PREFACE

This textbook is designed for an introductory course in molecular biology. But what is molecular biology? The definition of this elusive term depends on who is doing the defining. In this book, I consider molecular biology to be the study of genes and their activities at the molecular level.

When I was a student in college and graduate school I found that I became most excited about science, and learned best, when the instructor emphasized the experimental strategy and the data that led to the conclusions, rather than just the conclusions themselves. Thus, when I began teaching an introductory molecular biology course in 1972, I adopted that teaching strategy and have used it ever since. I have found that my students react as positively as I did.

One problem with this approach, however, was that no textbook placed as great an emphasis on experimental data as I would have liked. So I tried assigning reading from the literature in lieu of a textbook. Although this method was entirely appropriate for an advanced course, it was a relatively inefficient process and not practical for a first course in molecular biology. To streamline the process, I augmented the literature readings with hand-drawn cartoons of the data I wanted to present, and I made transparencies of figures from the journal articles. But I really wanted a textbook that presented the concepts of molecular biology, along with experiments that led to those concepts. I wanted clear explanations that showed students the relationship between the experiments and the concepts. So, I finally decided that the best way to get such a book would be to write it myself. I had already coauthored a successful introductory genetics text in which I took an experimental approach—as much as possible with a book at that level. That gave me the courage to try writing an entire book by myself and to treat the subject as an adventure in discovery.

Organization

The book begins with a four-chapter sequence that should be a review for most students. Chapter 1 is a brief history of genetics. Chapter 2 discusses the structure and chemical properties of DNA. Chapter 3 is an overview of gene expression, and Chapter 4 deals with the nuts and bolts of gene cloning. All these are topics that the great majority of molecular biology students have already learned in an introductory genetics course. Still, students of molecular biology need to have a grasp of these concepts and may need to refresh their understanding of them. I do not deal specifically with these chapters in class; instead, I suggest students consult them if they need more work on these topics. These chapters are written at a more basic level than the rest of the book.

Chapter 5 describes a number of common techniques used by molecular biologists. It would not have been possible to include all the techniques described in this book in one chapter, so I tried to include the most common or, in a few cases, valuable techniques that are not mentioned elsewhere in the book. When I teach this course, I do not present Chapter 5 as such. Instead, I refer students to it when we first encounter a technique in a later chapter. I do it that way to avoid boring my students with technique after technique. I also realize that the concepts behind some of these techniques are rather sophisticated, and the students' appreciation of them is much deeper after they've acquired more experience in molecular biology.

Chapters 6–9 describe transcription in bacteria. Chapter 6 introduces the basic transcription apparatus, including promoters, terminators, and RNA polymerase, and shows how transcripts are initiated, elongated, and terminated. Chapter 7 describes the control of transcription in three different operons, then Chapter 8 shows how bacteria and their phages control transcription of many genes at a time, often by providing alternative sigma factors. Chapter 9 discusses the interaction between bacterial DNA-binding proteins, mostly helix-turn-helix proteins, and their DNA targets.

Chapters 10–13 present control of transcription in eukaryotes. Chapter 10 deals with the three eukaryotic RNA polymerases and the promoters they recognize. Chapter 11 introduces the general transcription factors that collaborate with the three RNA polymerases and points out the unifying theme of the TATA-box-binding protein, which participates in transcription by all three polymerases. Chapter 12 explains the functions of gene-specific transcription factors, or activators. This chapter also illustrates the structures of several representative activators and shows how they interact with their DNA targets. Chapter 13 describes the structure of eukaryotic chromatin and shows how activators can interact with histones to activate or repress transcription.

Chapters 14–16 introduce some of the posttranscriptional events that occur in eukaryotes. Chapter 14 deals with RNA splicing. Chapter 15 describes capping and polyadenylation, and Chapter 16 introduces a collection of

fascinating "other posttranscriptional events," including rRNA and tRNA processing, *trans*-splicing, and RNA editing. This chapter also discusses three kinds of posttranscriptional control of gene expression: (1) RNA interference; (2) modulating mRNA stability (using the transferrin receptor gene as the prime example); and (3) control by microRNAs.

Chapters 17–19 describe the translation process in both bacteria and eukaryotes. Chapter 17 deals with initiation of translation, including the control of translation at the initiation step. Chapter 18 shows how polypeptides are elongated, with the emphasis on elongation in bacteria. Chapter 19 provides details on the structure and function of two of the key players in translation: ribosomes and tRNA.

Chapters 20–23 describe the mechanisms of DNA replication, recombination, and translocation. Chapter 20 introduces the basic mechanisms of DNA replication and repair, and some of the proteins (including the DNA polymerases) involved in replication. Chapter 21 provides details of the initiation, elongation, and termination steps in DNA replication in bacteria and eukaryotes. Chapters 22 and 23 describe DNA rearrangements that occur naturally in cells. Chapter 22 discusses homologous recombination and Chapter 23 deals with translocation.

Chapter 24 presents concepts of genomics, proteomics, and bioinformatics. The chapter begins with an old-fashioned positional cloning story involving the Huntington disease gene and contrasts this lengthy and heroic quest with the relative ease of performing positional cloning with the human genome (and other genomes).

New to the Fourth Edition

One of the most significant developments in molecular biology since the last edition has been the recognition of the importance of noncoding RNAs. Some of these, like microRNAs (miRNAs), control gene expression in eukaryotes. Accordingly, a new section has been added to **Chapter 16**, showing how miRNAs can act like siRNAs in causing target mRNA destruction, or can depress the translation of target mRNAs without destroying them.

This edition also introduces riboswitches, segments of mRNAs that bind to small molecules and respond by changing RNA conformation, thereby controlling translation. In **Chapter 7**, we will see one example in which a small molecule binds to a riboswitch in a nascent mRNA, forcing it to form a terminator that aborts mRNA synthesis. In **Chapter 17**, we will see another example in which binding of a small molecule to a riboswitch in an mRNA causes the mRNA to change conformation in such a way as to hide the ribosome-binding site, blocking translation.

All but the introductory chapters of this fourth edition have been updated and include new information. Here are a few highlights:

Chapter 10: X-ray crystallography of RNA polymerase II in complex with DNA has revealed the role of the polymerase in separating the nascent RNA from its DNA template strand. Also, for the first time, the 12-subunit RNA polymerase II has been crystallized. This complete crystal structure shows a prominent clamp in the closed position, leaving room only for a single strand of DNA to enter the enzyme. This implies that the DNA must melt in order for the template strand to descend by itself into the active site.

Chapter 11: Structural studies on a TFIIB-polymerase II complex show that TFIIB binds to TBP at the TATA box via its C-terminal domain, and to polymerase II via its N-terminal domain. This bridging action effects a coarse positioning of the polymerase active center about 25 bp downstream of the TATA box.

Chapter 13: We still do not understand chromatin remodeling in detail, but we now know that catalyzed remodeling of nucleosomes involves the formation of distinct conformations of the nucleosomal DNA with respect to the core histones. This contrasts with uncatalyzed DNA exposure in nucleosomes, or simple sliding of nucleosomes along a stretch of DNA. We also have new insight into the structure of the 30-nm chromatin fiber. The crystal structure of a tetranucleosome consists of dual stacks of nucleosomes, with a zigzag of linker DNAs between the two stacks. This arrangement is incompatible with the solenoid model, and most other models for the 30-nm fiber, but it is consistent with a crossed-linker, two-start helix. That is, the two stacks of nucleosomes could coil around each other in a left-handed double helix.

Chapter 14: Alternative splicing is a very common phenomenon in higher eukaryotes. It represents a way to get more than one protein product out of the same gene, and a way to control gene expression in cells. Control of alternative splicing is exerted by splicing factors that bind to the splice sites and branchpoint, and also by proteins that interact with exonic splicing enhancers (ESEs), exonic splicing silencers (ESSs), and intronic silencing elements.

Chapter 16: In addition to the new material on miRNAs, this chapter includes new information on the mechanism of RNAi. In particular, the protein known as slicer, which cleaves target mRNAs, is now identified as Argonaute2, and a minimal RISC composed only of the Argonaute protein Ago2 and siRNA is described. New information on the assembly of the RISC is also presented.

Chapter 17: In addition to the new material on the control of prokaryotic translation initiation by a riboswitch, this chapter presents a model for the control of eukaryotic translation initiation by a microRNA.

Chapter 18: This chapter shows how the unusual amino acids selenocysteine and pyrrolysine are incorporated into growing polypeptides in response to the

termination codons UGA and UAG, respectively. It also discusses a prokaryotic protein called trigger factor, which associates with the ribosome in such a way as to catch nascent polypeptides as they emerge from the ribosome's exit tunnel. In this way, hydrophobic regions of a nascent polypeptide are protected from inappropriate associations until the appropriate partner is available.

Chapter 19: This chapter presents new evidence for the participation of the 2′-hydroxyl group of the P site tRNA in the chemistry of transpeptidation, the formation of peptide bonds by the ribosome. It also discusses the first crystal structure of the *E. coli* ribosome, which reveals a twisting of the head of the 30S particle that may be related to translocation.

Chapter 20: This chapter introduces two mechanisms for a new kind of base excision repair—the repair of 8-oxoguanine sites in DNA.

Chapter 22: This chapter discusses new evidence that suggests that Spo11, after causing and binding covalently to double-strand breaks, is released from the DSBs in a complex with oligonucleotides ranging from about 12–37 nt long. The cleavage by Spo11 appears to be asymmetric, which has implications for the mechanism of subsequent Holliday junction formation.

Chapter 24: The fields of genomics, proteomics, and bioinformatics are growing at an explosive rate. Accordingly, Chapter 24 is the most thoroughly updated chapter in the fourth edition, and it has been renamed to include bioinformatics in the title, reflecting the importance of this new subdiscipline. Some of the highlights of the updates in this chapter are: (1) Whole chromosome transcriptional mapping, which is high resolution mapping of transcripts to specific sites in chromosomes. Such studies have revealed the abundance of noncoding RNAs, including nonpolyadenylated RNAs, in eukaryotic cells. (2) RNAi analysis, in which RNAi is used to knock down the expression of genes on a genome-wide basis and the effects of these knockdowns are analyzed. (3) Tissue-specific expression profiling, which can be done by examining the spectrum of mRNAs whose levels are decreased by an exogenous miRNA, and comparing that to the natural spectrum of mRNA levels in various tissues. (4) The International HapMap Consortium has published a haplotype map including over 1 million human SNPs, discovered by genotyping 269 DNA samples from four distinct human populations. (5) Using bioinformatic techniques, scientists have discovered highly conserved sequence motifs in the promoter regions and 3′-UTRs of four mammalian species, including humans. The motifs in the promoter regions probably represent binding sites for transcription factors. Most of the motifs in the 3′-UTRs probably represent binding sites for miRNAs. (6) The finished draft of the human genome is accomplished. The importance of this achievement is discussed and the human X chromosome is featured because of its role in human disease.

Supplements

- A presentation CD-ROM contains digital files for most of the line art, tables, and photographs in the text in an easy-to-use format. This format is compatible with either PC or Macintosh.
- Text-Specific Website

  The following website, specific to this text, provides access to digital image files, updates, and web links for both students and instructors. Separate message boards for both instructor and student discussion are also available:

  www.mhhe.com/weaver4