

Preface

"The map is not the territory."

Alfred Korzybski

The last 50 years in biology have seen an explosion of both data and understanding that rivals the fertile period between Tycho Brahe's definitive naked-eye investigations of the heavens and Newton's introduction of the "System of the World." One of the consequences of these stunning advances is the danger of becoming overwhelmed by the vast quantities of data coming at us from quarters ranging from next-generation sequencing to quantitative microscopy. For example, at the time of this writing, there are in excess of two million ribosomal RNA sequences deposited on publically accessible databases. But what does it all mean? A central role of scientific textbooks is to attempt to come to terms with broad areas of progress and to organize and distill the vast amounts of available information in a conceptually useful manner. In our view, an effective textbook can act as a map to help curious people discover unfamiliar territories. As with real maps, different purposes are served by different kinds of abstraction. Some maps show roads, some show topography, with each being useful in its own context.

A number of textbook writers have undertaken the formidable task of writing excellent, comprehensive surveys of cell and molecular biology, although each one of these books serves as a slightly different kind of map for the same overlapping territory. Although we cover some of the same material as a typical molecular and cell biology book, our goal in this book is fundamentally different. There is no single, correct way to construct a conceptually simplified map for a huge and complex field such as cell and molecular biology. Most modern biology textbooks organize ideas, facts, and experimental data based on their conceptual proximity for some particular biological function. In contrast, this book examines the same set of information from the distinct perspective of *physical biology*. We have therefore adopted an organization in which the proximity of topics is based on the physical concepts that unite a given set of biological phenomena, instead of the cell biology perspective. By analogy to a map of the United States, a cell biology textbook might describe the plains of Eastern Colorado in the same chapter as the mountains of Western Colorado, whereas our physical biology book would group Eastern Colorado with the rolling fields of Iowa, and Western Colorado with mountainous West Virginia.

This book does not assume extensive prior knowledge on the part of the reader, though a grounding in both calculus and elementary physics is essential. The material covered here is appropriate for a first course in physical biology or biophysics for either undergraduates or graduate students. It is also intended for any scientist interested in learning the basic principles and applications of physical modeling for research in biology, and aims to provide a novel perspective even to scientists who are already familiar with some of the material. Throughout the book, our organization of ideas and data based on proximity in physical biology space juxtaposes topics that are not obviously related in cell biology space. For example, DNA

wrapping around nucleosomes in the eukaryotic nucleus, DNA looping induced by the binding of transcriptional repressors in the context of bacterial gene regulation, and DNA packing into the narrow confines of bacteriophage capsids all appear in the same chapter because they are related by the mechanical rules governing the bending of DNA. Next, the physical and mathematical treatment we derive for DNA bending is directly applied to other kinds of long, thin, biological structures, including the filaments of the cytoskeleton. This organizational principle brings into focus the central thesis of this book, namely, that the appropriate application of a relatively small number of fundamental physical models can serve as the foundation of whole bodies of quantitative biological intuition, broadly useful across a wide range of apparently unrelated biological problems.

During the 12-year journey that led to this book, we benefited immeasurably from the generosity and enthusiasm of hundreds of scientific colleagues who graciously shared their data, ideas, and perspectives. Indeed, in many respects, we view our book as an exercise in *quantitative journalism*, based upon extensive "interviews" with these various scientists in a wide range of disciplines. We offer this book as a report from the front, to share some of the most interesting things that we have learned from our colleagues with any and all inquiring individuals who wish to think both deeply and broadly about the connections between biology and the physical sciences. Our imagined audience spans the range from 18-year-old mechanical engineering undergraduates curious about the application of their discipline to medicine, to 40-year-old string theorists wishing to apply their mathematical and physical talents to living matter, to 70-year-old renowned biologists wondering whether their insights into living systems might be improved by a mathematical treatment.

Although the claim that a handful of simple physical models can shed more than superficial light on complex biological processes might seem naive, the biological research literature is teeming with examples where important quantitative insight into questions of pressing interest has been gained by the application of such models. In every chapter, we have chosen specific examples from classic and current research papers where quantitative measurements on biological systems can be largely understood by recourse to simple, fundamental, physical ideas. In cases where the simplest possible physical models fail to fit the data, the specific quantitative nature of the disparities can often lead to testable new biological hypotheses. For example, a simple calculation estimating the amount of time it would take for a newly synthesized protein to diffuse from the cell body of a motor neuron in the spinal cord to the synapse formed by the same neuron in the foot proves that diffusion is far too slow to get the job done, and an active transport process must occur. Inevitably, researchers performing experiments on biological systems must have physical models explicitly or implicitly in mind, whether imagining how changes in the rate of transcription initiation for a particular gene will lead to changes in the overall amount of the gene product in the cell, or picturing the ways that signaling molecules move through cellular space to encounter their targets, or envisioning how cell movements during embryogenesis lead to the final three-dimensional structures of organs and limbs. In this book, we aim to provide a physical and mathematical toolkit so that people used to thinking deeply about biological problems can make this kind of quantitative intuition explicit; we also hope to provide a perspective on biology that may inspire people from a background more heavily based in physics or

mathematics to seek out new biological problems that are particularly appropriate for this kind of quantitative analysis.

Our general approach follows four steps. First, we introduce a biological phenomenon; second, we perform simple order-of-magnitude estimates to develop a "feeling for the numbers" involved in that process; third, we demonstrate the application of an extremely simple first-pass model; and finally, where possible, we present a refinement of the oversimplified model to better approximate biological reality. Our goal is to share the pleasure in seeing the extent to which simple models can be tailored to reveal the *complexity* of observed phenomena. For our examples, we have chosen particular biological cases that we believe to be worthy illustrations of the concepts at hand and that have captured our imaginations, often because of particularly elegant or clever experiments that were designed to generate intriguing sets of quantitative data. While we have been conscientious in our exploration of these facts and in our construction of simple models, it is inevitable that we will have made errors due to our ignorance and also due to the fact that, in many cases, new discoveries may change the particulars of our case studies. (A list of errors and their corrections will be posted on the book's website as well as the website of one of the authors (R.P.).) Nevertheless, because our goal is to demonstrate the power of applying simple models to complex systems, even when some details are wrong or missing, we hope that any particular lapses will not obscure the overall message. Furthermore, in many cases, we have described phenomena that are still awaiting a satisfying physical model. We hope that many of our readers will seize upon the holes and errors in our exploration of physical biology and take these as challenges and opportunities for launching exciting original work.

Our second edition builds upon the foundations laid in the previous edition, with the addition of two new chapters that focus on central themes of modern biology, namely, light and life and the emergence of patterns in living organisms. The new Chapter 18 focuses on several key ways in which light is central in biology. We begin with an analysis of photosynthesis that illustrates the quantum mechanical underpinnings of both the absorption of light and the transfer of energy and electrons through the photosynthetic apparatus. The second part of our story in that chapter considers the rich and beautiful subject of vision. The new Chapter 20 uses insights garnered throughout the book to ask how it is that organisms ranging from flies to plants can build up such exquisite patterns. Here we explore Turing's famed model of several interacting chemical species undergoing chemical reactions and diffusion and other more recent advances in thinking about problems ranging from somitogenesis to phyllotaxis.

The book is made up of four major parts. Part 1, *The Facts of Life*, largely focuses on introducing biological phenomena. For biology readers already familiar with this material, the hope is that the quantitative spin will be enlightening. For physics readers, the goal is to get a sense of the biological systems themselves. Part 2, *Life at Rest*, explores those problems in biology that can be attacked using quantitative models without any explicit reference to time. Part 3, *Life in Motion*, tackles head-on the enhanced complexity of time-dependent systems exhibiting dynamic behavior. Finally, Part 4, *The Meaning of Life*, addresses various kinds of information processing by biological systems.

Because our hope is that you, our readers, represent a broad diversity of backgrounds and interests, throughout the book we try as much as possible to introduce the origin of the facts and principles

that we exploit. We are reluctant to ever simply assert biological “facts” or physical “results,” and would not expect you to blindly accept our assertions if we did. Therefore, we often describe classical observations by biologists over the past centuries as well as the most recent exciting results, and illustrate how current thinking about complex biological problems has been shaped by a progression of observations and insights. Extended discussions of this kind are separated from the main text in sections labeled *Experiments Behind the Facts*. In a complementary way, whenever we find it necessary to derive mathematical equations, we proceed step by step through the derivation and explain how each line leads to the next, so that readers lacking a strong background in mathematics can nevertheless follow every step of the logic and not be forced to take our word for any result. Specific sections labeled *The Math Behind the Models* and *The Tricks Behind the Math* provide summaries for the mathematical techniques that are used repeatedly throughout the book; many readers trained in physics will already be familiar with this material, but biologists may benefit from a brief refresher or introduction. In addition, we include sections labeled *Estimate* that help to develop a “feeling for the numbers” for particularly interesting cases.

Another critical new element in our second edition is a feature called *Computational Exploration*. The idea of these excursions is to show how simple computer analyses can help us attack problems that are otherwise inaccessible. In the first edition, we underemphasized “computation” because we wanted to combat the spurious idea that theory in biology is synonymous with computation. While we made this exaggeration to make a point, we did so at a price, because computation is not only useful, but downright indispensable in some problems. Further, one of the beauties of turning a model into a specific numerical computation is that to get a computer to produce a meaningful number, nothing can be left unspecified. The Computational Explorations are offered as a way for the reader to develop a particular habit of mind, and none of them should be viewed as illustrating the state of the art for making such calculations. Matlab and Mathematica code related to most of these explorations is provided on the book’s website.

Although we review the basic information necessary to follow the exposition of each topic, you may also find it useful to have recourse to a textbook or reference book covering the details of scientific areas among biology, physics, chemistry, and mathematics, with which you consider yourself less familiar. Some references that are among our favorites in these fields are suggested at the end of each chapter. More generally, our references to the literature are treated in two distinct ways. Our suggestions for *Further Reading* reflect our own tastes. Often, the choices that appear at a chapter’s end are chosen because of uniqueness of viewpoint or presentation. We make no attempt at completeness. The second class of *References* reflect work that has explicitly touched the content of each chapter, either through introducing us to a model, providing a figure, or constructing an argument.

At the end of each chapter, we include a series of problems that expand the material in the chapter or give the opportunity to attempt model-building for other case studies. In the second edition, we have considerably expanded the scope of the end-of-chapter problems. These problems can be used within formal courses or by individual readers. A complete *Solutions Manual*, covering all problems in the book, is available for instructors. There are several different types

of problems. Some, whose goal is to develop a “feeling for the numbers,” are arithmetically simple, and primarily intended to develop a sense of order-of-magnitude biology. Others request difficult mathematical derivations that we could not include in the text. Still others, perhaps our favorites, invite the readers to apply quantitative model-building to provocative experimental data from the primary research literature. In each chapter, we have loosely identified the different problems with the aforementioned categories in order to assist the reader in choosing which one to attack depending on particular need. The book’s website also includes *Hints for the Reader* for some of the more difficult problems.

Our book relies heavily on original data, both in the figures that appear throughout the book and in the various end-of-chapter problems. To make these data easily accessible to interested readers, the book’s website includes the original experimental data used to make all the figures in the book that are based upon published measurements. Similarly, the data associated with the end-of-chapter problems are also provided on the book’s website. It is our hope that you will use these data in order to perform your own calculations for fitting the many models introduced throughout the book to the relevant primary data, and perhaps refining the models in your own original work.

Student and Instructor Resources

Figures and PowerPoint® Presentations

The figures from the book are available in two convenient formats: PowerPoint and JPEG. There is one PowerPoint presentation for each chapter, and the JPEGs have been optimized for display on a computer.

Data Sets

The original data used to create both the figures and homework problems are available in Excel® spreadsheets. With this data, the reader can extend the theoretical tools developed in the book to fit experimental data for a wide range of problems. The data files contain explicit statement of all relevant units, and include references to the original sources.

Hints for Problems

This PDF provides both hints and strategies for attacking some of the more difficult end-of-chapter problems. In some cases, the hints provide intuition about how to set up the problem; in other cases, the hints provide explicit mathematical instructions on how to carry through more tricky manipulations.

Matlab® and Mathematica® Code

These files contain code for the Computational Explorations sidebars located throughout the book.

Movies

The movies complement the figures and discussion from the book by illustrating the rich dynamics exhibited by living organisms and the molecules that make them tick.

Solutions Manual

This PDF contains solutions to all problems in the book. It is available only to qualified instructors.

With the exception of the *Solutions Manual*, these resources are available on the *Physical Biology of the Cell*, 2nd Edition, media website:

<http://microsite.garlandscience.com/pboc2>

Access to the Solutions Manual is available to qualified instructors by emailing science@garland.com.

PowerPoint and Excel are registered trademarks of Microsoft Corporation in the United States and/or other countries.

MATLAB® is a trademark of The MathWorks, Inc.

Mathematica® is a trademark of Wolfram Research, Inc.