



# Contents

Preface xxvii

## CHAPTER 1

### **Introduction to Molecular Biology 1**

#### **1.1 Intellectual Foundation 2**

Two studies performed in the 1860s provided the intellectual underpinning for molecular biology. 2

#### **1.2 One Gene–One Polypeptide Hypothesis 4**

Each gene is responsible for the synthesis of a single polypeptide. 4

#### **1.3 Introduction to Nucleic Acids 5**

DNA contains the sugar deoxyribose and RNA contains the sugar ribose. 5

A nucleoside is formed by attaching a purine or pyrimidine base to a sugar. 6

A nucleotide is formed by attaching a phosphate group to the sugar group in a nucleoside. 7

DNA is a linear chain of deoxyribonucleotides. 8

#### **1.4 DNA: Hereditary Material 11**

Transformation experiments led to the discovery that DNA is the hereditary material. 11

Chemical experiments also supported the hypothesis that DNA is the hereditary material. 14

#### **1.5 Watson-Crick Model 15**

Rosalind Franklin and Maurice Wilkins obtained x-ray diffraction patterns of extended DNA fibers. 15

James Watson and Francis Crick proposed that DNA is a double-stranded helix. 15

#### **1.6 Central Dogma 19**

The central dogma provides the theoretical framework for molecular biology. 19

## **1.7 Introduction to Recombinant DNA Technology 20**

Recombinant DNA technology allows us to study complex biological systems. 20

A great deal of molecular biology information is available on the Internet. 22

Questions and Problems 22

Suggested Reading 23

## **CHAPTER 2**

## **Protein Structure and Function 24**

### **2.1 $\alpha$ -Amino Acids 25**

$\alpha$ -Amino acids have an amino group and a carboxyl group attached to a central carbon atom. 25

Amino acids are represented by three-letter and one-letter abbreviations. 28

### **2.2 Peptide Bond 29**

$\alpha$ -Amino acids are linked by peptide bonds. 29

Peptide chains have directionality. 30

### **2.3 Protein Purification 31**

Protein mixtures can be fractionated by chromatography. 31

Proteins and other charged biological polymers migrate in an electric field. 35

### **2.4 Primary Structure of Proteins 36**

The amino acid sequence or primary structure of a purified protein can be determined. 36

Polypeptide sequences can be obtained from nucleic acid sequences. 39

The BLAST program compares a new polypeptide sequence with all sequences stored in a data bank. 41

### **2.5 Introduction to Protein Structure 41**

Proteins with just one polypeptide chain have primary, secondary, and tertiary structures, whereas those with two or more chains also have quaternary structures. 41

### **2.6 Weak Noncovalent Bonds 43**

The polypeptide folding pattern is determined by weak noncovalent interactions. 43

### **2.7 Secondary Structures 46**

The  $\alpha$ -helix is a compact structure that is stabilized by hydrogen bonds. 46

The  $\beta$ -conformation is also stabilized by hydrogen bonds. 49

Loops and turns connect different peptide segments, allowing polypeptide chains to fold back on themselves. 49

Certain combinations of secondary structures, called supersecondary structures or folding motifs, appear in many different proteins. 50

We cannot yet predict secondary structures with absolute certainty. 50

### **2.8 Tertiary Structure 50**

X-ray crystallography studies reveal three-dimensional structures of many different proteins. 50

Intrinsically disordered proteins lack an ordered structure under physiological conditions. 53

The primary structure of a polypeptide determines its tertiary structure. 53  
Molecular chaperones help proteins to fold inside the cell. 56

## **2.9 Myoglobin, Hemoglobin, and the Quaternary Structure Concept 56**

Differences in myoglobin and hemoglobin function are explained by differences in myoglobin and hemoglobin structure. 56  
Hemoglobin subunits bind oxygen in a cooperative fashion. 60

## **2.10 Immunoglobulin G (IgG) and the Domain Concept 63**

Large polypeptides fold into globular units called domains. 63

## **2.11 Enzymes 65**

Enzymes are proteins that catalyze chemical reactions. 65  
Molecular biologists often use colorimetric and radioactive methods to measure enzyme activity. 65  
Enzymes lower the energy of activation but do not affect the equilibrium position. 66  
All enzyme-catalyzed reactions proceed through an enzyme-substrate complex. 68  
Molecular details for ES complexes have been worked out for many enzymes. 71  
Regulatory enzymes control committed steps in biochemical pathways. 72  
Regulatory enzymes exhibit sigmoidal kinetics and are stimulated or inhibited by allosteric effectors. 73

**BOX 2.1. IN THE LAB: Strategies for Protein Purification 33**

**BOX 2.2. CLINICAL APPLICATION: Sickle Cell Anemia 58**

**BOX 2.3. CLINICAL APPLICATION: Clinical Significance of 2,3-Bisphosphoglycerate 61**

**BOX 2.4. IN THE LAB: Radioactive Tracers 67**

Questions and Problems 74

Suggested Reading 77

## **CHAPTER 3 Nucleic Acid Structure 80**

### **3.1 DNA Size and Fragility 81**

Most eukaryotic DNA is located in a defined cell nucleus, whereas bacterial and archaeal DNA is present in the cytoplasm. 81  
DNA molecules vary in size and base composition. 81  
DNA molecules are fragile. 82

### **3.2 Recognition Patterns in the Major and Minor Grooves 82**

Enzymes can recognize specific patterns at the edges of the major and minor grooves. 82

### **3.3 DNA Bending 83**

Some base sequences cause DNA to bend. 83

<b>3.4</b>	<b>DNA Denaturation and Renaturation</b>	<b>84</b>
	DNA can be denatured.	84
	Hydrogen bonds stabilize double-stranded DNA.	86
	Base stacking also stabilizes double-stranded DNA.	87
	Ionic strength influences DNA structure.	87
	The DNA molecule is in a dynamic state.	87
	Alkali denatures DNA without breaking phosphodiester bonds.	88
	Complementary single strands can anneal to form double-stranded DNA.	88
<b>3.5</b>	<b>Helicases</b>	<b>89</b>
	Helicases are motor proteins that use the energy of nucleoside triphosphates to unwind DNA.	89
<b>3.6</b>	<b>Single-Stranded DNA Binding Proteins</b>	<b>90</b>
	Single-stranded DNA binding proteins stabilize single-stranded DNA.	90
<b>3.7</b>	<b>Topoisomers</b>	<b>90</b>
	DNA can exist as a circular molecule.	90
	Bacterial DNA usually exists as a covalently closed, circular double-stranded DNA molecule.	91
	Plasmid DNA molecules are used to study the properties of covalently closed, circular double-stranded DNA in vitro.	92
	Covalently closed, circular double-stranded DNA molecules often have superhelical structures.	92
<b>3.8</b>	<b>Topoisomerases</b>	<b>97</b>
	Topoisomerases catalyze the conversion of one topoisomer into another.	97
<b>3.9</b>	<b>Non-B DNA Conformations</b>	<b>98</b>
	A-DNA is a right-handed double helix with a deep major groove and very shallow minor groove.	98
	Z-DNA has a left-handed conformation.	100
	Several other kinds of non-B DNA structures appear to exist in nature.	102
<b>3.10</b>	<b>RNA Structure</b>	<b>104</b>
	RNA performs a wide variety of functions in the cell.	104
	RNA secondary structure is dominated by Watson-Crick base pairs.	104
	RNA tertiary structures are stabilized by interactions between two or more secondary structure elements.	105
<b>3.11</b>	<b>RNA World Hypothesis</b>	<b>105</b>
	The earliest forms of life on earth may have used RNA as both the genetic material and the biological catalysts needed to maintain life.	105
	<b>BOX 3.1. CLINICAL APPLICATION: Plasmids and Disease</b>	<b>93</b>
	<b>BOX 3.2. LOOKING DEEPER: Linking Number</b>	<b>95</b>
	Questions and Problems	107
	Suggested Reading	108

## CHAPTER 4 **Molecular Biology Technology 112**

### **4.1 Nucleic Acid Isolation 113**

The method used to isolate DNA depends on the DNA source. 113

Great care must be taken to protect RNA from degradation during its isolation. 114

### **4.2 Physical Techniques Used to Study**

#### **Macromolecules 115**

Different physical techniques are used to study macromolecules. 115

Electron microscopy allows us to see nucleic acids and nucleoprotein complexes. 115

Centrifugation can separate macromolecules and provide information about their size and shape. 117

Gel electrophoresis separates nucleic acids based on their rate of migration in an electric field. 119

SDS-PAGE can be used to determine a polypeptide's molecular mass. 120

### **4.3 Enzymatic Techniques Used to Manipulate DNA 122**

Nucleases are useful tools in DNA investigations. 122

Restriction endonucleases cleave within specific nucleotide sequences in DNA. 123

Restriction endonucleases can be used to construct a restriction map of a DNA molecule. 125

DNA fragments can be inserted into plasmid DNA vectors. 128

The Southern blot procedure is used to detect specific DNA fragments. 130

Northern and Western blotting are used to detect specific RNA and polypeptide molecules, respectively. 132

DNA polymerase I requires a template-primer. 132

DNA polymerase I has both 3'→5' and 5'→3' exonuclease activities. 133

DNA polymerase I can catalyze nick translation. 136

The polymerase chain reaction (PCR) is used to amplify DNA. 137

Site-directed mutagenesis can be used to introduce a specific base change within a gene. 138

The chain termination method for sequencing DNA uses dideoxynucleotides to interrupt DNA synthesis. 139

DNA sequences can be stitched together by using information obtained from a restriction map. 141

Shotgun sequencing is used to sequence long DNA molecules. 141

A new generation of DNA sequencing techniques is now being used for whole genome shotgun sequencing. 142

The human genome sequence provides considerable new information. 142

**BOX 4.1. IN THE LAB: Gel Electrophoresis and Topoisomer Separation 121**

**BOX 4.2. IN THE LAB: Capillary Electrophoresis 122**

**BOX 4.3. IN THE LAB: New Generation DNA Sequencers 143**

**Questions and Problems 146**

**Suggested Reading 148**

## CHAPTER 5

### Chromosomes 151

#### 5.1 Bacterial Chromatin 153

Bacterial DNA is located in the nucleoid. 153

#### 5.2 Introduction to Eukaryotic Chromatin 155

Eukaryotic chromatin is visible under the light microscope during certain stages of the cell cycle. 155

#### 5.3 Mitosis 155

The animal cell life cycle alternates between interphase and mitosis. 155

Mitosis allows cells to maintain the chromosome number after cell division. 156

#### 5.4 Meiosis 159

Meiosis reduces the chromosome number in half. 159

#### 5.5 Karyotype 162

Chromosome sites are specified according to nomenclature conventions. 162

A karyotype shows an individual cell's metaphase chromosomes arranged in pairs and sorted by size. 163

Fluorescent *in situ* hybridization (FISH) provides a great deal of information about chromosomes. 163

#### 5.6 The Nucleosome 166

Five major histone classes interact with DNA in eukaryotic chromatin. 166

The first level of chromatin organization is the nucleosome. 167

X-ray crystallography reveals the atomic structure of nucleosome core particles. 168

The precise nature of the interaction between H1 and the core particle is not known. 171

We do not know how nucleosomes are arranged in chromatin. 171

Condensins and topoisomerase II help to stabilize condensed chromosomes. 172

The scaffold model was proposed to explain higher order chromatin structure. 172

#### 5.7 The Centromere 173

The centromere is the site of microtubule attachment. 173

#### 5.8 The Telomere 175

The telomere, which is present at either end of a chromosome, is needed for stability. 175

**BOX 5.1. CLINICAL APPLICATION: Karyotype and Diagnosis 164**

**Questions and Problems 179**

**Suggested Reading 180**

## CHAPTER 6

### Genetic Analysis in Molecular Biology 185

#### 6.1 *Drosophila melanogaster* 186

Many fundamental genetic principles were discovered by studying the common fruit fly. 186

#### 6.2 *Escherichia coli* 189

*E. coli* is a gram-negative bacterium. 189



	Bacteria can be cultured in liquid or solid media.	191
	Plating can be used to detect auxotrophs.	194
	Specific notations, conventions, and terminology are used in bacterial genetics.	194
	Cells with altered genes are called mutants.	195
	Some mutants display the mutant phenotype under all conditions, whereas others display it only under certain conditions.	195
	Mutants can be classified on the basis of the changes in the DNA.	196
	Mutations have many uses in molecular biology.	197
	A genetic test known as complementation can be used to determine the number of genes responsible for a phenotype.	200
	<i>E. coli</i> cells can exchange genetic information by conjugation.	200
	The F plasmid can integrate into a bacterial chromosome and carry it into a recipient cell.	204
	Bacterial mating experiments can be used to produce an <i>E. coli</i> genetic map.	206
	F' plasmids contain part of the bacterial chromosome.	209
	Plasmid replication control functions are usually clustered in a region called the basic replicon.	210
<b>6.3</b>	<b>Budding Yeast (<i>Saccharomyces cerevisiae</i>)</b>	<b>211</b>
	Yeasts are unicellular eukaryotes.	211
	Specific notations, conventions, and terminology are used in yeast genetics.	212
	Yeast cells exist in haploid and diploid stages.	213
<b>6.4</b>	<b>New Tools for Studying Human Genetics</b>	<b>214</b>
	Different humans may exhibit DNA sequence variations within the same DNA region.	214
	Restriction fragment length polymorphisms facilitate genetic analysis in humans and other organisms.	216
	Somatic cell genetics can be used to map genes in higher organisms.	219
	<b>BOX 6.1. LOOKING DEEPER: Plasmid Functions Required for Mating</b>	<b>204</b>
	<b>BOX 6.2. IN THE LAB: DNA Fingerprint</b>	<b>218</b>
	Questions and Problems	219
	Suggested Reading	220

## CHAPTER 7 **Viruses in Molecular Biology** 223

### **7.1 Introduction to the Bacteriophages** 225

	Bacteriophages were of interest because they seemed to have the potential to serve as therapeutic agents to treat bacterial diseases.	225
	Investigators belonging to the “Phage Group” were the first to use viruses as model systems to study fundamental questions about gene structure and function.	226
	Bacteriophages come in different sizes and shapes.	227
	Bacteriophages have lytic, lysogenic, and chronic life cycles.	228
	Bacteriophages form plaques on a bacterial lawn.	228

## 7.2 Virulent Bacteriophages 230

The T-even phages are tailed viruses that inject their DNA into bacteria. 230

Bacterial and phage RNA polymerase help to reel phage T7 DNA into the bacterial cell. 234

*E. coli* phage  $\phi$ X174 contains a single-stranded circular DNA molecule. 238

Some phages have single-stranded RNA as their genetic material. 239

## 7.3 Temperate Phages 241

*E. coli* phage  $\lambda$  DNA can replicate through a lytic or lysogenic life cycle. 241

*E. coli* phage P1 can also replicate through a lytic or lysogenic life cycle. 245

## 7.4 Chronic Phages 248

A chronic phage programs the host cell for continued virion particle release without killing the cell. 248

## 7.5 Animal Viruses 249

Animal viruses can serve as excellent models for addressing fundamental questions in molecular biology. 249

Polyomaviruses contain circular double-stranded DNA. 250

Adenoviruses have linear blunt-ended, double-stranded DNA with an inverted repeat at each end. 253

Retroviruses use reverse transcriptase to make a DNA copy of their RNA genome. 255

BOX 7.1. LOOKING DEEPER: T-Even Phage Assembly 234

BOX 7.2. IN THE LAB: Generalized Transduction 246

BOX 7.3. IN THE LAB: Phage M13 Cloning Vector 249

BOX 7.4. IN THE LAB: Complementary DNA Synthesis 259

Questions and Problems 260

Suggested Reading 260

# CHAPTER 8 DNA Replication 264

## 8.1 General Features of DNA Replication 265

DNA replication is semiconservative. 265

Bacterial and eukaryotic DNA replication is bidirectional. 267

The DNA strand that grows in an overall 3'→5' direction is formed by joining short fragments. 267

DNA ligase connects adjacent Okazaki fragments. 271

RNA serves as a primer for Okazaki fragment synthesis. 272

## 8.2 Bacterial DNA Replication 273

The bacterial replication machinery has been isolated and examined *in vitro*. 273

Mutant studies provide important information about the enzymes involved in bacterial DNA replication. 275

The replicon model proposes that an initiator protein must bind to a DNA sequence called a replicator at the start of replication. 275

*E. coli* chromosomal replication begins at *oriC*. 276

Several enzymes act together at the replication fork. 280

DNA polymerase III is required for bacterial DNA replication. 280



DNA polymerase III holoenzyme has three distinct kinds of subassemblies. 281

The core polymerase has one subunit with 5'→3' polymerase activity and another with 3'→5' exonuclease activity. 283

The sliding clamp forms a ring around DNA, tethering the remainder of the polymerase holoenzyme to the DNA. 285

The clamp loader places the sliding clamp around DNA. 286

The replisome catalyzes coordinated leading and lagging strand DNA synthesis at the replication fork. 287

*E. coli* DNA replication terminates when the two growing forks meet in the terminus region, which is located 180 degrees around the circular chromosome from the origin. 291

The terminus utilization substance binds to *Ter* sites. 291

Topoisomerase IV and recombinase separate newly formed sister chromosomes. 291

### 8.3 Eukaryotic DNA Replication 293

The SV40 DNA replication system is a good model for *in vitro* eukaryotic DNA replication. 293

Eukaryotic replication machinery must replicate long linear duplexes with multiple origins of replication. 295

Eukaryotic chromosomes have many origins of replication. 296

Yeast origins of replication, called autonomously replicating sequences, determine the site of DNA chain initiation. 297

Pol δ and Pol ε are primarily responsible for copying the lagging- and leading-strand templates, respectively. 298

Studies of the *Tetrahymena* and yeast telomeres suggest that a terminal transferase-like enzyme is required for telomere formation. 300

Telomerase uses an RNA template to add nucleotide repeats to chromosome ends. 303

Telomerase plays an important role in solving the end-replication problem. 305

### 8.4 Replication Coupled Chromatin Synthesis 307

Chromatin disassembly and reassembly are tightly coupled to DNA replication. 307

**BOX 8.1. LOOKING DEEPER: Replication Regulation at the Initiation Stage** 279

**BOX 8.2. IN THE LAB: Measurement of Processivity** 282

**BOX 8.3. LOOKING DEEPER: Helicase and Primer Coordination** 288

**BOX 8.4. CLINICAL APPLICATION: Telomerase in Aging and Cancer** 304

Questions and Problems 310

Suggested Reading 311

## CHAPTER 9 DNA Damage and Repair 317

### 9.1 Radiation Damage 319

Ultraviolet light causes cyclobutane pyrimidine dimer formation and (6-4) photoproduct formation. 319

X-rays and gamma rays cause many different types of DNA damage. 321

<b>9.2</b>	<b>DNA Instability in Water</b>	<b>321</b>
	DNA is damaged by hydrolytic cleavage reactions. 321	
<b>9.3</b>	<b>Oxidative Damage</b>	<b>324</b>
	Reactive oxygen species damage DNA. 324	
<b>9.4</b>	<b>Alkylation Damage by Monoadduct Formation</b>	<b>325</b>
	Alkylating agents damage DNA by transferring alkyl groups to centers of negative charge. 325	
	Many environmental agents must be modified by cell metabolism before they can alkylate DNA. 325	
<b>9.5</b>	<b>Chemical Cross-Linking Agents</b>	<b>327</b>
	Chemical cross-linking agents block DNA strand separation. 327	
<b>9.6</b>	<b>Mutagen and Carcinogen Detection</b>	<b>331</b>
	Mutagens can be detected based on their ability to restore mutant gene activity. 331	
<b>9.7</b>	<b>Direct Reversal of Damage</b>	<b>333</b>
	Photolyase reverses damage caused by cyclobutane pyrimidine dimer formation. 333	
	O <sup>6</sup> -Alkylguanine, O <sup>4</sup> -alkylthymine, and phosphotriesters can be repaired by direct alkyl group removal by a suicide enzyme. 335	
<b>9.8</b>	<b>Base Excision Repair</b>	<b>337</b>
	The base excision repair pathway removes and replaces damaged or inappropriate bases. 337	
<b>9.9</b>	<b>Nucleotide Excision Repair</b>	<b>340</b>
	Nucleotide excision repair removes bulky adducts from DNA by excising an oligonucleotide bearing the lesion and replacing it with new DNA. 340	
<b>9.10</b>	<b>Mismatch Repair</b>	<b>345</b>
	The DNA mismatch repair system removes mismatches and short insertions or deletions that are present in DNA. 345	
<b>9.11</b>	<b>SOS Response and Translesion DNA Synthesis</b>	<b>347</b>
	Error-prone DNA polymerases catalyze translesion DNA synthesis. 347	
	RecA and LexA regulate the <i>E. coli</i> SOS response. 348	
	The SOS signal induces the synthesis of DNA polymerases II, IV, and V. 351	
	Human cells have at least 14 different template-dependent DNA polymerases. 352	
	<b>BOX 9.1. CLINICAL APPLICATION: The Interstrand Cross-Linking Agents Cisplatin and Psoralen</b> 328	
	<b>BOX 9.2. CLINICAL APPLICATION: Xeroderma Pigmentosum</b> 343	
	Questions and Problems 353	
	Suggested Reading 354	

## CHAPTER 10 **Double-Strand Break Repair and Homologous Recombination 359**

### **10.1 Replication Restart After Bacterial Replication Fork Collapse 360**

Cells use homologous recombination to restart replication after a replication fork collapses. 360

The RecBCD protein processes the broken ends of double-strand break to form a single-strand 3'-OH tail. 361

RecA acts as a recombinase to catalyze strand invasion. 362

The RuvAB protein complex catalyzes branch migration in *E. coli*. 363

RuvC helps convert the DNA with the Holliday junction into a DNA fork. 364

PriA helps to load DnaB on the reestablished DNA fork to restart replication. 364

### **10.2 Mitotic Recombination 364**

Eukaryotic cells use homologous recombination to repair double-strand breaks during the late S stage or the G<sub>2</sub> stage of the cell cycle. 364

### **10.3 Gene Knockouts 368**

Homologous recombination can be used to create gene knockouts in mice. 368

### **10.4 Nonhomologous End-Joining 371**

Nonhomologous end-joining connects broken ends that have no homology. 371

### **10.5 Meiotic Recombination 372**

Homologous recombination during meiosis can result in gene conversion. 372

Homologous recombination during meiosis requires some enzymes that are unique to it. 373

Questions and Problems 376

Suggested Reading 377

## CHAPTER 11 **Transposable Elements 380**

### **11.1 Bacterial Transposable Elements 381**

Insertion sequences (IS elements) are simple mobile genetic elements. 381

A composite transposon consists of a pair of IS elements on either side of one or more genes. 382

Tn5 and Tn10 are composite transposons, which move via a cut-and-paste transposition mechanism. 383

The Tn3 transposon uses replicative transposition to move to a new site while maintaining a copy at the original site. 387

Transposons are useful tools for bacterial studies. 389

### **11.2 Eukaryotic Transposable Elements 391**

Some eukaryotic transposable elements move through a DNA-only mechanism and others require an RNA intermediate. 391

The eukaryotic “*hAT*” mobile elements move by a DNA only cut-and-paste transposition mechanism. 392

Antibody and T-cell receptor genes are formed by a mechanism that resembles the Hermes transposition mechanism. 392

A long dormant transposable element in fish called *Sleeping Beauty* has been restored to full activity by recombinant DNA technology. 396

LTR retrotransposons copy an RNA intermediate to make cDNA, which then is integrated into the genome by a cut-and-paste mechanism. 397

Non-LTR elements move using a target-primed reverse transcription mechanism. 398

*Alu* is the most common SINE in the human genome. 401

There is considerable debate concerning what benefits, if any, retrotransposons make to the host cells. 401

BOX 11.1. IN THE LAB: Reporter Genes 391

BOX 11.2. LOOKING DEEPER: Barbara McClintock and the Discovery of “Jumping Genes” 393

Questions and Problems 402

Suggested Reading 403

## CHAPTER 12 Bacterial Transcription and Its Regulation 406

### 12.1 Introduction to the Bacterial RNA Polymerase Catalyzed Reaction 408

RNA polymerase requires a DNA template and four nucleoside triphosphates to synthesize RNA. 408

Bacterial RNA polymerases are large multisubunit proteins. 410

### 12.2 Initiation Stage 412

Bacterial RNA polymerase holoenzyme consists of a core enzyme and sigma factor. 412

A transcription unit must have an initiation signal called a promoter for accurate and efficient transcription to take place. 413

Active  $\sigma^{70}$  promoters have -10 and -35 boxes. 416

Genetic and biochemical studies provide additional information about bacterial promoters. 418

DNA footprinting shows that  $\sigma^{70}$  RNA polymerase holoenzyme binds to promoter DNA to form a closed and an open complex. 418

Bacterial RNA polymerase crystal structures show how the enzyme is organized and provide insights into how it works. 420

Members of the  $\sigma^{70}$  family have four conserved domains. 424

The  $\sigma^{70}$  RNA polymerase holoenzyme goes through several rounds of abortive initiation before promoter escape. 425

RNA polymerase “scrunches” DNA during transcription initiation. 425

### 12.3 Transcription Elongation Complex 427

The transcription elongation complex is a highly processive molecular motor. 427

Pauses influence the overall transcription elongation rate. 432

RNA polymerase can detect and remove incorrectly incorporated nucleotides. 432

### 12.4 Transcription Termination 432

Bacterial transcription machinery releases RNA strands at intrinsic and Rho-dependent terminators. 432

## 12.5 Messenger RNA 434

Bacterial mRNA may be monocistronic or polycistronic. 434  
Bacterial mRNA usually has a short lifetime compared with other kinds of bacterial RNA. 435

Controlling the rate of mRNA synthesis can regulate the flow of genetic information. 436

Messenger RNA synthesis can be controlled by negative and positive regulation. 436

## 12.6 Lactose Operon 437

The *E. coli* genes *lacZ*, *lacY*, and *lacA* code for  $\beta$ -galactosidase, lactose permease, and  $\beta$ -galactoside transacetylase, respectively. 437

The *lac* structural genes are regulated. 438

Genetic studies provide information about the regulation of *lac* mRNA. 439

The operon model explains the regulation of the lactose system. 440

Allolactose is the true inducer of the lactose operon. 442

The Lac repressor binds to the *lac* operator *in vitro*. 443

The *lac* operon has three *lac* operators. 444

The Lac repressor is a dimer of dimers, where each dimer binds to one *lac* operator sequence. 445

## 12.7 Catabolite Repression 448

*E. coli* uses glucose in preference to lactose. 448

The cAMP receptor protein combines with 3',5'-cyclic adenylylate to form a positive regulator or activator. 449

The cAMP • CRP complex binds to an activator site (AS) upstream from the *lac* promoter and activates *lac* operon transcription. 450

Glucose causes catabolite repression through cAMP modulation and inducer exclusion mechanisms. 451

cAMP • CRP activates more than 100 operons. 453

## 12.8 Tryptophan Operon 453

The tryptophan (*trp*) operon is regulated at the levels of transcription initiation, elongation, and termination. 453

BOX 12.1. LOOKING DEEPER: Alternate Sigma Factors 414

BOX 12.2. CLINICAL APPLICATION: Rifamycin 422

BOX 12.3. LOOKING DEEPER: RNA Polymerase

Movement 430

Questions and Problems 459

Suggested Reading 460

## CHAPTER 13 Eukaryotic Transcription 466

### 13.1 Introduction to Eukaryotic Nuclear RNA

#### Polymerases 468

The eukaryotic cell nucleus has three different kinds of RNA polymerase. 468

RNA polymerases I, II, and III can be distinguished by their sensitivities to inhibitors. 470

Each nuclear RNA polymerase has some subunits that are unique to it and some it shares with one or both of the two other nuclear RNA polymerases. 471

### **13.2 RNA Polymerase II Structure 472**

Yeast RNA polymerase II crystal structures help explain how the enzyme works. 472

The crystal structure has been determined for the complete 12-subunit yeast RNA polymerase II bound to a transcription bubble and product RNA. 474

Nuclear RNA polymerases have limited synthetic capacities. 474

### **13.3 Core Promoter for Protein-Coding Genes 476**

The core promoter for protein-coding genes extends from 40 bp upstream of the transcription start site to 40 bp downstream from this site. 476

### **13.4 General Transcription Factors:**

#### **Basal Transcription 478**

RNA polymerase II requires the assistance of general transcription factors to transcribe naked DNA from specific transcription start sites. 478

TFIID or its TBP subunit must bind to a TATA core promoter before other general transcription factors can do so. 480

General transcription factors and RNA polymerase interact at the promoter to form a preinitiation complex. 482

### **13.5 Transcription Elongation 486**

The C-terminal domain of the largest RNA polymerase subunit must be phosphorylated for chain elongation to proceed. 486

A variety of transcription elongation factors helps to suppress transient pausing during elongation. 487

Elongation factor SII reactivates arrested RNA polymerase II. 487

### **13.6 Regulatory Promoters, Enhancers, and Silencers 487**

Linker-scanning mutagenesis reveals the regulatory promoter's presence just upstream from the core promoter. 487

Enhancers stimulate transcription and silencers block transcription. 490

The upstream activating sequence regulates genes in yeast. 492

### **13.7 Introduction to Transcription Activator Proteins for Protein-Coding Genes 492**

Transcription activator proteins help to recruit the transcription machinery. 492

A combinatorial process determines gene activity. 493

DNA affinity chromatography can be used to purify transcription activator proteins. 494

A transcription activator protein's ability to stimulate gene transcription can be determined by a transfection assay. 495

### **13.8 DNA-Binding Domains in Transcription Activator Proteins 496**

Transcription activator proteins are commonly grouped according to the structures of their DNA-binding domains. 496

### **13.9 Activation Domains in Transcription Activator Proteins 511**

Activation domain tend to be intrinsically disordered. 511

Gal4 has DNA-binding, dimerization, and activation domains. 511

Other proteins act along with Gal4 to regulate GAL genes. 512  
The activation domain must associate with a DNA-binding domain to stimulate transcription. 514

### **13.10 Mediator 517**

Squelching occurs when transcription activator proteins compete for a limiting transcription machinery component. 517  
Mediator is required for activated transcription. 517  
The yeast Mediator complex associates with activators at the UAS in active yeast genes. 521

### **13.11 Epigenetic Modifications 523**

Cells remodel or modify chromatin to make the DNA in chromatin accessible to the transcription machinery. 523  
Histone modification influences transcription of protein-coding genes. 525  
DNA methylation plays an important role in determining whether chromatin will be silenced or actively expressed in vertebrates. 527  
Epigenetics is the study of inherited changes in phenotype caused by changes in chromatin other than changes in DNA sequence. 527

### **13.12 RNA Polymerase I Catalyzed Transcription 528**

RNA polymerase I is required to synthesize 5.8S, 18S, and 28S rRNA. 528  
The rRNA transcription unit promoter consists of a core promoter and an upstream promoter element (UPE). 531  
5.8S, 18S, and 23S rRNA syntheses take place in the nucleolus. 532  
RNA polymerase I is a multisubunit enzyme with a structure similar to that of RNA polymerase II. 532  
The upstream binding factor and selectivity factor, working together, recruit RNA polymerase I to the rDNA promoter to form a preinitiation complex. 533  
RNA polymerase I forms a transcription elongation complex, leaving UBF and SL1/TIF-1B behind. 533  
RNA polymerase I transcription termination requires the assistance of a termination factor and a release protein. 534

### **13.13 RNA Polymerase III Catalyzed Transcription 535**

RNA polymerase III transcripts are short RNA molecules with a variety of biological functions. 535  
RNA polymerase III transcription units have three different types of promoters. 535  
RNA polymerase III does not appear to require additional factors for transcription elongation or termination. 537

**BOX 13.1. LOOKING DEEPER: The Archaeal RNA Polymerase 475**

**BOX 13.2. LOOKING DEEPER: TFIIB Assists in Open Promoter Formation 484**

**BOX 13.3. CLINICAL APPLICATION: The AP-1 Family of Transcription Activator Proteins 507**

**BOX 13.4. IN THE LAB: Yeast Two-Hybrid Assay 518**

**BOX 13.5. LOOKING DEEPER: Genomic Imprinting 529**

Questions and Problems 538

Suggested Reading 541



## CHAPTER 14 RNA Polymerase II: Cotranscriptional and Posttranscriptional Processes 551

### 14.1 Pre-mRNA 552

Eukaryotic cells synthesize large heterogeneous nuclear RNA molecules. 552

mRNA and hnRNA both have poly(A) tails at their 3'-ends. 552

### 14.2 Cap Formation 553

mRNA molecules have 7-methylguanosine caps at their 5'-ends. 553

5'-m<sup>7</sup>G caps are attached to nascent pre-mRNA chains when the chains are 20 to 30 nucleotides long. 556

All eukaryotes use the same basic pathway to form 5'-m<sup>7</sup>G caps. 556

CTD must be phosphorylated on Ser-5 to target a transcript for capping. 558

### 14.3 Split Genes 558

Viral studies revealed that some mRNA molecules are formed by splicing pre-mRNA. 558

Amino acid coding regions within eukaryotic genes may be interrupted by noncoding regions. 559

Exons tend to be conserved during evolution, whereas introns usually are not conserved. 564

A single pre-mRNA can be processed to produce two or more different mRNA molecules. 566

Combinations of the various splicing patterns within individual genes lead to the formation of multiple mRNAs. 568

Pre-mRNA requires specific sequences for precise splicing to occur. 569

Two splicing intermediates resemble lariats. 571

Splicing consists of two coordinated transesterification reactions. 574

### 14.4 Spliceosomes 575

Aberrant antibodies, which are produced by individuals with certain autoimmune diseases, bind to small nuclear ribonucleoprotein particles (snRNPs). 575

snRNPs assemble to form a spliceosome, the splicing machine that excises introns. 577

RNA and protein may both contribute to the spliceosome's catalytic site. 577

Cells use a variety of mechanisms to regulate splice site selection. 579

Splicing begins as a cotranscriptional process and continues as a posttranscriptional process. 583

mRNA splicing and export are coupled processes. 584

### 14.5 Cleavage/Polyadenylation and Transcription Termination 584

Poly(A) tail synthesis and transcription termination are coupled, cotranscriptional processes. 584

Transcription units often have two or more alternate polyadenylation sites. 587

Transcription termination takes place downstream from the poly(A) site. 588

RNA polymerase II transcription termination appears to involve allosteric changes and a 5'→3' exonuclease. 589

#### **14.6 RNA Editing 591**

RNA editing permits a cell to recode genetic information. 591

#### **14.7 The Gene Reconsidered 592**

The human proteome contains a much greater variety of proteins than would be predicted from the human genome. 592

Cotranscriptional and posttranscriptional processes force us to reconsider our concept of the gene. 592

**BOX 14.1. LOOKING DEEPER: Self-Splicing RNA 580**

Questions and Problems 593

Suggested Reading 594

### **CHAPTER 15 Small Silencing RNAs 600**

#### **15.1 RNA Interference (RNAi) Triggered by Exogenous Double-Stranded RNA 601**

The roundworm *Caenorhabditis elegans* is an attractive organism for molecular biology studies. 601

RNAi was discovered in *C. elegans*. 602

*In vitro* studies helped to elucidate the RNAi pathway. 605

Dicer cleaves long double-stranded RNA into fragments of discrete size. 607

RISC loading complex is required for siRISC formation. 609

RNAi blocks virus replication and prevents transposon activation. 609

#### **15.2 Transitive RNAi 610**

In some organisms, RNAi that starts at one site spreads throughout the entire organism. 610

SID-1, an integral membrane protein in *C. elegans*, assists in the systemic spreading of the silencing signal. 612

ERI-1, a 3'→5' exonuclease in *C. elegans*, appears to be a negative regulator of RNAi. 612

#### **15.3 RNAi as an Investigational Tool 612**

RNAi is a powerful tool for investigating functional genomics. 612

#### **15.4 MicroRNA Pathway 614**

The miRNA pathway blocks mRNA translation or causes mRNA degradation. 614

#### **15.5 Piwi Interacting RNAs (piRNAs) 617**

piRNAs help to maintain germ line stability in animals. 617

**BOX 15.1. LOOKING DEEPER: Dicer from *Giardia intestinalis* 607**

Questions and Problems 618

Suggested Reading 619

### **CHAPTER 16 Protein Synthesis 623**

#### **16.1 Introduction to the Ribosome 625**

Protein synthesis takes place on ribosomes. 625

Bacterial ribosomes are made of a large subunit with a 23S and 5S RNA and a small subunit with 16S RNA. 625

A eukaryotic ribosome also has a small and a large ribonucleoprotein subunit. 625  
Eukaryotic ribosomes exist free in the cytoplasm or attached to the endoplasmic reticulum. 627

## 16.2 Transfer RNA 628

An amino acid must be attached to a transfer RNA before it can be incorporated into a protein. 628  
All tRNA molecules have CCA<sub>OH</sub> at their 3'-ends. 631  
An amino acid attaches to tRNA through an ester bond between the amino acid's carboxyl group and the 2'- or 3'-hydroxyl group on adenosine. 633  
Yeast tRNA<sup>Ala</sup> was the first naturally occurring nucleic acid to be sequenced. 635  
tRNAs have cloverleaf secondary structures. 635  
tRNA molecules fold into L-shaped three-dimensional structures. 638

## 16.3 Aminoacyl-tRNA Synthetases 639

Some aminoacyl-tRNA synthetases have proofreading functions. 639  
Ile-tRNA synthetase has a proofreading function. 639  
Ile-tRNA synthetase can hydrolyze valyl-tRNA<sup>Ile</sup> and valyl-AMP. 640  
Each aminoacyl-tRNA synthetase can distinguish its cognate tRNAs from all other tRNAs. 641  
Selenocysteine and pyrrolysine are building blocks for polypeptides. 645

## 16.4 mRNA and the Genetic Code 647

mRNA programs ribosomes to synthesize proteins. 647  
Three adjacent bases in mRNA that specify an amino acid are called a codon. 648  
The discovery that poly(U) directs the synthesis of poly(Phe) was the first step in solving the genetic code. 650  
Protein synthesis begins at the amino terminus and ends at the carboxyl terminus. 652  
mRNA is read in a 5' to 3' direction. 653  
Trinucleotides promote the binding of specific aminoacyl-tRNA molecules to ribosomes. 654  
Synthetic messengers with strictly defined base sequences confirmed the genetic code. 654  
Three codons, UAA, UAG, and UGA, are polypeptide chain termination signals. 656  
The genetic code is nonoverlapping, commaless, almost universal, highly degenerate, and unambiguous. 657  
The coding specificity of an aminoacyl-tRNA is determined by the tRNA and not the amino acid. 658  
Some aminoacyl-tRNA molecules bind to more than one codon because there is some play or wobble in the third base of a codon. 659

## 16.5 Ribosome Structure 660

Bacterial 30S (small) subunits and 50S (large) subunits each have unique structures and functions. 660  
Bacterial ribosome structure has been determined at atomic resolution. 661

## **16.6 Four Stages of Protein Synthesis 662**

Protein synthesis can be divided into four stages. 662

## **16.7 Initiation Stage 664**

Each bacterial mRNA open reading frame has its own start codon. 664

Bacteria have an initiator methionine tRNA and an elongator methionine tRNA. 664

The 30S subunit is an obligatory intermediate in polypeptide chain initiation. 665

Initiation factors participate in the formation of 30S and 70S initiation complexes. 667

The Shine-Dalgarno sequence in mRNA interacts with the anti-Shine-Dalgarno sequence in the 16S rRNA. 668

Eukaryotic initiator tRNA is charged with a methionine that is *not* formylated. 671

Eukaryotic translation initiation proceeds through a scanning mechanism. 673

Translation initiation factor phosphorylation regulates protein synthesis in eukaryotes. 676

The translation initiation pathway in archaea appears to be a mixture of the eukaryotic and bacterial pathways. 676

## **16.8 Elongation Stage 677**

Polypeptide chain elongation requires elongation factors. 677

The elongation factors act through a repeating cycle. 677

An EF-Tu • GTP • aminoacyl-tRNA ternary complex carries the aminoacyl-tRNA to the ribosome. 677

Specific nucleotides in 16S rRNA are essential for sensing the codon-anticodon helix. 679

EF-Ts is a GDP-GTP exchange protein. 679

The ribosome is a ribozyme. 680

The hybrid-states translocation model offers a mechanism for moving tRNA molecules through the ribosome. 684

## **16.9 Termination Stage 685**

Bacteria have three protein release factors. 685

Mutant tRNA molecules can suppress mutations that create termination codons within a reading frame. 687

## **16.10 Recycling Stage 687**

The ribosome release factor is required for the bacterial ribosomal complex to disassemble. 687

## **16.11 Nascent Polypeptide Processing and Folding 688**

Ribosomes have associated enzymes that process nascent polypeptides and chaperones that help to fold the nascent polypeptides. 688

## **16.12 Signal Sequence 689**

The signal sequence plays an important role in directing newly synthesized proteins to specific cellular destinations. 689

**BOX 16.1. LOOKING DEEPER: Structural Basis for Posttransfer Editing 642**

**BOX 16.2. CLINICAL APPLICATION: A Neurodegenerative Disease Caused by an Editing Defect in Alanyl-tRNA Synthetase 643**

BOX 16.3. LOOKING DEEPER: Bacterial Pathway for Selenocysteyl-tRNA <sup>Sec</sup> Formation	647
BOX 16.4. LOOKING DEEPER: Structures of Eukaryotic and Archaeal Ribosomes	663
BOX 16.5. LOOKING DEEPER: Riboswitches	672
BOX 16.6. LOOKING DEEPER: Translocation Mechanism	686
Questions and Problems	691
Suggested Reading	693

## **GLOSSARY 703**

## **INDEX 733**