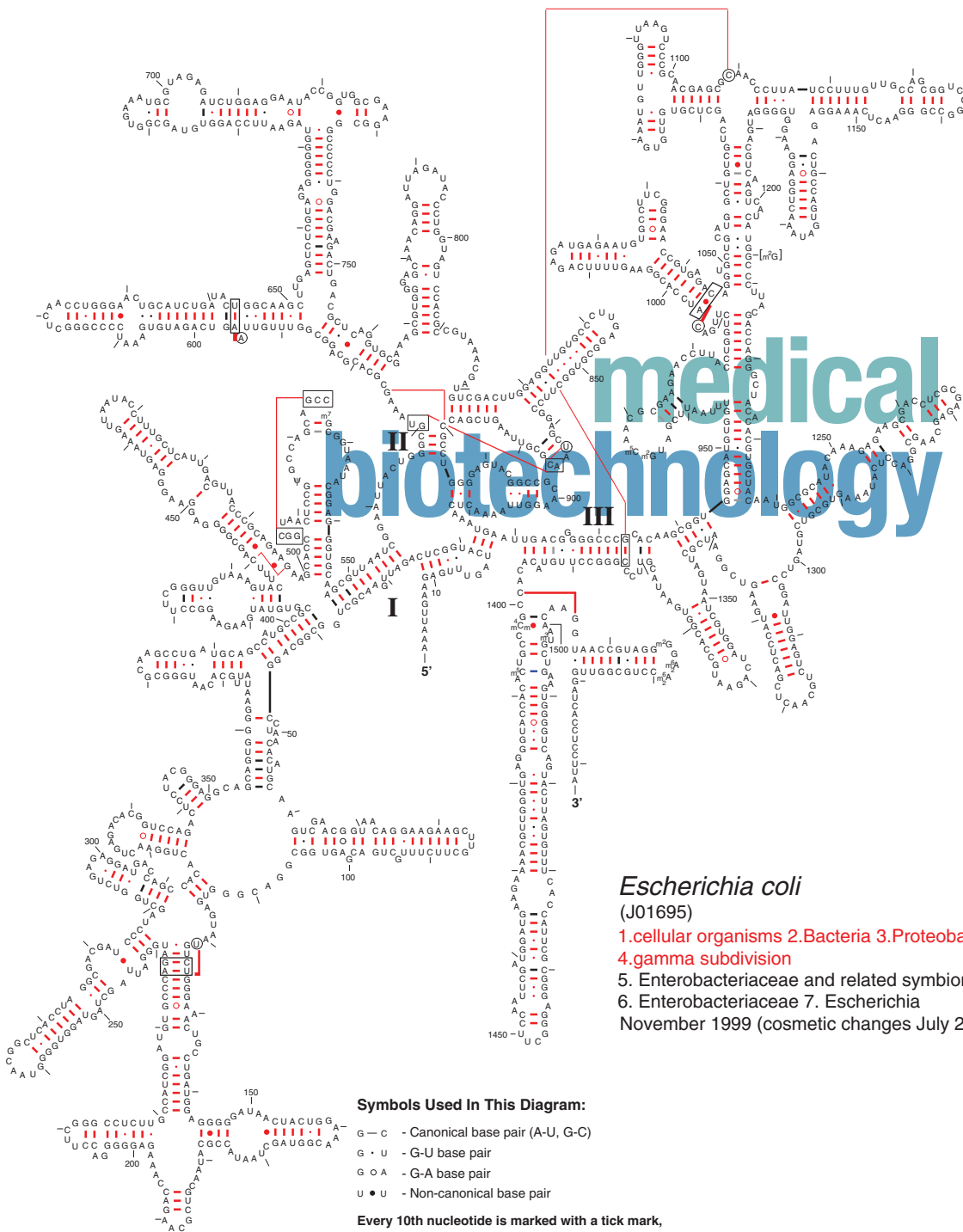


medical biotechnology



Bernard R. Glick • Terry L. Delovitch • Cheryl L. Patten



Escherichia coli
 (J01695)
 1. cellular organisms 2. Bacteria 3. Proteobacteria
 4. gamma subdivision
 5. Enterobacteriaceae and related symbionts
 6. Enterobacteriaceae 7. Escherichia
 November 1999 (cosmetic changes July 2001)

Symbols Used In This Diagram:
 — c — Canonical base pair (A-U, G-C)
 - - u - G-U base pair
 o o a - G-A base pair
 u • u - Non-canonical base pair

Every 10th nucleotide is marked with a tick mark,
 and every 50th nucleotide is numbered.
 Tertiary interactions with strong comparative data are connected by
 solid lines.

Citation and related information available at <http://www.rna.icmb.utexas.edu>

Figure 3.2 The *Escherichia coli* SSU rRNA secondary structure. (Courtesy of Robin Gutell. Adapted from Cannone JJ, Subramanian S, Schnare MN, Collett JR, D'Souza LM, Du Y, Feng B, Lin N, Madabusi LV, Müller KM, Pande N, Shang Z, Yu N, Gutell RR, *BMC Bioinformatics* 3:2, 2002. doi:10.1186/1471-2105-3-2 doi:10.1128/9781555818517.ch1.f1.11B)

This page intentionally left blank

medical biotechnology

Bernard R. Glick

Department of Biology
University of Waterloo
Waterloo, Ontario, Canada

Terry L. Delovitch

Department of Microbiology and Immunology
Western University
London, Ontario, Canada

Cheryl L. Patten

Biology Department
University of New Brunswick
Fredericton, New Brunswick, Canada



**ASM
PRESS**

WASHINGTON, DC

Copyright © 2014 American Society for Microbiology. All rights reserved. No part of this publication may be reproduced or transmitted in whole or in part or reused in any form or by any means, electronic or mechanical, including photocopying and recording, or by any information storage and retrieval system, without permission in writing from the publisher.

Disclaimer: To the best of the publisher's knowledge, this publication provides information concerning the subject matter covered that is accurate as of the date of publication. The publisher is not providing legal, medical, or other professional services. Any reference herein to any specific commercial products, procedures, or services by trade name, trademark, manufacturer, or otherwise does not constitute or imply endorsement, recommendation, or favored status by the American Society for Microbiology (ASM). The views and opinions of the author(s) expressed in this publication do not necessarily state or reflect those of ASM, and they shall not be used to advertise or endorse any product.

Library of Congress Cataloging-in-Publication Data

Glick, Bernard R., author.

Medical biotechnology / Bernard R. Glick, Department of Biology, University of Waterloo, Waterloo, Ontario, Canada; Terry L. Delovitch, Department of Microbiology and Immunology, Western University, London, Ontario, Canada; Cheryl L. Patten, Biology Department, University of New Brunswick, Fredericton, New Brunswick, Canada.

pages cm

Includes bibliographical references and index.

ISBN 978-1-55581-705-3 (hardcover)—ISBN 978-1-55581-889-0 (e-book) 1. Biotechnology.

2. Medical technology. I. Delovitch, T. L., author. II. Patten, Cheryl L., author. III. Title.

TP248.2.G57 2014

660.6—dc23

2013027259

10 9 8 7 6 5 4 3 2 1

Printed in the United States of America

Address editorial correspondence to ASM Press, 1752 N St. NW,
Washington, DC 20036-2904, USA

E-mail: books@asmusa.org

Send orders to ASM Press, P.O. Box 605, Herndon, VA 20172, USA

Phone: (800) 546-2416 or (703) 661-1593; Fax: (703) 661-1501

Online: <http://www.asmscience.org>

doi:10.1128/9781555818890

Cover and interior design: Susan Brown Schmidler

Illustrations: Patrick Lane, ScEYence Studios

Image credits for cover and section openers

Cell image on cover: A dendritic cell infected with human immunodeficiency virus (HIV), showing projections called filopodia (stained red) with HIV particles (white) at their ends. Reproduced from the cover of *PLoS Pathogens*, June 2012. Courtesy of Anupriya Aggarwal and Stuart Turville (Kirby Institute, University of New South Wales).

DNA image on cover: majcot/Shutterstock

DNA image on chapter and section opener pages: Mopic/Shutterstock

Section opener images: mouse, Sergey Galushko/Shutterstock; lab equipment,

Vasiliy Koval/Shutterstock; vaccine, Nixx Photography/Shutterstock

Some figures and tables in this book are reprinted or modified from Glick et al., *Molecular Biotechnology: Principles and Applications of Recombinant DNA*, 4th ed. (ASM Press, Washington, DC, 2010).

*To our spouses, Marcia Glick, Regina Delovitch, and Patrick Patten,
for their omnipresent love and tolerance, support, wisdom, and humor*

This page intentionally left blank

Contents

Preface xvii

About the Authors xviii

SECTION I

The Biology behind the Technology 1

1 Fundamental Technologies 3

Molecular Cloning 3

Preparation of DNA for Cloning 3

Insertion of Target DNA into a Plasmid Vector 7

Transformation and Selection of Cloned
DNA in a Bacterial Host 12

Cloning Eukaryotic Genes 15

Recombinational Cloning 19

Genomic Libraries 21

Amplification of DNA Using PCR 24

DNA Sequencing Technologies 28

Dideoxynucleotide Procedure 31

Pyrosequencing 33

Sequencing Using Reversible Chain Terminators 35

Sequencing by Ligation 36

Sequencing Whole Genomes 38

Shotgun Cloning Strategy 39

High-Throughput Next-Generation Sequencing Strategies 41



Genomics 42

Transcriptomics 46

Proteomics 51

Metabolomics 63

SUMMARY 67

REVIEW QUESTIONS 68

REFERENCES 69

2 Fundamental Concepts in Immunology 71

The Immune Response 71

Overview of Infection and Immunity 71

Functions of the Immune System 73

Innate Immunity 75

Adaptive Immunity 81

Cells of the Immune System: Cell-Mediated Immunity 88

Bone Marrow Precursor Cells 88

Myeloid Cells Mediate Innate Immunity 90

Lymphocytes 92

Antigen-Presenting Cells 96

Effector Cells 97

Tissues of the Immune System 97

Peripheral Lymphoid Organs 97

Lymphocyte Recirculation and Migration into Tissues 101

Antigen Recognition by T Cells 103

Humoral Immunity 114

Structure of Immunoglobulins 116

Functions of Immunoglobulins 118

Types of Antibodies: Applications 119

Polyclonal Antibodies 119

Antisera 120

Monoclonal Antibodies 120

Recombinant Antibodies 120

Immunological Techniques 122

Enzyme-Linked Immunosorbent Assay 123

Enzyme-Linked Immunospot Assay 125

Flow Cytometry 126

Mass Cytometry 128
Two-Photon Intravital Cell Imaging 129
SUMMARY 131
REVIEW QUESTIONS 132
REFERENCES 133

3 The Genetic Basis of Disease 135

Chromosomal Disorders and Gene Mapping 135

Chromosomes and Chromosome Abnormalities 135
Human Genome Mapping 146
Genome-Wide Association Studies 159

Single-Gene Disorders 164

Mode of Inheritance 164
Thalassemia 166
Sickle-Cell Anemia 167
Hemophilia 168
Cystic Fibrosis 169
Tay–Sachs Disease 170
Fragile X Syndrome 171
Huntington Disease 173

Polygenic Disorders and Gene Clustering 174

GWAS Strategies To Map Genes for Polygenic Disease 176
Breast Cancer 178
Alzheimer Disease 182
Type 1 Diabetes 185
Cardiovascular Disease 191

Mitochondrial Disorders 193

Disorders 193
Genetics 194
Mitochondrial Homeostasis and Parkinson Disease 195
Prevalence 199
Diagnosis and Prognosis 200
Treatment 201
SUMMARY 201
REVIEW QUESTIONS 202
REFERENCES 203

4 Immune Pathogenesis 207

Models of Immune System Lesions 207

Immunological Tolerance 207

Failure of Immune Tolerance and Development of Autoimmune Disease 218

Immune Surveillance against Tumors 224

Immune Evasion by Tumors 229

Inflammation and Immune Hypersensitivity Disorders 229

Types of Inflammation and Associated Immune Hypersensitivity Reactions 230

Immediate Hypersensitivity 230

Therapy for Immediate Hypersensitivity 234

Antibody- and Antigen–Antibody Complex-Induced Disease 236

T-Cell-Mediated Diseases 239

Immunodeficiency Disorders and Defects in Development of the Immune System 241

Primary Immunodeficiencies 242

Secondary Immunodeficiencies 245

AIDS 245

SUMMARY 252

REVIEW QUESTIONS 253

REFERENCES 254

5 Microbial Pathogenesis 257

Introduction 257

Bacterial Infections 258

Attachment to Host Cells 258

Invasion and Dissemination 261

Evasion of Host Defenses and Proliferation 266

Damage to Host Tissues 272

Identification of Bacterial Virulence Factors 279

Evolution of Bacterial Pathogens 285

Treatment of Bacterial Infections 290

Viral Infections 293

Attachment and Entry 295

Viral Gene Expression and Replication 300

Virus Assembly and Release	311
General Patterns of Viral Infections of Humans	315
Targets for Treatment of Viral Infections	320
SUMMARY	323
REVIEW QUESTIONS	324
REFERENCES	325

SECTION II

Production of Therapeutic Agents 327

6 Modulation of Gene Expression 329

Manipulating Gene Expression in Prokaryotes 330

Promoters	331
Translational Regulation	334
Codon Usage	335
Protein Stability	336
Fusion Proteins	338
Metabolic Load	341
Chromosomal Integration	343
Increasing Secretion	346
Overcoming Oxygen Limitation	349
Reducing Acetate	350
Protein Folding	352

Heterologous Protein Production in Eukaryotic Cells 354

Eukaryotic Expression Systems	354
<i>Saccharomyces cerevisiae</i> Expression Systems	356
Other Yeast Expression Systems	360
Baculovirus–Insect Cell Expression Systems	362
Mammalian Cell Expression Systems	368

Directed Mutagenesis 374

Oligonucleotide-Directed Mutagenesis with M13 DNA	374
Oligonucleotide-Directed Mutagenesis with Plasmid DNA	377
PCR-Amplified Oligonucleotide-Directed Mutagenesis	377
Error-Prone PCR	379
Random Mutagenesis	381



DNA Shuffling 383
Examples of Modified Proteins 384
SUMMARY 386
REVIEW QUESTIONS 388
REFERENCES 389

7 Genetic Engineering of Plants 393

Plant Transformation with the Ti Plasmid of *A. tumefaciens* 396

Physical Transfer of Genes to Plants 401

Chloroplast Engineering 403

Transient Gene Expression 405

Molecular Pharming 408

Therapeutic Agents 408

Antibodies 411

Edible Vaccines 412

SUMMARY 418

REVIEW QUESTIONS 418

REFERENCES 419

SECTION III

Diagnosing and Treating Human Disease 421

8 Molecular Diagnostics 423

Immunological Approaches To Detect Protein Biomarkers of Disease 424

Enzyme-Linked Immunosorbent Assays 424

Measuring Disease-Associated Proteins by Sandwich ELISA 428

Diagnosing Autoimmune Diseases by an Indirect ELISA 429

Immunoassays for Infectious Disease 430

Protein Arrays To Detect Polygenic Diseases 432

Immunoassays for Protein Conformation-Specific Disorders 435

DNA-Based Approaches to Disease Diagnosis 437

Hybridization Probes 437



Allele-Specific Hybridization	439
Oligonucleotide Ligation Assay	439
Padlock Probes	441
Allele-Specific PCR	442
TaqMan PCR	445
Real-Time PCR To Detect Infectious Disease	447
Detection of Multiple Disease-Associated Mutations Using Microarrays	450
Detection of Epigenetic Markers	451
Detection of SNPs by Mass Spectrometry	454
Detecting RNA Signatures of Disease	457
Detection of Disease-Associated Changes in Gene Expression Using Microarrays	457
Detection of RNA Signatures of Antibiotic Resistance in Human Pathogens	457
Detection of miRNA Signatures of Cancers	460
SUMMARY	461
REVIEW QUESTIONS	462
REFERENCES	463

9 Protein Therapeutics 465

Pharmaceuticals 466

Interferon	466
Human Growth Hormone	471
Tumor Necrosis Factor	473
Targeting Mitochondria	474
Extending Protein Half-Life	476
Engineered Bacteriophages	477

Recombinant Antibodies 479

Preventing Rejection of Transplanted Organs	479
Hybrid Human–Mouse Monoclonal Antibodies	480
Human Monoclonal Antibodies	482
Antibody Fragments	484
Combinatorial Libraries of Antibody Fragments	488
Anticancer Antibodies	491
Antianthrax Antibodies	493
Antiobesity Antibodies	495
Enhanced Antibody Half-Life	496

Enzymes 497

- DNase I 498
- Alginate Lyase 499
- Phenylalanine Ammonia Lyase 502
- α_1 -Antitrypsin 504
- Glycosidases 505

Lactic Acid Bacteria 507

- Interleukin-10 507
- Leptin 510
- An HIV Inhibitor 511
- Insulin 512

SUMMARY 514

REVIEW QUESTIONS 515

REFERENCES 515

10 Nucleic Acid Therapeutic Agents and Human Gene Therapy 519

Treating Genetic and Nongenetic Disorders 519

Targeting Specific mRNAs and DNAs 522

- Antisense RNA 522
- Aptamers 525
- Ribozymes 528
- DNAzymes 530
- Interfering RNA 530
- Zinc Finger Nucleases 534

Viral Delivery Systems 535

- Gammaretrovirus 535
- Lentivirus 542
- Adeno-Associated Virus 545
- Adenovirus 549
- Herpes Simplex Virus 1 555

Nonviral Nucleic Acid Delivery Systems 561

- Direct Injection 561
- Lipids 561
- Bacteria 563
- Dendrimers 565
- Antibodies 566
- Aptamers 566

Transposons 567
Prodrug Activation Therapy 569
Gene Therapy 571
Severe Combined Immunodeficiency 571
Cancer 572
Eye Disorders 576
Muscle Disorders 578
Neurological Disorders 580
SUMMARY 587
REVIEW QUESTIONS 587
REFERENCES 588

11 Vaccines 593

Vaccination: Overview 593

Advantages 593
Limitations 595
Current Vaccine Design 598

Subunit Vaccines 601

Herpes Simplex Virus 602
Cholera 604
Severe Acute Respiratory Syndrome 605
Staphylococcus aureus 607
Human Papillomavirus 609

Peptide Vaccines 612

Malaria 612
Cancer 616
Autoimmune Disease 618
Allergy 620

Dendritic Cell Vaccines 622

Human Immunodeficiency Virus 622
Cancer 623

DNA Vaccines 625

Delivery and Immune Mechanisms of Action 625
Advantages and Disadvantages 626
Improved Efficacy and Immunogenicity 627

Attenuated Vaccines 634

Cholera 634
Salmonella Species 637

Leishmania Species 639

Poliovirus 639

Influenza Virus 640

Dengue Virus 643

Vector Vaccines 645

Vaccines Directed against Viruses 645

Vaccines Directed against Bacteria 651

Bacteria as Antigen Delivery Systems 655

Adjuvants 660

Systems Biology and Evaluation of Vaccines 663

SUMMARY 666

REVIEW QUESTIONS 667

REFERENCES 668

12 Societal Issues 671

Safety and Ethical Issues 672

Regulation of New Drugs 672

Regulation of Genetic and Genomic Testing 676

Patenting Biotechnology 678

Patenting 678

Patenting in Different Countries 680

Patenting DNA Sequences 681

Patenting Living Organisms 683

Patenting and Fundamental Research 684

Economic Issues 684

SUMMARY 687

REVIEW QUESTIONS 688

REFERENCES 689

Glossary 691

Index 715

Preface

From the very beginning of the biotechnology revolution in the early 1970s, many scientists understood that this new technology would radically change the way that we think about health care. They understood early on, well before any products were commercialized, that medical science was about to undergo a major paradigm shift in which all of our previous assumptions and approaches would change dramatically. Forty years later, biotechnology has delivered on much of its early promise. Hundreds of new therapeutic agents, diagnostic tests, and vaccines have been developed and are currently available in the marketplace. Moreover, it is clear that we are presently just at the tip of a very large iceberg, with many more products in the pipeline. It is likely that, in the next 10 to 15 years, biotechnology will deliver not only new products to diagnose, prevent, and treat human disease but also entirely new approaches to treating a wide range of hitherto difficult-to-treat or untreatable diseases.

We have written *Medical Biotechnology* with the premise that it might serve as a textbook for a wide range of courses intended for premedical and medical students, dental students, pharmacists, optometrists, nurses, nutritionists, genetic counselors, hospital administrators, and other individuals who are stakeholders in the understanding and advancement of biotechnology and its impact on the practice of modern medicine. The book is intended to be as jargon-free and as easy to read as possible. In some respects, our goal is to demystify the discipline of medical biotechnology. This is not a medical textbook per se. However, a discussion of some salient features of selected diseases is presented to illustrate the applications of many biotechniques and biochemical mechanisms. Thus, this book may be considered a biomedical road map that provides a fundamental understanding of many approaches being pursued by scientists to diagnose, prevent, and treat a wide range of ailments. Indeed, this presents a large challenge, and the future is difficult to predict. Nevertheless, we hope that this volume will provide a useful introduction to medical biotechnology for a wide range of individuals.

About the Authors



Bernard R. Glick is a professor of biology at the University of Waterloo in Waterloo, Ontario, Canada, where he received his PhD in biochemistry in 1974. His current research is focused on the biochemical and genetic mechanisms used by plant growth-promoting bacteria to facilitate plant growth. In addition to his nearly 300 research publications, Dr. Glick is a coauthor of the textbook *Molecular Biotechnology: Principles and Applications of Recombinant DNA*, published by ASM Press. According to Google Scholar, his work has been cited more than 15,000 times. In addition to having served two terms as chair of the Department of Biology at Waterloo, Dr. Glick has taught 10 courses in five countries on various aspects of biotechnology.



Terry L. Delovitch obtained his BSc in chemistry (1966) and PhD in chemistry/immunology (1971) at McGill University. He received postdoctoral training at the Massachusetts Institute of Technology and Stanford University and then joined the faculty of the University of Toronto. In 1994, he was appointed senior scientist and director of the Autoimmune Disease Group on Type 1 Diabetes at the Robarts Research Institute, Western University, London, Ontario, Canada. After a 45-year research career, he retired from Western in 2011. He has received several academic awards and published about 200 research papers, review articles, and book chapters. He is the former chief scientific advisor to the Juvenile Diabetes Research Foundation Canada and past president of the International Immunology of Diabetes Society, and he is a consultant or advisor for several biotechnology and pharmaceutical

companies, granting agencies, and journal editorial boards and a national allergy, asthma, and immunology research network. He and his wife, Regina, live in Toronto.



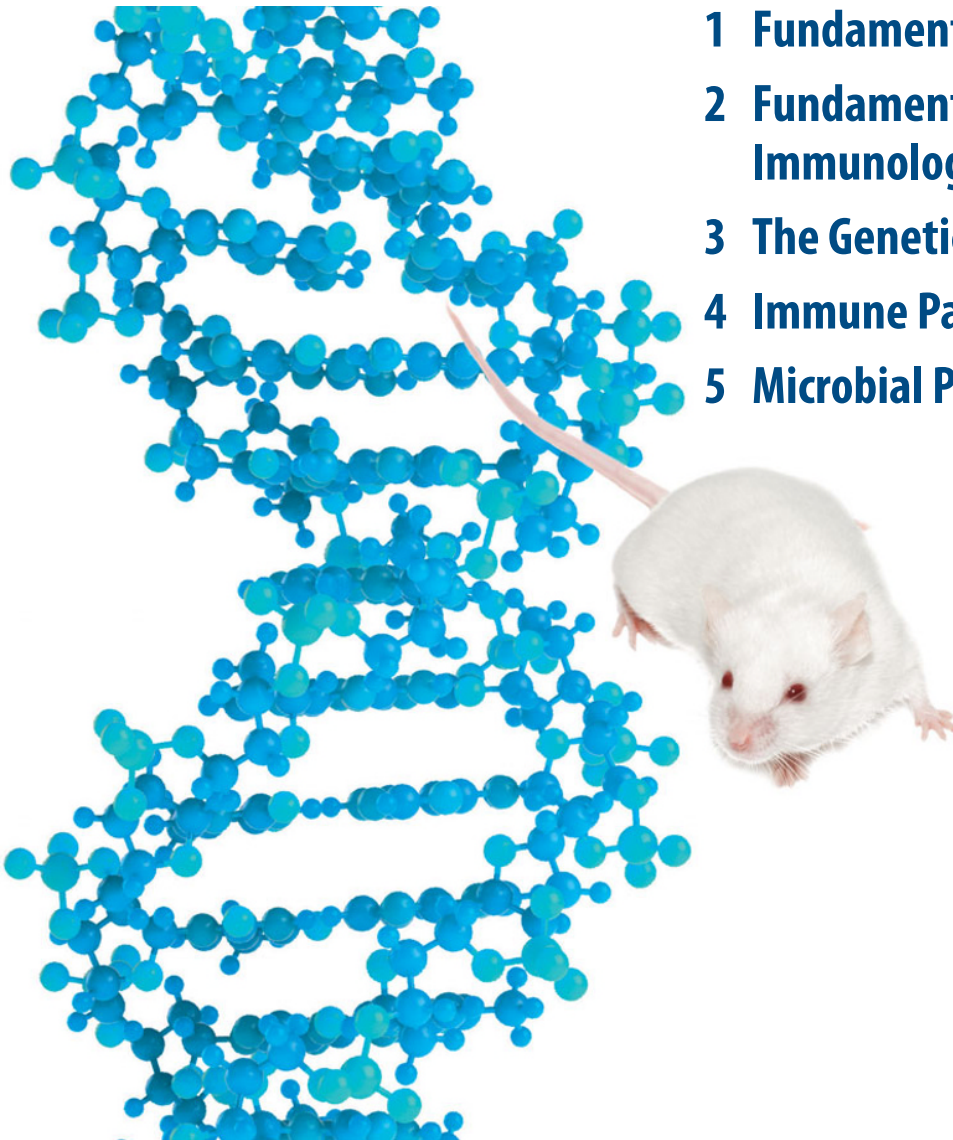
Cheryl L. Patten is an associate professor of microbiology and associate chair of the Department of Biology at the University of New Brunswick (UNB) in Fredericton, New Brunswick, Canada. Dr. Patten received her PhD from the University of Waterloo in 2001 and did postdoctoral work at McMaster University before joining the UNB faculty in 2004. Her research aims to understand how bacteria respond to the host environment at the biochemical and genetic levels. In particular, she is interested in secreted bacterial metabolites that may impact host health. As well as teaching introductory and advanced courses in microbiology, she enjoys introducing first-year science students to the wonders of biochemistry and molecular biology. She is a coauthor of another ASM Press textbook, *Molecular Biotechnology: Principles and Applications of Recombinant DNA*.

This page intentionally left blank

SECTION I

The Biology behind the Technology

- 1 Fundamental Technologies
- 2 Fundamental Concepts in Immunology
- 3 The Genetic Basis of Disease
- 4 Immune Pathogenesis
- 5 Microbial Pathogenesis



This page intentionally left blank

1

Molecular Cloning

- Preparation of DNA for Cloning
- Insertion of Target DNA into a Plasmid Vector
- Transformation and Selection of Cloned DNA in a Bacterial Host
- Cloning Eukaryotic Genes
- Recombinational Cloning

Genomic Libraries

Amplification of DNA Using PCR

DNA Sequencing Technologies

- Dideoxynucleotide Procedure
- Pyrosequencing
- Sequencing Using Reversible Chain Terminators
- Sequencing by Ligation

Sequencing Whole Genomes

- Shotgun Cloning Strategy
- High-Throughput Next-Generation Sequencing Strategies

Genomics

- Transcriptomics
- Proteomics
- Metabolomics

SUMMARY

REVIEW QUESTIONS

REFERENCES

Fundamental Technologies

Molecular Cloning

Molecular biotechnology uses a variety of techniques for isolating genes and transferring them from one organism to another. At the root of these technologies is the ability to join a sequence of deoxyribonucleic acid (DNA) of interest to a vector that can then be introduced into a suitable host. This process is known as **recombinant DNA technology** or molecular cloning. A vast number of variations on this basic process has been devised. Development of the core technologies depended on an understanding of fundamental processes in molecular biology, bacterial genetics, and nucleic acid enzymology (Box 1.1). The beginning of the application of these technologies for the purpose of manipulating DNA has been credited to Stanley Cohen of Stanford University, Stanford, California, who was developing methods to transfer plasmids, small circular DNA molecules, into bacterial cells, and Herbert Boyer at the University of California at San Francisco, who was working with enzymes that cut DNA at specific nucleotide sequences. They hypothesized that Boyer's enzymes could be used to insert a specific segment of DNA into a plasmid and then the recombinant plasmid could be introduced into a host bacterium using Cohen's method. Within a few years, the method was used successfully to produce human insulin, which is used in the treatment of diabetes, in *Escherichia coli*. In the 25 years since the first commercial production of recombinant human insulin, more than 200 new drugs produced by recombinant DNA technology have been used to treat over 300 million people for diseases such as cancer, multiple sclerosis, cystic fibrosis, and cardiovascular disease and to provide protection against infectious diseases. Moreover, over 400 new drugs are in the process of being tested in human trials to treat a variety of serious human diseases.

Preparation of DNA for Cloning

In theory, DNA from any organism can be cloned. The target DNA may be obtained directly from genomic DNA, derived from messenger ribonucleic acid (mRNA), subcloned from previously cloned DNA, or synthesized in vitro. The target DNA may contain the complete coding sequence

box 1.1 The Development of Recombinant DNA Technology

Most important technologies are developed in small steps, and recombinant DNA technology is no exception. The ability to join DNA molecules from different sources to produce life-changing therapeutic agents like human insulin depends on the contributions of many researchers. The early 1970s were ripe for the development of recombinant DNA technology following the milestone discoveries of the structure of DNA by Watson and Crick (Watson and Crick, 1953) and the cracking of the genetic code by Nirenberg, Matthaei, and Jones (Nirenberg and Matthaei, 1961; Nirenberg et al., 1962). Building on this, rapid progress was made in understanding the structure of genes and the manner in which they are expressed. Isolating and

preparing genes for cloning would not be possible without type II restriction endonucleases that cut DNA in a sequence-specific and highly reproducible manner (Kelly and Smith, 1970). Advancing the discovery by Herbert Boyer and colleagues (Hedgpeth et al., 1972), who showed that the RI restriction endonuclease from *E. coli* (now known as EcoRI) made a staggered cut at a specific nucleotide sequence in each strand of double-stranded DNA, Mertz and Davis (Mertz and Davis, 1972) reported that the complementary ends produced by EcoRI could be rejoined by DNA ligase in vitro. Of course, joining of the restriction endonuclease-digested molecules required the discovery of DNA ligase (Gellert et al., 1968). In the meantime, Cohen and Chang (Cohen and Chang,

1973) had been experimenting with constructing plasmids by shearing large plasmids into smaller random pieces and introducing the mixture of pieces into the bacterium *E. coli*. One of the pieces was propagated. However, the randomness of plasmid fragmentation reduced the usefulness of the process. During a now-legendary lunchtime conversation at a scientific meeting in 1973, Cohen and Boyer reasoned that EcoRI could be used to splice a specific segment of DNA into a plasmid, and then the recombinant plasmid could be introduced into and maintained in *E. coli* (Cohen et al., 1973). Recombinant DNA technology was born. The potential of the technology was immediately evident to Cohen and others: “It may be possible to introduce in *E. coli*, genes specifying metabolic or synthetic functions such as photosynthesis, or antibiotic production indigenous to other biological classes.” The first commercial product produced using this technology was human insulin.

DNA
deoxyribonucleic acid

mRNA
messenger ribonucleic acid

for a protein, a part of the protein coding sequence, a random fragment of genomic DNA, or a segment of DNA that contains regulatory elements that control expression of a gene. Prior to cloning, both the source DNA that contains the target sequence and the cloning vector must be cut into discrete fragments, predictably and reproducibly, so that they can be joined (ligated) together to form a stable molecule. Bacterial enzymes known as type II **restriction endonucleases**, or (more commonly) restriction enzymes, are used for this purpose. These enzymes recognize and cut DNA molecules at specific base pair sequences and are produced naturally by bacteria to cleave foreign DNA, such as that of infecting bacterial viruses (**bacteriophage**). A bacterium that produces a specific restriction endonuclease also has a corresponding system to modify the sequence recognized by the restriction endonuclease in its own DNA to protect it from being degraded.

A large number of restriction endonucleases from different bacteria is available to facilitate cloning. The sequence and length of the recognition site vary among the different enzymes and can be four or more nucleotide pairs. One example is the restriction endonuclease HindIII from the bacterium *Haemophilus influenzae*. HindIII is a homodimeric protein (made up of two identical polypeptides) that specifically recognizes and binds to the DNA sequence $\begin{matrix} \text{AAGCTT} \\ \text{TTCGAA} \end{matrix}$ (Fig. 1.1A). Note that the recognition sequence is a **palindrome**, that is, the sequence of nucleotides in each of the two strands of the binding site is identical when either is read in the same polarity, i.e., 5' to 3'. HindIII cuts within the DNA-binding site between

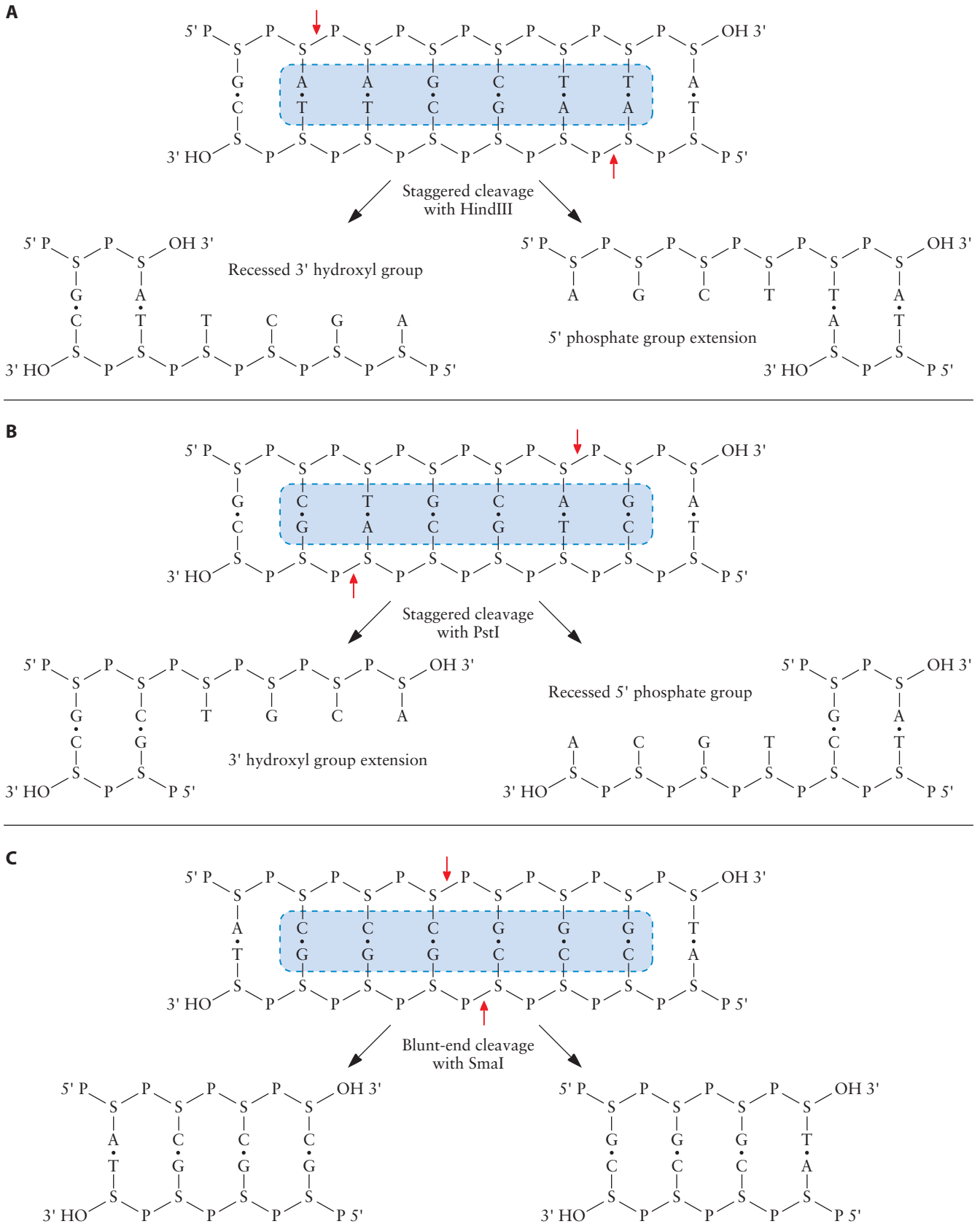


Figure 1.1 Type II restriction endonucleases bind to and cut within a specific DNA sequence. (A) HindIII makes a staggered cut in the DNA strands producing single-stranded, complementary ends (sticky ends) with a 5' phosphate group extension. (B) PstI also makes a staggered cut in both strands but produces sticky ends with a 3' hydroxyl group extension. (C) Cleavage of

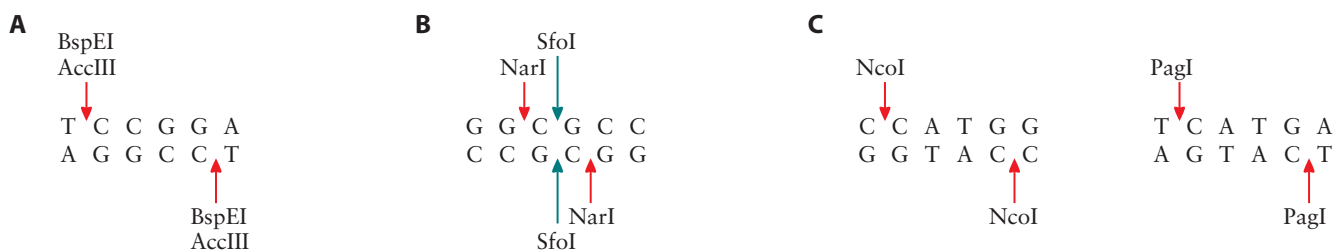
DNA with SmaI produces blunt ends. Arrows show the sites of cleavage in the DNA backbone. S, deoxyribose sugar; P, phosphate group; OH, hydroxyl group; A, adenine; C, cytosine; G, guanine; T, thymine. The restriction endonuclease recognition site is shaded. doi:10.1128/9781555818890.ch1.f1.1

the adjacent adenine nucleotides on each strand (Fig. 1.1A). Specifically, it cleaves the bond between the oxygen attached to the 3' carbon of the sugar of one nucleotide and the phosphate group attached to the 5' carbon of the sugar of the adjacent nucleotide. The symmetrical staggered cleavage of DNA by HindIII produces two single-stranded, complementary ends, each with extensions of four nucleotides, known as sticky ends. Each single-stranded extension terminates with a 5' phosphate group, and the 3' hydroxyl group of the opposite strand is recessed (Fig. 1.1A). Some other restriction enzymes, such as PstI, leave 3' hydroxyl extensions with recessed 5' phosphate ends (Fig. 1.1B), while others, such as SmaI, cut the backbone of both strands within a recognition site to produce blunt-ended DNA molecules (Fig. 1.1C).

Restriction enzymes isolated from different bacteria may recognize and cut DNA at the same site (Fig. 1.2A). These enzymes are known as **isoschizomers**. Some recognize and bind to the same sequence of DNA but cleave at different positions (**neoschizomers**), producing different single-stranded extensions (Fig. 1.2B). Other restriction endonucleases (**isocaudomers**) produce the same nucleotide extensions but have different recognition sites (Fig. 1.2C). In some cases, a restriction endonuclease will cleave a sequence only if one of the nucleotides in the recognition site is methylated. These characteristics of restriction endonucleases are considered when designing a cloning experiment.

Many other enzymes are used to prepare DNA for cloning. In addition to restriction endonucleases, nucleases that degrade single-stranded extensions, such as S1 nuclease and mung bean nuclease, are used to generate blunt ends for cloning (Fig. 1.3A). This is useful when the recognition sequences for restriction enzymes that produce complementary sticky ends are not available on both the vector and target DNA molecules. Blunt ends can also be produced by extending 3' recessed ends using a DNA polymerase such as **Klenow polymerase** derived from *E. coli* DNA polymerase I (Fig. 1.3B). **Phosphatases** such as calf intestinal alkaline phosphatase cleave the 5' phosphate groups from restriction enzyme-digested

Figure 1.2 Restriction endonucleases have been isolated from many different bacteria. (A) Isoschizomers such as BspEI from a *Bacillus* species and AccIII from *Acinetobacter calcoaceticus* bind the same DNA sequence and cut at the same sites. (B) Neoschizomers such as NarI from *Nocardia argentinensis* and SfoI from *Serratia fonticola* bind the same DNA sequence but cut at different sites. (C) Isocaudomers such as NcoI from *Nocardia corallina* and PagI from *Pseudomonas alcaligenes* bind different DNA sequences but produce the same sticky ends. Bases in the restriction enzyme recognition sequence are shown. Arrows show the sites of cleavage in the DNA backbone. doi:10.1128/9781555818890.ch1.f1.2



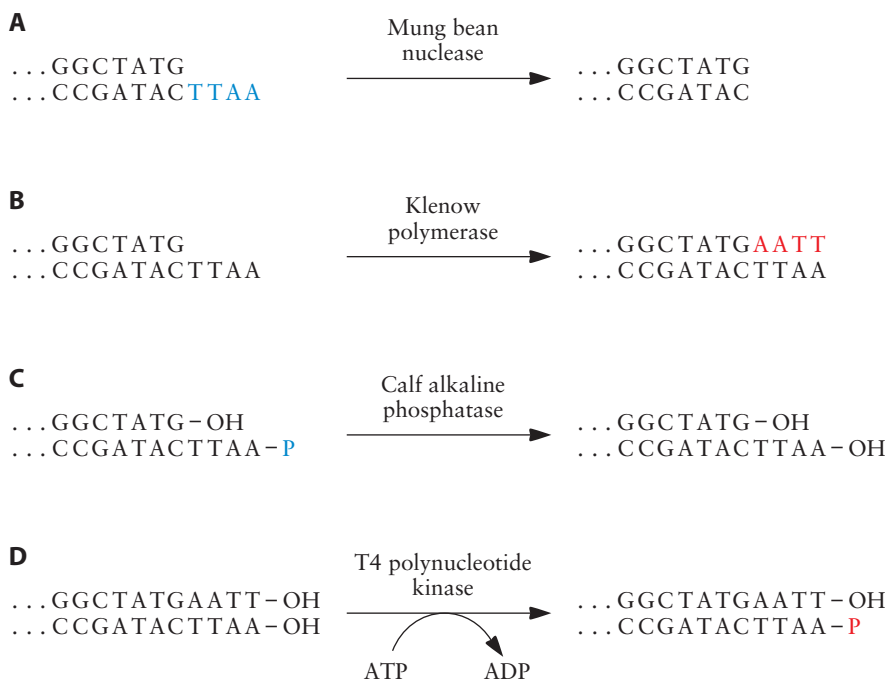


Figure 1.3 Some other enzymes used to prepare DNA for cloning. (A) Mung bean nuclease degrades single-stranded 5' and 3' extensions to generate blunt ends. (B) Klenow polymerase extends 3' recessed ends to generate blunt ends. (C) Calf alkaline phosphatase removes the 5' phosphate group from the ends of linear DNA molecules. (D) T4 polynucleotide kinase catalyzes the addition of a 5' phosphate group to the ends of linear DNA fragments. Dotted lines indicate that only one end of the linear DNA molecule is shown. doi:10.1128/9781555818890.ch1.f1.3

DNA (Fig. 1.3C). A 5' phosphate group is required for formation of a phosphodiester bond between nucleotides, and therefore, its removal prevents recircularization (self-ligation) of vector DNA. On the other hand, **kinases** add phosphate groups to the ends of DNA molecules. Among other activities, T4 polynucleotide kinase catalyzes the transfer of the terminal (γ) phosphate from a nucleoside triphosphate to the 5' hydroxyl group of a polynucleotide (Fig. 1.3D). This enzyme is employed to prepare chemically synthesized DNA for cloning, as such DNAs are often missing a 5' phosphate group required for ligation to vector DNA.

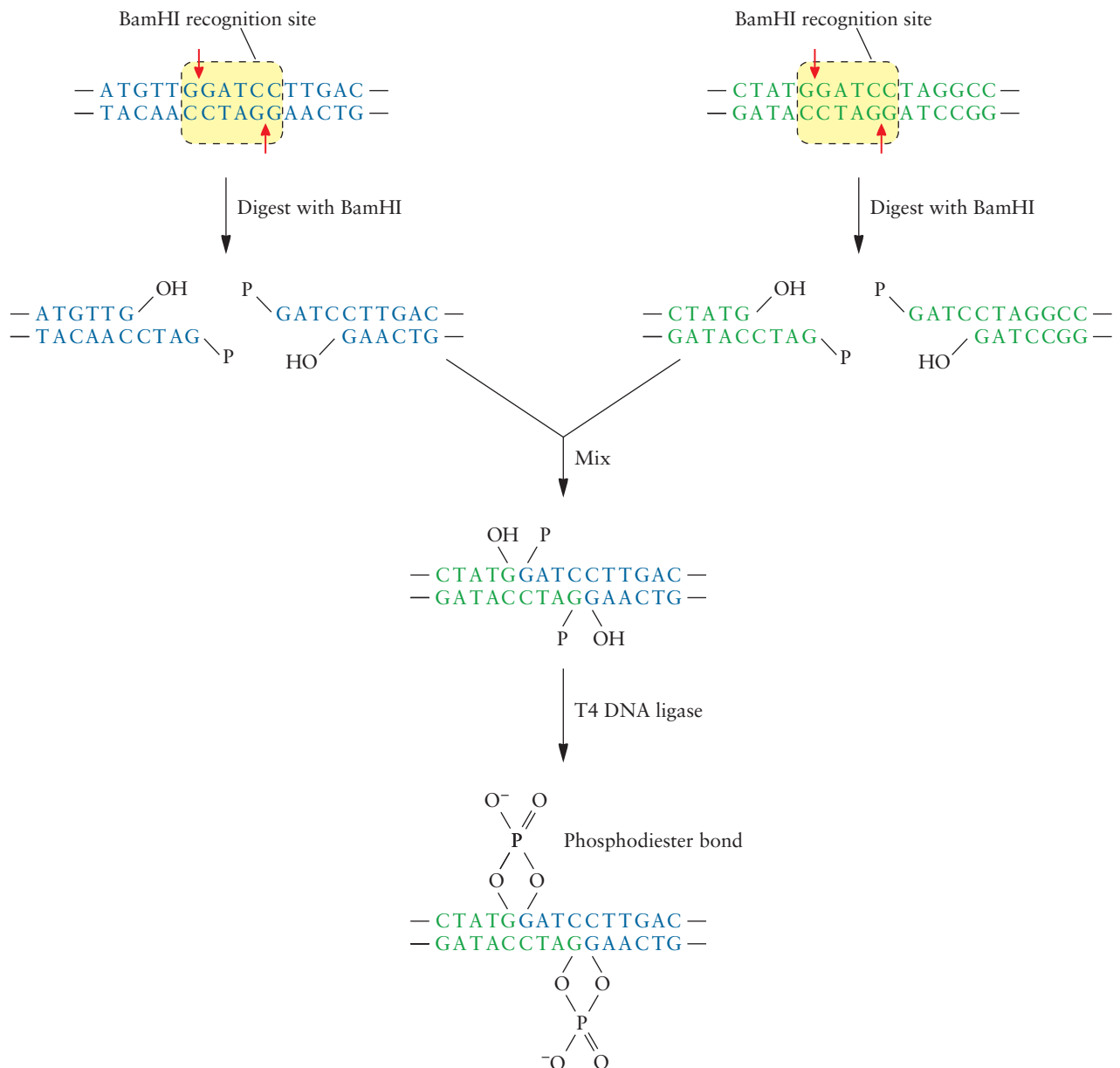
Insertion of Target DNA into a Plasmid Vector

When two different DNA molecules are digested with the same restriction endonuclease, that produces the same sticky ends in both molecules, and then mixed together, new DNA combinations can be formed as a result of base-pairing between the extended regions (Fig. 1.4). The enzyme **DNA ligase**, usually from the *E. coli* bacteriophage T4, is used to reform the phosphodiester bond between the 3' hydroxyl group and the 5' phosphate group at the ends of DNA strands that are already held together by the hydrogen bonds between the complementary bases of the extensions

(Fig. 1.4). DNA ligase also joins blunt ends, although this is generally much less efficient.

Ligation of restriction enzyme-digested DNA provides a means to stably insert target DNA into a vector for introduction and propagation in a suitable host cell. Many different vectors have been developed to act as carriers for target DNA. Most are derived from natural gene carriers, such as genomes of viruses that infect eukaryotic or prokaryotic cells and

Figure 1.4 Ligation of two different DNA fragments after digestion of both with restriction endonuclease BamHI. Complementary nucleotides in the single-stranded extensions form hydrogen bonds. T4 DNA ligase catalyzes the formation of phosphodiester bonds by joining 5' phosphate and 3' hydroxyl groups at nicks in the backbone of the double-stranded DNA. doi:10.1128/9781555818890.ch1.f1.4



integrate into the host genome or plasmids that are found in bacterial or fungal cells. Others are synthetically constructed **artificial chromosomes** designed for delivery of large pieces of target DNA (>100 kilobase pairs [kb]) into bacterial, yeast, or mammalian host cells. Many different vectors that carry sequences required for specific functions, for example, for expression of foreign DNA in a host cell, are described throughout this book. Here, vectors based on bacterial **plasmids** are used to illustrate the basic features of a cloning vector.

Plasmids are small, usually circular, double-stranded DNA molecules that are found naturally in many bacteria. They can range in size from less than 1 kb to more than 500 kb and are maintained as extrachromosomal entities that replicate independently of the bacterial chromosome. While they are not usually essential for bacterial cell survival under laboratory conditions, plasmids often carry genes that are advantageous under particular conditions. For example, they may carry genes that encode resistance to antibiotics or heavy metals, genes for the degradation of unusual organic compounds, or genes required for toxin production. Each plasmid has a sequence that functions as an origin of DNA replication that is required for it to replicate in a host cell. Some plasmids carry information for their own transfer from one cell to another.

The number of copies of a plasmid that are present in a host cell is controlled by factors that regulate plasmid replication and are characteristic of that plasmid. High-copy-number plasmids are present in 10 to more than 100 copies per cell. Other, low-copy-number plasmids are maintained in 1 to 4 copies per cell. When two or more different plasmids cannot coexist in the same host cell, they are said to belong to the same **plasmid incompatibility group**. But plasmids from different incompatibility groups can be maintained together in the same cell. This coexistence is independent of the copy numbers of the individual plasmids. Some microorganisms have been found to contain as many as 8 to 10 different plasmids. In these instances, each plasmid can carry out different functions and have its own unique copy number, and each belongs to a different incompatibility group. Some plasmids can replicate in only one host species because they require very specific proteins for their replication as determined by their origin of replication. These are generally referred to as **narrow-host-range plasmids**. On the other hand, **broad-host-range plasmids** have less specific origins of replication and can replicate in a number of bacterial species.

As autonomous, self-replicating genetic elements, plasmids are useful vectors for carrying cloned DNA. However, naturally occurring plasmids often lack several important features that are required for a good cloning vector. These include a choice of unique (single) restriction endonuclease recognition sites into which the target DNA can be cloned and one or more selectable genetic markers for identifying recipient cells that carry the cloning vector–insert DNA construct. Most of the plasmids that are currently used as cloning vectors have been genetically modified to include these features.

An example of a commonly used plasmid cloning vector is pUC19, which is derived from a natural *E. coli* plasmid. The plasmid pUC19 is