and death of some cells matched by the production of new ones. Some cells, such as red blood cells, have a long lifespan (c. 120 days) while others, such as platelets, turn over in a matter of 1–2 days. In the case of proteins, those with very short half-lives have amino acid sequences that favour rapid proteolysis by the range of enzymes designed to hydrolyse proteins. Equally, those with longer half-lives have a more proteolytic-resistant structure.

A major advantage of this process of turnover is that the body is able to respond rapidly to a change in metabolic state by altering both synthesis and degradation to achieve the necessary response. One consequence of this turnover is the high energy cost of continuing synthesis. There is also the potential for nutrient imbalance and metabolic dysfunction if the rates of synthesis and degradation do not match.

The consequences of a reduction in substrate synthesis will vary between the nutrients, depending on the half-life of the nutrient. The half-life is the time taken for half of the substrate to be used up, and is dependent on the rate of utilisation of the nutrient. Thus, if synthesis of a nutrient with a short half-life is stopped, the level of that nutrient will fall quickly. By contrast, a nutrient with a long half-life will disappear more slowly.

Since proteins have the most complex structures undergoing very significant turnover, it is worth dwelling on the mechanism of this turnover. Synthesis is fairly straightforward. Each protein has its own gene and the extent to which that gene is expressed will vary according to metabolic needs. In contrast to synthesis, a reasonably small array of lysosomal enzymes is responsible for protein degradation.

1.4 Flux

The flux of a nutrient through a metabolic pathway is a measure of the activity of the pathway. If one considers the flux of glucose from the blood to the tissues, the rate of utilisation is approximately 2 mg/kg body weight per minute at rest.

However, this does not normally lead to a fall in blood glucose because it is balanced by an equivalent rate of glucose production by the liver, so the net flux is zero. This concept of flux can be applied at the cellular, tissue/organ or whole-body level, and can also relate to the conversion of one substrate/nutrient to another (i.e. the movement between metabolic pathways).

Flux is not necessarily related to the size of the pool or pathway through which the nutrient or metabolite flows. For example, the membrane of a cell will have several phospholipids present and each will have some level of arachidonic acid. The rate at which arachidonic acid enters one of the phospholipid pools and exits from that phospholipid pool is often higher in the smaller pools.

1.5 Metabolic pools

An important aspect of metabolism is that the nutrients and metabolites are present in several pools in the body (Figure 1.2). At the simplest level, for a given metabolite there are three pools, which will be illustrated using the role of dietary essential fatty acids in eicosanoid synthesis.

In the *functional pool*, the nutrient/metabolite has a direct involvement in one or more bodily functions. In the chosen example, intracellular free arachidonic acid, released from membrane-bound stores on stimulation with some extracellular signal, is the functional pool. It will be acted on by the key enzyme in eicosanoid synthesis, cyclo-oxygenase.

The *storage pool* provides a buffer of material that can be made available for the functional pool when required. Membrane phospholipids store arachidonic acid in the sn-2 position at quite high concentrations, simply to release this fatty acid when prostaglandin synthesis is needed. In the case of platelets, the eicosanoid thromboxane A_2 is synthesised from arachidonic acid released into the cytoplasm by stimuli such as collagen.

The *precursor pool* provides the substrate from which the nutrient/metabolite can be synthesised. Linoleic acid represents a good example of a precursor pool. It is elongated and desaturated in the liver to yield arachidonic acid. Thus, the hepatic pool of linoleic acid is the precursor pool in this regard.

The precursor, storage and functional pool model does not apply to all nutrients outlined above. The essential nutrients and the minerals and trace elements do not have a precursor pool. Nevertheless, no nutrient exists in a single homogeneous pool and an awareness of the existence of metabolic pools is essential to understand human metabolism. For example, one might expect that a fasted individual would show a fall in all essential nutrient levels in the plasma pool. In many instances, this is not the case initially because of the existence of storage pools, such as liver stores of iron or vitamin A. In the case of folic acid, fasting causes a rise in blood folic acid levels and this is explained by the concept of metabolic pools. A considerable amount of folic acid enters the gut via the bile duct and is reabsorbed further down the digestive tract. Thus, there is an equilibrium between the blood folate pool and the gut folate pool. Fasting stops gall bladder contraction and thus the flow of folate to the gut, and hence folate is redistributed from one pool to another.

Another example of how an awareness of metabolic pools helps us to understand nutrition and metabolism is the intracellular free amino acid pool. This is the functional pool from which protein is synthesised. As this pool is depleted in the process of protein synthesis, it must be repleted, otherwise protein synthesis stops. Moreover, it is not just the intracellular pool of amino acids that matters but the intracellular pool of essential amino acids or, more precisely, the intracellular pool of the most limiting essential amino acid. Calculations show that if the pool of the most limiting amino acid in mammalian cells was not

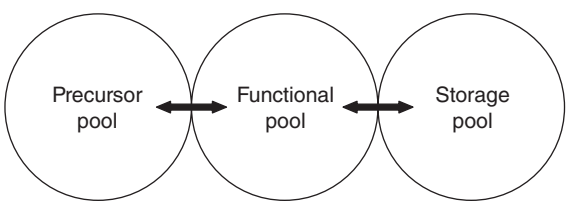


Figure 1.2 The pools in the body in which nutrients and metabolites may exist.

replenished, protein synthesis would cease in under one hour. This highlights the need to transfer the limiting amino acid across the cell membrane, which raises the question of how that pool is repleted. Effectively, it can only be repleted if there is a comparable rate of protein degradation to provide the key amino acid, assuming the balance is zero. Thus, there are links between the protein pool of amino acids and the extra- and intracellular pools of amino acids.

The size of these pools varies substantially for different nutrients and metabolites. When studying the activities of metabolic processes within the body, it is often necessary to measure or estimate the size of the various pools in order to derive quantitative information about the overall rates of the processes. In addition, the actual situation may be more complex than the simple three-pool model described above.

Nutritional assessment often involves some biochemical assessment of nutritional status. Blood is frequently the pool that is sampled and even there, blood can be separated into:

- Erythrocytes, which have a long lifespan and are frequently used to assess folic acid status
- Cells of the immune system, which can be used to measure zinc or ascorbic acid status
- Plasma, which is used to ascertain the levels of many biomarkers
- Fractions of plasma, such as cholesteryl esters used to ascertain long-term intake of polyunsaturated fatty acids.

In addition to sampling blood, nutritionists may take muscle or adipose tissue biopsies, or samples of saliva, buccal cells, hair and even toenails. A knowledge of how a nutrient behaves in different metabolic pools is critically important in choosing the correct tissue to sample to measure or judge nutritional status. For example, the level of folic acid in plasma is determined by the most recent intake pattern and thus is subject to considerable fluctuation. However, since erythrocytes remain in the circulation for about 120 days, a sample of erythrocytes will represent very recently synthesised cells right through to erythrocytes ready for recycling through the turnover mechanism previously described. As erythrocytes do not have a nucleus, they cannot switch on genes that might influence folate levels, and so the cell retains the level of folate that prevailed at the time of synthesis. Thus, erythrocyte folate is a good marker of long-term intake. The free form of many minerals and trace elements is potentially toxic, and for this reason their level in the plasma is strictly regulated. Hence, blood levels are not used to assess long-term intake of selenium, but toenail clippings can be used.

1.6 Adaptation to altered nutrient supply

In many circumstances, the body can respond to altered metabolic and nutritional states in order to minimise the consequences of such alterations. For example, the brain has an obligatory requirement for glucose as a substrate for energy and it accounts for a significant part of resting energy expenditure. During undernutrition, where glucose input does not match glucose needs, the first adaptation to the altered metabolic environment is to increase the process of gluconeogenesis, which involves the diversion of amino acids into glucose synthesis. That means less amino acid entering the protein synthesis cycle of protein turnover. Inevitably, protein reserves begin to fall. Thus, two further adaptations are made. The first is that the brain begins to use less glucose for energy (replacing it by ketones as an alternative metabolic fuel). The second is that overall, resting energy expenditure falls to help sustain a new balance if possible. Stunting in infants and children, reflected in a low height for age, can be regarded as an example of successful adaptation to chronic low energy intake. If the period of energy deprivation is not too long, the child will subsequently exhibit a period of accelerated or catch-up growth (see Chapter 7). If it is protracted, the stunting will lead to a permanent reprogramming of genetic balance.

In many instances, the rate of absorption of nutrients may be enhanced as an adaptive mechanism to low intakes. Some adaptations appear to be unsuccessful but work for a period, effectively buying time in the hope that normal intakes will be resumed. In essential fatty acid deficiency, the normal processes of elongation and desaturation of fatty acids take place but the emphasis is on the wrong fatty acid, that is, the non-essential 18-carbon monounsaturated fatty acid (oleic acid, C18:1 n-9) rather than the deficient dietary essential 18-carbon polyunsaturated fatty acid (linoleic acid, C18:2 n-6). The resultant

20-carbon fatty acid does not produce a functional eicosanoid. However, the body has significant reserves of linoleic acid which are also used for eicosanoid synthesis and so the machinery of this synthesis operates at a lower efficiency than normal. Eventually, if the dietary deficiency continues then pathological consequences ensue. In effect, adaptation to adverse metabolic and nutritional circumstances is a feature of survival until the crisis abates. The greater the capacity to mount adaptations to adverse nutritional circumstances, the greater the capacity to survive.

1.7 Perspectives on the future

These core concepts of nutrition will remain forever, but they will be refined by the emerging concept of nutrigenomics (see Chapter 2). We will develop a greater understanding of how changes in the nutrient content of one pool will alter gene expression to influence events in another pool and how this influences the flux of nutrients between pools. We will better understand how common single nucleotide polymorphisms will determine the level of nutrient intake to achieve nutrient balance in different individuals.

There is growing interest in the role of the gut microbiome in promoting good health in people. Care is needed in the use of only the term 'microbiome' as it covers a wide variety of different locations in the body (e.g. skin, respiratory system, genitourinary system, different stages of the gastrointestinal tract) and not just the large intestine. The potential for the composition of the microbiome to be influenced by the diet has generated great interest in identifying dietary components that may improve or diminish gut health, and in turn have broader effects on general health. The possibility that manipulation of the gut microbiome may influence signalling in the gut–brain axis is likely to be a major topic of investigation going forward, and the influence of dietary components on the composition of the microbiome as well as providing the substrates for bacteria could be major aspects of nutrition and health to be incorporated into the core concepts in the future.

A final topic to consider is the issue of personalisation of diet and whether it is a valid biological concept which requires much greater flexibility in the food supply. Clearly, the genomic differences and variability in epigenetic changes between people imply potential differences in nutrient requirements. The possibility that the gut microbiome varies between people is likely to add to these differences but whether it is possible to truly personalise the diet at the level of the individual based on these factors is likely to need a substantial amount of research in the future.

Further reading

Frayn, K.N. (2003). Metabolic Regulation: A Human Perspective, 2e. Oxford: Blackwell Publishing.

Websites

health.nih.gov/search.asp?category_id=29
http://themedicalbiochemistrypage.org
www.nlm.nih.gov/medlineplus/foodnutritionandmetabolism.html

2
Molecular Aspects of Nutrition

Herman E. Popeijus, Helen M. Roche, and Ronald P. Mensink

Key messages

- The genome forms the information or blueprint to build up an organism and contains the full complement of genes (genotype). An individual's genotype interacts with the environment (of which nutrition is one element), which are collectively expressed as an individual's phenotype.
- The specific order of nucleotides within DNA forms the basic units of genetic information. It is organised into chromosomes. Typically, every human cell contains 46 chromosomes (or 23 chromosome pairs). The chromosomes are a series of individual genes, wherein diploid cells contain two copies of each.
- Genetic variation reflects DNA sequence alterations, due to random mating or damage. This variation is referred to as genetic polymorphisms or mutations. Genetic polymorphisms are common forms of genetic heterogeneity, whereby different forms of a given allele are common in a population. Genetic mutations refer to more uncommon alterations.
- Transcription (or gene expression) refers to the process whereby information encoded in the genes is converted into mRNA sequence. Translation refers to converting mRNA sequence information into a protein, which may in turn affect metabolism. Some, but not all, transcriptional and translational elements are converted into observable phenotypes.
- There are several tools used to investigate molecular aspects of nutrition: animal models, organoids, cell/tissue culture models, molecular cloning, CRISPR/Cas9, RNAi, gene expression analysis (polymerase chain reaction [PCR], DNA microarrays and sequencing), proteomics, stable isotopes, and metabolomics.
- Genetic variability is relatively simple to characterise. Understanding the interaction with the environment, of which food is one complex element, is the challenge. The resultant gene–environment interactions can determine nutrient requirements, the metabolic response to food and nutrient intake and/or susceptibility to diet-related diseases.
- Nutrients and non-nutrient food components can interact with the genome to modulate gene, protein and metabolite expression. These interactions between nutrition and the genome are referred to as *molecular nutrition* or *nutrigenomics*.
- Personalised nutrition seeks to understand if/how food or nutrient intake could be used to manipulate an individual's metabolic response and/or to reduce their predisposition to diet-related diseases, based on their genetic and/or metabolic phenotype.

2.1 Introduction

Our genes determine every characteristic of life: gender, physical characteristics, metabolic functions, life stage and responses to external or environmental factors, which include nutrition. Nutrients have the ability to interact with the human genome to alter gene, protein and metabolite expression, which in turn can affect normal growth, health and disease. The Human Genome Project has provided an enormous amount of genetic information and thus a greater understanding of our genetic background. It is well known that nutrients, and non-nutritive food components, can interact with the genome. This aspect of nutritional science is known as *molecular nutrition or nutrigenomics*.

Molecular nutrition looks at the relationship between the human genome and nutrition from two perspectives. First, the genome determines every individual's genotype (or genetic background), which in turn can determine their

Nutrition and Metabolism, Third Edition. Edited on behalf of The Nutrition Society by Helen M. Roche, Ian A. Macdonald, Annemie M.W.J. Schols and Susan A. Lanham-New.

Companion Website: www.wiley.com/go/nutrition/metabolism3e

nutrient state, metabolic response and/or genetic predisposition to diet-related diseases. Second, nutrients can interact with the genome and alter gene, protein or metabolite expression. Gene expression (or transcription) is only the first stage of the whole-body or metabolic response to a nutrient and a number of post-translational events (e.g. enzyme activity, protein half-life, co-activators, co-repressors), but metabolomic events can also modify the ability of nutrients to alter an individual's phenotype.

This chapter will review the core concepts in molecular biology, introduce the genome and discuss how we can characterise the effect of nutrition on gene, protein and metabolite expression using state-of-the-art transcriptomic, proteomic and metabolomic technologies, identifying some important research tools used to investigate these molecular aspects of nutrition, such as characterising how genetic background can determine nutrition and health. Chapter 13 'Application of "Omics" Technology', in the Nutrition Research Methodologies textbook) provides detailed information in relation to each technology. This chapter provides some examples of how nutrients regulate gene, protein and metabolite expression, within the context of human nutrition and metabolism.

Overall, the principal aim of nutrigenomics/molecular nutrition is to understand how the genome interacts with food, nutrients and non-nutrient food components, within the context of diet-related diseases. It attempts to determine nutrients that enhance the expression of gene, protein and metabolic pathways/networks that are associated with health and suppress those that predispose to disease. While it is unrealistic to assume that food intake and good nutrition can overcome our genetic fate, good nutrition can improve health and quality of life. Therefore, it is essential that we extend our understanding of the molecular interplay between the genome, food and nutrients to facilitate a greater understanding of the mechanistic relationship between diet, health and disease.

2.2 Core concepts in molecular biology

The genome, DNA, and the genetic code

The *genome* refers to the total genetic information carried by a cell or organism. In simple terms, the genome (or DNA sequence) contains the full complement of genes. The expression of each gene leads in general to the formation of a protein which, together with many other proteins coded by other genes, forms tissues, organs and systems, to constitute the whole organism. In complex multicellular organisms, the information carried within the genome gives rise to multiple tissues (muscle, bone, adipose tissue, etc.).

The characteristics of each cell type and tissue are dependent on differential gene expression by the genome, whereby only those genes are expressed that code for specific proteins to confer the individual characteristics of the cells that constitute each organ. For example, gene expression in muscle cells results in the formation of muscle-specific proteins, that are critical for the differentiation, development and maintenance of muscle tissue. These genes are completely different from those expressed in osteoblasts, osteoclasts and osteocytes, which form bone. These differentially expressed proteins can have a wide variety of functions: as structural components of the cell or as regulatory proteins, including enzymes, hormones, receptors and intracellular signalling proteins that confer tissue specificity.

It is very important to understand the molecular basis of cellular metabolism because incorrect expression of genes at the cellular level can disrupt whole-body metabolism and lead to disease. Aberrant gene expression can lead to cellular disease when proteins are produced in the wrong place, at the wrong time, at abnormal levels or as a malfunctioning isoform that can compromise whole-body health. Furthermore, different nutritional states and intervention therapies can modulate the expression of cellular genes and thereby the formation of proteins. The ultimate goal of molecular nutrition is to understand how nutrients interact with the genome to alter the expression of genes, determine the formation and function of proteins and modulate metabolite profiles, all of which play a role in health and disease.

Deoxyribonucleic acid (DNA) is the most basic unit of genetic information, as the DNA sequence codes for the amino acids that form cellular proteins. Two individual DNA molecules are packaged as the chromosomes within the nucleus of animal and plant cells. The basic structure and composition of DNA are illustrated in Figure 2.1. DNA is composed of large polymers, with a linear backbone composed of residues of

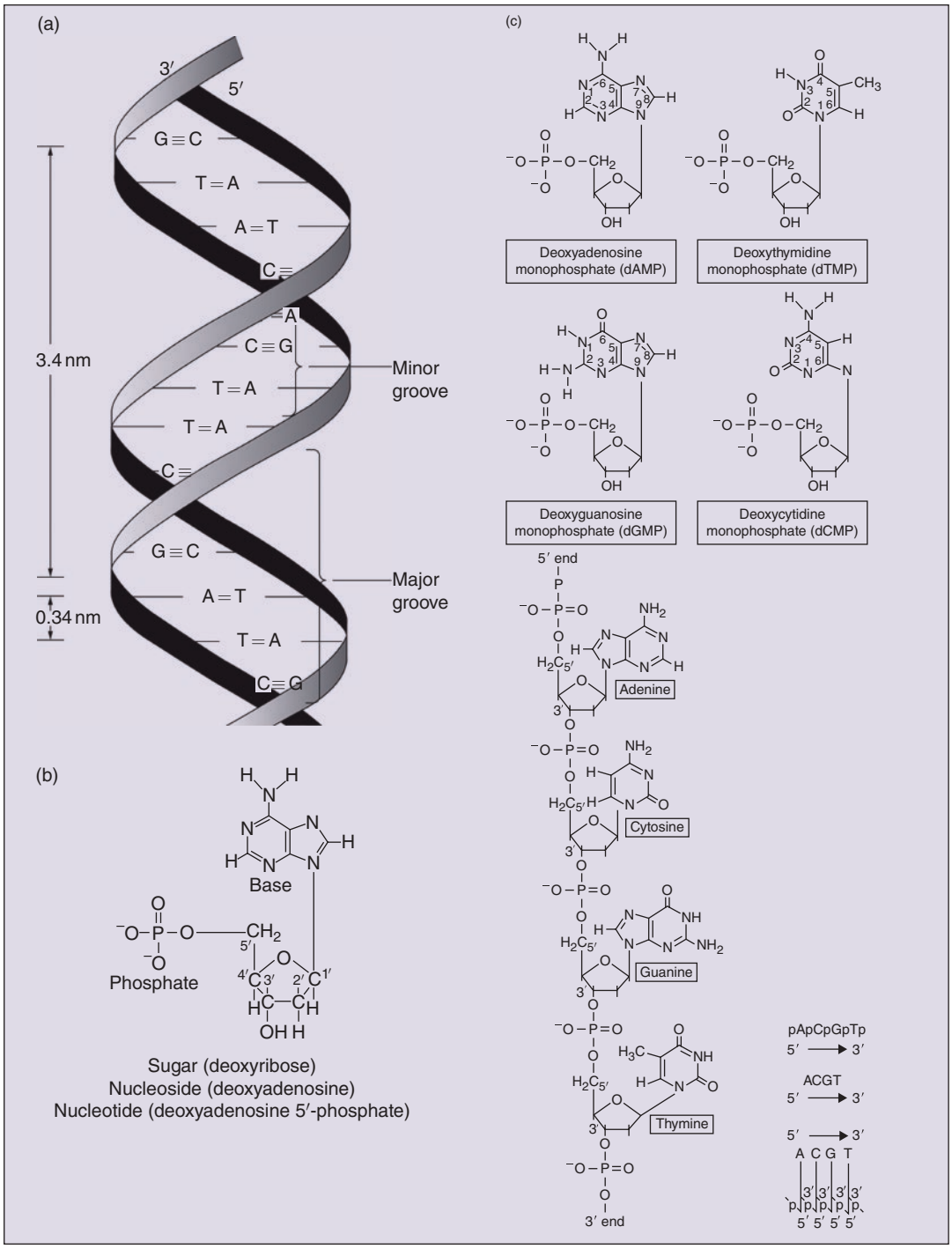


Figure 2.1 Structure and composition of DNA. DNA contains deoxynucleotides consisting of a specific heterocyclic nitrogenous base (adenine, guanine, cytosine or thymine) joined to a deoxyribose phosphate moiety. Adjacent deoxynucleotides are linked through their phosphate groups to form long polynucleotide chains. (a) The DNA double helix; (b) a nucleotide; (c) the purine and pyrimidine bases. *Source:* Cox TM and Sinclair J 1997/With permission from John Wiley & Sons.

the five-carbon sugar residue deoxyribose, which are successively linked by covalent phosphodiester bonds. A nitrogenous base, either a *purine* (*adenine* [A] or *guanine* [G]) or a *pyrimidine* (*cytosine* [C] or *thymine* [T]), is attached to each deoxyribose. DNA forms *a double-stranded helical structure*, in which the two separate DNA polymers wind around each other. The two strands of DNA run *antiparallel*, such that the deoxyribose linkages of one strand runs in the 5′–3′ direction and the other strand in the opposite 3′–5′ direction. The double helix is mainly maintained by hydrogen bonds between nucleotide pairs. According to the *base-pair rules*, adenine always binds to thymine via two hydrogen bonds and guanine binds to cytosine via three hydrogen bonds. This complementary base-pair rule ensures that the sequence of one DNA strand specifies the sequence of the other.

The *nucleotide* is the basic repeat unit of the DNA strand and is composed of deoxyribose, a phosphate group and a nitrogenous base. The 5′–3′ sequential arrangement of the nucleotides in the polymeric chain of DNA contains the *genetic code* for the arrangement of amino acids in proteins. The genetic code is the universal language that translates the information stored within the DNA of genes into proteins. It is universal between all species known so far. The genetic code is read in groups of three nucleotides. These three nucleotides, called a *codon*, are specific for one particular amino acid. Table 2.1 shows the 64 possible codons, of which 61 specify for 22 different amino acids, while three sequences (TAA, TAG, TGA) are stop codons (i.e. do not code for an amino acid). Most amino acids are coded for by more than one codon; this is referred to as *redundancy*. For example, the amino acid isoleucine may be coded by the DNA sequence ATT, ATC or ATA. Each amino acid sequence of a protein always begins with a methionine residue because the start codon (ATG) codes for methionine. The three stop codons signal the end of the coding region of a gene and the resultant polypeptide sequence.

Chromosome (karyotype)

In eukaryotic cells, DNA is packaged into *chromosomes* and every cell contains a set of chromosomes (Figure 2.2). Each chromosome has a narrow waist known as the *centromere*, which divides each chromosome into a short and a long arm, labelled p and q, respectively. The beginning and end of each double-stranded helix consists of a specific repetitive sequence of several

Table 2.1 Conversion of the genetic code into amino acids.

	Second base				
First base	T	C	A	G	Third base
T	TTT (Phe)	TCT (Ser)	TAT (Tyr)	TGT (Cys)	T
	TTC (Phe)	TCC (Ser)	TAC (Tyr)	TGC (Cys)	C
	TTA (Leu)	TCA (Ser)	TAA (Stop)	TGA (Stop)	A
	TTG (Leu)	TCG (Ser)	TAG (Stop)	TGG (Trp)	G
C	CTT (Leu)	CCT (Pro)	CAT (His)	CGT (Arg)	T
	CTC (Leu)	CCC (Pro)	CAC (His)	CGC (Arg)	C
	CTA (Leu)	CCA (Pro)	CAA (Gln)	CGA (Arg)	A
	CTG (Leu)	CCG (Pro)	CAG (Gln)	CGG (Arg)	G
A	ATT (Ile)	ACT (Thr)	AAT (Asn)	AGT (Ser)	T
	ATC (Ile)	ACC (Thr)	AAC (Asn)	AGC (Ser)	C
	ATA (Ile)	ACA (Thr)	AAA (Lys)	AGA (Arg)	A
	ATG (Met)	ACG (Thr)	AAG (Lys)	AGG (Arg)	G
G	GTT (Val)	GCT (Ala)	GAT (Asp)	GGT (Gly)	T
	GTC (Val)	GCC (Ala)	GAC (Asp)	GGC (Gly)	C
	GTA (Val)	GCA (Ala)	GAA (Glu)	GGA (Gly)	A
	GTG (Val)	GCG (Ala)	GAG (Glu)	GGG (Gly)	G

Codon table; from genetic code into amino acids: Ala, alanine; Arg, arginine; Asn, asparagine; Asp, aspartic acid; Cys, cysteine; Gln, glutamine; Glu, glutamic acid; Gly, glycine; His, histidine; Ile, isoleucine; Leu, leucine; Lys, lysine; Met, methionine; Phe, phenylalanine; Pro, proline; Ser, serine; Thr, threonine; Trp, tryptophan; Tyr, tyrosine; Val, valine, and ''stop" for no amino acid. Furthermore, there are three Stop codons -TAA; TG; TGA-, a sequence that does not code for an amino acid as there is not tRNA to bind to that codon.

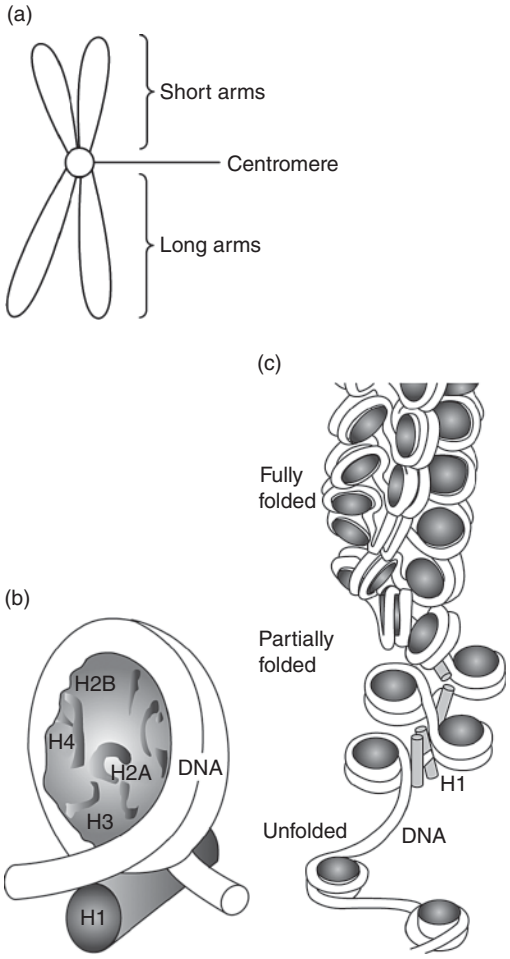


Figure 2.2 Structure of (a) a chromosome, (b) the nucleosome and (c) chromatin. *Source:* Cox TM and Sinclair J 1997/With permission from John Wiley & Sons.

kilobases, also known as the *telomere*. DNA is packaged in a very compact structure within the nucleus. Condensing of DNA is essential because the human cell contains approximately 4×10^9 nucleotide pairs, termed *base pairs* (*bp*) of DNA, whose extended length would approach more than 1 m. The DNA of each chromosome is wrapped approximately 2.5 times around a set of eight histone proteins and stored in the nucleus. The *nucleosome* is the most basic unit of the chromosome. It is composed of a 145 bp linear strand of double-stranded DNA wound around a complex of *histone proteins* (two of each of the four histone proteins H2a, H2b, H3 and H4). Nucleosomes are linked together by the histone protein H1 to form *chromatin*. During cell division, this is then further compacted with the aid of non-histone chromosomal proteins to generate a chromosome. The structure of DNA in chromatin is important because it has profound effects on the ability of DNA to be transcribed.

The chromosomal complement or karyotype refers to the number, size and shape of the chromosomes as seen during mitosis, a specific phase during the cell cycle. The human karyotype is composed of 22 pairs of autosomes and a pair of sex chromosomes: XX in the female and XY in the male. Most human cells contain 46 chromosomes, the *diploid* number. Chromosomal disorders are characterised by abnormalities of chromosomal number or structure. They may involve the autosomes or the sex chromosomes and may be the result of a *germ cell mutation* in the parent (or a more distant ancestor) or a *somatic mutation* in which only a proportion of cells will be affected (mosaicism). The normal chromosome number is an exact multiple of the *haploid* number (23) and is referred to as the diploid number. A chromosomal number that exceeds the diploid number (46) is called *polyploidy*, and one that is not an exact multiple number is *aneuploidy*. Aneuploidy usually occurs when the pair of chromosomes fails to segregate (non-disjunction) during meiosis, which results in an extra copy of a chromosome (*trisomy*) or a missing copy of a chromosome (*monosomy*). Down syndrome is a common example of trisomy – it is due to the presence of three copies of chromosome 21 (trisomy 21).

Structural abnormalities of chromosomes also occur. A *translocation* is the transfer of chromosomal material between chromosomes. Chronic myeloid leukaemia results from the translocation of genetic material between chromosome 8 and chromosome 22. This results in an abnormal chromosome, known as the Philadelphia chromosome, the expression of which results in leukaemia. *Chromosomal deletions* arise from the loss of a portion of the chromosome between two break points. *Inversions* arise from two chromosomal breaks with inversion through 180° of the chromosomal segment between the breaks.

Genotype, phenotype, and allelic expression

The *genotype* of an organism is the total number of genes that make up a cell or organism. The term, however, is also used to refer to alleles present at one locus. Each diploid cell contains two

copies of each gene; the individual copies of the gene are called *alleles*. The definition of an allele is one of two (or more) alternative forms of a gene located at the corresponding site (*locus*) on homologous chromosomes. One allele is inherited from the maternal gamete and the other from the paternal gamete, therefore the cell can contain the same or different alleles of every gene. *Homozygous* individuals carry two identical alleles of a particular gene. *Heterozygotes* have two different alleles of a particular gene. The term *haplotype* describes a cluster of alleles that occur together on a DNA segment and/or are inherited together. *Genetic linkage* is the tendency for alleles located close together to be transmitted together through meiosis and hence be inherited together.

Genetic polymorphisms are different forms of the same allele in the population. The 'normal' allele is known as the *wild-type allele*, whereas the variant is known as the polymorphic or mutant allele. A polymorphism differs from a mutation, in that it occurs in a population more frequently than a recurrent mutation. By convention, a polymorphic locus is one at which there are at least two alleles, each of which occurs with frequencies greater than 1%. Alleles with frequencies less than 1% are considered as a recurrent mutation. The alleles of the ABO blood group system are examples of genetic polymorphisms. The *single nucleotide polymorphism* (*SNP*) is a common pattern of inherited genetic variation (or common mutation) that involves a single base change in the DNA. Another common form of genetic variation includes *copy number variations* (*CNV*). It is estimated that about 0.4% of the human genome differs with respect to CNV.

There are several ways to characterise genetic variation. Traditionally, genetic variants were identified using methods including restriction fragment length polymorphism (RFLP), positional cloning and gene sequencing. These approaches tended to identify specific candidate genes or haplotypes. More recently, whole-genome and transcriptome sequencing technologies that allow full genetic sequence determination of all expressed proteins have become available. Sequencing now provides the opportunity to identify multiple genetic variants. The biggest challenge is to translate genetic data into knowledge, wherein genetic data are coupled with physiology data to make sense out of them. Complex bioinformatics using neural networks, machine learning and artificial intelligence facilitate analysis of the large amount of data generated.

Epigenetics is a relatively new field of research which refers to changes in gene expression due to mechanisms other than changes in the underlying DNA sequence. The molecular basis of epigenetics is complex; put simply, it refers to altered DNA structure. It involves modifications of the activation of certain genes, but not the basic DNA sequence. For example, DNA methylation refers to the addition of methyl groups to the DNA, which in turn affects transcriptional activity. Folate status can affect DNA methylation, which in turn can affect gene expression through mechanisms that are being actively researched.

Still, there is a considerable amount of research needed that investigates relationships between common genetic polymorphisms or epigenetics with disease because certain genetic variations may predispose an individual to a greater risk of developing a disease. The effect of genetic variation in response to dietary change is also of great interest because some polymorphisms/epigenetic states may determine an individual's response to dietary changes. Hence, genetic variation can determine the efficacy of nutritional approaches, which may in turn determine the outcome of certain disease states. The inter-relationship between diet, disease and genetic variation will be discussed in more detail in Section 2.5.

The *phenotype* is the observable biochemical, physiological or morphological characteristics of a cell or individual resulting from the expression of the cell's genotype, within the environment in which it is expressed. Allelic variation and expression can affect the phenotype of an organism. A *dominant allele* is the allele of a gene that contributes to the phenotype of a heterozygote. The non-expressing allele that makes no contribution to the phenotype is known as the *recessive allele*. The phenotype of the recessive allele is only demonstrated in homozygotes who carry both recessive alleles. *Co-dominant alleles* contribute equally to the phenotype. The ABO blood groups are an example of co-dominant alleles, where both alleles are expressed in an individual. In the case of *partial dominance*, a combination of alleles is expressed simultaneously and the phenotype of the heterozygote is intermediate between that of the two homozygotes. For

example, in the case of the snapdragon, a cross between red and white alleles will generate heterozygotes with pink flowers. *Genetic heterogeneity* refers to the phenomenon whereby a single phenotype can be caused by different allelic variants.

DNA damage, genetic mutations, polymorphisms, and heritability (monogenic and polygenic disorders)

Human DNA sequence changes slowly under normal physiological conditions over time or with age, due to inevitable errors occurring during normal DNA synthesis associated with cellular renewal. In addition, many agents contribute to DNA changes, as they cause DNA damage, including ionising radiation, ultraviolet light, chemical mutagens, viruses, free radicals produced during many cellular processes and nutrition. A change in the nucleotide sequence is known as either a *mutation* or a *polymorphism*, depending on the frequency of the alteration. Polymorphisms are more frequent sequence variants or simply 'common mutation'.

A mutation may be defined as a permanent transmissible change in the nucleotide sequence of a chromosome, usually in a single gene, which may lead to loss or change of the normal function of the gene. A mutation can have a significant effect on protein production or function because it can alter the amino acid sequence of the protein that is coded by the DNA sequence in a gene. *Point mutations* include *insertions*, *deletions, transitions* and *transversions*. Two types of events can cause a point mutation: chemical modification of DNA, which directly changes one base into another, or a mistake during DNA replication that causes, for instance, the insertion of the wrong base into the polynucleotide during DNA synthesis. Transitions are the most common type of point mutations and result in the substitution of one pyrimidine (C–G) or one purine (A–T) by the other. Transversions are less common, where a purine is replaced by a pyrimidine or vice versa.

The functional outcome of mutations can vary significantly. Most of the DNA is non-coding and mutations may therefore not necessarily result in a change of our proteins. However, non-coding regions may be important in regulation, controlling transcription. On the other hand, a single-point mutation in a coding region can change the third nucleotide in a codon and not change the amino acid that is translated, or it may cause the incorporation of another amino acid into the protein – this is known as a *missense mutation*. The functional effect of a missense mutation varies greatly depending on the site of the mutation and the importance of the protein in relation to health. A missense mutation can have no apparent effect on health or it can result in a serious medical condition. For example, sickle cell anaemia is due to a missense mutation of the β-globin gene: a glutamine is changed to valine in the amino acid sequence of the protein. This has drastic effects on the structure and function of the β-globin protein, which causes aggregation of deoxygenated haemoglobin and deformation of the red blood cell. A nucleotide change can also result in the generation of a stop codon (*nonsense mutation*) and no functional protein will be produced. *Frameshift mutations* refer to small deletions or insertions of bases that alter the reading frame of the nucleotide sequence; hence, the different codon sequence will affect the expression of amino acids in the peptide sequence.

Heritability refers to how much a disease can be ascribed to genetic rather than environmental factors. It is expressed as a percentage, a high value indicating that the genetic component is important in the aetiology of the disease. Genetic disorders can simply be classified as *monogenic* (or single-gene disorders) or *polygenic diseases* (multifactorial diseases). In general, we have a far greater understanding of the *single-gene disorders* because they are due to one or more mutant alleles at a single locus and most follow simple *Mendelian inheritance*. Examples of such disorders are:

- autosomal dominant: familial hypercholesterolaemia, von Willebrand disease, achondroplasia
- autosomal recessive: cystic fibrosis, phenylketonuria, haemochromatosis, α-thalassaemia, β-thalassaemia
- X-linked dominant: vitamin D-resistant rickets
- X-linked recessive: Duchenne muscular dystrophy, haemophilia A, haemophilia B, glucose-6-phosphate dehydrogenase deficiency.

Some single-gene disorders show non-Mendelian patterns of inheritance, which are explained by different degrees of penetrance and variable gene expression. *Penetrance* means that a genetic

lesion is expressed in some individuals but not in others. For example, people carrying a gene with high penetrance have a high probability of developing any associated disease. A low-penetrance gene will result in only a slight increase in disease risk. *Variable expression* occurs when a genetic mutation produces a range of phenotypes. *Anticipation* refers to the situation when a Mendelian trait manifests as a phenotype with decreasing age of onset and often with greater severity as it is inherited through subsequent generations (e.g. Huntington chorea, myotonic dystrophy). *Imprinting* refers to the differential expression of a chromosome or allele depending on whether the allele has been inherited from the male or female gamete. This is due to selective inactivation of genes according to the paternal or maternal origin of the chromosomes. Although there are only a few examples of diseases that arise as a result of imprinting (e.g. Prader–Willi and Angelman syndromes), it is thought that this form of gene inactivation may be more important than previously realised.

Polygenic (or multifactorial) diseases are those due to a number of genes (e.g. cancer, coronary heart disease, diabetes, obesity). Even though polygenic disorders are more common than monogenic disorders, we still do not understand the full genetic basis of many of these conditions. This reflects the fact that there is interaction between many candidate genes, that is, those genes that are thought to play an aetiological role in multifactorial conditions. Furthermore, in polygenic inheritance, a trait is in general determined by a combination of both the gene and the environment.

The Human Genome Project

Historically, the Human Genome Project (HGP) was an important source of genetic information to advance human genetics and molecular nutrition, especially in terms of providing the opportunity to understand the interaction between nutrition and the human genome. Interestingly, the HGP showed that the human genome was apparently less complex than was anticipated. For example, the size of the human genome is only 30 times greater than that of the fruit fly and 250 times larger than yeast. Also importantly, only 3% of the DNA in the human genome constitutes coding regions and codes for approximately 25 000 genes. Humans only have 2–3 times as many genes as a fruit fly. Compared with the fruit fly, the human genome has many more non-coding or intronic regions. The 2–3-fold difference between humans and fruit flies is largely accounted for by the greater number of control genes (e.g. transcription factors) in the human genome. Although these intronic regions do not code for genes as such, many have important functional roles, including promoter and/or repressor regions, contributing to isoform variations of genes, and interacting with regulatory RNA molecules, known as miRNA or long non-coding RNAs.

The real challenge is not sequencing the genetic code as such but understanding how the human genome interacts with the environment. The HGP showed that humans are very alike; it estimated that humans are 99.8% genetically similar. Nevertheless, it is very apparent that human phenotypes are very diverse. Thus, the implications of the HGP in relation to molecular nutrition are not yet fully known. The challenge is to identify the proportion of human genetic variation that is relevant to nutrition. The term 'gene mining' refers to the process that identifies new genes involved in nutrition, health and disease. At the most basic level, we already know that the human genome determines nutrient requirements; for example, gender determines iron requirements – the iron requirement of menstruating women is greater than that of men of the same age. In the case of folate, research would suggest that the methylene tetrahydrofolate reductase (MTHFR) polymorphism could determine an individual's folate requirements. Since there are fewer genes than anticipated, it has been proposed that different isoforms of the same gene with different functionality are important. Already there are examples of this; for example, the three isoforms of the apolipoprotein E (Apo E) gene determine the magnitude of postprandial triacylglycerol metabolism. Furthermore, it has been proposed that the interaction between the human genome and the environment is an important determinant of within- and between-individual variation.

Within the context of molecular nutrition, we will have to determine how nutrients alter gene expression and determine the functional consequences of genetic polymorphisms. With the information generated from the HGP, we will have a more complete understanding and information in relation to the relevance of genetic

variation, and how alterations in nutrient intake or nutritional status affect gene expression in a way that is relevant to human health and disease processes. In essence, the challenge is to bridge the gap between the genome sequence and whole-organism biology, nutritional status and intervention.

2.3 Gene and protein expression: transcription and translation

Gene structure and function

Gene expression refers to the process whereby the information encoded in the DNA of a gene is converted into a final product, which confers the observable phenotype upon the cell. A *gene* may be defined as the nucleic acid sequence that is necessary for the synthesis of a functional transfer RNA (tRNA), ribosomal RNA (rRNA), a peptide or protein in a temporal and tissue-specific manner. In the case of a peptide or protein, a gene is not directly translated into a protein; it is expressed via a nucleic acid intermediary called *messenger RNA* (*mRNA*). The transcriptional unit of every gene is the sequence of DNA transcribed into a single mRNA molecule, starting at the promoter and ending at the terminator regions.

The essential features of a gene and mRNA are presented in Figure 2.3, including key elements such as the two non-coding (or untranslated) promoter and terminator regions at the beginning and end of the gene. Also, the coding regions are a mix of exons and introns. Each gene contains DNA sequences that code for the amino acid sequence of the protein, which are called *exons*. These exons are interrupted by non-coding DNA sequences, which are called *introns*. Whilst introns are non-coding regions, they may contain critical regulatory elements that may alter gene expression. The *promoter region* is located immediately upstream of the gene-coding region; it contains DNA sequences, known as the TATA and CAAT boxes, which define the DNA binding sites at which general transcription factors bind to start. The promoter region also contains enhancer and inhibitor elements where specific transcription factors can bind that enhance or inhibit the possibility of the general transcription factors to start gene expression. This promoter region can be sensitive to nutritional and hormonal cues, for example insulin response elements. The last exon ends with a *stop codon* (TAA, TAG or TGA), which represents the end of the gene-coding region and is followed by the terminator DNA sequence that defines the end of the gene-coding region.

Ribonucleic acid

Ribonucleic acid, like DNA, carries genetic information. The composition of RNA is very similar to DNA, and it plays a key role in all stages of gene expression. RNA is also a linear polynucleotide, but it differs from DNA in that it is mostly single-stranded and the sugar of the bases is polymers of ribose instead of deoxyribose. Furthermore, the pyrimidine base *uracil* (*U*) is used as a substitute for thymine (T) in DNA, and it is relatively unstable when compared to DNA. There are at least five different types of RNA in eukaryotic cells and all are involved in gene expression.

- Messenger RNA (mRNA) molecules are long, linear, single-stranded polynucleotides that are direct copies of DNA. mRNA is formed by transcription of DNA.
- Small nuclear RNA (snRNA) is a short, ± 150 nucleotide (nt) RNA molecule that forms, together with a protein, the small nuclear ribonucleoprotein (snRNP). Several of these

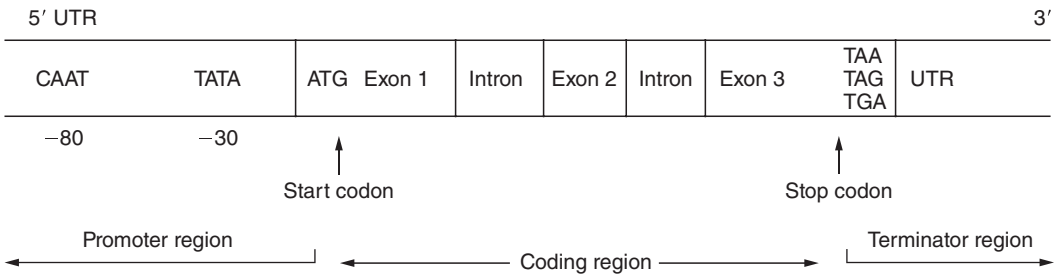


Figure 2.3 Essential features of the gene.

snRNPs form with other proteins the spliceosome that facilitates the removal of introns from the precursor mRNA.

- Ribosomal RNA (rRNA) is a structural and functional component of ribosomes. Ribosomes, which are present in the cytoplasm and on the rough endoplasmic reticulum (rER), are the machines that synthesise mRNA into amino acid polypeptides. The ribosomes are composed of rRNA and ribosomal proteins.
- Transfer RNA (tRNA) is a small RNA molecule that donates amino acids during translation or protein synthesis.
- A group of small RNAs such as microRNAs (miRNA), small interfering RNAs (siRNAs), Piwi-interacting RNAs (piRNAs) and repeat-associated siRNAs (rasiRNAs). These small RNA molecules have various functions at the transcriptional and/or post-transcriptional level, including the breakdown of mRNA, repression of transcription and chromatin remodelling.

RNA is about 10 times more abundant than DNA in eukaryotic cells; ~80% is rRNA, ~15% is tRNA and ~1–5% is mRNA. In the cell, mRNA is normally found associated with protein complexes called *messenger ribonucleoprotein* (*mRNP*), which package mRNA and aid its transport into the cytoplasm, where it is decoded into a protein.

Transcription

RNA transcription is the process whereby the genetic information encoded in DNA is transferred into a *heterogeneous nuclear RNA* (*hnRNA*) because of its considerable variation in size. Instead of hnRNA, often the term *precursor mRNA* (*pre-mRNA*) is used to designate that it refers to a just produced, unprocessed 'raw' RNA molecule. Formation of the pre-mRNA is the first step in the process of gene expression and occurs in the nucleus of the cell. It can be divided into four stages: template recognition, initiation, elongation and termination.

Transcription is catalysed by DNA-dependent *RNA polymerase* enzymes (Figure 2.4). In eukaryotic cells, RNA polymerase II synthesises mRNA, while RNA polymerase I and RNA polymerase III synthesise tRNA and rRNA. The strand of DNA that directs synthesis of mRNA

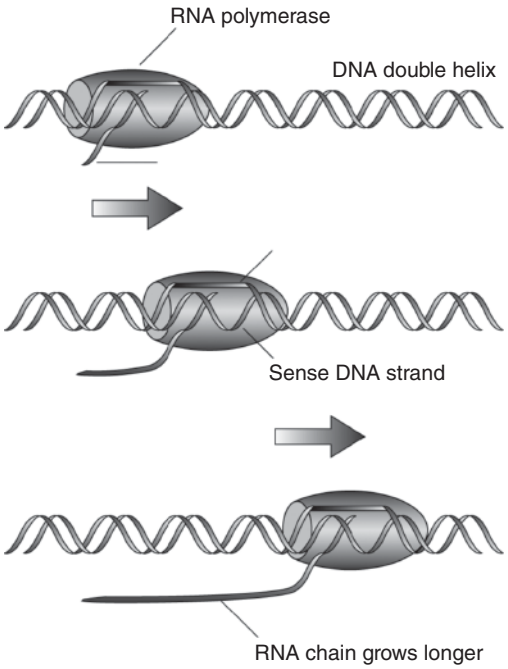


Figure 2.4 RNA polymerase II transcribes the information in DNA into RNA.

via complementary base pairing is called the template or *antisense strand*. The other DNA strand that bears the same nucleotide sequence is called the coding or *sense strand*. Therefore, RNA represents a copy of DNA.

Transcription is a multistep process. When the specific transcription factors switch on the promoter, transcription is initiated by the assembly of an initiation complex formed by the general transcription factors (TFIID, TFIIA, TFIIB, TFIIF, TFIIE, TFIIH) at the promoter. They then 'recruit' RNA polymerase II that binds to these general transcription factors. More specifically, the general transcription factor, TFIID, recognises the promoter and binds to the TATA box at the start of the gene. TFIIH unwinds double-stranded DNA to expose the unpaired DNA nucleotide and the DNA sequence is used as a template from which RNA is synthesised. TFIIE and TFIIH are required for promoter clearance, allowing RNA polymerase II to commence movement away from the promoter. Initiation describes the synthesis of the first nine nucleotides of the RNA transcript. Elongation describes the phase during which RNA polymerase moves along the DNA and extends the growing RNA molecule. The RNA molecule is synthesised by adding nucleotides to the free 3′-OH end of the

growing RNA chain. As new nucleotides can only be attached at this free 3′-OH end, RNA synthesis always takes place in the 5′–3′ direction. Growing of the RNA chain is ended by a process called *termination*, which involves recognition of the *terminator sequence* that signals the dissociation of the polymerase complex.

Post-transcriptional processing of RNA

After transcription of DNA into RNA, the newly synthesised pre-mRNA is modified. This process is called post-transcriptional processing of RNA (Figure 2.5). The primary transcript is a pre-mRNA molecule that represents a full copy of the gene extending from the promoter to the terminator region of the gene and includes introns and exons. While still in the nucleus, the newly synthesised RNA is capped, polyadenylated and spliced. *Capping* refers to the addition of a modified guanine (G) nucleotide (7-methylguanosine) at the 5′ end of the mRNA. This 7-methylguanosine cap has several functions: it protects the synthesised RNA from enzymatic attack, it aids pre-mRNA splicing, helps to transport the mRNA out of the nucleus, and enhances translation of the mRNA. *Polyadenylation* involves the addition of a string of adenosine (A) residues (a *poly-A tail*) to the 3′ end of the pre-mRNA. Then the pre-mRNA is spliced, which is performed by the spliceosome with the help of small nuclear RNA (snRNA). This RNA–protein complex recognises the consensus sequences at each end of the intron (5′-GU and AG-3′) and excises the introns so that the remaining exons are spliced together to form the mature mRNA molecule. After capping, polyadenylation and splicing, the mRNA is transported from the nucleus to the cytoplasm for translation.

Alternative splicing describes the process whereby other, alternative splice recognition sites lead to the removal of certain exons or retention of (parts of) introns, thereby generating different mature mRNA molecules and ultimately different proteins. As illustrated in Figure 2.6, one primary RNA transcript can be spliced in several different ways, in this example leading to three mRNA isoforms, one with five exons or two with four exons. Upon translation, the different mRNA isoforms will give rise to different isoforms of the protein product of the gene. This is a relatively common phenomenon. Importantly, those variants can have different functionality.

Although the physiological or metabolic relevance of the different isoforms of many proteins is not fully understood, it surely will be relevant to molecular nutrition. For example, the peroxisome proliferator activator receptor-γ (PPARγ) gene can produce seven different isoforms of mRNA ($PPAR\gamma_{1\text{-}7}$), because of different promoters and alternative splicing. To illustrate, the $PPAR\gamma_2$ mRNA isoform is responsive to nutritional states. $PPAR\gamma_2$ mRNA expression is increased in the fed state; its expression is

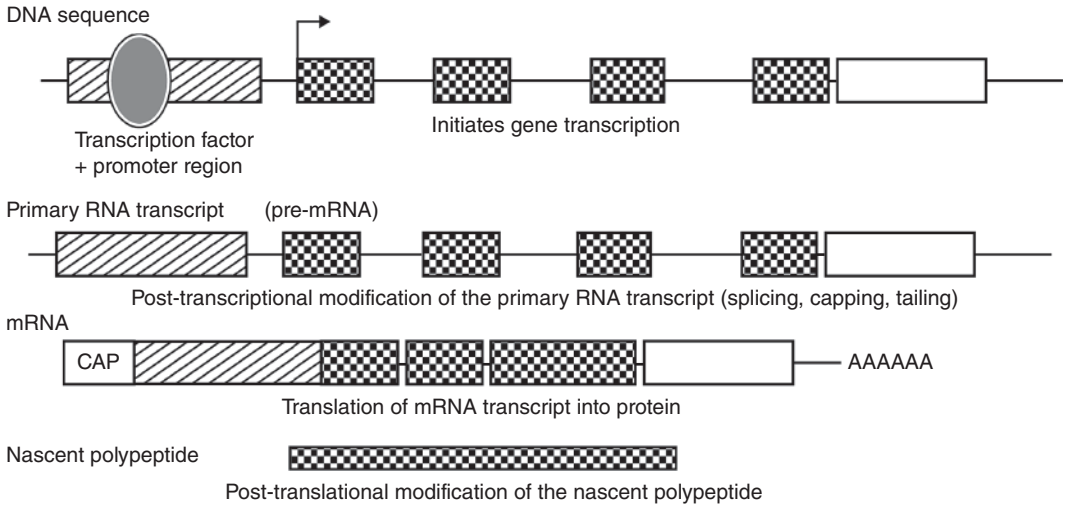


Figure 2.5 Transcription and processing of mRNA. Primary RNA transcript (pre-mRNA). Post-transcriptional modification (splicing, capping and tailing) of the primary RNA transcript (pre-mRNA).