Thermodynamics and Statistical Mechanics of Macromolecular Systems

The structural mechanics of proteins that fold into functional shapes, polymers that aggregate and form clusters, and organic macromolecules that bind to inorganic matter can be understood only through statistical physics and thermodynamics.

This book reviews the statistical mechanics concepts and tools necessary for the study of structure formation processes in macromolecular systems that are essentially influenced by finite-size and surface effects. Readers are introduced to molecular modeling approaches, advanced Monte Carlo simulation techniques, and systematic statistical analyses of numerical data. Applications to folding, aggregation, and substrate adsorption processes of polymers and proteins are discussed in great detail. Particular emphasis is placed on the reduction of complexity by coarse-grained modeling, which allows for the efficient, systematic investigation of structural phases and transitions.

Providing insight into modern research at this interface between physics, chemistry, biology, and nanotechnology, this book is an excellent reference for graduate students and researchers.

Michael Bachmann is Associate Professor in the Department of Physics and Astronomy at the University of Georgia. His major fields of interest include theoretical physics, computational physics, statistical physics, biophysics, and chemical physics.

Thermodynamics and Statistical Mechanics of Macromolecular Systems

MICHAEL BACHMANN

The University of Georgia

CAMBRIDGE UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781107014473

First published 2014

Printed in the United Kingdom by T. J. International Ltd, Padstow

A catalog record for this publication is available from the British Library

Library of Congress Cataloging-in-Publication Data
Bachmann, Michael
Thermodynamics and statistical mechanics of macromolecular systems / Michael Bachmann.
pages cm
ISBN 978-1-107-01447-3 (hardback)
1. Macromolecules–Thermodynamics. 2. Biomolecules–Structure.
3. DNA–Structure. 4. Statistical mechanics.
I. Title.
QP801.P64B33 2014
547′.7–dc23 2013032171

ISBN 978-1-107-01447-3 Hardback

Dedicated to my family

Contents

Preface and outline

The idea to write this book unfolded when I more and more realized how equally frustrating and fascinating it can be to design research projects in molecular biophysics and chemical physics – frustrating for the sheer amount of inconclusive and contradicting literature, but fascinating for the mechanical precision of the complex interplay of competing interactions on various length scales and constraints in conformational transition processes of biomolecules that lead to functional geometric structures. Proteins as the "workhorses" in any biological system are the most prominent examples of such biomolecules.

The ability of a "large" molecule consisting of hundreds to tens of thousands of atoms to form stable structures spontaneously is typically called "cooperativity." This term is not well defined and could easily be replaced by "emergence" or "synergetics" – notions that have been coined in other research fields for the same mysterious feature of macroscopic ordering effects. There is no doubt that the origin of these net effects is of "microscopic" (or better nanoscopic) quantum nature. By noting this, however, we already encounter the first major problem and the reason why heterogeneous polymers such as proteins have been almost ignored by theoretical scientists for a long time. From a theoretical physicist's point of view, proteins are virtually "no-no's." Composed of tens to thousands of amino acids (already inherently complex chemical groups) linearly lined up, proteins reside in a complex, aqueous environment under thermal conditions. They are too large for a quantum-chemical treatment, but too small and too specific for a classical, macroscopic description. They do not at all fulfill the prerequisites of the thermodynamic limit and do not scale. In consequence, the standard statistical theory of phase transitions is not directly applicable, although many aspects of molecular structure formation processes resemble those known from phase transitions. Since 20 types of amino acids occur frequently in bioproteins, the number of possible compositions is astronomically large, but only of the order of 100 000 highly specific types of bioproteins are functional in the human cell system. Beside this obviously elementary evolutionary aspect, the heterogeneous composition (which causes glass-like behavior) and their high specialization level raise the question, to what extent folding properties can be generic at all. This is actually one of the key questions. A negative answer is not very likely; nature has always proven that even the most complex structures possess symmetries (in a more general context), which explain their stability. Stability is necessary, because these molecular systems exist and function in a thermal environment. It is even appropriate to formulate the whole problem in the following way: it is the interplay and balance between system and environment that stabilizes the structure of the system. Having said that, there is no reasonable way to try to understand any structure formation process without including thermodynamics and, therefore, statistical mechanics.

Another apparent problem is that analytical approaches virtually fail to explain processes of heterogeneous systems, leaving computer simulations the only available tool for theoretical studies. Since protein folding is a relatively slow process (microseconds to seconds), it is almost impossible to use molecular dynamics simulations, operating on nanosecond timescales, for folding studies. Alternatively, Monte Carlo methods are inefficient, if the surrounding water molecules are explicitly simulated. The models are generally not well defined and computer simulations on atomic scales often require large-scale supercomputing resources. The abovementioned key question of generics affects the possibility and limitation of using much more efficient coarse-grained models. For these reasons, studies of biomolecular systems remain a true challenge to theoretical and computational biologists, chemists, and physicists. However, the fact that, among others, neurodegenerative diseases such as Alzheimer's and all virus infections are associated with structural properties of biomolecules makes it worth the efforts to research macromolecular systems of such scale.

This book is a "research book" for the interdisciplinary community. This means it offers many approaches to deal with molecular systems by means of statistical mechanics and computer simulation, yet it will give no precise answers to the above questions. It shall provide young scientists from all affected disciplines of natural and technological sciences with the background to get started, but it also addresses senior scientists by promoting alternative views. The book could also be of value as a compendium as it includes widely accepted research results, in particular for homopolymer systems.

More specifically, we are going to discuss thermodynamic properties of conformational transitions for single- and multiple-chain polymer and protein systems, with particular focus dedicated to molecular folding, aggregation, and adsorption processes to solid substrates. In most of the presented examples, we will investigate the structural transitions by statistical analyses of simplified models. This is based on the idea that in cooperative processes like structural transitions, the collective action of the mechanical degrees of freedom allows for a reduction of the phase space. In other words, the essential features of these transitions are expected to be described qualitatively correctly by models in which a strongly reduced number of effective degrees of freedom is considered only. This reduction of mechanical complexity is called coarse-graining and has proven to be extremely successful in the understanding of complex phenomena and phase transitions of macroscopic systems. If the analogy of structural transitions in rather small molecular systems and phase transitions of macroscopic systems holds true, then coarse-grained approaches can also be valuable tools for the description of molecular behavior. Coarse-grained modeling and simulation will, therefore, play a vital role in this book.

The finiteness of molecular systems, the geometric nature of the structural transitions, and the constraints (e.g., stiff bonds) that affect the mechanical motion render a theoretical treatment typically very difficult. For this reason, the design of efficient algorithms is inevitable for unraveling the properties of structural transitions of molecular systems. We will discuss various examples throughout this book, where sophisticated computer simulation methodologies were employed to obtain the statistical information needed for a thermodynamic analysis of such transitions. Therefore, a short review of modern

simulational methods is also included, as it is considered to be beneficial for readers who wish to get started or who would simply like to know where the results discussed in this book originated from.

In the first chapter of this book, we begin with an introduction of the molecular structure and the modeling of linear macromolecules. Fundamental aspects of thermodynamics and statistical mechanics, with emphasis on finite-size effects and their statistical analysis, are reviewed in Chapter 2. In Chapter 3, properties of the complete sequence and conformation space are systematically analyzed for short lattice proteins by exact enumeration of a minimalistic hydrophobic–polar heteropolymer model. Computer simulations of larger systems require efficient algorithms. Such algorithms are reviewed in Chapter 4 and important aspects of analyses of finitely long time series of data generated by these algorithms are discussed. As a first application, the study of homopolymer freezing and collapse transitions on regular lattices is the subject of Chapter 5. In this regard, the influence of surface and finite-size effects upon crystallization of elastic flexible and semiflexible polymers is addressed in detail in Chapters 6 and 7, respectively. Returning to proteins, characteristic folding properties of proteins and the classification of folding channels are investigated in Chapters 8 and 9. Generic local geometries like secondary structures induced by constraints that effectively reflect many-body effects are discussed in Chapter 10 by introducing tube-like polymers. The extension of coarse-grained modeling to multiple-chain systems is described in Chapter 11, where also analyses of aggregation transitions of short heteropolymers in different statistical ensembles are presented. In Chapter 12, we unravel the hierarchical nature of phase transitions by discussing the exemplified aggregation transition of homopolymers. Pseudophase diagrams of adsorption processes of lattice and off-lattice homopolymers to solid substrates are investigated in detail in Chapter 13. An introductory, simple example for substrate-specific binding of peptides to solid substrates is studied in detail in Chapter 14.

I am indebted to Wolfhard Janke for many years of successful cooperation, advice, and support. For constant support and extremely helpful discussions, I am also thankful to David P. Landau, Kurt Binder, Hagen Kleinert, and Gerhard Gompper. The collaboration with Thomas Neuhaus, Anders Irbäck, Joan Adler, Axel Pelster, and Qianqian Cao is also very much appreciated. It is a particular pleasure to thank my long-term collaborators Thomas Vogel, Stefan Schnabel, Karsten Goede, Monika Möddel, Christoph Junghans, Jonathan Groß, Daniel T. Seaton, and Tristan Bereau for the joint successful work, without which it would not have been possible to write this book. Many other people have also actively and passively helped streamline my thoughts in the exciting field of structural biophysics, which I am also quite thankful for. I would also like to thank Érica de Mello Silva and Paulo H. L. Martins at the Universidade Federal de Mato Grosso, Cuiabá, for their kind hospitality during my current visit to Brazil.

Michael Bachmann
Athens, GA (USA)
June 2013