PREFACE

Overburdening necessarily leads to superficiality. Teaching should be such that what is offered is perceived as a valuable gift, and not as a hard duty.

—Albert Einstein

This is a textbook on concepts in biochemistry: concepts from physics, from chemistry, from biology. I have chosen proteins as the subject of the book because they provide the richest illustrations of those concepts. However, the book is not a description of the physical properties of proteins and their biological functions. Rather, my goal is to communicate a way of thinking about problems in biochemistry—in science, really. In this sense, I use proteins as a vehicle. But proteins are much more than a vehicle for this process: they are the most fascinating molecules that exist.

This book was born from teaching. Confronted with the need to choose a textbook for a one-semester course on proteins, I realized that none was suitable for what I was trying to accomplish. Having taught biochemistry for more than 20 years, I came to appreciate that students struggle with the sheer volume of material in introductory courses. They are asked to know an ever-growing amount of information (yet what a human being can reasonably learn and comprehend has not increased) at the expense of conceptual understanding. More than 50 undergraduate students have worked in my research laboratory. I have noticed that the amount of material learned in the classroom that is retained and usable is quite small. This is true even of students who completed the courses with excellent results. Why?

Biochemistry textbooks have become increasingly concentrated on biological aspects and on large amounts of information. Conversely, understanding chemical, physical, and mathematical aspects of biochemical topics has received less attention. Yet already about 20 years ago the National Research Council (NRC) of the National Academies reached the conclusion that "life science majors must acquire a much stronger foundation in the physical sciences. . . than they now get"[2] and recommended increasing connections between biological and physical sciences as a means to bridge the communication gap that exists between researchers in different disciplines. The NRC also emphasized the importance of raising interest for research in undergraduate students as early as possible. I wanted to do something to help change this state of affairs.

This book was written for students. My hope was to endow students with a frame of thought that would help them tackle problems they encounter in the course of their careers. Last year, I received an email from a student. He wrote, "We have never met, but you have helped me immensely. I am a Biophysics PhD candidate at Johns Hopkins, and I just passed my qualifying exam due in no small part to your book *Proteins: Concepts in Biochemistry*. The book has become a kind of bible for grad students in our program, passed down from generation to generation." This was the most gratifying message I have received about the book—thank you. Finally, whereas the book is not a monograph for researchers, I believe it will be helpful for all those who seek a foundation in more physical aspects of biochemistry.

For instructors, this textbook contains a manageable amount of material for a one-semester course on proteins for advanced undergraduates and beginning graduate students in chemistry, biochemistry, biophysics, and biology. At the PhD level, courses on proteins often have a heterogeneous population of students, with varied backgrounds. Those courses are often taught on the basis of journal articles from the scientific literature. In these cases too, I think this book will provide the necessary foundation for many

[2]National Research Council. 2003. *Bio 2010: Transforming Undergraduate Education for Future Research Biologists*. National Academic Press, Washington, DC.

students. The recommended preparation is one semester of calculus and a knowledge of introductory physical chemistry. An undergraduate general biochemistry course is helpful but not essential, as the book covers the basics of protein and DNA composition and structure.

I designed this book to present the essential concepts in biochemistry without overburdening students with too many facts or too much formalism. The book covers protein structure, evolution, stability, folding, ligand binding, and enzyme catalysis. This second edition includes the latest developments on protein folding concerning prediction of structure from sequence. The chapter on ligand binding has been reorganized. And the chapter on enzyme catalysis has been completely rewritten. It now concludes with a new section on the role of enzyme dynamics in catalysis, including hydrogen tunneling.

I wanted to approach these topics from a perspective of probabilistic (statistical) and quantitative thinking. Therefore, the book begins with an introduction to statistical thermodynamics, applied to biological macromolecules. The concept of the partition function is introduced early and used throughout. With the exception of Chapter 3 (evolution), all chapters use the partition function method to solve various problems. Mathematics are essential in science, and many topics in biochemistry require mathematics. However, I have tried to keep the mathematical formalism to a minimum. When mathematics are used, I have tried to explain the meaning of formulas and procedures in plain English. Mathematics tends to scare biochemistry students. The barrier, however, is mainly one of language, not of concepts. I have found that most students can follow a mathematical argument if it is explained in words. I have also sacrificed formalism for intuition. In so doing, I have taken approaches that purists may frown upon. I think this was a risk worth taking.

The book starts simple and becomes progressively more complex as the topics themselves become more complex and interconnected. Later chapters build on the previous ones: the book is meant to be read in sequence; it is not a reference text to look things up. Because of that, some sections may be assigned as homework reading, rather than covered in class. Each chapter ends with a set of problems of variable degrees of difficulty, not necessarily in order. A solution manual is available for instructors. The overall complexity of the problems tends to increase as the book progresses. Finally, each chapter includes a list of references, organized by topics. The references are listed under the topic for which they are most relevant, but of course some references are relevant for more than one topic. The references are not cited in-line to facilitate reading, but it should be fairly evident which references were used for a particular topic.

Some additional resources are highly recommended. One is the Protein Data Bank (PDB), which is the source of the protein structures that illustrate this book. The second is a means of visualization and rendering of those structures, such as the PyMol software. All protein structures in this book were produced with PyMol, which I am pleased to acknowledge. The open-source version is available for free to everyone.

Finally, I would like to acknowledge the people who have helped me in this endeavor. I am grateful to all the reviewers of the manuscript. In particular, I want to thank those whom I personally bothered and who have gone far beyond the normal task of a reviewer to help me. Namely, in alphabetical order, Antje Almeida, Robert Baldwin, Steve Benkovic, Arieh Ben-Naim, Alan Fersht, Karen Fleming, Martin Gruebele, Chris Halkides, Phil Hanoian, Michael Harms, Andrew Herr, Anne Hinderliter, Sandro Keller, Judith Klinman, Hee-Seung Lee, Alesia McKeown, Mike Messina, Adam Offenbacher, Bill Pearson, Stephen Rader, George Rose, and Shozo Yokoyama. One of the most delightful aspects of writing this book was the opportunity to meet, in person or by email, many scientists whom I have admired throughout my career. What is more, my requests for help and advice from them were always met with kindness and generosity. Their criticism (sometimes harsh but deserved) and suggestions were essential to improve the book. The views presented here, however, and any mistakes or errors that may remain, are of course, my exclusive responsibility.

One of my motivations to write this book was to pass to students what I have learned. Therefore, I want to take this chance to thank the people from whom I myself have learned the most: my first mentor in science, Winchil Vaz, for the many hours and years of great discussions; my PhD advisor, Tom Thompson, from whom I learned so many things, among which is not to be afraid of starting to work on something I know nothing about; and my postdoctoral advisor, Rod Biltonen, from whom I learned to use the partition function in biochemical problems and to whom, in this and many other senses, this book is a tribute.

Finally, I thank Antje for her support and patience with me during the many years of spoiled weekends and vacations while the writing of this book took place.

Paulo Almeida,
Wilmington, January 12, 2025