# ESSENTIAL GENETICS

PETER J. RUSSELL PhD

# GENETICS

PETER J. RUSSELL PhD
*Professor of Biology*
*Reed College, Portland, Oregon*

BLACKWELL SCIENTIFIC PUBLICATIONS
OXFORD LONDON EDINBURGH
BOSTON PALO ALTO MELBOURNE

Blackwell Scientific Publications
Editorial offices:
Osney Mead, Oxford, OX2 0EL
8 John Street, London, WC1N 2ES
23 Ainslie Place, Edinburgh, EH3 6AJ
52 Beacon Street, Boston
Massachusetts 02108, USA
667 Lytton Avenue, Palo Alto
California 94301, USA
107 Barry Street, Carlton
Victoria 3053, Australia



First published 1980
Reprinted 1982
Second edition 1987

Typeset by Enset (Photosetting),
Midsomer Norton, Bath, Avon
Printed in Great Britain by
Butler & Tanner Ltd,
Frome and London

DISTRIBUTORS

USA and Canada
Blackwell Scientific Publications Inc
PO Box 50009, Palo Alto
California 94301

Australia
Blackwell Scientific Publications
(Australia) Pty Ltd
107 Barry Street,
Carlton, Victoria 3053

British Library
Cataloguing in Publication Data

Russell, Peter J.
Essential genetics—2nd ed.
1. Genetics
I. Title
575.1 QH430

ISBN 0-632-01602-7

Library of Congress
Cataloging-in-Publication Data

Russell, Peter J.
Essential genetics.
Updated ed. of: Lecture notes on genetics. 1980.
Includes index.
1. Genetics. I. Russell, Peter J. Essential genetics. II. Title.
QH430.R867 1987 575.1 87–734

ISBN 0-632-01602-7

# ESSENTIAL

SECOND EDITION

# Essential Genetics

# Preface

*Essential Genetics*, the second edition of *Lecture Notes on Genetics*, is a brief but balanced review of the important aspects of genetics. It is intended to be used in a one-semester college course in genetics either as the primary text or as a supplementary text. Prior completion of college courses in general biology and chemistry is assumed, although some material from general biology is reviewed in this text to make the discussion of topics complete.

*Essential Genetics* is organized as a series of concise chapters that present the principles of genetics in an up-to-date, readable way. Many of the classical and modern genetics experiments are described so that the reader can gain an appreciation of the methodology of genetics as well as the facts and concepts of the subject. It is hoped that this book will give the reader a solid understanding of genetics as well as a sense of the excitement that pervades this rapidly progressing field. For those readers who are stimulated to seek further information about genetics, historical and up-to-date references are provided at the end of each chapter.

The sequence of chapters in this book is the same as in *Lecture Notes on Genetics*, beginning with the molecular aspects of genetics, progressing to transmission genetics, and then discussing the regulation of gene expression, extranuclear inheritance, and population genetics. All chapters of the book have been revised, with particular attention directed toward updating the material on molecular genetics, while maintaining the book at a reasonable length. Significant revision of the molecular genetics sections of the book include: addition of a discussion of other forms of DNA, including Z-DNA (Chapter 1); updating the presentations of chromatin structure and repetitive DNA sequences (Chapter 2) and DNA replication and topoisomerases; addition of a discussion of transposable genetic elements in prokaryotes and eukaryotes (Chapter 2); expanding the discussion of the individual steps in mitosis and meiosis (Chapter 5); addition of a description of the molecular structure of centromeres and telomeres (Chapter 5); updating and additional detailing of the transcription process, particularly promoters, RNA synthesis, and processing (Chapter 7); updating of the material on protein synthesis and the signal hypothesis for the secretion of proteins (Chapter 8); presentation of new material on the molecular aspects of mitochondria and chloroplasts (Chapter 17); and updating the chapters on gene regulation in bacteria

(Chapter 19) and in eukaryotes (Chapter 20). In addition, Chapter 12, "Recombinant DNA," was revised extensively to reflect the advances that have been made in that area since the first edition. Chapter 12 now includes, for example, a discussion of rapid DNA sequencing and more complete discussions of gene cloning techniques. Chapter 18, "Biochemical Genetics," was revised to include a new section on the genetic control of protein structure. Chapter 20, "Regulation of Gene Expression in Eukaryotes," was revised to include a more complete discussion of the general aspects of gene regulation, the roles of nonhistones, and the action of steroid hormones. Chapter 21, "Population Genetics," has been greatly expanded for this edition. Another significant revision is the addition of a number of new illustrations and photographs throughout the text to reinforce the text discussions.

Two pedagogical features have been added to the text of this second edition to aid students and to enhance their understanding and appreciation of genetics principles. First, a glossary of terms has been added so that the reader can quickly learn the key genetics terms and their meanings. These terms are in bold face in the text. Second, a set of questions and problems has been provided for each chapter so that the reader can test his or her understanding of the concepts presented in the chapters. The answers to the questions and problems are provided at the end of the book.

**Acknowledgements**

I acknowledge my wife, Jenny, and my children, Steven and Kristie, for their support during this project. I am grateful to all of the people at Blackwell Scientific Publications, Inc. who have worked with me on the book, especially John Staples. I thank Little, Brown & Co. for giving permission to use the "Questions and Problems" from my text, GENETICS, © 1986. A number of those questions and problems were graciously provided by Rowland H. Davis (University of California, Irvine), H. Branch Howe (University of Georgia, Athens), and David D. Perkins (Stanford University).

I thank the following people for providing new photographs for this edition: M. Meyer (Universiteit van Amsterdam, The Netherlands); P. Oudet (Laboratoire de Génétique Moleculaire des Eucaryotes du C.N.R.S., Institut de Chimie Biologique, Strasbourg, France); M. Meselson (Harvard University); O. L. Miller, Jr (University of Virginia, Charlottesville); G. Stoffler (Max-Planck Institut für Molekulare Genetik, Berlin-Dahlem, Federal Republic of Germany); and G. Morgan (University of Washington, Seattle).

I am grateful to the Literary Executor of the late Sir Ronald A. Fisher, FRS, to Dr. Frank Yates, FRS, and to Longman Group Ltd, London, for

permission to reprint part of a table of chi-square probabilities from their book *Statistical Tables for Biological, Agricultural, and Medical Research* (6th edition, 1974).

Last, but by no means least, I thank the reviewers of the second edition manuscript.

Peter J. Russell
Reed College

# Contents

# Chapter 1

# The Genetic Material

**Outline**



The central theme of this book is the genetic material: its nature, structure, organization, replication, expression, etc. The approach used will be to discuss the salient facts in the context of the current literature and the analytical methods used in the areas under discussion. The emphasis of the book will be on the interrelationships between genetics, molecular biology, and biochemistry.

## Requirements for the genetic material

The genetic material is of central importance to cell function and therefore must fulfill a number of basic requirements:

1. It must contain the information for cell structure, function, and reproduction in a stable form. This information is encoded in the sequence of basic building blocks of the genetic material.
2. It must be possible to replicate the genetic material accurately such that the same genetic information is present in descendant cells and in successive generations.
3. The information coded in the genetic material must be able to be decoded to produce the molecules essential for the structure and function of cells.
4. The genetic material must be capable of (infrequent) variation. Specifically, mutation and recombination of the genetic material are the foundations for the evolutionary process.

The nucleic acids, **deoxyribonucleic acid** (DNA) and **ribonucleic acid** (RNA), meet all these requirements.

## Nucleic acid structure

Both DNA and RNA are linear polymeric macromolecules. The monomeric unit is called a **nucleotide**; it is **deoxyribonucleotide** in the case of DNA and **ribonucleotide** in RNA. A nucleotide consists of three components: a nitrogenous base (which is a derivative of either purine or pyrimidine), a pentose sugar, and a phosphate group (Fig. 1.1).

The carbon positions in the pentose sugar ring are labeled 1′ to 5′ to distinguish them from the numbering of the positions in the ring structure of the bases. The phosphoryl groups may be attached to any hydroxyl group of

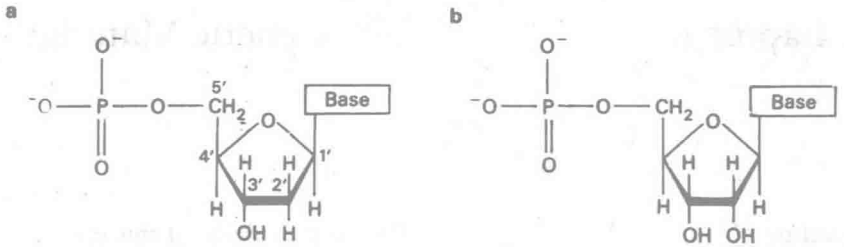
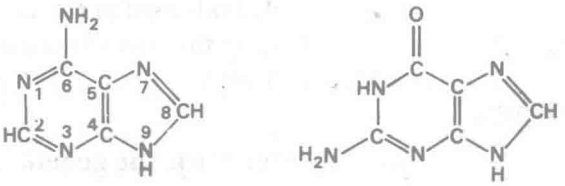


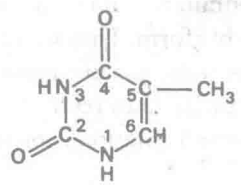

Fig. 1.1. Structure of the nucleotide components of DNA and RNA. (a) Deoxyribonucleoside 5′-monophosphate (monomeric unit of DNA). (b) Ribonucleoside 5′-monophosphate (monomeric unit of RNA).



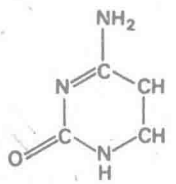
Adenine (A)



Guanine (G)

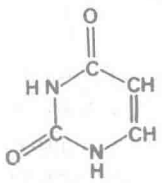


Thymine (T)
(5-methyluracil)



Cytosine (C)

Fig. 1.2. Purine and pyrimidine nitrogenous bases found in DNA.



Uracil (U)

Fig. 1.3. The structure of uracil, the nitrogenous base found in RNA instead of thymine.

the sugar and the 5′- and 3′-nucleotides are of particular importance for the structure and function of DNA and RNA.

Four different deoxyribonucleotides are the major components of DNA. These are distinguished by the type of nitrogenous base they contain. The four bases characteristic of the deoxyribonucleotides are the purine derivatives, adenine (A) and guanine (G), and the pyrimidine derivatives, thymine (T) and cytosine (C) (Fig. 1.2).

Similarly, RNA is characterized by four different ribonucleotides that, like the monomeric units of DNA, contain the bases adenine, guanine, and cytosine. However, instead of thymine, RNA contains the pyrimidine derivative uracil (U), which has chemical and physical properties similar to those of thymine (Fig. 1.3).

In both DNA and RNA, the bases are attached to the pentose moiety by a covalent bond between the 1′ carbon of the sugar and the 9-position nitrogen of the purines or the 3-position nitrogen of the pyrimidines.

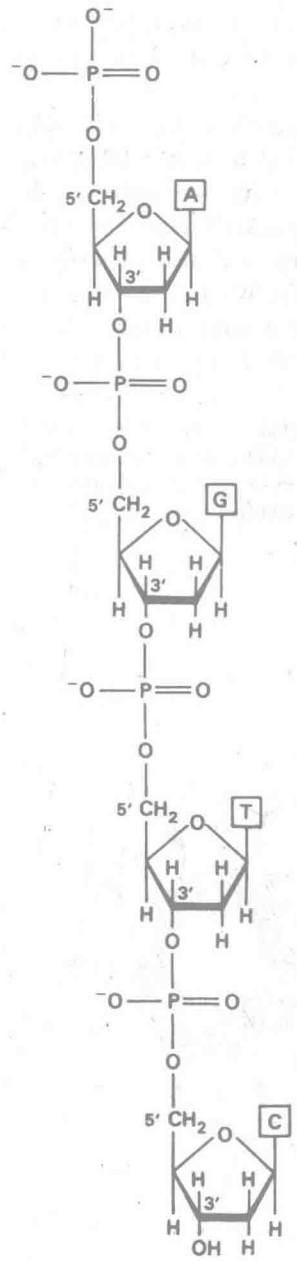



Fig. 1.4. An oligodeoxyribonucleotide chain showing the linkages between the monomeric units in a single DNA chain.

Another distinction between DNA and RNA is the nature of the pentose sugar each contains. Deoxyribonucleotides contain deoxyribose, whereas ribonucleotides contain ribose. As a result the two nucleic acids have different chemical properties, which are biologically important (e.g. enzymes can be specific for DNA or RNA) and may be exploited to separate the two molecules in the laboratory.

In DNA and RNA the mononucleotides are linked together by 3′, 5′-phosphodiester bonds. Thus the backbone of both molecules consists of alternating phosphate and pentose groups. The bases are not part of the backbone structure. An example of an oligodeoxyribonucleotide is shown in Fig. 1.4.

Polynucleotides have polarity. The pentose sugar at one end of the chain has a 5′-hydroxyl or phosphoryl group (5′ end), and the sugar at the other end has a 3′-hydroxyl group (3′ end). A shorthand way to represent a polynucleotide strand is depicted in Fig. 1.5.

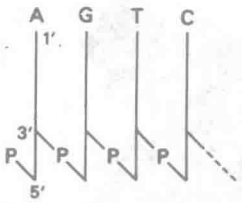



Fig. 1.5. A shorthand way to represent a polynucleotide chain.

## The DNA double helix

In 1953 James D. Watson and Francis H. C. Crick proposed that DNA is in the form of a double-stranded, right-handed helix. (A right-handed helix is one that winds clockwise when viewed from the end.) The evidence for their hypothesis was as follows:

1. The DNA molecule consists of bases, sugars, and phosphoryl groups linked together as a polynucleotide chain as discussed earlier.
2. E. Chargaff analyzed the nucleotides released by chemical hydrolysis and found that the total amount of purines present is always equal to the total amount of pyrimidines present. More specifically, adenine always equals thymine (A = T), and guanine always equals cytosine (G = C). Thus the following equations hold for double-stranded DNAs:

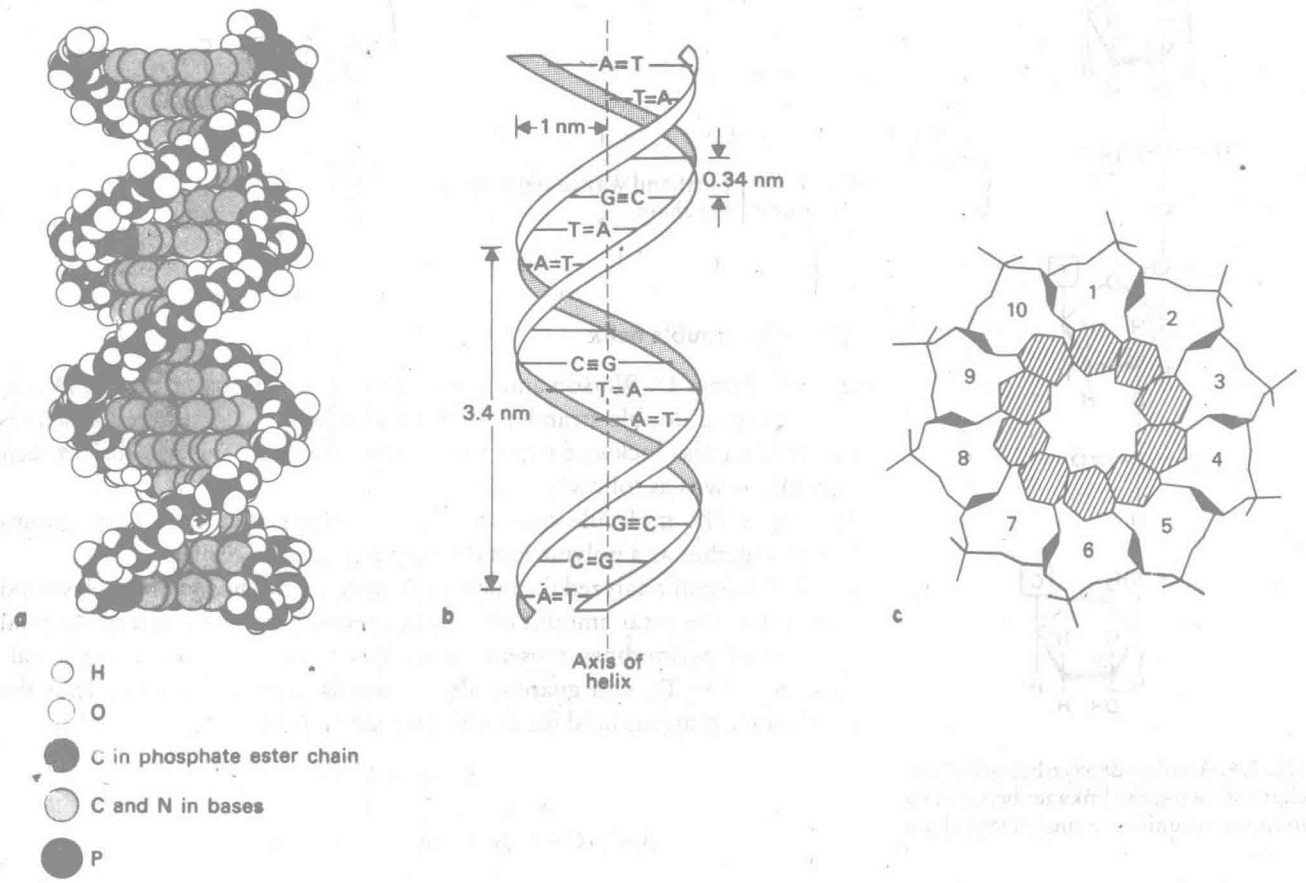
$$A+G = C+T$$
$$A+G/C+T = 1$$
$$A+T/G+C \text{ does not } = 1 \text{ (in most cases)}$$

The last is called the base ratio of the DNA and is usually expressed as % GC. The base ratio varies widely among organisms but it remains constant for any one species.

R. Franklin, M. H. F. Wilkins and coworkers analyzed fibers of DNA by x-ray diffraction. The patterns they obtained indicated that DNA was a helical structure consisting of two or more chains wound around each other.

Watson and Crick fitted the chemical and physical data into a symmetric structure that was compatible with all the facts and also possessed the properties one would expect of genetic material. The Watson–Crick model of DNA involves two polynucleotide chains that are wound around each other to form a *double helix* (Fig. 1.6). The two chains are joined together by

**Fig. 1.6.** The Watson–Crick model of DNA. (a) Molecular model of DNA double helix. (After M. Feughelman et al. 1955. *Nature*, 175:834, courtesy of M. H. F. Wilkins). (b) Diagrammatic representation of the DNA double helix. (c) Diagram of one of the strands of the helix, viewed down the central axis. The sugar-phosphate backbone is outside, the bases (all pyrimidines here; in hatching) are inside. The tenfold symmetry is evident, thus enforcing a repeat in the helix after 10 nucleotide residues.



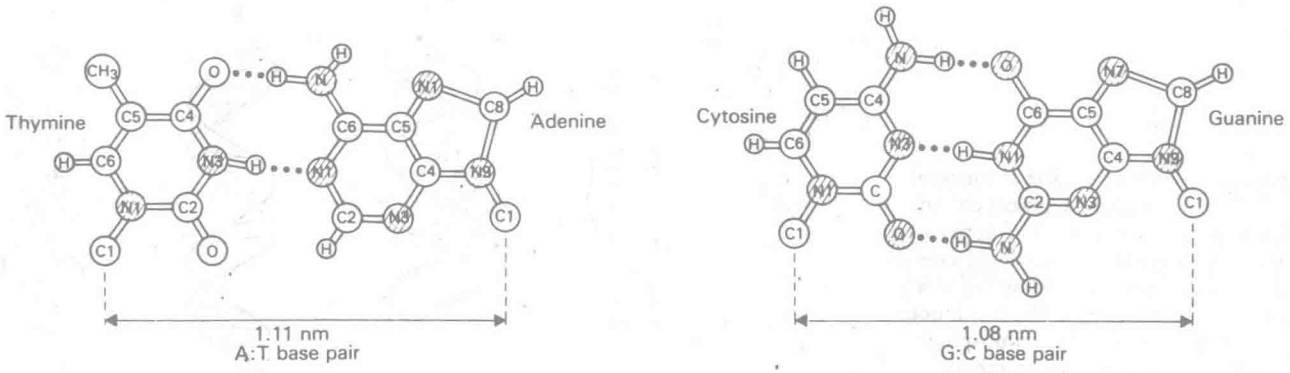



**Fig. 1.7.** Structures of the complementary base pairs in DNA—adenine:thymine and guanine:cytosine. In each case the bases are attached to the deoxyribose of the sugar-phosphate backbone by a covalent bond to the 1′ carbon of the sugar, labeled C1′ in the diagrams.

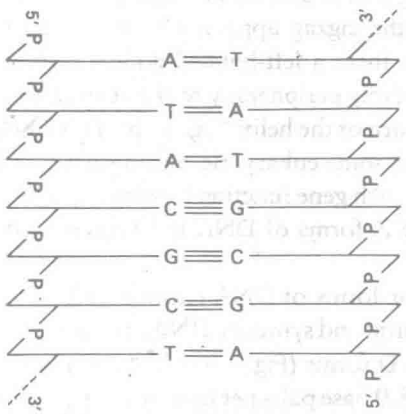



**Fig. 1.8.** A diagrammatic representation of double-stranded DNA showing the opposite polarity of the two strands.

hydrogen bonding between the bases, which are flat structures stacked like coins and arranged at right angles to the long axis of the polynucleotide chain. The sugar-phosphate backbones are on the outside of the helix.

From the model it is possible to show that there are 10 base pairs per each complete turn of the polynucleotide chain. Since the distance between adjacent base pairs is 0.34 nm, it follows that the DNA helix has one turn each 3.4 nm of length. (1 nm = 1 nanometer = $10^{-9}$ m).

The most important feature of the model is the specific pairing of the bases. Only two complementary base pairs, A-T and G-C, can form stable bonds in the double-helical structure (Fig. 1.7). As a result of this, the nucleotide sequence in one strand dictates the nucleotide sequence of the other. In other words, the two strands are complementary. The A-T base pair has two hydrogen bonds, and the G-C pair has three hydrogen bonds. This specific complementary base pairing is of central importance for many functions of the nucleic acids (for example, DNA replication transcription, and translation).

Another property of the model is that the two chains of the double helix are oriented with opposite (antiparallel) polarity in terms of the 3′, 5′-phosphate-deoxyribose linkages (Fig. 1.8).

## Different DNA forms

The Watson and Crick model is a right-handed double helix of DNA. This is designated B DNA and is the most common of the configurations that DNA can assume. A stylized diagram of B DNA is shown in Fig. 1.9a. There are 10.0 base pairs per turn of the helix, and the base pairs are more-or-less perpendicular to the helix axis.

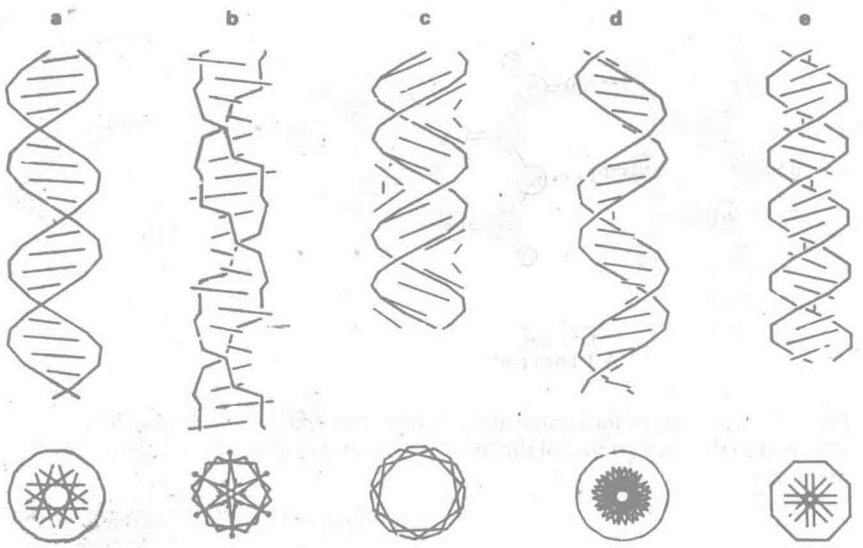



Fig. 1.9. Models for various forms of DNA. In each case a segment containing 20 base pairs is shown. The upper views are perpendicular to the helical axes and the lower views are "end on", looking along the helical axes. The continuous helical lines are the sugar-phosphate backbones, and the line segments indicate the positions of the base pairs: (a) B DNA; (b) Z DNA; (c) A DNA; (d) C DNA; (e) D DNA. (Reproduced with permission, from the *Annual Review of Biochemistry*, Vol. 51. 1982 by Annual Reviews Inc.)

Other forms of DNA have been described that differ from B DNA with respect to the direction of helical coiling and/or spacing and tilting of the base pairs. In Z DNA (named because of the zigzag appearance of the sugar-phosphate backbones), the backbones form a left-handed double helix in which the base pairs are arranged relatively peripherally to the central helix axis, and there are 12.0 base pairs per turn of the helix (Fig. 1.9b). This DNA form has been shown to exist in at least some eukaryotic chromosomes and may occur in some DNA locations in which gene function is being regulated. In vitro, the interconversion of B and Z forms of DNA has been demonstrated.

While B and Z DNA are the major forms of DNA encountered, other DNA forms have been described in natural and synthetic DNA. These are the right-handed double-helical A, C, and D forms (Fig. 1.9c, d, and e, respectively), which have 11.0, 7.9–9.6, and 8.0 base pairs per turn, respectively. In the A form the base pairs are arranged toward the outside of the helix when viewed along the helical axes (Fig. 1.9b, bottom), and they are inclined at about a 19° angle to the helix axis (the base pairs are perpendicular to the helix axis in B DNA). In the C form, the base pairs are arranged towards the middle of the helix (as is the case in B DNA) and, like A DNA, are at an inclined angle to the helix axis, although the angle of inclination is less extreme than in A DNA. In D DNA the helix as viewed along the helix axis is not circular but hexagonal in cross section. In this form of DNA, the base pairs are arranged toward the middle of the helix and are at an inclined angle to the helix axis. Note that the angles of inclination of the base pairs are approximately opposite in the C and D forms of DNA (compare Figs. 1.9d and e).

## Evidence that DNA is the genetic material

Many lines of evidence strongly indicate that DNA is the genetic material in may organisms. Three examples are given here:

1. The nucleic acids show maximal absorbance of ultraviolet light at a wavelength of 260 nm (Fig. 1.10), and this correlates exactly with the wavelength at which maximal mutagenesis of cells can be achieved by ultraviolet irradiation. This observation provided further evidence that nucleic acids and not proteins (which show maximal absorbance of light at 280 nm) are the genetic material.

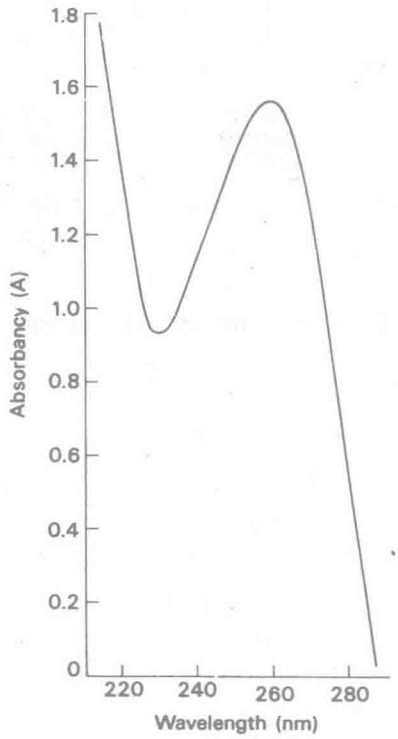



Fig. 1.10. Ultraviolet light absorbancy spectrum of DNA showing the maximal absorbancy.

2. In 1928 F. Griffith discovered that the S strain of the bacterium *Diplococcus pneumoniae* (pneumococcus), when injected into mice, causes death by septicemia (blood poisoning). Another strain, the R strain, had no effect on the same mice. The distinction between the two strains lies in the fact that the S strain bacteria have a polysaccharide capsule around them, resulting in a smooth colony appearance when they grow on solid medium in a culture dish (and hence the S designation). The R strain produces rough-appearing colonies, owing to the lack of the capsule. Griffith showed that the

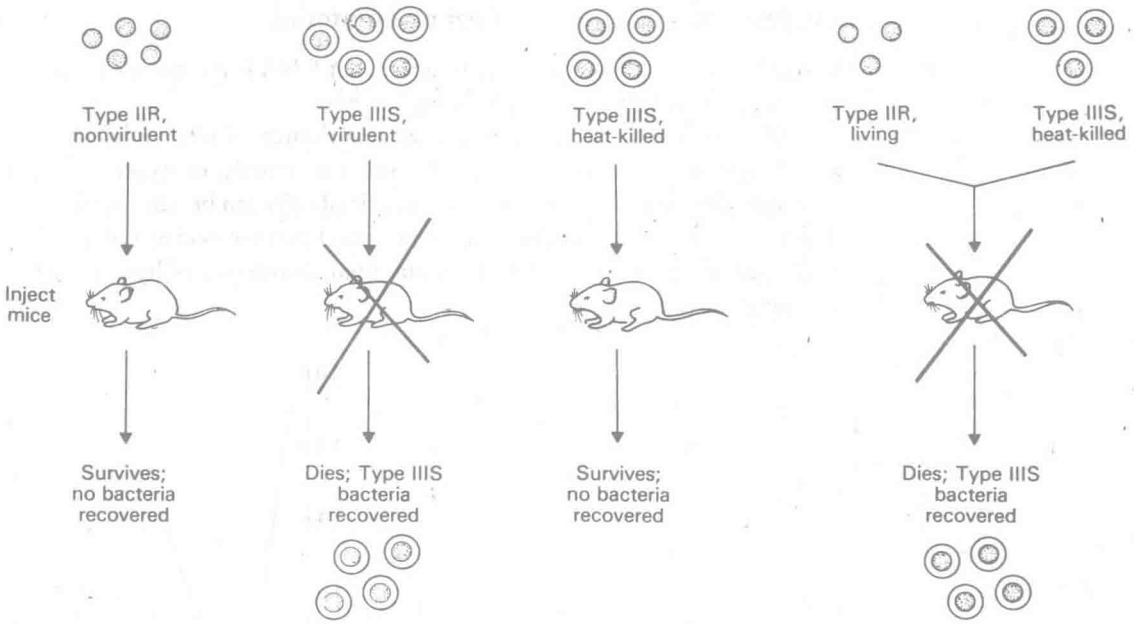



**Fig. 1.11.** Transformation experiment of Griffith. (After M. W. Strickberger, 1976. *Genetics*. Macmillan, New York.)

S bacteria could mutate spontaneously to give rise to the R type. Moreover, when mice were injected with a combination of live R bacteria and heat-killed S bacteria, the mice died from septicemia and live S bacteria could be isolated from their blood (Fig. 1.11). Thus something from the dead bacteria converted the R bacteria into S-type cells; this process is called transformation.

The transformation phenomenon received further scrutiny from O. T. Avery, C. M. Macleod, and M. McCarty in 1944. In some classic experiments they set out to determine the chemical nature of the substance (the so-called *transforming principle*) that induced the specific transformation of the pneumococcal types. They showed that a DNA fraction isolated from the S strain was capable of transforming unencapsulated R-type bacteria into fully encapsulated S-type cells. None of the other cell fractions, such as RNA, protein, lipid, carbohydrates, etc., was able to effect the transformation. Further, the transforming activity of the DNA fraction could be abolished by treatment with **deoxyribonuclease** (DNase), a DNA-degrading enzyme, but not by **ribonuclease** (RNase), an RNA-degrading enzyme. These results strongly indicated that DNA was the genetic material. However, Avery's work was criticized because the nucleic acids isolated were not completely pure and contained proteins.