BIOMEDICAL ENGINEERING

Real-Time Biomolecular Simulations

The Behavior of Biological Macromolecules from a Cellular Systems Perspective

MICHAEL PETERS

Q7
P482

Real-Time Biomolecular Simulations

The Behavior of Biological Macromolecules from a Cellular Systems Perspective

Michael H. Peters, Ph.D.

New York Chicago San Francisco
Lisbon London Madrid Mexico City
Milan New Delhi San Juan
Seoul Singapore Sydney Toronto

The McGraw-Hill Companies

McGraw-Hill books are available at special quantity discounts to use as premiums and sales promotions, or for use in corporate training programs. For more information, please write to the Director of Special Sales, Professional Publishing, McGraw-Hill, Two Penn Plaza, New York, NY 10121-2298. Or contact your local bookstore.

Copyright ©2007 by The McGraw-Hill Companies. All rights reserved. Printed in the United States of America. Except as permitted under the Copyright Act of 1976, no part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written permission of publisher, with the exception that the program listings may be entered, stored, and executed in a computer system, but they may not be reproduced for publication.

1234567890 DOC DOC 019876

ISBN-13: 978-0-07-146071-2
ISBN-10: 0-07-146071-3

Sponsoring Editor	Kenneth P. McCombs
Editorial Supervisor	Janet Walden
Project Manager	Vasundhara Sawhney
Acquisitions Coordinator	Laura Hahn
Copy Editor	Beatrice Sussman
Proofreader	Ojen Yumnam
Indexer	Robert Swanson
Production Supervisor	James Kussow
Composition	International Typesetting and Composition
Illustration	International Typesetting and Composition
Art Director, Cover	Handel Low

Information has been obtained by McGraw-Hill from sources believed to be reliable. However, because of the possibility of human or mechanical error by our sources, McGraw-Hill, or others, McGraw-Hill does not guarantee the accuracy, adequacy, or completeness of any information and is not responsible for any errors or omissions or the results obtained from the use of such information.

About the Author

Michael H. Peters, Ph.D., is a teacher/researcher in biomedical engineering and molecular biophysics. His particular focus is on the computational molecular biophysics of cell signal molecules (ligands) and their interactions with specific cellular targets, and clinical applications of ligands-receptor systems. The author of a textbook and over 30 journal articles, he is a frequent speaker at conferences and universities throughout the United States.

Preface

Living cells and tissues are highly dynamic entities that carry out complex interconnected processes and functions that depend on both space and time. Dynamic analysis is required to connect cellular structure to function and to predict cellular behavior under different "environmental" conditions and stresses. Furthermore, many important cellular processes, such as chemotaxis, feed-back control loops, and clock reactions, are well-known to be non-linear dynamic events that further complicate analytical descriptions.

It is clear that a quantitative analysis of cellular behavior that connects structure to function requires a systematic analysis of the particular unit processes and their interactions in the living cell. This book was written in order to lay out the basic systems engineering approach to quantitatively describe the dynamic cell. An introduction to the essential biochemistry in living systems is given in Chapter 1, followed by a description of the particular engineering systems of the living cell in Chapter 2. These systems are

- Cellular communications and transport
- Biochemical production
- Motility, mechanical work, and structural support
- Reproduction or division
- Information systems, checkpointing, and reliability

The majority of cellular processes are carried out by so-called macromolecules, such as peptides and proteins, lipids, carbohydrates, and nucleotides and oligonucleotides. Although this may appear to complicate the description of the living cell, in some ways it simplifies things. In short, the overall dynamical

此为试读,需要完整PDF请访问: www.ertongbook.com

description of cellular processes involves large spatial and time scale separations between the behavior of the cellular solvent, i.e., water, dissolved ions, and small molecules, and that of the "workhorse" biological macromolecules. Chapters 3 and 4 develop the so-called multiple time scale methods that are used to exploit this disparity and simplify real-time macromolecular dynamics.

Finally, the last chapter of this book demonstrates the integration of system pathways to describe a cascade of cellular events that may lead to a particular phenotype. Although a full characterization of cellular processes along such rigorous lines is still a formidable goal, the last chapter explores some simplified examples of interconnected cellular system processes in order to illustrate the essential features of a rigorous approach in a revealing, analytical way.

In the future, large-scale computations will allow for the addition of more pathways and interconnectivity between systems in order to approach a more meaningful or useful total simulation of the dynamic living cell.

Acknowledgments

Thanks to Mr. Jack Schreiber at Virginia Commonwealth University for proofreading the final copy chapters of this book. Special thanks to my wife, Karen, and daughter, Emily, for letting me work a bit on weekends.

Contents

Chapter One

The World of Biological Macromolecules

The complexities associated with living systems are as far-reaching and as lacking understanding today as any field of science. Moreover, no single subject is as critical to our existence, including our interactions with our environment, as the dynamics and functionality of living systems.

Structure of Biological Molecules

For the most part, living systems can be broken down into the description of the molecules comprising them and, in particular, the dynamics, functionalities, and interactions of those molecules. Many molecules comprising living systems are actually large groups of smaller building-block molecules. For example,

peptides and proteins represent large groups of smaller building-block molecules called *amino acids,* and starches and carbohydrates (the polysaccharides) represent groups of molecules called *simple sugars.*

Molecules representing large groups of smaller subunits often are referred to as macromolecules. Some macromolecules can be relatively large, such as proteins that can have molecular weights averaging over 30,000 as compared to, say, water with a molecular weight of only 18 Daltons.

Biological macromolecules represent the "molecular workhorses" of the living cell. They carry out communication as signaling molecules and produce other biochemical molecules as reaction catalysts, provide mechanical structure and motility, store information via their particular molecular arrangements, direct reproduction, and function as cellular "police" eager to arrest and dispose of aberrant molecules. Their world is our world, and the totality of their behavior and environmental interaction manifests into who we are.

Our story begins deceptively simple. In general, there are just four basic classifications of biological molecules:

- amino acids
- carbohydrates
- nucleotides
- lipids and fats

These molecules, however, can have quite a number of molecular variations and combinations, making the world of biological molecules extremely complex. For example, there are over 20 different kinds of biological amino acids, and groups of amino acids often are combined with particular groups of carbohydrates, such as in cell membrane molecules called *glycoproteins* (Figure 1-1).

Each of the first three types of biological molecules can self-combine through covalent bonding to form bioploymers. For example, amino acids combine through peptide bonds to form proteins, carbohydrates combine to form polysaccharides, and nucleotides combine to form nucleic acids.

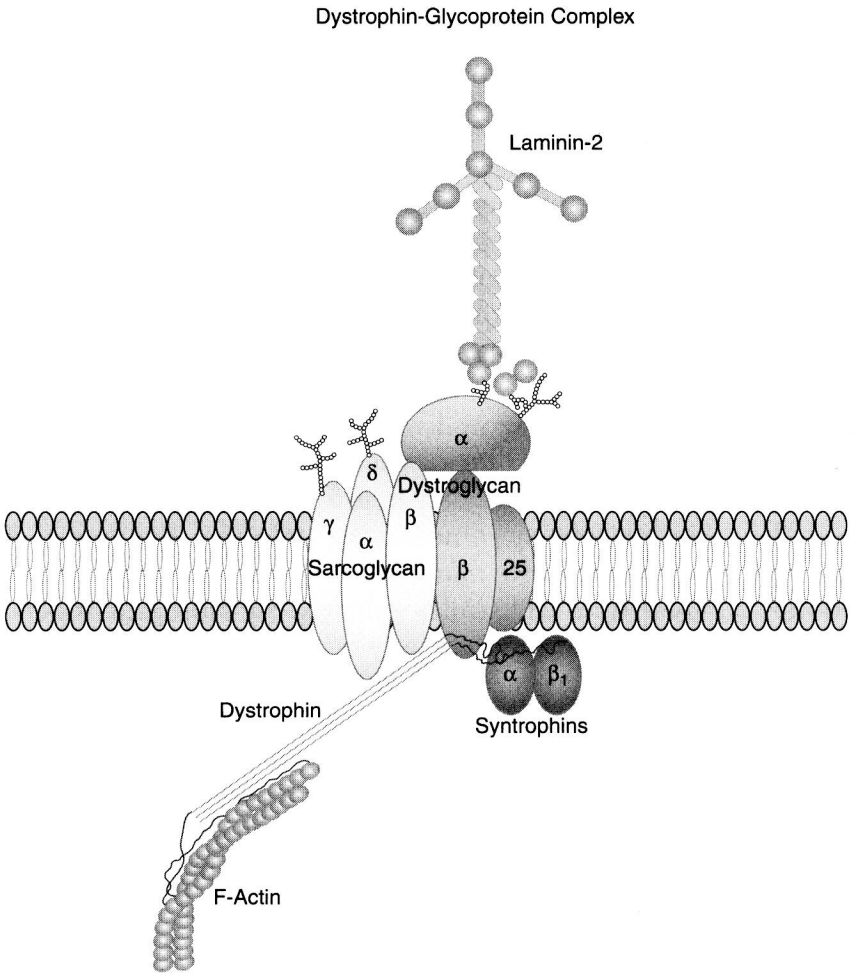


FIGURE 1-1 Dystrophin-glycoprotein complex found in the plasma membrane (sarcolemma) of muscle fibers. The disease Duchenne muscular dystrophy is an X-linked genetic disorder involving the production of the protein dystrophin. Image from Kevin Campbell at the University of Iowa, by permission.

The fourth classification of biological molecules, called *lipids* and *fats,* can undergo self-association, a description of noncovalent interactions, leading to particular "superstructures." For example, cell membranes are composed of associating lipid molecules in a type of structure called a *bilayer.*

Biological molecules have a variety of cellular functions. The biopolymers of amino acids, the peptides and proteins, function as enzymes, cell signal molecules, cellular support structures, and as a cellular energy source. The biopolymers of carbohydrates, the polysaccharides, store energy, function as support structures, and sometimes encode information through particular molecular structure arrangements. Nucleic acids are the primary molecules of information encoding in cells, but they also can act as enzymes and support units. Lipids function as both support structures (membranes) and are an important cellular energy source.

To predict the function and role of biological molecules requires a detailed description of their structure and dynamics. Here we briefly describe the structure of biological molecules and their associated macromolecules, and in Chapters 3, 4, and 5 we address their dynamics. Our focus throughout this book is on "real-time" dynamics—that is, the biological macromolecule dynamics that are actually occurring in the living cell. As we will see, cellular processes represent a well-orchestrated spatial and temporal sequencing of highly specific biomacromolecular interactions in order to carry out their various functions. These interactions are dynamic events that take place over a variety of time scales and length scales. Quantitatively, these interactions often are highly nonlinear, which manifests into a spectrum of behaviors, such as the so-called clock reactions of a cell cycle.

We will also find it useful to adopt a systems perspective of cellular processes. In Chapter 2, we identify five important subsystems of cellular behavior in order to organize the overall functions of the living cell. Within each subsystem, we show how biomolecular interactions dictate the behavior and performance of that particular subsystem. In Chapters 3 and 4, we demonstrate the physical and mathematical approaches necessary to describe the real-time dynamics of these interactions.

In Chapter 5, we demonstrate how the integration of subsystem dynamics allows for a comprehensive analysis of total cellular behavior. Although such a comprehensive analysis is far beyond the reach of current computational simulation

此为试读,需要完整PDF请访问: www.ertongbook.com

capabilities, the necessary approach via the connection of the cellular subsystem dynamics is nonetheless a promising route to the simulation of the living cell.

Chapter 1 provides a minimally necessary description of biological molecular structure for preparation to subsequent discussion and presentation of biomolecular dynamics, which is the central theme of this book. More detailed presentations of biological molecular structure can be found in texts on biochemistry.[1]

Amino Acids, Peptides, and Proteins

The basic molecular formula for amino acids is shown in Figure 1-2. The central carbon atom is called the alpha carbon, and the R group, called the side group, varies depending on the specific type of amino acid. Because of the central carbon tetrahedron, amino acids will exhibit stereo isomerism, resulting in the mirror image forms noted in Figure 1-2. The L-form is by far the most common in nature, but interestingly the D-form does make a rare appearance in some living systems.

Nature has provided us with 20 different forms of amino acids based on the specific side group. The side groups can be charged, polar or nonpolar, leading to a variety of different functions and interactions. The 20 different amino acids are listed in Table 1-1, according to the side group specification.

Amino acids can be covalently linked with one another through so-called peptide bonds, as shown in Figure 1-3, which can result in relatively low-molecular-weight macromolecules with simple linear geometric structures called *peptides* and higher-molecular-weight, more complex geometric structures called *proteins*. In protein synthesis ("translation"), discussed in more detail in Chapter 3, peptide bond formation is catalyzed by the enzyme peptidyl transferase, resulting in rapid production of peptides and proteins by cells.

---

[1] See, e.g., Pratt and Cornely (2004), and Voet et al. (2006).

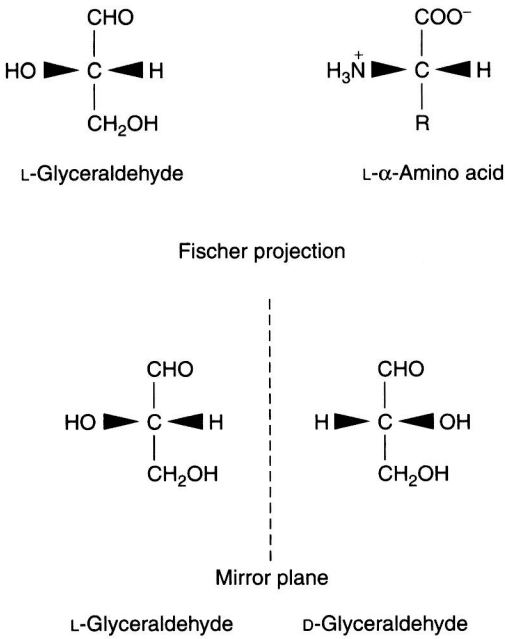
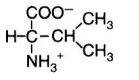
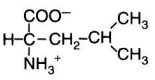
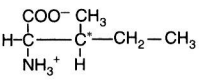
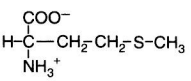
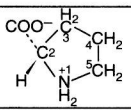
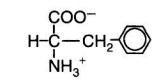
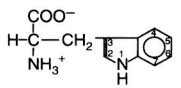


FIGURE 1-2 The basic structure of L-amino acids and comparison with another similar natural molecule, L-glyceraldehyde. The L- and D-forms are mirror images, or stereo isomers, of one another. The D-amino acids rarely are found in nature.

Protein structure is broadly divided into primary, secondary, tertiary, and quaternary classifications. *Primary structure* refers to the particular amino acid sequence in a continuous chain, such as LHLWR representing the chain sequence leucine-histidine-leucine-tryptophan-arginine. *Secondary structure* refers to the most elementary geometric structural patterns that include alpha helixes and beta sheets (see Figure 1-4). *Tertiary protein* structure is the particular geometric arrangement of secondary structures and represents the natural, or folded, state of proteins, as shown in Figure 1-5.

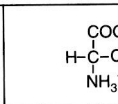
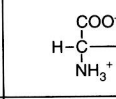
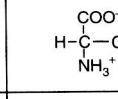
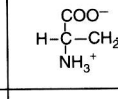
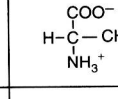
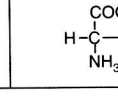
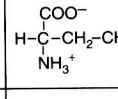
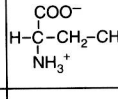
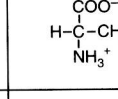
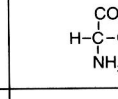
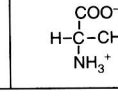
The atomic force interactions of secondary structural units are generally noncovalent with the exception of sulfide bridges that may arise from sulfur containing amino acid side groups. Tertiary structure often is stabilized by protein atom interactions with water. *Quaternary structure* applies only to proteins containing multiple amino acid chains or multiple protein subunits, as exemplified by hemoglobin shown in Figure 1-6. Quaternary

Table 1-1 The 20 naturally occurring amino acids classified according to their particular side chains. The mass is given in daltons. The pK of the side chain is important in establishing its ionization state in solution. For pKs less than the pH (~7) of the solution, the hydronium ions do not remain attached to their associated ionizable group (nonprotonated form). Histidine, with its pK nearly the same as the physiological pH, has an approximately equal molar amount of its protonated and nonprotonated side chains.

Name, Three-Letter Symbol and One-Letter Symbol	Structural Formula	Residue Mass (D)	Average Occurrence in Proteins (%)	pK_1 α-COOH	pK_2 α-NH_3	pK_R Side Chain
Amino acids with nonpolar side chains						
Glycine Gly G		57.0	7.2	2.35	9.78	
Alanine Ala A		71.1	7.8	2.35	9.87	
Valine Val V		99.1	6.6	2.29	9.74	
Leucine Leu L		113.2	9.1	2.33	9.74	
Isoleucine Ile I		113.2	5.3	2.32	9.76	
Methionine Met M		131.2	2.2	2.13	9.28	
Proline Pro P		97.1	5.2	1.95	10.64	
Phenylalanine Phe F		147.2	3.9	2.20	9.31	
Tryptophan Trp W		186.2	1.4	2.46	9.41	

*Data from Dawson, R.M.C., Elliott, D.C., Elliott, W.H., and Jones, K.M., *Data for Biochemical Research*, (3d ed.), pp. 1–31, Oxford Science Publications NY (1986).

Table 1-1 (*Continued*)

Name, Three-Letter Symbol and One-Letter Symbol	Structural Formula	Residue Mass (D)	Average Occurrence in Proteins (%)	pK_1 α-COOH	pK_2 α-NH_3	pK_R Side Chain
Amino acids with uncharged polar side chains						
Serine Ser S		87.1	6.8	2.19	9.21	
Threonine Thr T		101.1	5.9	2.09	9.10	
Asparagine Asn N		114.1	4.3	2.14	8.72	
Glutamine Gln Q		128.1	4.3	2.17	9.13	
Tyrosine Tyr Y		163.2	3.2	2.20	9.21	10.46 (phenol)
Cysteine Cys C		103.1	1.9	1.92	10.70	8.37 (sulfhydryl)
Amino acids with charged polar side chains						
Lysine Lys K		128.2	5.9	2.16	9.06	10.54 (ε-NH_3^+)
Arginine Arg R		156.2	5.1	1.82	8.99	12.48 (guanidino)
Histidine His H		137.1	2.3	1.80	9.33	6.04 (imidazole)
Aspartic acid Asp D		115.1	5.3	1.99	9.90	3.90 (β-COOH)
Glutamic acid Glu E		129.1	6.3	2.10	9.47	4.07 (γ-COOH)

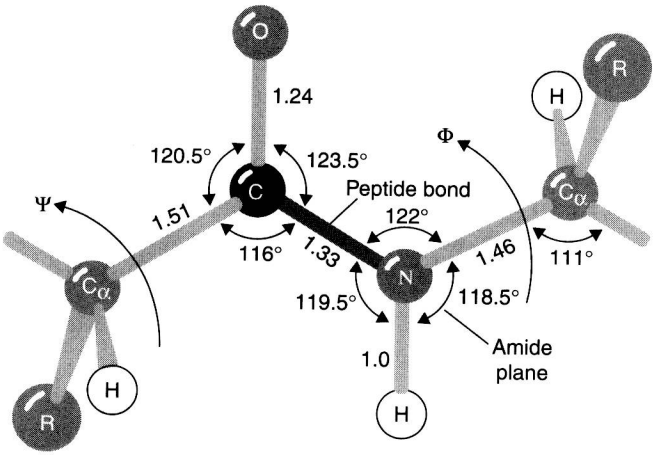


FIGURE 1-3 Illustration of the peptide bond that links amino acids (residues) together. The primary structure of proteins is defined as simply the order of the particular amino acids in a polypeptide chain, beginning at the N-terminal residue. The peptide bond lies in a rigid plane. Also shown are the torsional or dihedral angles around the C_α—N (Φ) and the C_α—C bond (Ψ), which give rise to a multitude of possible conformational states of proteins.

structure helps add geometrical macromolecular symmetry to highly specific protein subunits that generally lack it. The symmetry may help reduce orientational entropic barriers to protein function and still maintain a high degree of specificity via an intact tertiary structure. The natural drawback is the requirement of larger molecules and additional processing steps.

Active proteins are water soluble and typically have charged and polar amino acids on their exterior and more hydrophobic amino acids in their interior. These proteins often are referred to as *globular proteins* (L., *globus,* meaning globe-shaped) because of their compact spherical shape. Proteins can be unfolded or denatured using soaps that disrupt their water soluble structure. Many proteins, such as cell membrane spanning proteins, have both hydrophilic and hydrophobic exterior regions. The hydrophobic portions are found in the membrane region, and the hydrophilic parts are either within the cell (cytosol) or in the exterior fluid (extracellular region).

The wide variety of conformational states of proteins emanates from several fundamental molecular flexibilities

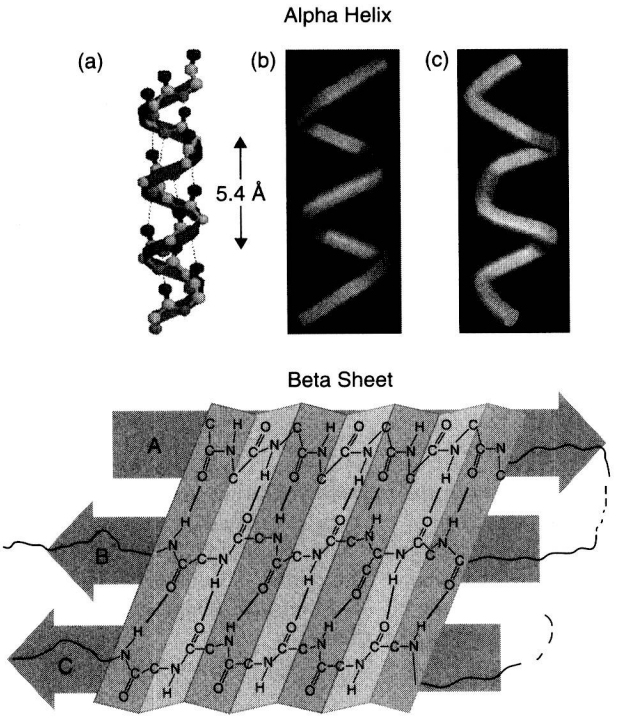


FIGURE 1-4 The two most common secondary structural forms found within polypeptide chains. Alpha helixes can be right-handed or left-handed, and beta sheets can be either parallel or antiparallel.

associated with peptide bonds. As shown in Figure 1-3, peptide bonding between two amino acids results in a rigid plane that is bonded on its vertices by the two associated alphacarbons, the carbonyl oxygen, and the amide hydrogen. The peptide bond (C—N) lies in the center of the rigid plane. Significant molecular rotation, however, around both the N—C_α bond and the C_α—C bond can take place. The angles associated with these rotations, ϕ and ψ, respectively, are known as either *dihedral* or *torsional angles,* and, as originally shown by Ramachandran, they assume particular, biased values for different secondary structures as shown in Figure 1-7.[2]

[2] See Voet et al. (2006), for more discussion.

此为试读,需要完整PDF请访问: www.ertongbook.com