

Preface

THREE SIGNIFICANT CHANGES HAVE TAKEN PLACE in molecular biology since the fifth edition of *Molecular Biology of the Gene* appeared four years ago. The first is in the cost and influence of genome sequencing; the second is the explosion in our appreciation of RNA as a regulatory molecule; and the third is the widespread touting of systems biology as the future of our field in the postgenomic era. The influence of all three developments is evident in this, the sixth edition.

Numerous species—together with one of the authors—have had their genomes sequenced in the past four years. And currently each genome can be done at a cost almost 500-fold lower than was possible in 2003. In that year we knew the sequences of the model systems—yeast, worm, *Arabidopsis*, fly, mouse, human—and perhaps two or three other animals. Now we have a much wider representation from the tree of life: we have the chimp, rat, and honeybee genomes; the cat and dog; several more insects; rice; the sea anemone; the opossum; and many others besides. And we have many more representatives of each group: as an example, 13 *Drosophila* species have by now been sequenced. And finally, as noted above, we have entered the era of the individual human genome—as well as James Watson's, Craig Venter's genome has been completely sequenced.

Along with this mushrooming in the amount of sequencing, the influence of sequencing is becoming ever more apparent and routine. There are several cases in this new edition where evolutionary and mechanistic arguments are based on the ability to compare genomes—in some cases between closely related species, in others more distantly.

Genome sequences have contributed to the second big development—the growing understanding of all that RNA does. The extent and importance of alternative splicing and the widespread use of regulatory RNAs represent areas where the text has been extensively revised.

Systems biology remains a rather ill-defined term, evoking somewhat different things for different people. For some it reflects high-throughput methodology, for example; for others it reflects mathematical modeling of biological systems. In this edition, we examine one aspect of this emerging field—the one with most obvious and immediate significance to topics covered in this book—namely, the representation and modeling of gene regulatory networks.

Organization of the Book

Although much has changed, many core values and organizing principles of the book have not. Thus, the new edition retains the structure of the last, being divided into five parts.

Part 1 lays the groundwork for what follows—it summarizes the history of genetics and molecular biology and explains the chemical principles that determine the structure and function of macromolecules.

Part 2 covers the organization and maintenance of the genetic material. Beginning with a description of the structures of DNA and RNA, it also covers the organization of genomes in chromosomes and explains DNA replication, recombination, and repair.

Part 3 deals with expression of the genome through transcription, splicing, and translation.

Part 4 is about how the expression of genes within the genome is regulated. This includes mechanisms of transcriptional regulation, the role of regulatory RNAs in controlling gene expression, gene regulation in the control of development, and how changes in regulation underlie much of evolution. Finally, we consider genome-wide studies of gene expression and how gene regulatory networks are modeled and analyzed in the emerging fields of systems and synthetic biology.

Part 5 comprises two chapters: one on the techniques of molecular biology, genomics, proteomics, and bioinformatics and the other on the model organisms whose study has revealed many of the underlying principles of molecular biology.

As well as revision of each chapter, there are significant additions to this edition. Most notably, there are two new chapters and a new range of boxed features.

New Chapters and Boxes

Chapter 18: Regulatory RNAs One of the striking developments since the fifth edition is the extent to which RNAs are found to be regulators of gene expression. In this chapter, we cover everything from riboswitches and small RNAs in bacteria, through RNA interference and microRNAs in eukaryotes, to the role of RNA regulators in X-chromosome inactivation. As well as mechanistic details of how these regulators work, we discuss how they were discovered and how they have afforded us new ways to manipulate gene expression artificially.

Chapter 20: Genome Analysis and Systems Biology The genomes of many animals have been sequenced and found to be strikingly similar. This finding emphasizes that it is largely differences in the regulation, rather than the identity, of the genes within a genome that determine phenotype. Techniques now exist to study genome-wide gene expression patterns; we discuss how these techniques work and what they have revealed. In the second half of the chapter, we learn about the representation of gene regulatory networks as so-called wiring diagrams (an example is shown on the cover of this book as is discussed below). These diagrams emphasize the information flow and types of decision-making steps within such pathways.

Boxes The sixth edition includes an array of new boxes, each assigned to one of four color-coded categories:

- **KEY EXPERIMENTS** describe critical experiments in each field, ranging from the historical to the very recent. Without disrupting the narrative flow of the main text, they give students insight into how scientific problems are solved.
- **TECHNIQUES** describe techniques of particular relevance to the chapter in question, complementing the more general techniques described in the last two chapters of the book.
- **MEDICAL CONNECTIONS** highlight links between the basic molecular biology described and a variety of human diseases. These boxes emphasize how basic research informs current medical understanding.
- **ADVANCED CONCEPTS** explore selected concepts in more depth to allow students a glimpse of cutting-edge ideas.

As well as being included in the detailed table of contents, there is a separate listing of all the Boxes by topic on page xxxi.

Media Icons New web references and icons within the text direct students to explore the animations and structural tutorials on the companion website.

Supplements

Companion Website www.aw-bc.com/watson

The Companion Website combines the contents from the previous edition's "Student Media" CD-ROM into one fully integrated resource that helps students better understand complex structures and molecular interactions. Two new structural tutorials on Argonaute and RecBCD have been added in the sixth edition in addition to the popular structural media tutorials of the previous edition.

III. RuvA Tetramer-Holliday Junction DNA

○ In tetrameric RuvA, each monomer is a lobe of the symmetrical tetramer. The domain I beta barrels of each monomer are positioned centrally, and domains II and III are located peripherally.

○ Examining the charge distribution of atoms over the surface of the RuvA tetramer, it is clear that the DNA binding surface is largely positively charged (basic), whereas the opposite surface is largely negatively charged (acidic). The orange, yellow, and red colors represent relatively negatively charged atoms and the green and blue colors represent more positively charged atoms (red and blue indicate acidic or basic residues). The relative positive charge of the DNA binding surface attracts the negatively charged DNA backbone.

○ An exception to the mostly basic surface of the DNA binding side of the RuvA tetramer is an eight residue acidic "central pin" containing glutamate⁹³ and aspartate⁹⁵ from each monomer. This negatively charged center may repel the negatively charged oxygens of the DNA backbone, driving the DNA away from the center of the Holliday junction.

○ A number of residues that interact directly or indirectly (via water molecules) with the DNA backbone line the channels in which Holliday junction DNA is bound.

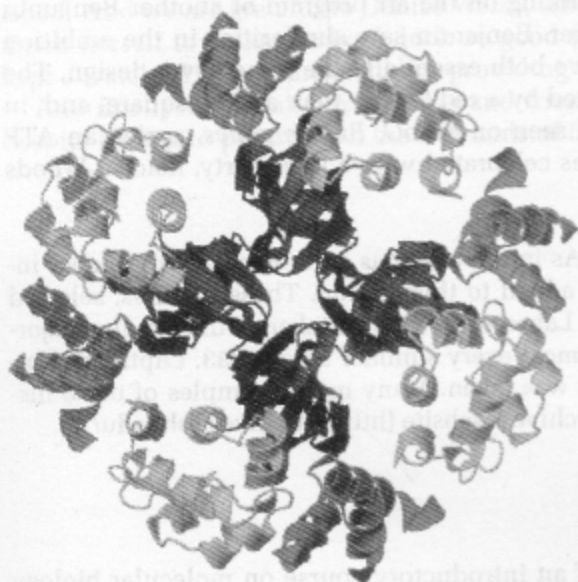
reset

return to beginning

IV. References

Ariyoshi, M., Nishino, T., Iwasaki, H., Shinagawa, H., Morikawa, K.: Crystal Structure of the Holliday Junction DNA in Complex with a Single RuvA Tetramer. *Proc. Natl. Acad. Sci.* **97**: 8257-8262 (2000).

Hargreaves, D., Rice, D. W., Sedelnikova, S. E., Artymiuk, P. J., Lloyd, R. G., Ruffert, J. B.: Crystal



Glossary

An online glossary defines terms that appear in bold in the text.

Instructor Resource CD-ROM 978-0-3215-2766-0/0-3215-2766-6

This dual-platform CD-ROM contains all art and tables from the book in jpeg and PowerPoint format in high-resolution (150-dpi) files. This CD-ROM also contains the answers to the end-of-chapter Critical Thinking questions posed to students on the Companion Website.

Transparency Acetates 0-3215-3639-8/978-0-3215-3639-6

The transparency package features approximately 300 four-color illustrations from the text selected by the authors.

Cover Design and Part-Opener Photographs

Cover image The cover is a representation from systems biology of a frequently observed motif in the circuitry governing gene expression (see Chapter 20). The motif, known as an incoherent feed-forward loop, consists of two regulatory genes, A and B, and a target gene C whose expression is indicated as the output. In response to an input, gene A is turned on. The gene A product is an activator both of the target gene C and of the second regulatory gene B. The gene B product is a repressor of gene C. The representation “AND” in the figure indicates that the circuit operates with the logic of a so-called AND gate, a term borrowed from electrical engineering, which means that the output requires (in the case shown) the presence of the gene A activator AND the absence of the gene B repressor. Incoherent feed-forward loops have the property of causing target genes to be expressed in a pulse. That is, gene C is turned on by the activator in response to the input and then switched off by the delayed accumulation of the repressor.

The cover design of this, the sixth edition, consciously mimics that of the first edition, published by Benjamin (a forerunner of Benjamin Cummings) in 1965. That cover, together with those of the second through fifth editions, is reproduced on the back cover of this edition. The first edition cover was designed by Bill Prokos, a New York City painter, who in 1965 was working on the art program of another Benjamin book, *Bioenergetics*, by Albert Lehninger. Benjamin saw similarities in the ambition and style of these two books and so gave both essentially the same cover design. The Lehninger book, however, was dominated by a red (rather than a blue) square, and, in place of the turn of DNA double helix seen on *MBoG*, *Bioenergetics* carried an ATP molecule. Publication of both books was celebrated with a joint party, held in Woods Hole that year.

Cold Spring Harbor Laboratory Photographs As in the previous edition, each part opener includes a few photographs, some newly added to this edition. These pictures, selected from the archives of Cold Spring Harbor Laboratory, were all taken at the lab, the majority during the Symposia hosted there almost every summer since 1933. Captions identify who is in each picture and when it was taken. Many more examples of these historic photos can be found at the CSHL archives website (<http://archives.cshl.edu/>).

Acknowledgments

Parts of the current edition grew out of an introductory course on molecular biology taught by one of us (RL) at Harvard University, and this author is grateful to Steve Harrison and Jim Wang who contributed to this course in past years and whose influence is reflected in Chapter 6 and elsewhere. We are also particularly grateful to Craig Hunter, who wrote the section on the worm for Chapter 21, and to Rob Martienssen, who wrote the section on the plant *Arabidopsis* for that same chapter.

We have shown sections of the manuscript to various colleagues and their comments have been most useful in ensuring the text and figures reflect current ideas and information. Specifically, we thank Katsura Asano, Jamie Cate, Amy Caudy, Richard Ebright, Mike Eisen, Chris Fromme, Brenton Graveley, Ann Hochschild, Jim Hu, David Jeruzalmi, Leemor Joshua-Tor, Sandy Johnson, Adrian Krainer, Karoline Luger, Julian Lewis, Sue Lovett, Rob Martienssen, Bill McGinnis, Matt Michael, Lily Mirels, Nipam Patel, Mark Ptashne, Danny Reinberg, and Bruce Stillman.

We also thank those who provided us with figures, or the wherewithal to create them: Sean Carroll, Seth Darst, Richard Ebright, Brenton Graveley, Ann Hochschild, Julian Lewis, Bill McGinnis, Phoebe Rice, Dan Rokhsar, Nori Satoh, Matt Scott, Peter Sorger, Tom Steitz, Andrzej Stasiak, Dan Voytas, and Steve West.

We are indebted to Leemor Joshua-Tor who rendered all the structure figures so beautifully. Her skill and patience were much appreciated. We are also grateful to

those who provided their software¹: Per Kraulis, Robert Esnouf, Ethan Merritt, Barry Honig, and Warren Delano. Coordinates were obtained from the Protein Data Bank (www.rcsb.org/pdb/); and citations to those who solved each structure are included in the figure legends.

Our art program was again executed by a talented team from the Dragonfly Media Group, led by Craig Durant and assisted by Helen Wortham. Denise Weiss and Ed Atkeson produced the cover design. We thank Clare Clark and the CSHL Archive for providing the photos for the part openers and for much help tracking them down.

We thank Susan Winslow at Benjamin Cummings who did much to ensure this new edition happened and even more to make things as easy as possible once it began in earnest. Gary Carlson took over late in the game, and we are grateful to him for letting things continue seamlessly. In overseeing the book's development, Jan Argentine managed endlessly to accommodate our demands, reorganizing our tasks to provide the time and support we needed, while still protecting the sacred schedule. Kaaren Janssen edited our text and redrew our messy sketches, generating momentum with her inexhaustible desire to help, and Inez Sialiano kept organized the resulting drafts and corrections, and Carol Brown the permissions. In production, the patience of Kathleen Bubbeo and Susan Schaefer was as tireless as it needed to be in accommodating the changes we continued to generate even after we were supposed to have finished. And we are grateful to Denise Weiss, who not only oversaw production, but did much herself to make the book look so good from page layout to design. John Inglis kept a watchful eye on proceedings, providing useful advice whenever needed.

And finally we thank our families and friends for again providing strong support, despite having to put up with the demands of the project just as much as we did.

JAMES D. WATSON
TANIA A. BAKER
STEPHEN P. BELL
ALEXANDER GANN
MICHAEL LEVINE
RICHARD LOSICK

TANIA A. BAKER is the Whitbread Professor of Biology at the Massachusetts Institute of Technology and an Investigator of the Howard Hughes Medical Institute. She received a B.S. in biochemistry from the University of Wisconsin, Madison, and a Ph.D. in biochemistry from Stanford University in 1988. Her graduate research was carried out in the laboratory of Professor Arthur Kornberg and focused on mechanisms of initiation of DNA replication. She did postdoctoral research in the laboratory of Dr. Kiyoshi Mizushima at the National Institute of Health, studying the mechanism and regulation of RNA transcription. Her current research explores mechanisms and regulation of genetic recombination, genome-wide protein unfolding, and ATP-dependent protein degradation. Professor Baker received the 2001 Eli Lilly Research Award from the American Society of Microbiology and the 2000 MIT School of Science Teaching Prize for Undergraduate Education. She is a fellow of the American Academy of Arts and Sciences since 2004, and was elected to the National Academy of Sciences in 2007. She is co-author (with Arthur Kornberg) of the book *DNA Replication*, Second Edition.

STEPHEN P. BELL is a Professor of Biology at the Massachusetts Institute of Technology and an Investigator of the Howard Hughes Medical Institute. He received B.A. degrees from the Department of Biochemistry, Molecular Biology, and Cell Biology and the Integrated Science Program at Northwestern University and a Ph.D. in biochemistry at the University of California, Berkeley in 1981. His graduate research was carried out in the laboratory of Dr.

¹Per Kraulis granted permission to use MolScript (Kraulis P.J. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**: 946–950). Robert Esnouf gave permission to use BobScript (Esnouf R.M. 1997. An extensively modified version of MolScript that includes greatly enhanced coloring capabilities. *J. Mol. Graph. Model.* **15**: 132–134). In addition, Ethan Merritt gave us use of Raster 3D (Merritt E.A. and Bacon D.J. 1997. Raster3D: Photorealistic molecular graphics. *Methods Enzymol.* **277**: 505–524), and Barry Honig granted permission to use GRASP (Nicholls A., Sharp K.A., and Honig B. 1991. Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**: 281–296). Warren DeLano agreed to the use of PyMOL (DeLano W.L. 2002. *The PyMOL Molecular Graphics System*. DeLano Scientific, Palo Alto, California).