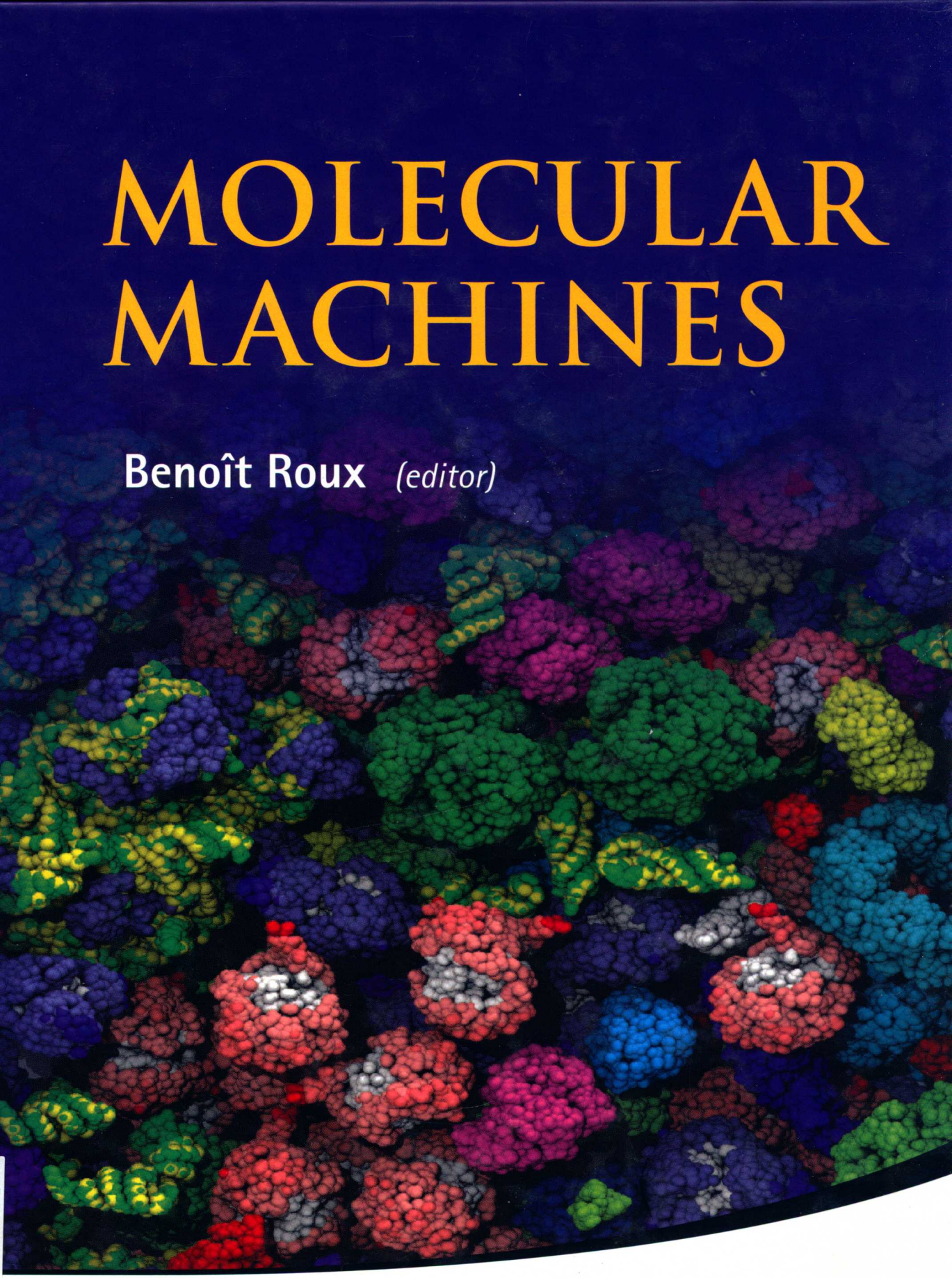
MOLECULAR MACHINES

Benoît Roux *(editor)*



World Scientific

MOLECULAR MACHINES

editor

Benoît Roux

University of Chicago, USA

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Cover photo adapted from McGuffee SR, Elcock AH (2010) Diffusion, Crowding & Protein Stability in a Dynamic Molecular Model of the Bacterial Cytoplasm. *PLoS Comput Biol* **6(3)**: e1000694.doi:10.1371/journal.pcbi.1000694.

MOLECULAR MACHINES

ISBN-13 978-981-4343-44-2
ISBN-10 981-4343-44-7

Typeset by Stallion Press
Email: enquiries@stallionpress.com

Printed in Singapore by Mainland Press Pte Ltd.

MOLECULAR MACHINES

Preface

In everyday language, a machine is generally an intricate device composed of different moving parts that is able to utilize energy to perform some useful work. Complicated machines can also be found in biology where they are complex macromolecular assemblies of proteins, nucleic acids, and carbohydrates that consume energy in order to perform specific functions. The concerted action of all those tiny "molecular machines" underlies all the activities of the living cell. At the present time, we know the structure of several biological macromolecular systems in great atomic detail. Yet this is generally insufficient to fully comprehend how they are able to accomplish specific tasks. For this, it is necessary to identify the different moving parts and understand how they work together. Despite some great progress, our understanding of the function of the biological molecular machines is clearly still in its infancy. Breaking new ground with these difficult problems is likely to require novel paradigms, permitting a seamless integration of structural, dynamical and functional data from experiments and theories. To some extent, what will be needed is not yet completely established. The goal of this volume is to provide an introduction to the world of complex biological molecular machines to a broad audience of students and researchers in the biosciences. Each chapter is written by leading experts to cover the results from cutting-edge research, while remaining broadly accessible. Although it is not possible to offer a comprehensive review of this vast subject, the volume covers a wide range of biological systems from both theoretical and experimental fronts. The volume starts with a review of the structure of bacterial cytoplasm in Chapter 1, thus setting the context for the function of any biological machine. Biological macromolecular systems must perform in the crowded and confusing intracellular environment. Chapter 2 follows by describing how the energy from sunlight is captured and stored by the chromatophore in photosynthethic bacteria. Plans and blueprints are required to know how to make complicated molecules. This information is encoded into the long DNA molecules and Chapter 3 reviews the function of polymerases, which serve to read and copy the DNA. Chapter 4 describes how the genetic information is decoded and translated by the ribosome into an amino acid sequence to form a protein. While the polypeptide chain emerging from the ribosome can sometimes fold spontaneously into a complicated shape, special help is required in some cases. Chapter 5 reviews how chaperonins assist the proper folding of proteins. Chapter 6 reviews the function of myosin and actin, and their role in the mechanism of muscle contraction. Chapter 7 reviews the function of protein kinases, which are part of a critical communication and signaling network inside the cell. The last six chapters are focused on the proteins residing in the cell membrane. Membrane proteins, such as transporters and ion channels (and the chromophore reviewed in Chapter 2), provide great examples of vectorial molecular machines exploiting concentration gradients of ions or the electrical potential difference across the membrane to perform useful work. Chapter 9 presents a comparative review of five membrane transporters, the leucine transporter (LeuT),

the glycerol-3-phosphate transporter (GlpT), the glutamate transporter (GlT), an ATP-binding cassette (ABC) transporter, and the Na^+-coupled galactose transporter (vSGLT), while Chapter 8 is focused on the Na^+/H^+ antiporter, Chapter 10 is focused on the ABC transporter family, and Chapter 11 is focused on Na^+-coupled secondary transporters. Chapter 12 describes the structure and function of the voltage-gated K^+-selective channel, and Chapter 13 provides a broad view of the role of ion channels in cardiac muscle contraction.

Lastly I would like to personally thank Jin-Yun Liang who first suggested the creation of this volume, and Shelley Chow and Jihan Abdat from World Scientific for their great help and guidance during the editorial process.

About the Authors

Chapter 1: Molecular Behavior in Biological Cells: The Bacterial Cytoplasm as a Model System
Adrian H. Elcock[*] and Andrew S. Thomas
Department of Biochemistry
University of Iowa
Iowa City, IA 52242
[*]Corresponding author: adrian-elcock@uiowa.edu

Chapter 2: The Light-Harvesting Apparatus in Purple Photosynthetic Bacteria
Johan Strümpfer,[1,3] Jen Hsin,[1,2] Melih Şener,[1,2] Danielle Chandler,[1,2] and Klaus Schulten[*,1,2,3]
[1]Beckman Institute, Urbana-Champaign
[2]Department of Physics
[3]Center for Biophysics and Computational Biology
University of Illinois at Urbana-Champaign, Urbana, IL 61801
[*]Corresponding author: kschulte@ks.uiuc.edu

Chapter 3: DNA Polymerases: Structure, Function and Modeling
Tamar Schlick[*]
Department of Chemistry and Courant Institute of Mathematical Sciences
New York University
251 Mercer Street, New York, NY 10012
[*]Corresponding author: schlick@nyu.edu

Chapter 4: Information Processing by Nanomachines: Decoding by the Ribosome
Karissa Y. Sanbonmatsu,[*,1] Scott C. Blanchard,[2] and Paul C. Whitford[1]
[1]Theoretical Biology and Biophysics
Theoretical Division, Los Alamos National Laboratory
MS K710, Los Alamos NM 87545
[2]Department of Physiology and Biophysics
Weill Cornell Medical College
1300 York Avenue, New York, NY 10065
[*]Corresponding author: kys@lanl.gov

Chapter 5: Chaperonins: The Machines Which Fold Proteins
Del Lucent,[1] Martin C. Stumpe,[2] and Vijay S. Pande*[,3]
[1]OCE Postdoctoral Fellow
Center for Materials Science and Engineering
343 Royal Parade, Parkville VIC, 3052, Australia
[2]Simbios
Stanford University
Stanford, CA 94305
[3]Department of Chemistry
Stanford University
Stanford, CA 94305
*Corresponding author: pande@stanford.edu

Chapter 6: Muscle and Myosin
Ronald S. Rock, Jr.*
Department of Biochemistry and Molecular Biology
The University of Chicago
929 E. 57th Street, Chicago, IL 60637 USA
*Corresponding author: rrock@uchicago.edu

Chapter 7: Protein Kinases: Phosphorylation Machines
***Elaine E. Thompson,[1] Susan S. Taylor,[1,2,3] and J. Andrew McCammon[1,2,3,4]**
[1]Department of Chemistry and Biochemistry
[2]Department of Pharmacology
[3]Howard Hughes Medical Institute
[4]Center for Theoretical Biological Physics
University of California at San Diego
9500 Gilman Drive, La Jolla, CA 92093, USA
*Corresponding author: elainet@mccammon.ucsd.edu

Chapter 8: Computational Studies of Na^+/H^+ Antiporter: Structure, Dynamics and Function
Assaf Ganoth, Raphael Alhadeff, Isaiah T Arkin*
The Hebrew University of Jerusalem,
Institute of Life Sciences, Department of Biological Chemistry,
Edmund J Safra Campus, Givat-Ram, Jerusalem, 91904, Israel
*Corresponding author: arkin@savion.huji.ac.il

Chapter 9: Membrane Transporters — Molecular Machines Coupling Cellular Energy to Vectorial Transport Across the Membrane
Zhijian Huang,[1,2] Saher A. Shaikh,[1,2] Po-Chao Wen,[1,2,3] Giray Enkavi,[1,2,3] Jing Li,[1,2,3] and Emad Tajkhorshid[*,1,2,3]
[1]Department of Biochemistry
[2]Beckman Institute
[3]Center for Biophysics and Computational Biology,
University of Illinois at Urbana-Champaign, Urbana, IL 61801
[*]Corresponding author: emad@life.illinois.edu

Chapter 10: ABC Transporters
E. P. Coll and D.P. Tieleman[*]
[1]Dept. of Biological Sciences, BIOL 415
University of Calgary
2500 University Dr. NW, Calgary, Alberta, Canada, T2N 1N4
[*]Corresponding author: tieleman@ucalgary.ca

Chapter 11: Sodium-coupled Secondary Transporters: Insights from Structure-based Computations
Elia Zomot, Ahmet Bakan, Indira H. Shrivastava, Jason DeChancie, Timothy R. Lezon and Ivet Bahar[*]
Department of Computational & Systems Biology
School of Medicine, University of Pittsburgh
3064 BST3, 3501 Fifth Ave, Pittsburgh, PA 15213
[*]Corresponding author: bahar@pitt.edu

Chapter 12: Voltage-gated Ion Channels: The Machines Responsible for the Nerve Impulse
Benoit Roux[*] and Francisco Bezanilla
Department of Biochemistry and Molecular Biology
The University of Chicago
929 E. 57th Street, Chicago, IL 60637, USA
[*]Corresponding author: roux@uchicago.edu

Chapter 13: Voltage-gated Channels and the Heart
Jonathan R. Silva[*,1] and Yoram Rudy[2]
[1]Department of Cell Biology and Physiology
Washington University School of Medicine
[2]Washington University in St. Louis
The Fred Saigh Distinguished Professor of Engineering,
Professor of Biomedical Engineering, Medicine, Cell Biology & Physiology, Radiology, and Pediatrics

Director, Cardiac Bioelectricity and Arrhythmia Center (CBAC)
Washington University in St. Louis
Cardiac Bioelectricity Center
290 Whitaker Hall, Campus Box 1097, One Brookings Drive, St. Louis, MO 63130-4899, USA
*Corresponding author: jonsilva@gmail.com

Contents

Molecular Behavior in Biological Cells

1

The Bacterial Cytoplasm as a Model System

Adrian H. Elcock and Andrew S. Thomas

---

1. Introduction

Given a choice, most, if not all of the molecular machines covered in this book would probably prefer to perform their functions in their intended physiological environment, rather than in a test-tube. This is in contrast to the wishes of most experimental and computational biophysicists, however, for whom there are compelling reasons for choosing to study molecules *in vitro*, not least of which is the fact that reconstituting interesting behavior with purified components (proteins, DNA etc.) can unambiguously establish the identities of the factors responsible for the observed behavior. But the *in vitro* environment, while being very well suited to providing clean, unequivocal insights, is in many respects very different from the actual environment encountered by most molecular machines, and this is especially true for those that function *inside* biological cells.[1] First, there is clearly a vast increase in environmental complexity: a protein *in vitro* may only be accompanied by a few other types of molecules (water molecules, dissolved salts, and buffer molecules that help to maintain pH), but the same protein inside, for example, the cytoplasm of a bacterial cell may encounter hundreds to thousands of different kinds of protein molecules, RNA molecules, metabolites (glucose, ATP etc.), lipid molecules, and even the bacterial chromosome (DNA). A number of these molecules may be required partners for the protein, e.g. cofactors or substrates of enzymes, or other proteins with which it may form important functional complexes. But the vast majority of them will not be, and each protein must therefore be able to find its intended partners from a very wide array of unsuitable alternatives.

Not only is the compositional complexity vastly greater *in vivo*, but so is the total concentration of macromolecules. The macromolecular concentration inside a human red blood cell (erythrocyte), for example, is around 350 mg/mL,[2,3] while that in the cytoplasm of the bacterium *Escherichia coli* is around 300 mg/mL, of which ~220 mg/mL is protein.[4] For comparison, the protein concentration inside typical crystals (of the type used to solve structures by X-ray crystallography) is of the order of 150 mg/mL; it is therefore interesting that the cellular interior can remain sufficiently like a fluid that macromolecules can still diffuse (albeit somewhat slowly) from one end of the cell to another. This is especially so given that most proteins, when concentrated to

such high levels will "crash out", i.e. precipitate from solution. As we shall see, this very high concentration of "stuff" can, in principle, have consequences not only for the kinetics of macromolecular processes but also for their thermodynamics.

2. A Simple Analogy for Molecular Behavior *in vivo*

This chapter will be concerned primarily with outlining the ways in which molecular behavior *in vivo* might be different from that observed *in vitro*. As a starting point, we will illustrate potential differences by examining a concrete example: the diffusional behavior of the protein CheY from *E. coli*; most of the same basic principles should apply, however, to more or less any protein. CheY is a small 128-residue (14 kDa) protein that shuttles back and forth between chemoreceptors (which sense chemical signals present in the exterior environment) and the flagellar motor (which turns the bacterium's flagellum either counterclockwise, driving the bacterium forward, or clockwise, causing it to tumble).[5]

How are we to think about the diffusional aspect of CheY's function at a molecular level? To answer this question we will try to construct a real world analogy starting from the moment that the protein becomes phosphorylated and released by the histidine kinase CheA (which is attached to the membrane-bound chemoreceptors) to the moment that it modulates the organism's swimming behavior by directly binding to the flagellar motor control protein, FliM. One possibility is to think of CheY as a courier, bearing a message that is important, perhaps vital, to the survival of the nation (i.e. organism). We will make life somewhat simpler for ourselves by restricting ourselves to a 2D setting: we will consider the cell as analogous to a large airport terminal. CheY's job, therefore, is simply to carry its message (phosphate group) from one end of the terminal to the other, where another person — more important than our courier — awaits to receive the message.

Before we even consider adding the rest of the intracellular milieu, it is useful to imagine how things might work if, aside from the chemoreceptors and the flagellar motor, an *E. coli* cell were to be filled only with water. In terms of our analogy, this would correspond to a situation in which the entire terminal is deserted. We might naively imagine that passing the message to the receiver requires only that the courier walk the length of the terminal. But since our courier and recipient represent molecules, they have neither eyes, ears, noses nor any other senses that might enable them to identify each other at long distances. Furthermore, it is important to remember that since they also do not have brains, they have no conception of what it is that they are "supposed" to do: the courier in our scenario, for example, doesn't know that he/she is supposed to deliver a message anymore than the recipient knows that he/she is supposed to receive a message. Because of these issues, it is more realistic to assume that our courier will not recognize our recipient until they literally bump — face to face — into each other. We can mimic this in our real world model by applying a blind-fold to *both* the courier and the recipient, and by assuming that they have no other way of communicating with one another other than by a sense of touch.

It is also important for us to recognize that our courier will not simply move across the platform in a straight line: molecules in aqueous solution do not move as if shot from a cannon (i.e. ballistically); instead, on the kinds of timescales and lengthscales that we are interested in here (nanoseconds and nanometers and beyond), they move *diffusively*, that is, they are subject to repeated collisions with solvent molecules that continually reorient and redirect them, leading

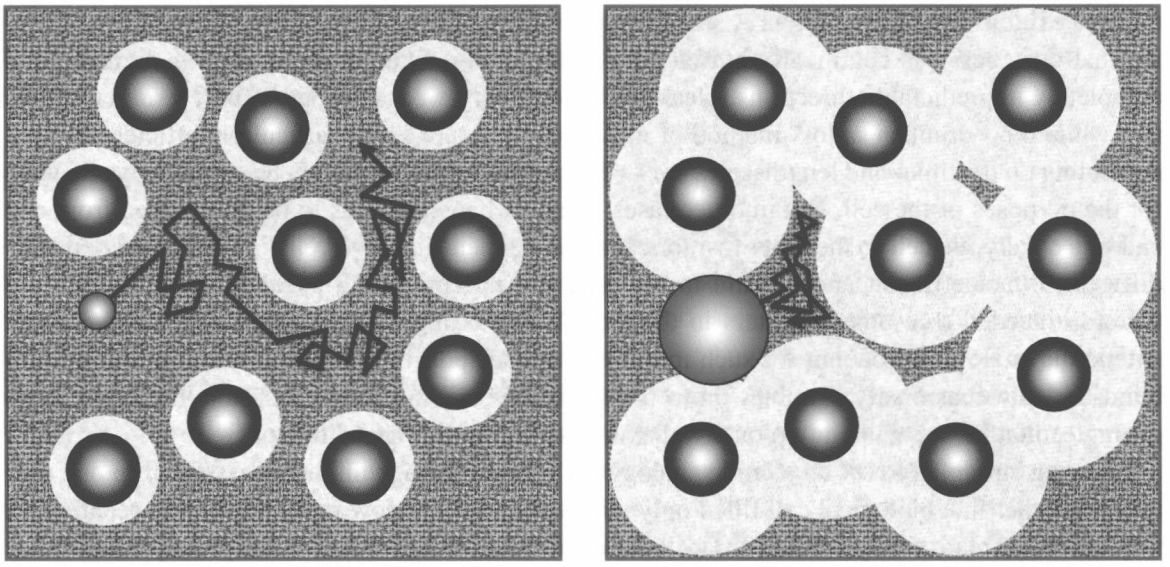
them to exhibit Brownian motion. A simple way to add this to our real world situation is to imagine that our courier is continually buffeted by sudden, short-lived gusts of wind that come from completely unpredictable directions (clearly, it's not a very well designed airport terminal). Note that while this "drunken sailor" method of movement may not seem particularly efficient, it may, depending on the time- and lengthscales over which movement is required, be completely adequate for the purposes of the cell. But in other cases — e.g. when cargo has to be transported from one end of a eukaryotic cell to the other (as, for example, happens in nerve cells in the human body) — diffusion is not efficient enough and motor proteins, moving along pseudo-static "tracks", are therefore needed (see other chapters in this book). In passing, it is worth noting that we could extend the analogy to account for such motor-driven transport by imagining that the courier, still blind-folded, occasionally stumbles on to the terminal's moving walkways; note that there is no guarantee that he/she will end up on a walkway moving in the right direction!

Having struggled to set up a crude analogy for how a message might be passed from one protein to another in a biological cell filled only with water, we can now rerun the entire scenario in a way that more realistically mimics intracellular conditions. As noted above, the principal difference is that there are very high concentrations of other macromolecules present; we can account for this by imagining that the platform is now extremely crowded with other people, some the size of small children, some the size of Shrek, but *all* of them blind-folded. Clearly, the situation will be chaotic: our courier, who was already having a difficult enough time navigating the deserted terminal is now also repeatedly jostled and blocked by the crowd. We can probably imagine that this will slow the courier considerably although the extent to which he/she is slowed may perhaps depend on his/her size: small children, for example, are often quite adept at finding their ways through crowded rooms, whereas we could imagine that Shrek might have a great deal more difficulty. A schematic illustration of these differences is shown in Figure 1. But we can also imagine a more complicated situation. What if the courier encounters a member of the crowd who happens to share some of the physical features of the message's recipient? In such cases, we must expect that the courier will tend to remain close to the "impostor" for some period of time (perhaps a very long time), with the result that his/her diffusive exploration of the environment might be significantly slowed. In molecular terms, such a situation could occur if CheY encounters macromolecules that have similar electrostatic potentials or similar patches of exposed hydrophobic residues, to those of the flagellar motor protein FliM. In passing, therefore, we can speculate that evolution might have acted to decrease the extent to which such non-specific (and unintended) interactions are likely to occur.

It is hopefully apparent from the above description that we *can* construct a real world analogy for the situation encountered by CheY in the *E. coli* cytoplasm, but we have to be prepared to alter it in some fundamental ways to make it work. While it is probably useful at an illustrative level — since it allows us to begin to frame our thinking about what life might be like inside a cell — it should not be thought that it describes all of the possible differences between the *in vitro* and *in vivo* environments (for example, Section 3).

3. Macromolecular Crowding Effects

One of the more subtle effects of the highly crowded environment faced by molecules *in vivo* can only be revealed by using some ideas from statistical thermodynamics.[6] This is the effect of



(a) (b)

Figure 1. A schematic illustration of how diffusion might be affected by highly crowded conditions for (a) a small protein, and (b) a large protein. The diffusing protein of interest is shown as a red circle, the other macromolecules comprising the crowded environment are shown as blue circles. The light blue "haloes" around each crowder molecule indicate the volume that is inaccessible to the diffusing protein due to steric interactions: note that the excluded volume experienced by the large protein (b) is considerably larger than that experienced by the small protein (a). The black lines indicate potential diffusional trajectories for the two proteins.

steric ("excluded-volume") interactions between a protein of interest (again, e.g. CheY) and its environment of surrounding "crowder molecules". As we will see, such interactions can, in principle, have very significant thermodynamic effects, and this realization has led, in recent years, to the development of an entire sub-field of biophysics devoted to understanding and predicting effects due to "macromolecular crowding".[7,8] An illustrative example of the crowding effect is shown in Figure 2 in which we consider an idealized protein folding equilibrium. In what follows, we will make a number of simplifications in order to ease our calculations, but it is important to bear in mind that the basic argument and conclusions are unaffected by these simplifications.

We consider a simplified "protein" that can exist in one of only two conformations: one, a highly compact native state, and another, an extended unfolded state. These are shown in Figures 2a and 2b respectively. As shown in these figures, we will assume that the protein exists within the confines of a 2D box consisting of 12×12 squares. We will further assume that the energies of both protein conformations are equal. This assumption makes our life considerably easier since, in the language of statistical thermodynamics,[6] it means that the two conformations have identical "Boltzmann weights"; all we have to do to estimate the equilibrium constant for the folding reaction, therefore, is take the ratio of the number of possible conformations of the folded protein and the unfolded protein. In the present case, therefore, $K_{eq} = 1/1 = 1$, which in turn means that the free energy of folding of the protein, which is given by $\Delta G = -RT \ln K_{eq} = -RT \ln (1) = 0$.

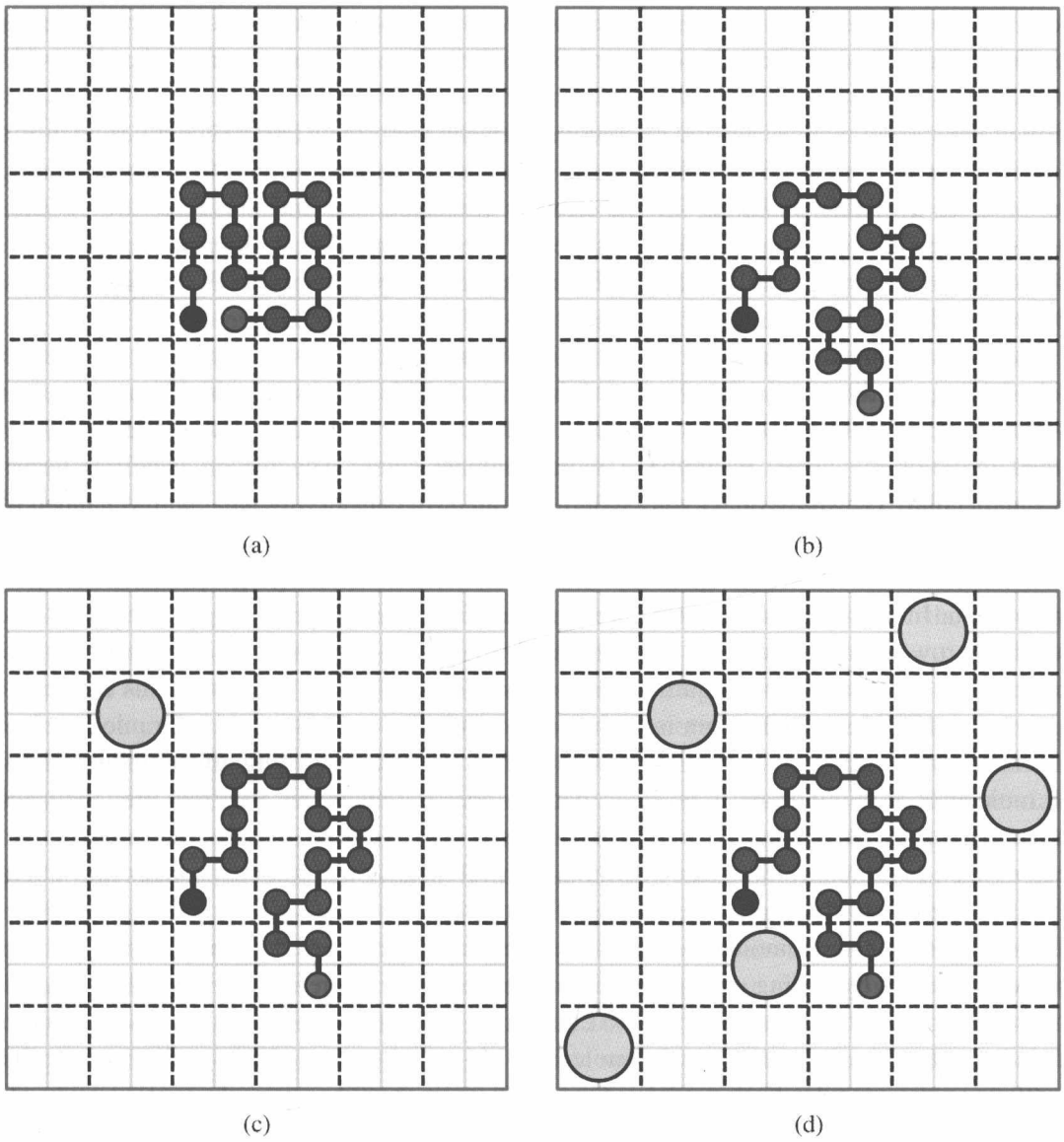


Figure 2. A simple 2D model of crowding effects on a protein Folding reaction. (a) Folded state structure of a model protein consisting of 16 "amino acids" (circles) connected by peptide bonds (thick lines); the blue and red circles represent, respectively, the N- and C- termini of the protein. (b) The same protein in a putative unfolded conformation. (c) One possible configuration of the same system with a single added crowder molecule (large grey circle); note that for simplicity we assume that the crowder can only be placed within one of the large dashed squares, of which there are 36 in total. (d) One possible configuration of the same system with five added crowder molecules.

Next we consider what happens when we add an idealized large macromolecule to the system (represented by the large filled circle in Figure 2c). We will assume, again just for the purposes of making the calculations simpler, that the "crowder molecule" can only occupy the larger, dashed squares on the grid (i.e. in the absence of the protein, it would be restricted to 6×6 possible positions).