

MOLECULAR BIOLOGY OF THE CELL

fourth edition

Bruce Alberts

Alexander Johnson

Julian Lewis

Martin Raff

Keith Roberts

Peter Walter



GS Garland Science
Taylor & Francis Group

Garland
Vice President: Denise Schanck
Managing Editor: Sarah Gibbs
Senior Editorial Assistant: Kirsten Jenner
Managing Production Editor: Emma Hunt
Proofreader and Layout: Emma Hunt
Production Assistant: Angela Bennett
Text Editors: Marjorie Singer Anderson and Betsy Dilernia
Copy Editor: Bruce Goatly
Word Processors: Fran Dependahl, Misty Landers and Carol Winter
Designer: Blink Studio, London
Illustrator: Nigel Orme
Indexer: Janine Ross and Sherry Granum
Manufacturing: Nigel Eyre and Marion Morrow

Cell Biology Interactive
Artistic and Scientific Direction: Peter Walter
Narrated by: Julie Theriot
Production, Design, and Development: Mike Morales

Oversize
QH581.2 .M718 2002
c.2
Molecular biology of the
cell

Bruce Alberts received his Ph.D. from Harvard University and is President of the National Academy of Sciences and Professor of Biochemistry and Biophysics at the University of California, San Francisco. **Alexander Johnson** received his Ph.D. from Harvard University and is a Professor of Microbiology and Immunology at the University of California, San Francisco. **Julian Lewis** received his D.Phil. from the University of Oxford and is a Principal Scientist at the Imperial Cancer Research Fund, London.

Martin Raff received his M.D. from McGill University and is at the Medical Research Council Laboratory for Molecular Cell Biology and Cell Biology Unit and in the Biology Department at University College London. **Keith Roberts** received his Ph.D. from the University of Cambridge and is Associate Research Director at the John Innes Centre, Norwich. **Peter Walter** received his Ph.D. from The Rockefeller University in New York and is Professor and Chairman of the Department of Biochemistry and Biophysics at the University of California, San Francisco, and an Investigator of the Howard Hughes Medical Institute.

© 2002 by Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter.
© 1983, 1989, 1994 by Bruce Alberts, Dennis Bray, Julian Lewis, Martin Raff, Keith Roberts, and James D. Watson.

All rights reserved. No part of this book covered by the copyright hereon may be reproduced or used in any format in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or information storage and retrieval systems—without permission of the publisher.

Library of Congress Cataloging-in-Publication Data
Molecular biology of the cell / Bruce Alberts ... [et al.].-- 4th ed.

p. cm

Includes bibliographical references and index.

ISBN 0-8153-3218-1 (hardbound) -- ISBN 0-8153-4072-9 (pbk.)

1. Cytology. 2. Molecular biology. I. Alberts, Bruce.

[DNLM: 1. Cells. 2. Molecular Biology.]

QH581.2 .M64 2002

571.6--dc21

2001054471 CIP

Published by Garland Science, a member of the Taylor & Francis Group, 29 West 35th Street, New York, NY 10001-2299

Printed in the United States of America

15 14 13 12 11 10 9 8 7 6 5 4 3 2 1

Preface

"There is a paradox in the growth of scientific knowledge. As information accumulates in ever more intimidating quantities, disconnected facts and impenetrable mysteries give way to rational explanations, and simplicity emerges from chaos."

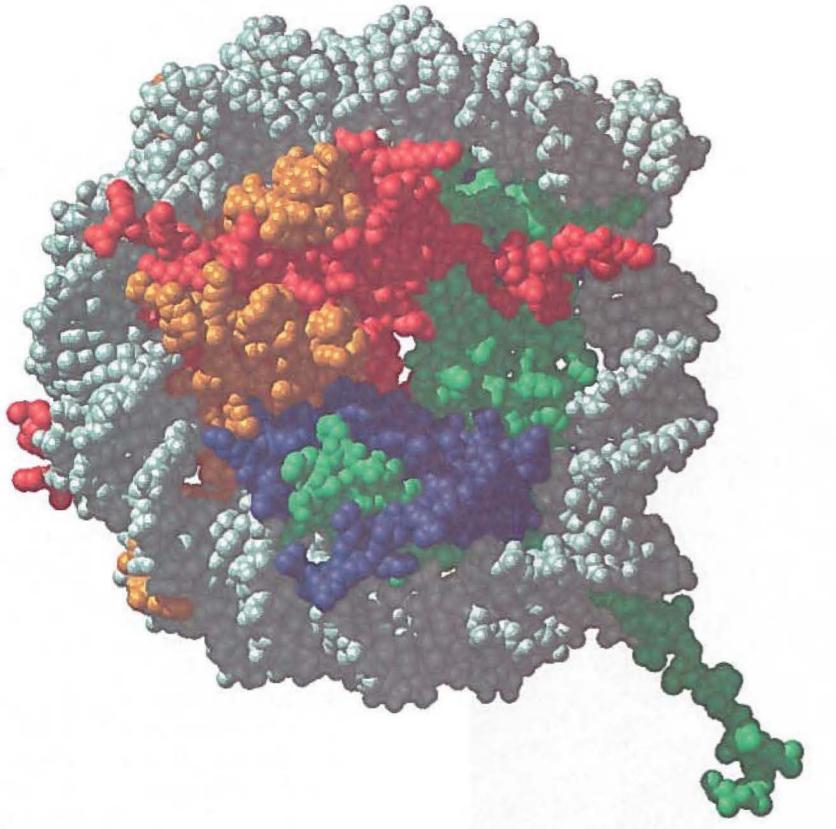
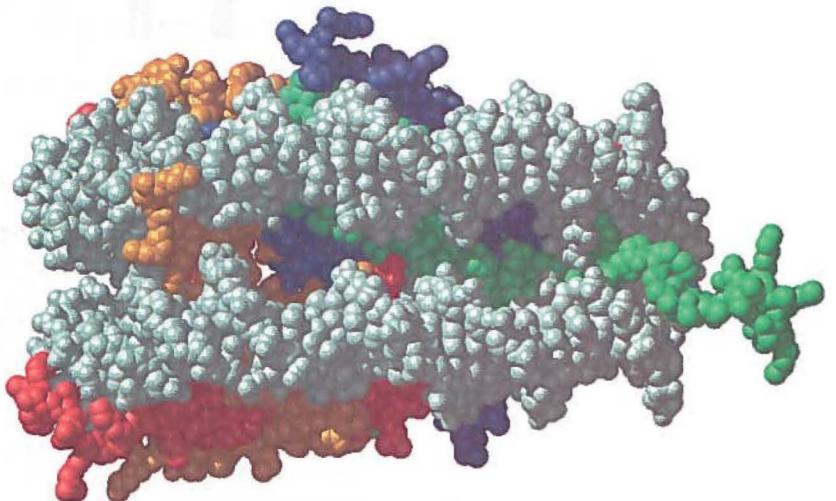
Thus began the preface of our first edition, written 18 years ago. Much of what we wrote in that preface holds for the present edition too. Our goals have not changed: we want to make cell biology comprehensible. We aim, as before, to give readers a perspective both on what is known, and on what is unknown. We have written the book for a wide range of students, but it should also prove useful for scientists wishing to follow progress outside their own specialized fields. As in past editions, each chapter is a joint composition of multiple authors and has been reviewed by a number of experts. This helps to explain why this edition, like the first, has been "a long time in gestation—three times longer than an elephant, five times longer than a whale."

Although information in the life sciences is expanding rapidly, human brain capacity is not. This discrepancy creates an increasing challenge for textbook writers, teachers, and especially students. We have been forced to think harder than ever in deciding what facts and concepts are essential. We have had to summarize—omitting much detail that, in overabundance, can prevent understanding. But specific examples are still needed to bring the subject to life. Thus, the most difficult part of the revision has been deciding what to leave out, while the most rewarding has been the opportunity that new discoveries provide to strengthen the conceptual framework.

What is new in the 4th edition?

Genomics has given us a new perspective that has demanded a complete recasting and expansion of the material on molecular genetics (Chapter 1 and Chapters 4 through 8). These six chapters can now be used as a stand-alone text in molecular biology. As before, the book ends with a major section on Cells in Their Social Context, but we have added a new chapter on Pathogens, Infection, and Innate Immunity. This reflects the remarkable advances in the understanding of the cell biology of infection, as well as the renewed awareness that infectious disease remains one of the greatest unconquered dangers in our world.

As our book makes clear, the complete sequencing of the genomes of hundreds of organisms, from bacteria to humans, has revolutionized our understanding of living things and the relationships between them. At last we can see what is there: the set of genes and proteins is finite, and we can list them. But we also recognize that these components are combined for use in marvelously subtle and complex ways, even in the simplest of organisms. Therefore, the traditional explanatory cartoons that we show on nearly every page of the book generally represent only the primitive first step toward an explanation. These drawings cannot capture the enormous complexity of the networks of protein-protein interactions that are responsible for most intracellular processes, whose understanding will require new and more quantitative forms of analysis. Thus, we are no longer as confident as we were 18 years ago that simplicity will eventually



The nucleosome. The basic structural unit of all eukaryotic chromosomes is the nucleosome. The DNA double helix (gray) is wrapped around a core particle of histone proteins (colored) to create the nucleosome. Nucleosomes are spaced roughly 200 nucleotide pairs apart along the chromosomal DNA. (Reprinted by permission from K. Luger et al., *Nature* 389:251–260, 1997. © Macmillan Magazines Ltd.)



DNA AND CHROMOSOMES

Life depends on the ability of cells to store, retrieve, and translate the genetic instructions required to make and maintain a living organism. This *hereditary* information is passed on from a cell to its daughter cells at cell division, and from one generation of an organism to the next through the organism's reproductive cells. These instructions are stored within every living cell as its **genes**, the information-containing elements that determine the characteristics of a species as a whole and of the individuals within it.

As soon as genetics emerged as a science at the beginning of the twentieth century, scientists became intrigued by the chemical structure of genes. The information in genes is copied and transmitted from cell to daughter cell millions of times during the life of a multicellular organism, and it survives the process essentially unchanged. What form of molecule could be capable of such accurate and almost unlimited replication and also be able to direct the development of an organism and the daily life of a cell? What kind of instructions does the genetic information contain? How are these instructions physically organized so that the enormous amount of information required for the development and maintenance of even the simplest organism can be contained within the tiny space of a cell?

The answers to some of these questions began to emerge in the 1940s, when researchers discovered, from studies in simple fungi, that genetic information consists primarily of instructions for making proteins. Proteins are the macromolecules that perform most cellular functions: they serve as building blocks for cellular structures and form the enzymes that catalyze all of the cell's chemical reactions (Chapter 3), they regulate gene expression (Chapter 7), and they enable cells to move (Chapter 16) and to communicate with each other (Chapter 15). The properties and functions of a cell are determined almost entirely by the proteins it is able to make. With hindsight, it is hard to imagine what other type of instructions the genetic information could have contained.

The other crucial advance made in the 1940s was the identification of **deoxyribonucleic acid (DNA)** as the likely carrier of genetic information. But the mechanism whereby the hereditary information is copied for transmission from cell to cell, and how proteins are specified by the instructions in the DNA, remained completely mysterious. Suddenly, in 1953, the mystery was solved when the structure of DNA was determined by James Watson and Francis Crick. As mentioned in Chapter 1, the structure of DNA immediately solved the problem of how the information in this molecule might be copied, or *replicated*. It also provided the first clues as to how a molecule of DNA might encode the instructions for making proteins. Today, the fact that DNA is the genetic material is so fundamental to biological thought that it is difficult to realize what an enormous intellectual gap this discovery filled.

Well before biologists understood the structure of DNA, they had recognized that genes are carried on **chromosomes**, which were discovered in the nineteenth century as threadlike structures in the nucleus of a eucaryotic cell that become visible as the cell begins to divide (Figure 4–1). Later, as biochemical analysis became possible, chromosomes were found to consist of both DNA and protein. We now know that the DNA carries the hereditary information of the cell (Figure 4–2). In contrast, the protein components of chromosomes function largely to package and control the enormously long DNA molecules so that they fit inside cells and can easily be accessed by them.

In this chapter we begin by describing the structure of DNA. We see how, despite its chemical simplicity, the structure and chemical properties of DNA make it ideally suited as the raw material of genes. The genes of every cell on Earth are made of DNA, and insights into the relationship between DNA and genes have come from experiments in a wide variety of organisms. We then consider how genes and other important segments of DNA are arranged on the long molecules of DNA that are present in chromosomes. Finally, we discuss how eucaryotic cells fold these long DNA molecules into compact chromosomes. This packing has to be done in an orderly fashion so that the chromosomes can be replicated and apportioned correctly between the two daughter cells at each cell division. It must also allow access of chromosomal DNA to enzymes that repair it when it is damaged and to the specialized proteins that direct the expression of its many genes.

This is the first of four chapters that deal with basic genetic mechanisms—the ways in which the cell maintains, replicates, expresses, and occasionally improves the genetic information carried in its DNA. In the following chapter (Chapter 5) we discuss the mechanisms by which the cell accurately replicates and repairs DNA; we also describe how DNA sequences can be rearranged through the process of genetic recombination. Gene expression—the process through which the information encoded in DNA is interpreted by the cell to guide the synthesis of proteins—is the main topic of Chapter 6. In Chapter 7, we describe how gene expression is controlled by the cell to ensure that each of the many thousands of proteins encrypted in its DNA is manufactured only at the proper time and place in the life of the cell. Following these four chapters on basic genetic mechanisms, we present an account of the experimental techniques used to study these and other processes that are fundamental to all cells (Chapter 8).

THE STRUCTURE AND FUNCTION OF DNA

Biologists in the 1940s had difficulty in accepting DNA as the genetic material because of the apparent simplicity of its chemistry. DNA was known to be a long polymer composed of only four types of subunits, which resemble one another chemically. Early in the 1950s, DNA was first examined by x-ray diffraction analysis, a technique for determining the three-dimensional atomic structure of a molecule (discussed in Chapter 8). The early x-ray diffraction results indicated that DNA was composed of two strands of the polymer wound into a helix. The observation that DNA was double-stranded was of crucial significance and

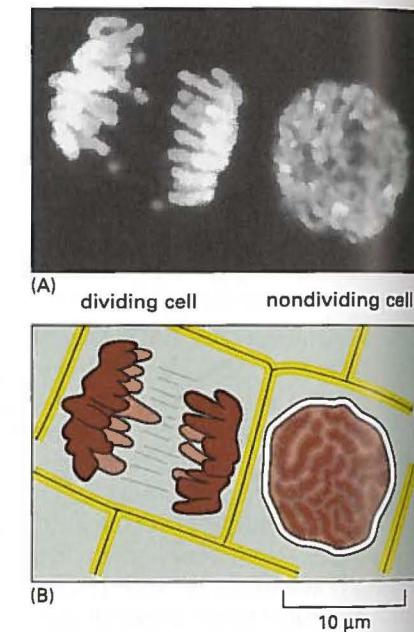


Figure 4-1 Chromosomes in cells. (A) Two adjacent plant cells photographed through a light microscope. The DNA has been stained with a fluorescent dye (DAPI) that binds to it. The DNA is present in chromosomes, which become visible as distinct structures in the light microscope only when they become compact structures in preparation for cell division, as shown on the left. The cell on the right, which is not dividing, contains identical chromosomes, but they cannot be clearly distinguished in the light microscope at this phase in the cell's life cycle, because they are in a more extended conformation. (B) Schematic diagram of the outlines of the two cells along with their chromosomes. (A, courtesy of Peter Shaw.)

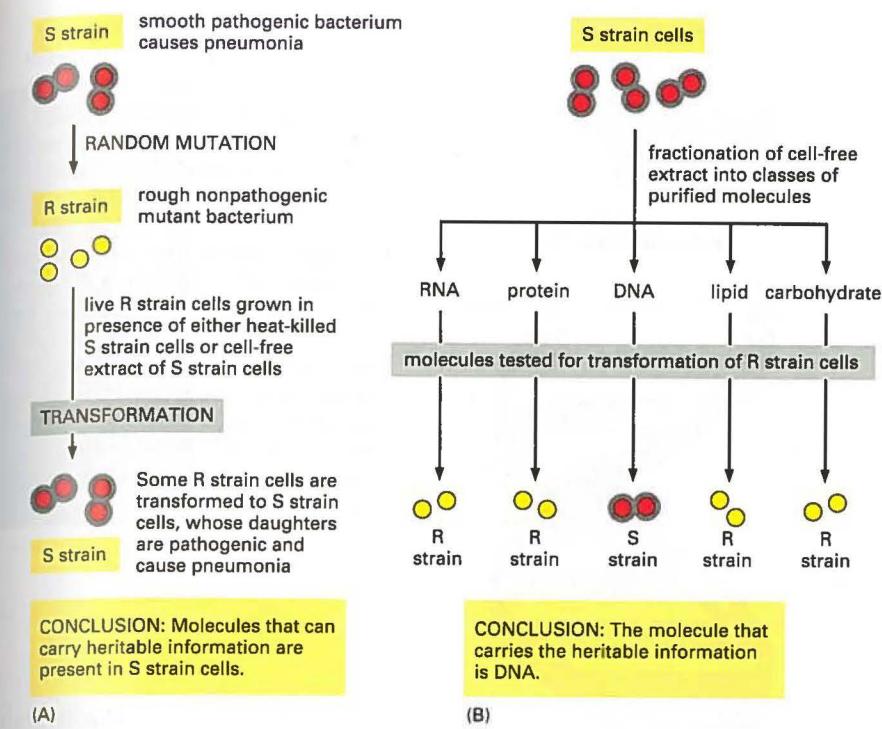


Figure 4-2 Experimental demonstration that DNA is the genetic material. These experiments, carried out in the 1940s, showed that adding purified DNA to a bacterium changed its properties and that this change was faithfully passed on to subsequent generations. Two closely related strains of the bacterium *Streptococcus pneumoniae* differ from each other in both their appearance under the microscope and their pathogenicity. One strain appears smooth (S) and causes death when injected into mice, and the other appears rough (R) and is nonlethal. (A) This experiment shows that a substance present in the S strain can change (or transform) the R strain into the S strain and that this change is inherited by subsequent generations of bacteria. (B) This experiment, in which the R strain has been incubated with various classes of biological molecules obtained from the S strain, identifies the substance as DNA.

provided one of the major clues that led to the Watson–Crick structure of DNA. Only when this model was proposed did DNA's potential for replication and information encoding become apparent. In this section we examine the structure of the DNA molecule and explain in general terms how it is able to store hereditary information.

A DNA Molecule Consists of Two Complementary Chains of Nucleotides

A DNA molecule consists of two long polynucleotide chains composed of four types of nucleotide subunits. Each of these chains is known as a **DNA chain**, or a **DNA strand**. **Hydrogen bonds** between the base portions of the nucleotides hold the two chains together (Figure 4–3). As we saw in Chapter 2 (Panel 2–6, pp. 120–121), nucleotides are composed of a five-carbon sugar to which are attached one or more phosphate groups and a nitrogen-containing base. In the case of the nucleotides in DNA, the sugar is deoxyribose attached to a single phosphate group (hence the name deoxyribonucleic acid), and the base may be either **adenine (A)**, **cytosine (C)**, **guanine (G)**, or **thymine (T)**. The nucleotides are covalently linked together in a chain through the sugars and phosphates, which thus form a “backbone” of alternating sugar-phosphate-sugar-phosphate (see Figure 4–3). Because only the base differs in each of the four types of subunits, each polynucleotide chain in DNA is analogous to a necklace (the backbone) strung with four types of beads (the four bases A, C, G, and T). These same symbols (A, C, G, and T) are also commonly used to denote the four different nucleotides—that is, the bases with their attached sugar and phosphate groups.

The way in which the nucleotide subunits are lined together gives a DNA strand a chemical polarity. If we think of each sugar as a block with a protruding knob (the 5' phosphate) on one side and a hole (the 3' hydroxyl) on the other (see Figure 4–3), each completed chain, formed by interlocking knobs with holes, will have all of its subunits lined up in the same orientation. Moreover, the two ends of the chain will be easily distinguishable, as one has a hole (the 3' hydroxyl) and the other a knob (the 5' phosphate) at its terminus. This polarity in a DNA chain is indicated by referring to one end as the **3' end** and the other as the **5' end**.

The three-dimensional structure of DNA—the **double helix**—arises from the chemical and structural features of its two polynucleotide chains. Because

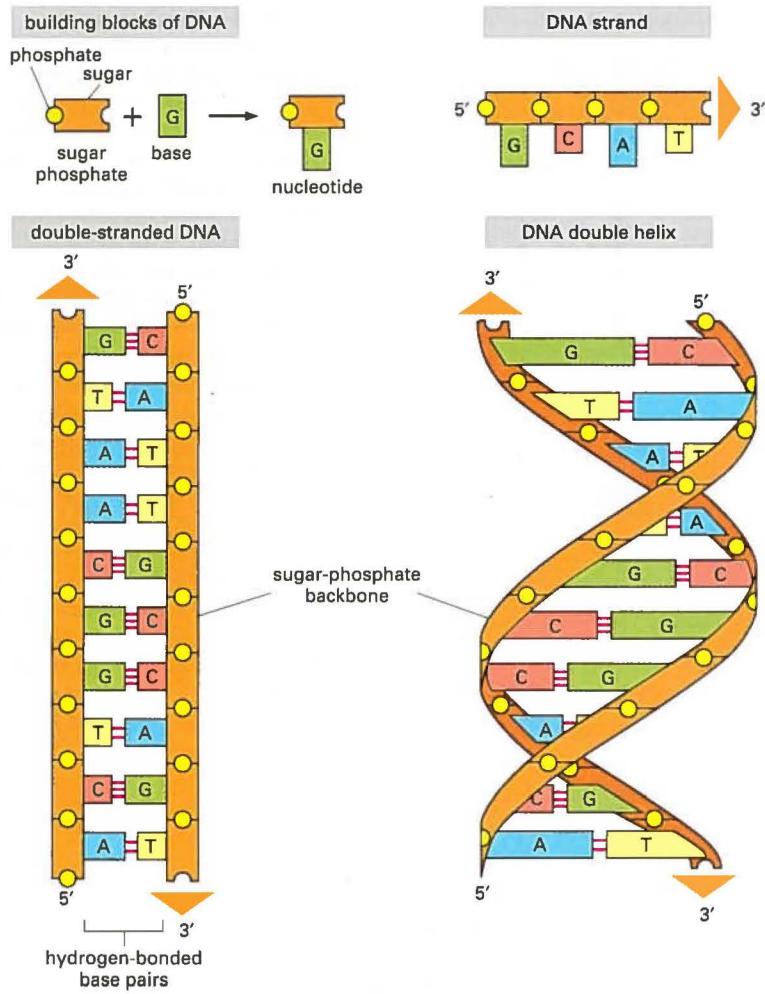


Figure 4-3 DNA and its building blocks. DNA is made of four types of nucleotides, which are linked covalently into a polynucleotide chain (a DNA strand) with a sugar-phosphate backbone from which the bases (A, C, G, and T) extend. A DNA molecule is composed of two DNA strands held together by hydrogen bonds between the paired bases. The arrowheads at the ends of the DNA strands indicate the polarities of the two strands, which run antiparallel to each other in the DNA molecule. In the diagram at the bottom left of the figure, the DNA molecule is shown straightened out; in reality, it is twisted into a double helix, as shown on the right. For details, see Figure 4-5.

these two chains are held together by hydrogen bonding between the bases on the different strands, all the bases are on the inside of the double helix, and the sugar-phosphate backbones are on the outside (see Figure 4-3). In each case, a bulkier two-ring base (a purine; see Panel 2-6, pp. 120–121) is paired with a single-ring base (a pyrimidine); A always pairs with T, and G with C (Figure 4-4). This *complementary base-pairing* enables the **base pairs** to be packed in the

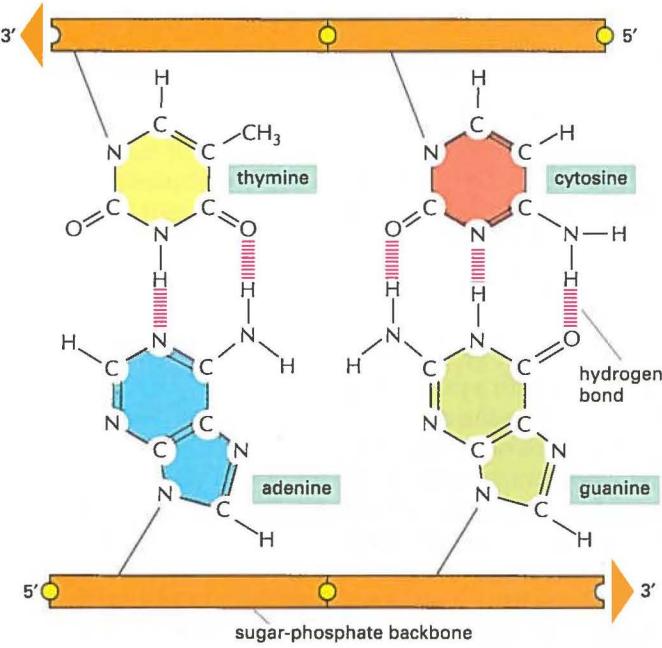


Figure 4-4 Complementary base pairs in the DNA double helix. The shapes and chemical structure of the bases allow hydrogen bonds to form efficiently only between A and T and between G and C, where atoms that are able to form hydrogen bonds (see Panel 2-3, pp. 114–115) can be brought close together without distorting the double helix. As indicated, two hydrogen bonds form between A and T, while three form between G and C. The bases can pair in this way only if the two polynucleotide chains that contain them are antiparallel to each other.

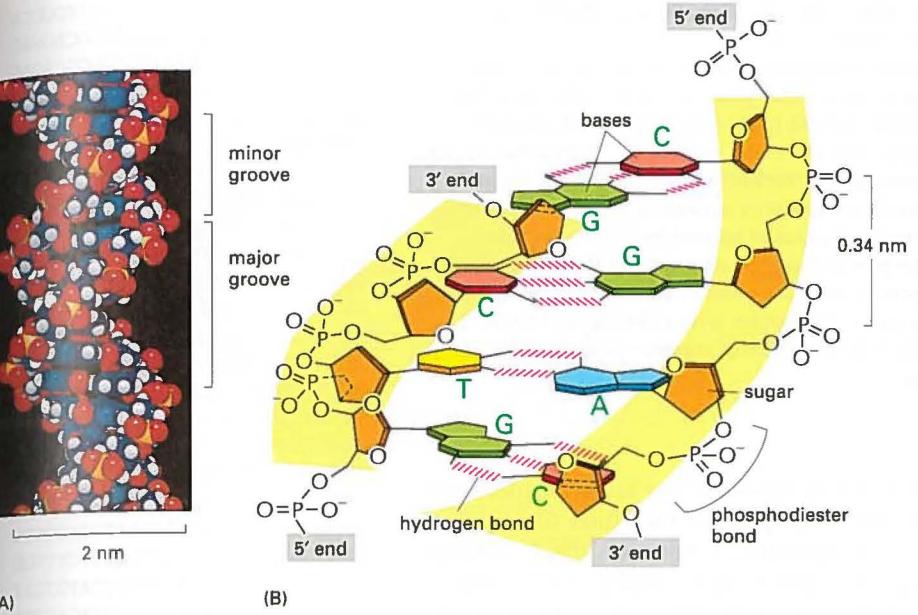


Figure 4-5 The DNA double helix. (A) A space-filling model of 1.5 turns of the DNA double helix. Each turn of DNA is made up of 10.4 nucleotide pairs and the center-to-center distance between adjacent nucleotide pairs is 3.4 nm. The coiling of the two strands around each other creates two grooves in the double helix. As indicated in the figure, the wider groove is called the major groove, and the smaller the minor groove. (B) A short section of the double helix viewed from its side, showing four base pairs. The nucleotides are linked together covalently by phosphodiester bonds through the 3'-hydroxyl (–OH) group of one sugar and the 5'-phosphate (P) of the next. Thus, each polynucleotide strand has a chemical polarity; that is, its two ends are chemically different. The 3' end carries an unlinked –OH group attached to the 3' position on the sugar ring; the 5' end carries a free phosphate group attached to the 5' position on the sugar ring.

energetically most favorable arrangement in the interior of the double helix. In this arrangement, each base pair is of similar width, thus holding the sugar-phosphate backbones an equal distance apart along the DNA molecule. To maximize the efficiency of base-pair packing, the two sugar-phosphate backbones wind around each other to form a double helix, with one complete turn every ten base pairs (Figure 4-5).

The members of each base pair can fit together within the double helix only if the two strands of the helix are **antiparallel**—that is, only if the polarity of one strand is oriented opposite to that of the other strand (see Figures 4-3 and 4-4). A consequence of these base-pairing requirements is that each strand of a DNA molecule contains a sequence of nucleotides that is exactly **complementary** to the nucleotide sequence of its partner strand.

The Structure of DNA Provides a Mechanism for Heredity

Genes carry biological information that must be copied accurately for transmission to the next generation each time a cell divides to form two daughter cells. Two central biological questions arise from these requirements: how can the information for specifying an organism be carried in chemical form, and how is it accurately copied? The discovery of the structure of the DNA double helix was a landmark in twentieth-century biology because it immediately suggested answers to both questions, thereby resolving at the molecular level the problem of heredity. We discuss briefly the answers to these questions in this section, and we shall examine them in more detail in subsequent chapters.

DNA encodes information through the order, or sequence, of the nucleotides along each strand. Each base—A, C, T, or G—can be considered as a letter in a four-letter alphabet that spells out biological messages in the chemical structure of the DNA. As we saw in Chapter 1, organisms differ from one another because their respective DNA molecules have different nucleotide sequences and, consequently, carry different biological messages. But how is the nucleotide alphabet used to make messages, and what do they spell out?

As discussed above, it was known well before the structure of DNA was determined that genes contain the instructions for producing proteins. The DNA messages must therefore somehow encode proteins (Figure 4-6). This relationship immediately makes the problem easier to understand, because of the chemical character of proteins. As discussed in Chapter 3, the properties of a protein, which are responsible for its biological function, are determined by its three-dimensional structure, and its structure is determined in turn by the linear

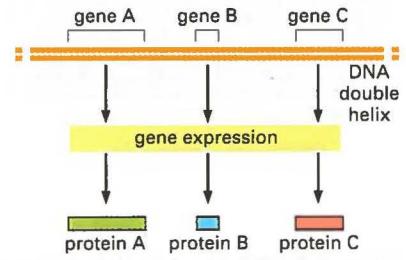


Figure 4-6 The relationship between genetic information carried in DNA and proteins.

Figure 4–7 The nucleotide sequence of the human β -globin gene. This gene carries the information for the amino acid sequence of one of the two types of subunits of the hemoglobin molecule, which carries oxygen in the blood. A different gene, the α -globin gene, carries the information for the other type of hemoglobin subunit (a hemoglobin molecule has four subunits, two of each type). Only one of the two strands of the DNA double helix containing the β -globin gene is shown; the other strand has the exact complementary sequence. By convention, a nucleotide sequence is written from its 5' end to its 3' end, and it should be read from left to right in successive lines down the page as though it were normal English text. The DNA sequences highlighted in yellow show the three regions of the gene that specify the amino sequence for the β -globin protein. We see in Chapter 6 how the cell connects these three sequences together to synthesize a full-length β -globin protein.

sequence of the amino acids of which it is composed. The linear sequence of nucleotides in a gene must therefore somehow spell out the linear sequence of amino acids in a protein. The exact correspondence between the four-letter nucleotide alphabet of DNA and the twenty-letter amino acid alphabet of proteins—the genetic code—is not obvious from the DNA structure, and it took over a decade after the discovery of the double helix before it was worked out. In Chapter 6 we describe this code in detail in the course of elaborating the process, known as *gene expression*, through which a cell translates the nucleotide sequence of a gene into the amino acid sequence of a protein.

The complete set of information in an organism's DNA is called its **genome**, and it carries the information for all the proteins the organism will ever synthesize. (The term genome is also used to describe the DNA that carries this information.) The amount of information contained in genomes is staggering: for example, a typical human cell contains 2 meters of DNA. Written out in the four-letter nucleotide alphabet, the nucleotide sequence of a very small human gene occupies a quarter of a page of text (Figure 4–7), while the complete sequence of nucleotides in the human genome would fill more than a thousand books the size of this one. In addition to other critical information, it carries the instructions for about 30,000 distinct proteins.

At each cell division, the cell must copy its genome to pass it to both daughter cells. The discovery of the structure of DNA also revealed the principle that makes this copying possible: because each strand of DNA contains a sequence of nucleotides that is exactly complementary to the nucleotide sequence of its partner strand, each strand can act as a **template**, or mold, for the synthesis of a new complementary strand. In other words, if we designate the two DNA strands as S and S', strand S can serve as a template for making a new strand S', while strand S' can serve as a template for making a new strand S (Figure 4–8).

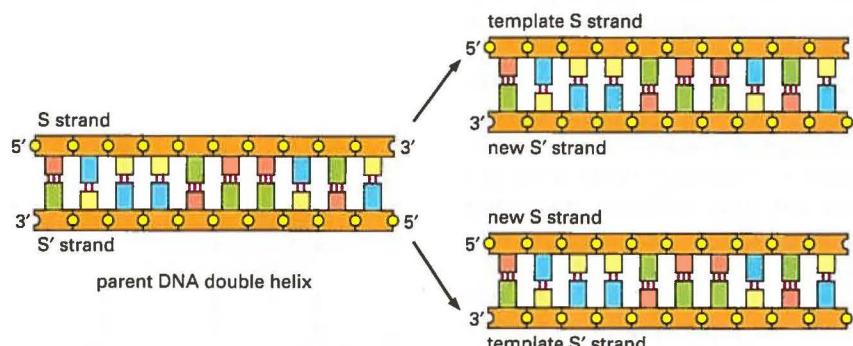


Figure 4–8 DNA as a template for its own duplication. As the nucleotide A successfully pairs only with T, and G with C, each strand of DNA can specify the sequence of nucleotides in its complementary strand. In this way, double-helical DNA can be copied precisely.

```

CCCTGTGGAGCCACACCTAGGGTGGCCA
ATCTACTCCCAGGAGCAGGGAGGCAGAG
CCAGGGCTGGCATAAAAGTCAGGGCAGAG
CCATCTATTGCTTACATTGCTCTGACAC
AACTGTGTTCACTAGCAACTCAAACAGACA
CCATGGTGCACCTGACTCCTGAGGAGAAGT
CTGCCGTACTGCCCTGTTGGGCAAGGTGA
ACGTGGATGAAGTGGTGGTGGGAGGCGCTGG
GCAGGTTGGTATCAAGGTTACAAGACAGGT
TTAAGGAGACCAATAGAACTGGCATGTG
GAGACAGAGAAGACTCTGGGTTCTGATA
GGCACTGACTCTCTGCCTATTGGTCTAT
TTTCCACCCCTTAGGCTGCTGGGGTCTAC
CCTGGACCCAGAGGTTCTTGAGTCCCTT
GGGGATCTGTCACCTGATGCTGTTAG
GGCACCCCTAAGGTGAAGGCTCATGGCAAG
AAAGTGTGGCTGGCTTTAGTGTGATGGCTG
GCTCACCTGGACAACCTCAAGGGCACCTT
GCCACACTGAGTGAAGTCAACTGTGACAAG
CTGCACTGGATCTGAGAACTTCAGGGTG
AGTCTATGGGACCCCTGATGTTTCTTCC
CCTCTTTCTATGTTAAGTTCATGTCAT
AGGAAGGGAGAAGTAACAGGTACAGTTT
AGAATGGGAAACAGACGAATGATTGCACTA
GTGTTGAAGTCTCAGGATCTGTTAGTTTC
TTTTATTGCTGTTCTAAACAATTGTTTC
TTTGTAAATTCTGTTCTTTTTTTT
CTTCTCGCAATTACTTAACTTAACTTAA
TATCTGAGATACTTAAAGTAACTTAA
AAAACATTACACAGTCTGCTAGTACATT
ACTATTGGAATATAATGTGTGCTTATTG
ATATTCTAAATCTCTACTTTATTTCCTT
TTATTTTAATTGATACATAATCATTATAC
ATATTATGGGTTAAAGTGTATGTTAA
TATGTGTACACATATTGACCAATCAGGT
AATTTGCATTGTAATTAAAAATGCT
TTCTCTTTAATATACTTTTGTGTTATC
TTATTCTAATACTTCCCTAATCTCTTC
TTTCAGGCAATAATGATACAATGTATCAT
GCCTCTTGACCAATTCTAAAGATAACAG
TGATAATTCTGGTAAAGCAATAGCAAT
ATTCTGCAATAATATTCTGCAATATAA
ATTGTAACGTGATGTAAGAGGTTCATATTG
CTAATAGCAGCTACAATCCAGCTACCATTC
TGCTTTATTGTTATGGTGGGATAAGGCTG
GATTATTCTGAGTCCAAGCTAGGGCTTTT
GCTAATCATGTTCATACCTCTTATCTCCT
CCCACAGCTCCTGGCAACGTGCTGGCTG
TGTGCTGGCCCATCTTGGCAAAGAATT
CACCCACCACTGAGGCTGCCTATCAGAA
AGTGGTGGCTGGTGGCTAATGCCCTGGC
CCACAAGTACTAAGCTGCTTCTTGC
TGTCCAATTCTATTAAGGTCCTTGTG
CCCTAAAGTCCAACACTAAACTGGGGATA
TTATGAGGGCTTGAGCATCTGATTCTG
CCTAATAAAAACATTATTCTGAAATTTT
TGATGTTATTAAATTATTCTGAAATTTT
ACTAAAAAGGGAATGTGGGAGGTCACTGCA
TTTAAACATAAGGAAATGATGAGCTGTT
AAACCTGGGAAATACACTATATCTTAA
CTCCATGAAAGAAGGTGAGGCTGCAACAG
CTAATGCACATTGGCAACAGCCCCCTGATGC
CTATGCCATTATCCTCCAGAAAAGGAT
TCTTGAGGGCTTGATTTGAGGTTAAAG
TTTGCTATGCTGTTACATACTTAT
TGTGTTAGCTGTCCTCATGAAATGCTTT

```

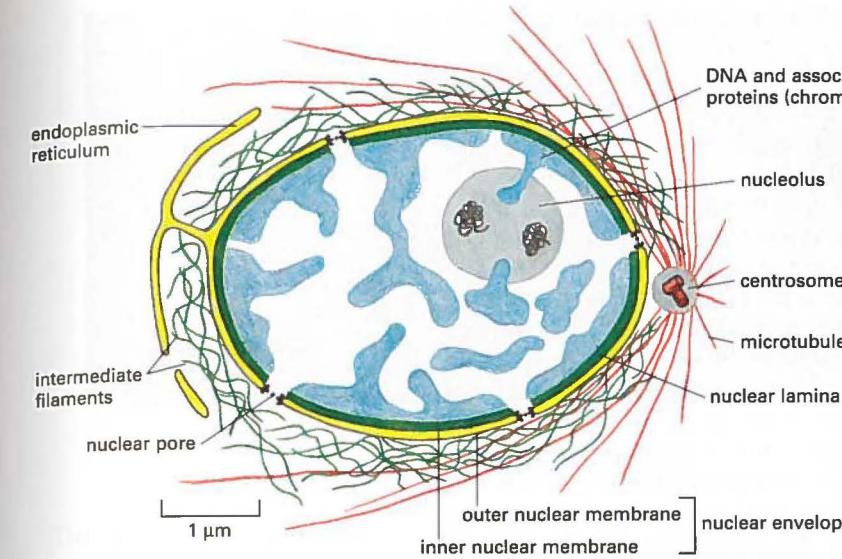


Figure 4–9 A cross-sectional view of a typical cell nucleus. The nuclear envelope consists of two membranes, the outer one being continuous with the endoplasmic reticulum membrane (see also Figure 12–9). The space inside the endoplasmic reticulum (the ER lumen) is colored yellow; it is continuous with the space between the two nuclear membranes. The lipid bilayers of the inner and outer nuclear membranes are connected at each nuclear pore. Two networks of intermediate filaments (green) provide mechanical support for the nuclear envelope; the intermediate filaments inside the nucleus form a special supporting structure called the nuclear lamina.

Thus, the genetic information in DNA can be accurately copied by the beautifully simple process in which strand S separates from strand S', and each separated strand then serves as a template for the production of a new complementary partner strand.

The ability of each strand of a DNA molecule to act as a template for producing a complementary strand enables a cell to copy, or *replicate*, its genes before passing them on to its descendants. In the next chapter we describe the elegant machinery the cell uses to perform this enormous task.

In Eucaryotes, DNA Is Enclosed in a Cell Nucleus

Nearly all the DNA in a eucaryotic cell is sequestered in a nucleus, which occupies about 10% of the total cell volume. This compartment is delimited by a **nuclear envelope** formed by two concentric lipid bilayer membranes that are punctured at intervals by large nuclear pores, which transport molecules between the nucleus and the cytosol. The nuclear envelope is directly connected to the extensive membranes of the endoplasmic reticulum. It is mechanically supported by two networks of intermediate filaments: one, called the **nuclear lamina**, forms a thin sheetlike meshwork inside the nucleus, just beneath the inner nuclear membrane; the other surrounds the outer nuclear membrane and is less regularly organized (Figure 4–9).

The nuclear envelope allows the many proteins that act on DNA to be concentrated where they are needed in the cell, and, as we see in subsequent chapters, it also keeps nuclear and cytosolic enzymes separate, a feature that is crucial for the proper functioning of eucaryotic cells. Compartmentalization, of which the nucleus is an example, is an important principle of biology; it serves to establish an environment in which biochemical reactions are facilitated by the high concentration of both substrates and the enzymes that act on them.

Summary

Genetic information is carried in the linear sequence of nucleotides in DNA. Each molecule of DNA is a double helix formed from two complementary strands of nucleotides held together by hydrogen bonds between G-C and A-T base pairs. Duplication of the genetic information occurs by the use of one DNA strand as a template for formation of a complementary strand. The genetic information stored in an organism's DNA contains the instructions for all the proteins the organism will ever synthesize. In eucaryotes, DNA is contained in the cell nucleus.