

Biomolecular Thermodynamics

From Theory to Application

Douglas E. Barrick



Foundations of Biochemistry and Biophysics

BIOCHEMISTRY AND BIOPHYSICS



CRC Press
Taylor & Francis Group
an informa business

ISBN 978-1-138-06884-1



9 781138 068841

www.crcpress.com • an informa business

BIOMOLECULAR THERMODYNAMICS
Douglas E. Barrick



Biomolecular Thermodynamics

From Theory to Application

Douglas E. Barrick



CRC Press

Taylor & Francis Group

Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

MATLAB® is trademarks of the Math Works, Inc. and are used with permission. The Mathworks does not warrant the accuracy of the text or exercises in this book. This book's use or discussion of MATLAB® software or related products does not constitute endorsement or sponsorship by the Math Works of a particular pedagogical approach or particular use of the MATLAB® software.

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2018 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works

Printed on acid-free paper

International Standard Book Number-13: 978-1-4398-0019-5 (Paperback)
978-1-138-06884-1 (Hardback)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging-in-Publication Data

Names: Barrick, Douglas, author.
Title: Biomolecular thermodynamics : from theory to application / Douglas Barrick.
Other titles: Foundations of biochemistry and biophysics.
Description: Boca Raton : Taylor & Francis, 2017. | Series: Foundations of biochemistry and biophysics
Identifiers: LCCN 2017005294 | ISBN 9781439800195 (pbk. : alk. paper) | ISBN 9781138068841 (hardback)
Subjects: | MESH: Thermodynamics | Biochemical Phenomena | Models, Theoretical
Classification: LCC QP517.T48 | NLM QU 34 | DDC 572/.436--dc23
LC record available at <https://lccn.loc.gov/2017005294>

Visit the Taylor & Francis Web site at
<http://www.taylorandfrancis.com>

and the CRC Press Web site at
<http://www.crcpress.com>

Biomolecular Thermodynamics

From Theory to Application

Foundations of Biochemistry and Biophysics

Founding Series Editor: John J. (Jack) Correia

Currently available

Quantitative Understanding of Biosystems: An Introduction to Biophysics

Thomas M. Nordlund

Biomolecular Thermodynamics: From Theory to Application

Douglas E. Barrick

Biomolecular Kinetics: A Step-by-Step Guide

Clive R. Bagshaw

Forthcoming

Physical Principles in Nucleic Acid Chemistry

David E. Draper

RNA Biophysics

Kathleen B. Hall

To Debbie and Timmy, who together are my role models for science,
love, and life.

Series Preface

Biophysics encompasses the application of the principles, tools, and techniques of the physical sciences to problems in biology, including determination and analysis of structures, energetics, dynamics, and interactions of biological molecules. Biochemistry addresses the mechanisms underlying the complex reactions driving life, from enzyme catalysis and regulation to the structure and function of molecules. Research in these two areas has a huge impact in pharmaceutical sciences and medicine.

These two highly interconnected fields are the focus of this book series. It covers both the use of traditional tools from physical chemistry such as nuclear magnetic resonance (NMR), x-ray crystallography, and neutron diffraction, as well as novel techniques including scanning probe microscopy, laser tweezers, ultrafast laser spectroscopy, and computational approaches. A major goal of this series is to facilitate interdisciplinary research by training biologists and biochemists in quantitative aspects of modern biomedical research and teaching core biological principles to students in physical sciences and engineering.

Proposals for new volumes in the series may be directed to Lu Han, senior publishing editor at CRC Press, Taylor & Francis Group (lu.han@taylorandfrancis.com).

Preface

This book introduces students to the concepts and skills necessary to understand the behavior of molecular systems in the biological and chemical sciences at a quantitative level, through the formalisms of classical and statistical thermodynamics. Although the application of thermodynamics to biomolecular and chemical problems is not new, most existing books written for traditional courses focus on the underlying theory and equations but do not provide adequate understanding of how to apply these equations in the analysis of experimental measurements. Thus, it is often difficult for students to properly leverage the knowledge obtained from traditional (bio)physical chemistry courses to “real world” physical chemical research. Moreover, many treatments of classical and statistical thermodynamics are complete and rigorous, but do not provide a lot of intuition either to newcomers or returning customers.

Thus, my goals with this book are twofold: provide a deep and intuitive understanding of the ideas and equations of thermodynamics, and provide the computational and statistical skills to analyze data using the thermodynamic framework. To achieve the first goal, I have developed a presentation that I hope is both intuitive and rigorous. In my experience, many of the great texts in physical chemistry, both at the undergraduate and graduate level (e.g., Hill, McQuarrie, Callan), are better at presenting to those who already understand thermodynamics. The equations are concise, the narrative is logical, and the endpoint is advanced. In reading these texts for the first (and second, and third) time, I learned to do the math, but I did not understand *why* I was doing the math. For example, why does analysis of simple heat engines lead to the most important equation in thermodynamics, the Clausius inequality? What is the statistical basis for spontaneous heat flow from hot to cold? Why does the Lagrange method take us from the rather arcane-seeming ensemble construct to the equilibrium (Boltzmann) distribution? What do the binding polynomials of Wyman and Gill represent, beyond the arithmetic?

To develop this intuition, I have emphasized on illustrations and examples, following the lead of Dill and Bromberg’s groundbreaking book on driving forces in physical chemistry. Using modern three-dimensional plotting software, the results of manipulations using multivariable calculus (such as the Legendre transform and Lagrange method in classical and statistical thermodynamics) can be directly visualized in space, rather than simply as equations. By combining step-by-step pictures of the math, it is hoped that students will come away with the concepts, rather than just learning derivations (such as the Boltzmann distribution) as a memorized progression of equations. At the same time, I have made efforts not to sacrifice the rigor of the equations, but to expand them, especially in areas that are typically left off. Though for some audiences, some of the details can be skipped without losing a working understanding, it is hoped that those with a deep interest in fundamentals will find what they are looking for.

My second major goal with this book is to teach students how to use computers to think about and solve problems in biological and chemical thermodynamics. Most physical chemistry books present equations, and draw plots of equations as con-

tinuous curves. I emphasize to my students at Hopkins that we do not interact with nature by measuring curves and lines, rather, we measure points. That is, our data are discrete. It is what we do with our measurements that tell us about our system, how it behaves, where equilibrium lies, and why. In this book, readers will learn to extract thermodynamic quantities from a broad spectrum of experimental measurements using computers to analyze data, test models, and extract thermodynamic parameters. It is my opinion that students learn more, for example, from fitting a set of p - V measurements for a gas with various models than from memorizing ideal and van der Waals equations of state. As an important additional step, several methods are presented to help students think about and quantify uncertainties in fitted parameters, to avoid the common pitfall of treating fitted quantities as error-free. Another emphasis with computer-aided learning is to use random number generation and simulation to generate statistical distributions and to simulate and analyze the dynamics of molecular systems. This approach greatly complements the equations and logic of statistical thermodynamics. When students simulate the energy distributions of a molecular system and compare these to theoretical distributions from statistical thermodynamics, the theory becomes practice.

To achieve this second goal, the material in this book has been designed to be taught with the aid of a high-level interpreted mathematics program such as Mathematica or MATLAB®. Having taught semester-long biophysical chemistry courses with both programs, I find Mathematica to be easier for students to both learn and use, and as a result, their efforts are directed more toward learning and doing physical chemistry than to debugging code. For this reason, examples in this book are based on Mathematica. Other software packages that could be used in conjunction with this book include python (and its modules numpy, scipy, sympy, matplotlib), especially when run in the Jupyter notebook platform, which has much of the look and functionality of a Mathematica notebook. A third package that could be used here is R, which lacks much of the symbolic math and visualization, but is strong in statistical analysis. A major advantage of both python (and its add-ons) and R is that they are available at no cost.

The first two chapters of this book focus on background and mathematical tools that are used throughout. Chapter 1 focuses on probabilities and statistics, including how probabilities for events combine, how outcomes build up over many events to give discrete distributions of various types (with emphasis on binomial and multinomial distributions), and how these distributions relate to continuous distributions (with emphasis on the Gaussian and exponential distributions). Chapter 2 reviews multivariable calculus, emphasizing partial differentiation, maximization of multivariable functions, exact differentials and their properties, and path integration. Chapter 2 then describes curve fitting using linear and nonlinear least squares, and presents methods (relating to the covariance matrix, bootstrap technique, chi-squared statistics, and the f -test) to estimate errors in fitted parameters, their correlations with each other, and the goodness of fit of different models.

The main part of this book progresses from mostly classical thermodynamic to mostly statistical thermodynamic material. Chapter 3 introduces some fundamental concepts relating to system, surroundings (and the “thermodynamic reservoir”), equilibrium, and reversibility. Next, the first law is presented and discussed, along with analysis of heat and work. Emphasis is placed on gasses to take advantage of their simple equations of state. After discussing reversible changes, irreversible heat transfers and expansions are analyzed for some simple systems. This section on irreversible thermodynamics can easily be skipped by beginning students.

Chapter 4 introduces the second law and entropy. Though emphasis is placed on a classical derivation and analysis, using heat engines to derive the Clausius inequality, a few statistical models are developed in parallel, to show how spontaneous change (and entropy increase) corresponds to increasing the number of configurations

available. One of these models, which describes heat transfer between two subsystems, is used throughout this book, especially when developing statistical thermodynamics in Chapters 9 and 10. As with Chapter 3, entropy changes are calculated for some irreversible transformations, and can be skipped by beginners.

Chapter 5 takes the idea of the entropy as a thermodynamic potential for an isolated system and transforms it into potentials for nonisolated systems. The most important of these is the Gibbs free energy potential at constant temperature and pressure. These transforms are done in a simple approach using differentials, but are also developed in a mathematically more rigorous way using Legendre transforms. This rigor is essential for those seeking a solid understanding of the fundamentals, but again it can be skipped by beginning students. Chapter 5 also introduces the concept of molar (intensive) quantities rather than the extensive quantities discussed in Chapters 3 and 4. Using these molar quantities, the thermodynamics of mixtures (and partial molar quantities) are introduced. Emphasis is placed on the partial molar free energies (chemical potentials) and their relationships (leading to the Gibbs–Duhem equation).

Chapter 6 uses the concept of chemical potential to discuss phase equilibrium in simple systems. Using experimental enthalpy, heat capacity, and volume data for water, a phase diagram is constructed using chemical potentials and their pressure and temperature dependences. Following this exercise, the Gibbs–Duhem equation is applied to different phases of material, and evaluated graphically. This graphical analysis provides a visual picture that leads directly to the Gibbs phase rule and the Clausius–Clapeyron equation.

Chapter 7 explores the concentration dependence of the chemical potential, progressing from the ideal gas to the ideal solution to ideal dilute solutions. Along the way, the concept of the standard state is developed. Using the mole-fraction standard state, thermodynamics of mixing is developed for an ideal solution (following Raoult’s law). Using a lattice mixing model, nonideality is built in, and the mixing thermodynamics of the regular (or mean-field) solution is developed. In addition to showing how complexity can be built into a simple model, the regular solution reveals liquid–liquid phase separation behavior. Finally, the thermodynamics of chemical reactions is developed using chemical potentials on a molar scale. Two approaches are taken. First, a difference approach is used to develop the familiar relationships between reaction free energies, concentrations, and equilibrium constants. Second, a calculus-based approach is developed to show the relation between the reaction free energy, the Gibbs energy, and the position of chemical equilibrium. Finally, the temperature and pressure dependence of chemical reactions are discussed.

Chapter 8 applies the reaction thermodynamics developed in Chapter 7 to a fundamental process in biology: the protein folding reaction. This “reaction” is spectacular in its structural complexity, yet in many cases, it can be modeled as a very simple reaction involving just two thermodynamic states. Emphasis in this chapter is placed on methods for studying the reaction, and on the importance of obtaining data (i.e., “melts”) where there is a measurable equilibrium between folded and unfolded conformations. Thermal and chemical denaturation approaches are described in which equilibrium is monitored with optical probes, and calorimetric methods are also described. Key thermodynamic signatures of the folding reaction are discussed, including large entropy and enthalpy change, and a significant heat capacity change.

Chapters 9 through 11 present a formal development of statistical thermodynamics. In Chapter 9, the motivations for the ensemble method (including the challenges of constructing detailed dynamic models, and the ergodic hypothesis) are discussed. Concepts and variables of ensembles are developed and illustrated for simple models. Using the heat flow model from Chapter 4, an ensemble is developed to model an isolated system (the “microcanonical” ensemble). A method is

presented to determine the equilibrium ensemble distribution in which the number of ensemble configurations is maximized, subject to a constraint of fixed ensemble size, using the Lagrange method. In this derivation, considerable use is made of visuals, showing how constraints interact with configurational variables, and how multivariable calculus provides the equilibrium solution. The resulting distribution defines a microcanonical “partition function,” which is fundamentally related to the entropy—the thermodynamic potential for the isolated system.

Chapter 10 extends the ensemble approach to systems that can exchange thermal energy (but not volume) with their surroundings. The equilibrium position (i.e., the Boltzmann distribution) for this “canonical” ensemble is again found by Lagrange’s method with an additional constraint on the total energy of the ensemble. The resulting canonical partition function is logarithmically proportional to the Helmholtz free energy—again, the thermodynamic potential for the constant volume system. This approach is then extended to systems that can expand and contract, leading to the equilibrium distribution for a system at constant temperature and pressure, and providing a direct connection to the Gibbs free energy potential.

With the fundamental ensembles in place, Chapter 11 makes some key modifications to the statistical thermodynamic approach. First, the ensemble approach is applied to a system with just one molecule, leading to molecular partition functions. We show how, as long as molecules do not interact, these simple molecular partition functions can be combined to give a partition function for systems with many molecules, and how they can also be used to extract thermodynamic quantities. One place where these molecular partition functions are particularly useful is in analyzing chemical reactions. By building an ensemble of reactive molecules, we generate a “reaction” partition function that is used in one form or another in the remaining chapters.

Chapter 12 applies the reaction partition function to the “helix-coil transition” in polypeptides. Unlike the folding of globular proteins described in Chapter 8, the helix-coil transition involves many partly folded states. We develop statistical thermodynamic models of increasing complexity, progressing from a model in which all residues are independent and identical to a model in which residues are not identical, to models where residues are energetically coupled to their nearest-neighbor. The nearest-neighbor or “Ising” approach is a workhorse model for describing cooperativity in chemical and biological systems. In addition to learning about this important type of reaction, this chapter shows how to build statistical thermodynamic models, starting simple and moving toward complex, as is demanded by experimental observations.

Chapters 13 and 14 describe the binding of multiple ligands to macromolecules. To describe such data, we develop a binding partition function referred to as a “binding polynomial.” In Chapter 13, we develop a macroscopic scheme for ligand binding which counts the number of ligands bound but not their arrangement among binding sites. Two different phrasings of binding constants are introduced (stepwise and overall), and their relative advantages and disadvantages are discussed. The relationship between the binding polynomial and various experimental representations of binding data are described, including the fraction of ligand bound, the binding capacity, and the (somewhat outdated but still useful) Hill plot, as are the effects of positive and negative cooperativity on these representations. Chapter 13 concludes by treating the binding of multiple different ligands to a macromolecule, and introduces the concept of thermodynamic linkage.

Chapter 14 analyzes binding from a microscopic perspective. Although the microscopic approach can be more complicated (in terms of parameters and equations), it provides a more mechanistic and more fundamental way of thinking about binding. Again, binding constants are phrased in terms stepwise and overall reactions,

and the relationships between these constants (and the macroscopic constants) are analyzed using basic topology and graph theory. Unlike our approach with the helix-coil transition, we start here with the most complex and general models, and then simplify these models to include macromolecular structure and symmetry. The result is a lean set of models for noncooperative and cooperative ligand binding. We conclude by combining conformational changes (ideas from Chapters 8 and 12) with ligand binding both at the macroscopic and microscopic level, leading to a thermodynamic description of macromolecular allostery.

MATLAB® is a registered trademark of The MathWorks, Inc. For product information, please contact:

The MathWorks, Inc.

3 Apple Hill Drive

Natick, MA 01760-2098 USA Tel: 508 647 7000

Fax: 508-647-7001

E-mail: HYPERLINK "mailto:info@mathworks.com"info@mathworks.com

Web: HYPERLINK "http://www.mathworks.com/"www.mathworks.com

