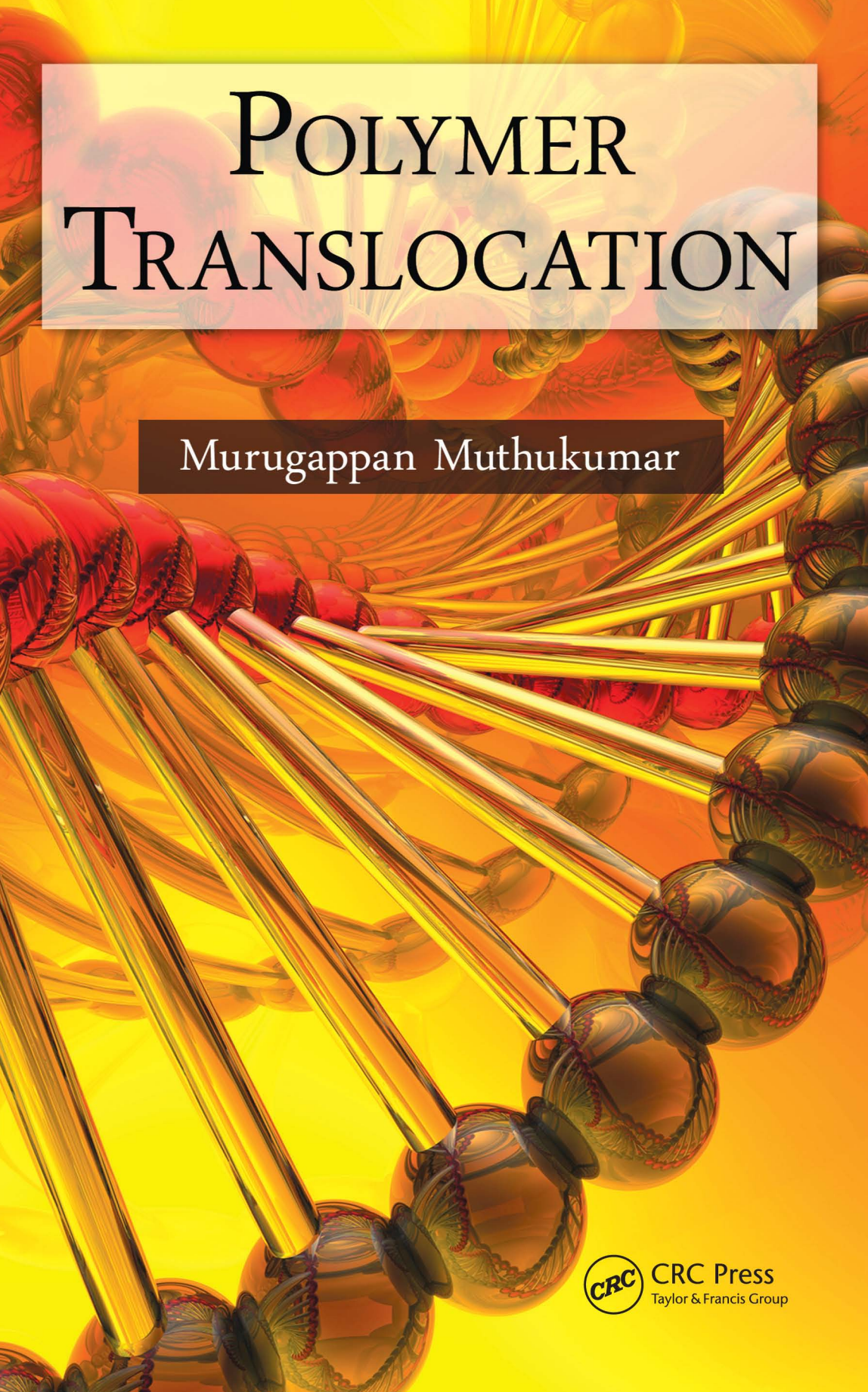
Polymer Translocation

Murugappan Muthukumar



CRC Press
Taylor & Francis Group

Polymer Translocation

Polymer Translocation

Murugappan Muthukumar

CRC Press
Taylor & Francis Group
Boca Raton London New York

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

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

Version Date: 20110708

International Standard Book Number-13: 978-1-4200-7517-5 (eBook - PDF)

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

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

DEDICATION

To Lalitha

CONTENTS

PREFACE

The process of polymer translocation occurs in many biological and biotechnological phenomena, where electrically charged polymer molecules, such as polynucleotides and proteins and their complexes, move from one region of space to another in crowded environments. Substantial research activities are currently being pursued in an effort to understand the macromolecular basis of polymer translocation. These activities are further stimulated by the societal need to sequence an enormous number of genomes immediately and inexpensively. Due to the inherent challenges of formulating the molecular basis of polymer translocation, this area has attracted a diverse set of researchers in biology, physics, chemistry, materials science, chemical engineering, and electrical engineering.

The thread that is central to polymer translocation is polyelectrolyte physics, which is perhaps one of the most challenging areas of modern research. The challenge in understanding the complex behavior of polyelectrolyte molecules arises from three long-range forces due to chain connectivity, electrostatic interactions, and hydrodynamic interactions. In addition, translocation of polyelectrolyte molecules through a protein pore or a solid-state nanopore becomes more complex by the polymer–pore interactions, confinement effects, and flow fields in the system. Unraveling the rich phenomenology of polymer translocation requires a grasp of modern concepts of polymer physics and polyelectrolyte behavior.

With this goal in mind, this book strives to present a summary of the key concepts of polyelectrolyte structures, electrolyte solutions, ionic flow, mobility of charged macromolecules, polymer capture by pores, and threading of macromolecules through pores. The main concepts and theoretical results are presented without formal derivations whereas the cited references provide adequate derivations. For situations where there is a lack of readily usable references, derivations are given. Every effort has been made to give the reader

an overview of basic concepts, established experimental facts, relevance of the concepts to real systems, ongoing challenges, and strategies for applying these ideas and summarized formulas to design new experiments. An attempt has also been made to avoid heavy mathematics and an exhaustive repetition of published literature.

ACKNOWLEDGMENTS

This book would not have been possible without the collaborations with my former and current students, as should be evident from the bibliography. I am sincerely grateful to all of them for their intellectual stimulation and friendship. I especially thank Stephen Mirigian, Mithun Kumar Mitra, and Jing Hua for many valuable suggestions for improving the presentation and for help with many figures. I am indebted to S. Bezrukov, E. DiMarzio, D. Hebert, R. Kumar, J. Machta, and C. Santangelo for their comments on one or more chapters.

I also gratefully acknowledge support from the National Institutes of Health and the National Science Foundation, without which the writing of this book and some of the advances presented here would have been almost impossible. I am also indebted to many collaborators outside my laboratory. In particular, it is a pleasure to acknowledge the tremendous encouragement and support from H. Bayley, S. Bezrukov, D. Branton, D. Deamer, E. DiMarzio, J. Kasianowicz, and M. Schmidt.

Finally, it is a great pleasure to thank Luna Han for her invitation to write a book on polymer translocation and her constant encouragement throughout the writing stage. It is also a pleasure to thank Kyle Meyer and Anto Aroshini Fernandez for their tremendous support during the production of the book.

AUTHOR

Murugappan Muthukumar is Wilmer D. Barrett Distinguished Professor at the University of Massachusetts, Amherst. After his undergraduate education at the University of Madras, India, he studied chemical physics at the University of Chicago, Illinois, followed by postdoctoral research at the Cavendish Laboratory, Cambridge University, Cambridge, U.K. He has taught polymer physics and biophysics at the University of Massachusetts for over 25 years. His research interests cover many areas of polymer science and physical biology, including polymer statistics, dynamics and thermodynamics of polymer solutions, and phase transitions of polymer systems. His current major interests concern the basic mechanisms of polymer crystallization, biomineralization, polyelectrolyte physics, assembly of viruses, and polymer translocation.

Muthukumar is a fellow of the American Physical Society and has received the Alfred P. Sloan fellowship, the Dillon Medal and the Polymer Physics Prize of the American Physical Society, the Chancellor's Medal of the University of Massachusetts, a Senior Humboldt Award, and the Gutenberg Lecture Award from the University of Mainz.

1

GENERAL PREMISE

Polymer translocation is one of the most fundamental macromolecular processes in life. This ubiquitous phenomenon deals with how electrically charged polymer molecules, such as polynucleotides and proteins, move from one region of space to another in crowded environments. Examples of biological phenomena for which polymer translocation is a crucial fundamental step include passage of mRNA through nuclear pore complexes, injection of DNA from a virus head into a host cell, gene swapping through pili, and protein translocation across biological membranes through channels (Lodish et al. 2007, Alberts et al. 2008). In addition, primarily due to societal and technological demands on DNA sequencing, there has recently been a tremendous effort to monitor and control the translocation of single macromolecules through a single pore made of proteins or synthetic solid-state materials. Although these apparently diverse phenomena emerge from different specific chemical details that are unique to each of these phenomena, we seek to identify the most common universal features behind these phenomena. The chemical details indeed decorate the basic universal feature of the passage of long macromolecules differently and impart specific directions and targets. We will first attempt to identify the common universal aspects of translocation and then to explore ways of incorporating the specific details relevant to different contexts. After illustrating the richness of the phenomenon with a few examples, we will offer an operating definition of the process and introduce the main concept, namely the entropic barrier idea (Muthukumar 2007), behind the polymer translocation. This will be followed by a brief outline of the various significant components, which need to be put together for a molecular understanding of the polymer translocation phenomenon.

1.1 BIOLOGICAL CONTEXTS

If we were to look into any volume element inside a eukaryotic cell (Figure 1.1), we are most likely to meet charged polymer molecules (such as proteins, RNA, and DNA) and electrolyte ions. In fact, the cell is a very crowded environment, and due to the nature of electrical (Coulomb) forces mediating the structures and functions of the constituent molecules inside the cell, it can be considered to be a thick "Coulomb soup." A comprehensive fundamental understanding of the structures, dynamics, and mobilities of single macromolecules and their complexes with other molecules in the *in vivo* environments, even in

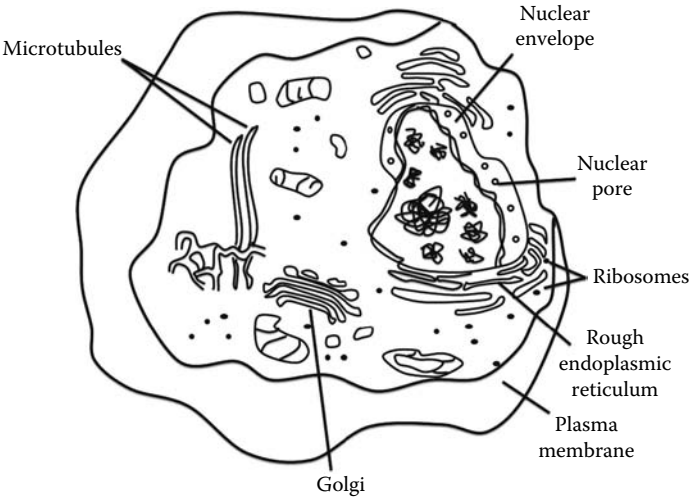


Figure 1.1 A cell is a crowded "Coulomb soup," with charged macromolecules and their assemblies moving between different compartments.

the absence of penetration through channels, is currently absent. Our expedition of trying to understand the translocation process by these molecules and their complexes, which themselves remain as poorly understood systems, becomes even more daunting. It is therefore necessary to investigate isolated translocation events before embarking on the coupled multiple translocation processes occurring simultaneously *in vivo*.

Even an isolated translocation process in *in vivo* is extremely rich in its phenomenology. As an illustration of the richness of details and complexities, consider the nuclear pore complex (Alberts et al. 2008), a crude sketch of which is given in Figure 1.2. The pore itself is apparently self-assembled by roughly a hundred different proteins with elaborate structural motifs: a basket-like cage with several openings capable of sieving different-sized molecules, a capillary in the middle of the passage with the capacity to dilate under macromolecular pressures, suspension of the passage into the double membrane of the nuclear envelope with a combination of hydrophobic and hydrophilic moieties, and charged polymer bristles protruding into the outside of the nucleus. The typical size of this assembly along the nuclear envelope is about 100 nm. It is through such an elaborate assembly, the mRNA, present as an mRNP complex of mRNA and more than 30 different carrier proteins inside the nucleus, is threaded into. In its own right, the mRNP complex is big with typical radius size of 50 nm. Thus, a structurally correlated object of about 50 nm is somehow pushed into the nuclear pore complex, and mRNA undergoes translocation. What is amazing is that this process is taking place all the time with fidelity, as instruction for synthesis of coded proteins would not occur without mRNA translocation. Indeed, we are yet to understand how this phenomenon takes

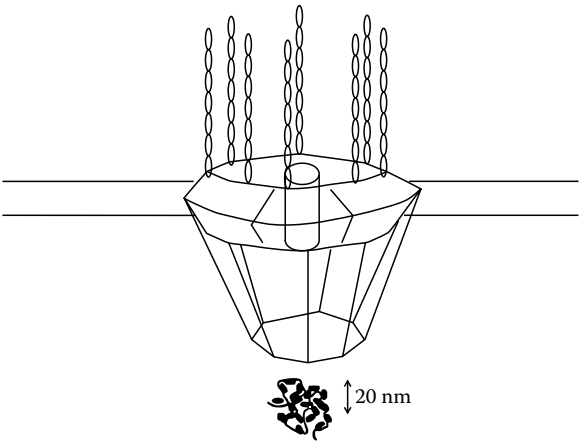


Figure 1.2 A sketch of the nuclear pore complex, and the mRNP. The typical feature sizes of these structures are 20–100 nm.

place. Nevertheless, it is evident that the phenomenon is manifest at very large length and timescales in comparison with atomistic scales. It is perhaps more fruitful to borrow ideas from polymer physics (de Gennes 1979), dealing with large-scale behaviors of macromolecules, to gain the "bird's eye view" and then to reckon the specific higher resolution features.

Not all contexts of polymer translocation are as complicated as in the nuclear pore complex. The passage of dsDNA from a virus head into a host cell as a single-file threading (Figure 1.3) and the transfer of DNA molecule from one bacterium to another (Figure 1.4) are examples of less complex situations. Again, the relevant length scales in the translocation phenomenon are much larger than atomistic length scales, calling for ideas from polymer physics.

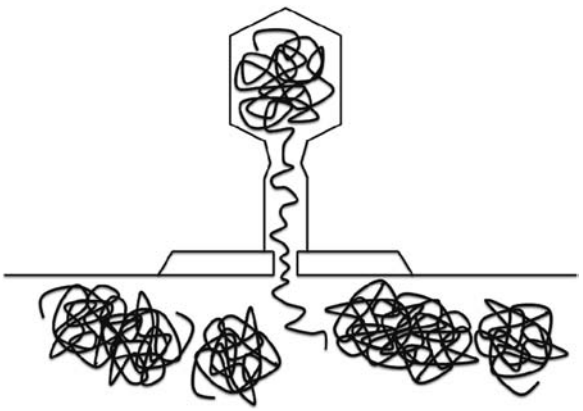


Figure 1.3 Cartoon of threading of DNA from a bacteriophage into a host cell.

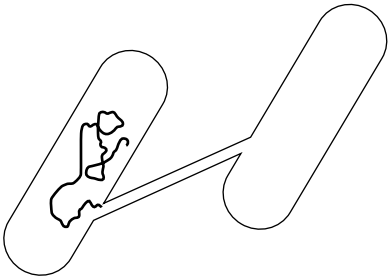


Figure 1.4 Cartoon of gene swapping between bacteria.

1.2 SINGLE-MOLECULE EXPERIMENTS

The above-mentioned *in vivo* biological phenomena are too complex to directly monitor one long macromolecule undergoing translocation in its totality. Fortunately, there have recently been many exciting single-molecule nanopore-based electrophysiology experiments, whereby the features of translocation by single polymer chains can be measured in great quantitative detail. Although these experiments were stimulated by the societal need of having to sequence enormous number of genomes immediately and inexpensively, they are serendipitously paving the way toward a fundamental molecular understanding of the phenomenon of polymer translocation.

In the single-molecule nanopore-based translocation experiments (Kasianowicz et al. 1996), a single nanopore is incorporated (either by a self-assembly of proteins or by ion-beam sculpting) into a membrane separating a donor (*cis*) chamber and an acceptor (*trans*) chamber. Each chamber contains a buffer solution with a strong electrolyte such as KCl. In many of the experiments, involving protein channels, the pore is a heptameric self-assembly of α-hemolysin (αHL) with a length of $\sim$10 nm and a narrow constriction of $\sim$1.4 nm, as sketched in Figure 1.5a (Song et al. 1996). In the case of solid-state nanopores (Figure 1.5b) (Chen et al. 2004a), the diameter is in the tunable range of 3–10 nm and the length is in the order of 10 nm or more. When an external voltage is applied across the membrane, the pore allows passage of small ions and the resulting ionic current is measured. When this experiment is repeated with ssDNA/RNA originally present in the *cis* chamber (with negative electrode), the measured ionic current decreases significantly whenever the polymer interferes with the pore. A typical trace of ionic current versus time for the passage of a polymer chain through αHL is given in Figure 1.6. Although every encounter with the pore is caused by identical polymer molecules, the resultant ionic response is stochastic. As marked in Figure 1.6, there are apparently three timescales. The time t_0 is the approach time between two successive events and we may define the inverse of the average t_0 as the capture rate R_c, independent of whether the polymer actually underwent the translocation process or not. Also, as indicated in Figure 1.6, there are partial blockades

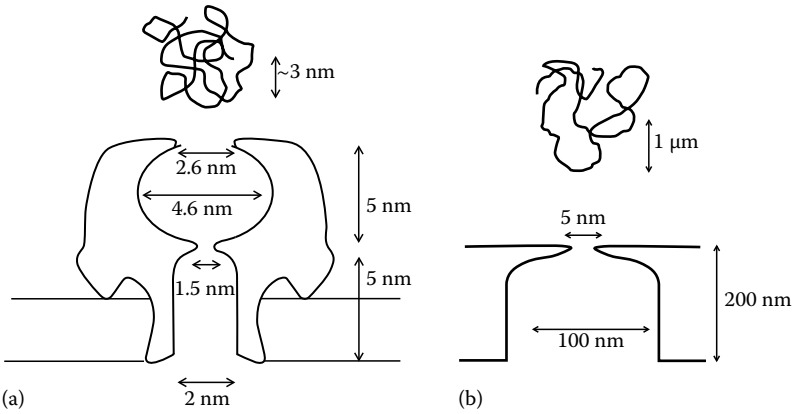


Figure 1.5 Sketches of (a) the α-hemolysin pore and (b) solid-state nanopore used in single-molecule electrophysiology experiments.

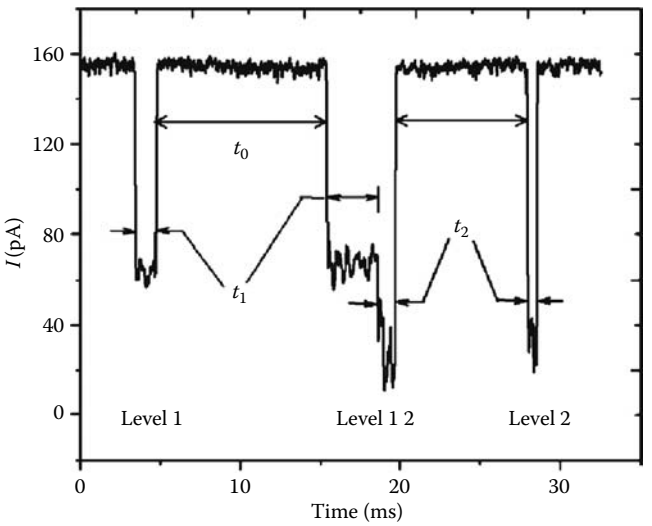


Figure 1.6 A typical ionic current associated with the encounter of polymer chains with an αHL pore.

(with duration t_1) and deep blockades (with duration t_2). Furthermore, the common feature of the experimental results is that the distributions of t_0, t_1, t_2, and the various blocked current levels are very broad.

The details of the time-dependence of the ionic current bear information on the manner in which polymer molecules attempt to translocate through a pore and the underlying molecular mechanism of polymer threading. Experiments show that the average translocation time, for single-file translocation processes,

is directly proportional to the polymer length and inversely proportional to the applied voltage, in spite of the fact that the translocation time generally has a broad distribution (Kasianowicz et al. 1996). The capture rate depends on the polymer concentration, the direction of the translocation (*cis*-to-*trans* vs. *trans*-to-*cis*), and the applied voltage above a threshold value (Henrickson et al. 2000). In addition, polymer sequences and their ability to spontaneously form secondary structures influence their migration through nanopores as manifest in the corresponding ionic current traces (Akeson et al. 1999, Meller et al. 2000). In fact, such distinguishing features in the ionic current traces, associated with translocation of different polymer sequences, raised high hopes for cultivating single-molecule electrophysiology technique into a fast sequencing technology (Branton et al. 2008).

As a complement to the threading of single-stranded DNA/RNA, and to avoid potential complications from the role of the vestibule of α-hemolysin pore in interpreting the ionic current traces, much experimental effort has gone into forming solid-state nanopores with diameters large enough to thread double-stranded DNA (Chen et al. 2004a). The ionic current traces associated with the passage of dsDNA through these solid-state nanopores are apparently more complex than the corresponding results for the α-hemolysin pore. Now, the polymer can translocate in quantized configurations such as a single file, chain with one hairpin, etc. Even the seemingly simplest situation of translocation of dsDNA through a nanopore of 15 nm diameter exhibits rich puzzles. Time-resolved fluorescence studies have revealed that a depletion (capture) region of about 3 μm (much larger than the pore size) develops in front of the pore (Chen et al. 2004a). The DNA molecules were found to diffuse slowly ($\sim 4s$) until they approach the capture region. Once the molecules reach the capture region, they were found to be depleted rapidly ($\sim$50 ms) by active pulling through the pore.

In addition to nanopores, nanoscopic channels have also been used to investigate the translocation of DNA molecules (Han and Craighead 2000). Consider a periodic alternation of deep ($\sim$2 μm) and shallow ($\sim$100 nm) wells, with the width of both of these wells being far wider than the size of a DNA molecule. Experimental measurement of the average time τ taken by one λDNA molecule to pass through a pair of adjacent deep and shallow wells showed that it takes shorter time for longer molecules in accordance with an empirical formula, $\tau \sim N^{-0.42}$, where N is proportional to the polymer length. This counterintuitive finding is in direct opposition to the linear increase of τ with N for single-file translocation through pores.

The general picture that emerges from the above selective description of the phenomenology of polymer translocation is that the phenomenon is quite complex even in simple experimental setups and is controlled by numerous factors. The translocation time of one macromolecule depends on the chain length, chemical sequence of the polymer, chain stiffness in terms of whether single-stranded or double-stranded, applied voltage, chemical nature of the pore, pore geometry, and flow fields in the experimental setups. In general, the

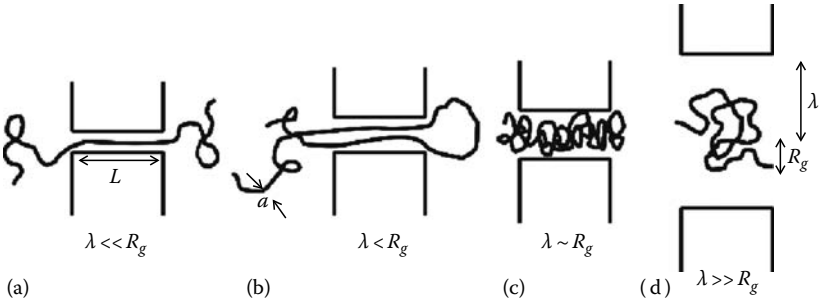


Figure 1.7 Different regimes (a–d) of polymer confinement by the pore. The process is translocation for $a \leq \lambda < R_g$.

translocation process is stochastic with broad distributions of various measures of the process, even though identical molecules are undergoing translocation.

1.3 NOMENCLATURE

It is perhaps useful to associate certain specific criteria in defining the process of polymer translocation, in order to distinguish it from the general transport of macromolecules. Consider a uniform pore with radius λ and length L, through which a chain of average radius of gyration R_g undergoes translocation. Let a be the radius of each of the monomers constituting the polymer. If λ is slightly larger than a but much smaller than R_g, the chain can undergo translocation only as a single file or as a hairpin (Figure 1.7a and b). If λ is much larger than a but smaller than R_g, the chain can be squeezed inside the pore as sketched in Figure 1.7c. On the other hand, if λ is much larger than R_g (Figure 1.7d), then the polymer undergoes transport through the capillary as in free solutions, except for the possible adsorption/depletion effects at the walls of the pore. In the nomenclature adopted here, the phenomenon of polymer translocation refers to the constrained motion of polymer chains where the size of the pore is smaller than the size of the polymer, $a \leq \lambda < R_g$.

1.4 ENTROPIC BARRIER IDEA

One of the inherent properties of an isolated polymer chain is its ability to assume a large number of conformations $\mathcal{N}$. As a result, the chain entropy ($k_B \ln \mathcal{N}$; k_B is the Boltzmann constant) can be high and its free energy F is given by

$$F = E - TS = E - k_B T \ln \mathcal{N}, \tag{1.1}$$

where E is the energy of interaction between the monomers and the surrounding solvent molecules, and T is the absolute temperature. There can be additional

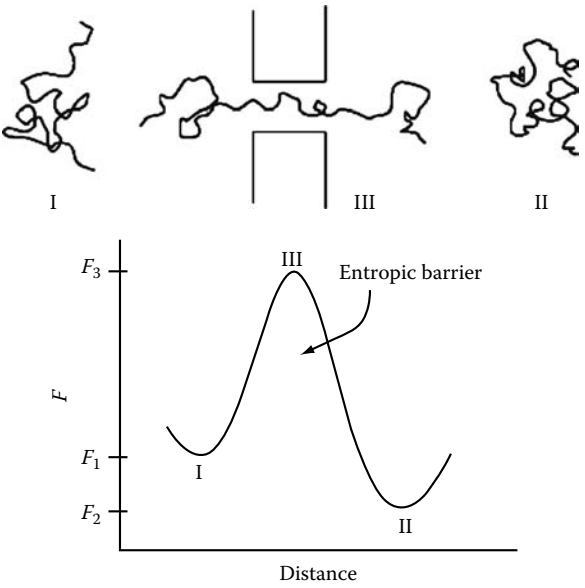


Figure 1.8 Genesis of the entropic barrier for polymer translocation.

entropic contributions to F due to a reorganization of solvent molecules accompanying conformational changes of the chain. When such a chain is exposed to a restricted environment such as a pore, the number of conformations that can otherwise be assumed by the chain is reduced, and as a result the chain entropy decreases and the chain free energy increases. This effect is depicted in Figure 1.8.

F_1, F_2, and F_3 are the free energies of the chain in regions I, II, and III, respectively. Owing to the reduction of conformations in region III, F_3 is higher than F_1 and F_2. We shall call $(F_3 - F_1)$ the entropic barrier to the passage of the chain out of region I. Although this barrier is called the entropic barrier, it is indeed a free energy barrier because additional enthalpic contributions to F_3 can arise from the interactions between the polymer and the pore. In general, the environment of the chain in region II can be different from that in region I (due to different electrochemical potentials in these regions), so that F_2 is not necessarily equal to F_1. The net driving potential for polymer translocation from region I to region II is $(F_1 - F_2)$. The polymer chain must negotiate the entropic barrier in order for it to successfully arrive at the opposite side of the pore.

It is important to recognize that the role of conformational entropy of polymer chains in various biological processes cannot be treated as only a minor factor. Since the temperature T is essentially fixed for a given physiological system and because only rather minor variations are permitted in E for a fixed T, the only way the free energy landscape can be dramatically modified must

be through the entropy S. The ability of polymer molecules to undergo large conformational changes, without losing their topological connectivity, makes them ideal candidates for large entropic changes. No wonder that life is made of polymer strings instead of, say, cubes or spheres. With the help of such entropic considerations, we will formulate the arguments for the structure, dynamics, mobility, and translocation of polymer chains in what follows in the book.

1.5 PHYSICS OF TRANSLOCATION

Given the rich features of the translocation phenomenon, the objective is to identify the various significant contributing factors and to assess their relative contributions to translocation. Even with the modern computational technologies, it is impossible to build from atomistic details and force fields at sub-nanometer resolution and calculate the behavior of the whole macromolecular assemblies of hundreds of nanometers in size. The nature of forces among all atoms of translocating polymer molecules, enzymes, and protein pores in salty and crowded aqueous solutions with highly heterogeneous dielectric function remains as a huge challenge to be unraveled. Nevertheless, it is worthwhile to explore theoretical possibilities where local details at the spatial resolution of amino acids and nucleotides are surrogated into coarse-grained parameters at multiple nanometer resolution. This will allow implementation of well-established concepts from polymer physics. With this attitude, we will present basic concepts, arguments, predictions, and comparison with experimental results related to polymer translocation in the following chapters.

The general scope of the translocation process may be divided into several separate parts. There are three essential steps associated with the transit of a polymer through a nanopore (illustrated in Figure 1.9 for a structureless blunt pore): (1) drift–diffusion, (2) capture, and (3) translocation.

1.5.1 Drift–Diffusion

In the first step (far away from the pore), the polymer undergoes a combination of drift, due to the externally imposed force fields, and diffusion arising from collisions with solvent molecules. The drift–diffusion of the polymer is established by the structure and size of the polymer, the nature of the background fluid (such as solvent quality and ionic strength), and influences from external forces (such as electric field and pressure gradient).

1.5.2 Capture

At the pore mouth, force fields may be generated by chemical decoration of the inside surface of the pore. More importantly, steep electric potential gradients may occur at the pore mouth due to the dielectric mismatch between the layer in which the pore is embedded and the rest of the system, in the presence of an applied voltage gradient. Furthermore, strong flow currents may arise at the

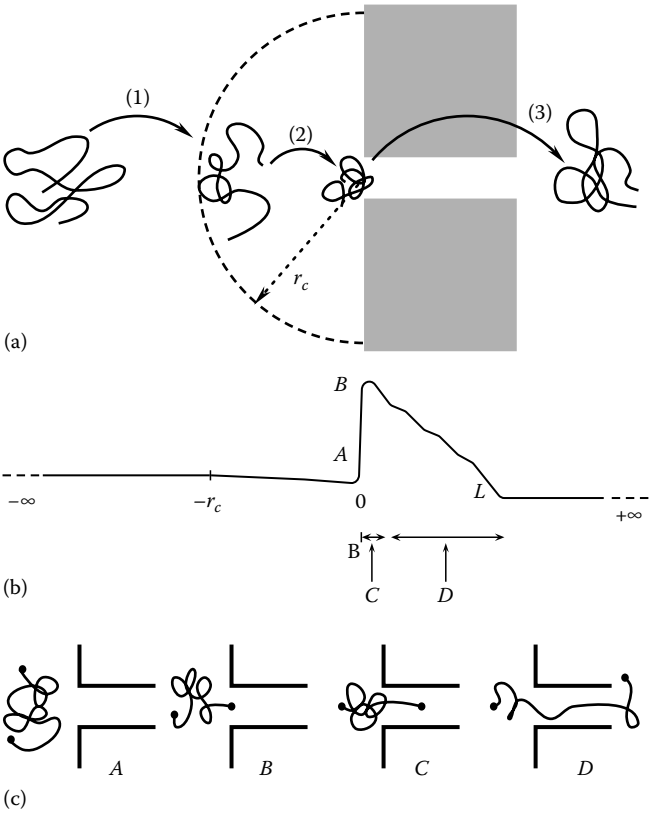


Figure 1.9 (a) Three main stages of polymer translocation process: (1) drift–diffusion, (2) capture, and (3) translocation; (b) free energy landscape; and (c) three stages in the third translocation step.

pore mouth due to the movement of water through the pore. In particular, for situations dealing with charge-bearing pores containing electrolyte solutions, this force, called the electroosmotic force, can be quite significant. Since the flux of the water flow must be continuous in the system, and since there is only a narrow passage for fluid flow inside the pore, strong velocity gradients may develop at the pore mouth. All of these effects generate an effective sucking force at the pore entrance, which in turn tries to capture the polymer to facilitate the translocation. Thus, near the pore, the flow field and the electric field can be significantly influenced by electroosmotic forces, dielectric mismatch between the pore wall and the aqueous medium, ionic strength gradients, and pressure gradients. Depending on the details of these contributing factors, the nature of the flow field within a range of r_c from the pore can be qualitatively different from that outside this range. r_c can vary between subnanometers to microns. Within the range of r_c, the polymer may undergo conformational deformation.

By experiencing such forces within r_c, the polymer approaches the pore mouth, designated as step (2) in Figure 1.9a. The capture of the polymer at the pore mouth is controlled by the strength of the sucking force at the pore entrance and by the range of the flow field in front of the pore where the velocity gradients are strong.

1.5.3 Translocation

In general, when the polymer is caught at the pore mouth, it is in a jammed state without any initial correlation between the chain ends and the pore mouth. The chain needs to unravel itself to place one of its ends at the pore mouth for the single-file translocation to occur, and then to thread through the pore. This step of translocation consists of three stages: (a) chain-end localization, (b) nucleation, and (c) threading. An entropic barrier must be overcome in placing one of the chain ends at the pore mouth from a jammed coil state (designated as $A \rightarrow B$ in Figure 1.9b and c), in order to enable the eventual single-file translocation. This is due to the requirement that one end must be at a specific spacial location, instead of all possible locations, whereby the chain end is losing its translational entropy. After the localization of one chain end, there is an additional entropic barrier for reducing the conformational degrees of freedom for the chain in order to be squeezed into the pore. The polymer chain is thus hung across the entropic barrier. As will be seen later, only if a sufficient number of monomers crosses this "nucleation barrier" can the chain undergo further translocation. The nucleation stage is $B \rightarrow C$ in Figure 1.9b and c. The final stage of translocation, $C \rightarrow D$, is a downhill threading process, which is in its own right a drift–diffusion process. The chain is finally kicked out of the pore into the receiver compartment as the ultimate step.

The shape of the pore can lead to additional complexities. As an example, the free energy landscape for a protein pore, such as αHL, containing a trap-like vestibule in front of the pore, is qualitatively different (Figure 1.10) from that for a blunt pore. Here, the jammed coil at the pore mouth is separated by two barriers, one for successful translocation into the *trans* side and the other for the return to the *cis* side.

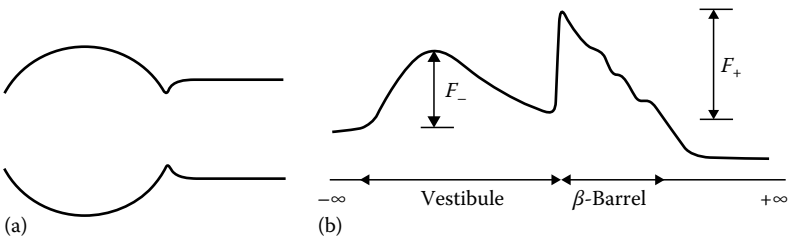


Figure 1.10 The polymer can get jammed in the metastable state at the pore mouth, for αHL, with two barriers for forward and backward movements: (a) sketch of αHL and (b) free energy profile.

1.6 OUTLOOK

The above description highlights only the generic physical aspects of the translocation process. In the translocation step, specific to a particular polymer and a particular pore, the fine details of the electrostatic and hydrophobic properties of the amino acids constituting the protein pores, charge decoration of the inside wall of the solid-state nanopores, geometry of the pore, and the polymer sequence contribute significantly. An accounting of these effects manifest at both subnanometer level and microscopic level is essential for a fundamental understanding of polymer translocation. In view of our approach to adopt a coarse-grained methodology, we will implement concepts from polymer physics cultivated over the past seven decades to explore this phenomenon. We shall first devote several chapters to discuss size, shape, and structure of isolated polymers in equilibrium and under flow, and their dynamics and mobility in free solutions containing a certain amount of strong electrolytes. We will address the various origins of the capture zone and the process of polymer capture. Quantitative descriptions of the entropic barrier and the free energy landscape associated with the translocation, and the kinetics of polymer threading into pores will be presented next. We will then put all of these components together in order to understand the various experimental results on translocation.

2

SIZE, SHAPE, AND STRUCTURE OF MACROMOLECULES

Polymer molecules are monomers contiguously connected by covalent bonds in a chain-like fashion. The monomers themselves are groups of atoms, and can be either identical repeat units (as in polyethylene or polyuridylic acid) or chemically different units (as in a protein molecule or a deoxyribonucleic acid (DNA) containing different bases). Depending on the chemical nature of the repeat units of the polymer, the number of monomers per chain, and the nature of the solvent in which the polymer is dispersed, the molecule can assume different sizes and shapes such as globular, coil-like, and rod-like. It might seem at the outset that it is necessary to treat each polymer in a given solvent condition as a unique case by accounting for the specific chemical nature of the polymer and solvent. However, it turns out that there are certain universal laws that can describe average polymer conformations. It is possible to surrogate the local degrees of freedom of chemical specificity into a few parameters and obtain useful coarse-grained models in order to understand the universal properties of polymer chains.

In this chapter, we shall introduce various measures of polymer conformations and some of their universal laws. We will give a summary of coarse-grained models of polymer conformations and discuss how local details are parametrized in these models. The basic vocabulary of polymer statistics, including concepts like persistence length, radius of gyration, hydrodynamic radius, size exponent, structure factor, fractal dimension, excluded volume parameter, and coil–globule transition, will be introduced. In experiments exploring the translocation phenomenon discussed in this book, the polymer is electrically charged and the solvent medium is an aqueous electrolyte solution. When electrical charges are present in a polar dielectric medium, there are additional significant concepts that are required to describe polymer statistics. This in itself constitutes a separate field of study and still remains to be fully understood. In view of this, we relegate the discussion of charged polymers and electrolyte solutions to the following chapters. In the present chapter, we shall consider only uncharged polymers.

2.1 MEASURES OF POLYMER CONFORMATIONS

When a polymer is dispersed in a solvent, there are generally three kinds of pairwise interactions at the local level: monomer–monomer, monomer–solvent,

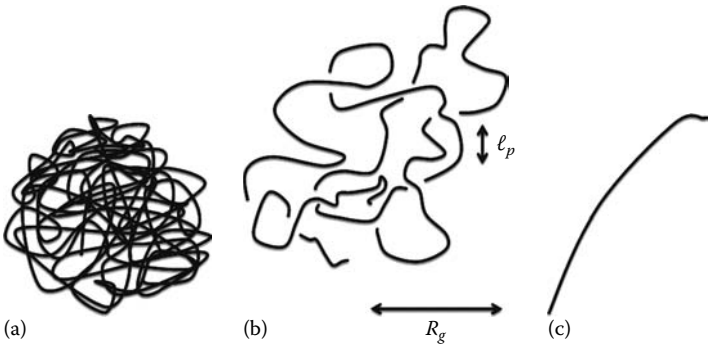


Figure 2.1 Major conformations of isolated polymer chains: (a) globule, (b) coil, and (c) rod-like. R_g is the radius of gyration and ℓ_p is the persistence length.

and solvent–solvent. If the hydrophobic interaction, due to van der Waals attractive forces, between the monomer units were to be dominant over the monomer–solvent interaction, then the monomers would aggregate together to form a globular structure (Figure 2.1a) by excluding the solvent molecules out of the globule. On the other hand, if the monomer–solvent interaction is preferable over the monomer–monomer attractive interaction, the solvent becomes a good solvent for the swelling of the polymer, which then adopts a swollen coil-like conformation (Figure 2.1b). In an average sense, the coil would look like a rough porous ball of wool, carving out a rough sphere of revolution with a radius R_g, called the radius of gyration (defined below). For some polymers, the chemical details associated with adjacent monomer units along the chain backbone are such that rotation of these monomers around the connecting chemical bond can be severely restricted. Furthermore, conformations of adjacent monomers could be locked together by hydrogen bonding, as in the helical conformations of polypeptides and double-stranded DNA molecules. As a result, the chain backbone can be locally stiff. If the contour length of the polymer is short enough, then it would look rod-like (Figure 2.1c), with the obvious shape anisotropy. If the contour length of the polymer is very long, the chain would be rod-like locally (for distances less than or comparable to the persistence length ℓ_p, defined below) but would bend and curve at longer distances appearing overall as a coil. Such polymers are called semiflexible polymers. Indeed, the chain can undergo conformational transitions between the coil and globular states when experimental conditions alter the relative weights of the monomer–monomer, monomer–solvent, and solvent–solvent energies. Also, the same polymer in identical experimental conditions could be either rod-like or coil-like, depending on its molecular length.

We shall now define some quantities, which are either measured experimentally or computed theoretically, to describe polymer conformations. These definitions are general, independent of the particular conformations taken by the polymer. As an example, consider a specified conformation of a polyethylene

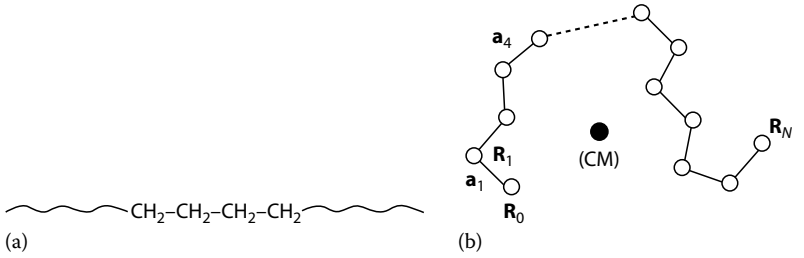


Figure 2.2 (a) Backbone structure of a polyethylene chain. (b) A typical conformation of a skeletal chain. $\mathbf{R}_i$ is the distance vector of the ith united atom from the center of mass (CM) of the conformation, and $\mathbf{a}_i$ is the bond vector connecting $(i-1)$th and ith united atoms.

chain (Figure 2.2a) with $(N+1)$ methylene monomers. Denoting each repeat unit (a methylene group in this case) as a united atom (skeletal atom), the skeletal structure of a conformation can be drawn as in Figure 2.2b. Here $\mathbf{R}_i$ is the position coordinate of the ith skeletal atom from the center of gravity (CM) of the specified chain conformation, and $\mathbf{a}_i$ is the bond vector of the ith skeletal bond. For each of such conformations, quantities like the end-to-end distance $\mathbf{R}$ and radius of gyration R_g can be defined as follows. Since the chain can adopt many conformations during the typical measurement times, we construct averages of these quantities over all possible conformations. These averages constructed in equilibrium are time-independent. We also assume that the repeat units are identical with identical bond lengths connecting them.

1. *Mean square end-to-end distance*, $\langle R^2 \rangle$:

$$\langle R^2 \rangle = \langle (\mathbf{R}_N - \mathbf{R}_0)^2 \rangle = \sum_{i=1}^{N} \sum_{j=1}^{N} \langle \mathbf{a}_i \cdot \mathbf{a}_j \rangle, \tag{2.1}$$

where the angular brackets denote the averages over allowed conformations at a given experimental condition.

2. *Radius of gyration*, R_g: This is defined as the root mean square radius of gyration, where the mean square radius of gyration is given by

$$R_g^2 = \frac{1}{(N+1)} \sum_{i=0}^{N} \left\langle R_i^2 \right\rangle. \tag{2.2}$$

This quantity is measured in static scattering techniques using light, x-rays, and neutrons.

3. *Hydrodynamic radius*, R_h:

$$R_h = \left(\frac{1}{(N+1)^2} \sum_{i=0}^{N} \sum_{j>i} \left\langle \frac{1}{| \mathbf{R}_i - \mathbf{R}_j |} \right\rangle \right)^{-1}. \tag{2.3}$$

This quantity is measured in dynamic light-scattering technique, and its origin lies in the hydrodynamic interactions in the solution.

4. *Size exponent*, ν: It is evident from the above definitions that after taking the conformational averages, $\langle R^2 \rangle, R_g$, and R_h depend only on the number of monomers $(N + 1)$ in the chain and the bond length (a). As will be derived in the following sections, it can be shown that each of these three quantities is directly proportional to an exponent ν of the number of monomers per chain. The only difference between the three relations of $\langle R^2 \rangle^{1/2}, R_g$, and R_h is the numerical prefactor of order unity. Also, the difference between the number of monomers $(N + 1)$ and the number of bonds (N) can be omitted for the large values of N for polymers typically dealt with in translocation experiments. By suppressing the proportionality constant for the root mean square end-to-end distance, average radius of gyration and the hydrodynamic radius, and using the generic symbol R to represent any of these three quantities, we write

$$R \sim aN^{\nu}, \tag{2.4}$$

where R is taken to represent the average "size" of the polymer, and ν is called the size exponent. The bond length a is used in the above equation to make both sides of the equation to have the same dimension of length. Furthermore, it is sometimes useful to think of the polymer coils as statistically fractal objects with their own fractal dimensions embedded in the space of three dimensions of the solution (or two dimensions corresponding to a membrane). Note that N is directly proportional to the mass of the polymer and that a compact object (with its own dimension being three) obeys the relation, $R^d \sim N$, where d is the space dimension ($d = 3, 2$, and 1, for a sphere, disk, and line, respectively). Analogous to this geometric relation, we define a dimension for the average polymer conformation by rewriting the above equation as

$$R^{d_f} \sim N, \tag{2.5}$$

where d_f is called the fractal dimension of the polymer and is defined according to the above two equations as

$$d_f \equiv \frac{1}{\nu}. \tag{2.6}$$

The fractal dimension of the polymer is different from the space dimension d in which the polymer is present.

5. *Shape factor*, R_g/R_h: The ratio of R_g to R_h is sometimes used to remark on the anisotropy of the shape of the molecule. Since the dependence of R_g and R_h on N is the same with identical size exponent, the ratio is only a numerical factor reflecting the different values of the proportionality factors in their relations to N. For example, the shape factor is 0.77 for a compact sphere and increases to values of about 4 for rod-like conformations.

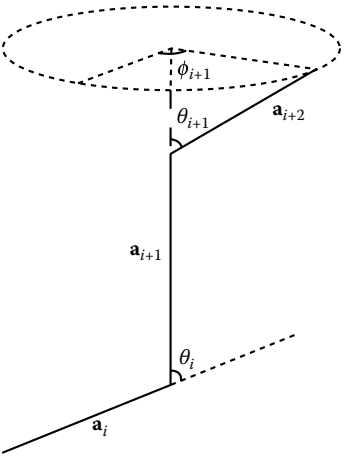


Figure 2.3 Sketch of local conformations defining the bond angle θ and the dihedral angle ϕ.

6. *Persistence length*, ℓ_p: The bond angles between the contiguous chemical bonds along the chain backbone cannot be arbitrary and are restricted by quantum mechanical properties. This feature results in the persistence of the direction of a bond over a certain distance along the chain contour. There can be additional reasons for this orientational persistence, due to hydrogen bonding among consecutive monomers as well. The relative orientation of a bond next to the preceding two bonds is defined by two angles, namely the bond angle θ and the dihedral angle ϕ, as illustrated in Figure 2.3. Here, the bond angle between the ith and $(i + 1)$th bond vectors is defined as $180 - \theta_i$. The dihedral angle ϕ_{i+1} is the angle between the plane of the bond vectors $\mathbf{a}_{i+2}$ and $\mathbf{a}_{i+1}$ and the plane of the bond vectors $\mathbf{a}_{i+1}$ and $\mathbf{a}_i$. The chemical nature of the atoms constituting the united atoms influences the allowed values of the dihedral angles, which then are manifest as the persistence length of the polymer. There are two ways of defining the persistence length. In one way, the average of the product of the orientation (that is the bond vector) of the ith bond and that of the jth bond is monitored as a function of the distance along the chain backbone $| i - j |$. This correlation function obeys the typical formula,

$$\langle \mathbf{a}_i \cdot \mathbf{a}_j \rangle = \mathrm{a}^2 \exp\left(-\frac{| i - j |}{\ell_p} \right), \tag{2.7}$$

where ℓ_p is defined as the persistence length. For distances $| i - j |$ less than ℓ_p, the bond orientations are correlated and hence the conformation is rod-like. For distances $| i - j |$ larger than ℓ_p, the bond orientations are uncorrelated and hence the conformation can become coil-like. Another way to define the persistence length of the polymer is by projecting the end-to-end distance vector of the chain on the first bond in the limit of large values of n. We shall later

introduce a model in order to extract the persistence lengths of polymers from experimental data.

7. *Monomer density distribution*, $\rho(r)$: For coil-like conformations, as sketched in Figure 2.1b, the density of monomers at the center is expected to be high and progressively decreasing with the radial distance until reaching the coil's boundary. This is to be contrasted with a compact object where the density is uniform until the radius at which it discontinuously becomes zero. Let $\rho(r)$ be the number of monomers inside a spherical volume of radius r around a tagged monomer inside a large polymer coil. It can be shown (de Gennes 1979), based on general arguments for self-similar fractal objects, that the number density of monomers depends on the radial distance according to

$$\rho(r) \sim \frac{1}{r^{3-d_f}}, \tag{2.8}$$

where d_f (reciprocal of the size exponent ν) is the fractal dimension of the coil. This result is valid as long as the radial distance is not too close to the monomeric dimensions or not larger or comparable to the polymer radius. Basically, this relation is obtained by constructing the ratio of number of monomers in a volume of radius r to this volume. If the coil can be assumed to be self-similar inside the polymer coil, then the numerator of this ratio is proportional to r^{d_f} (see Equation 2.5) and the denominator is proportional to r^3, thus leading to the above equation. Since d_f is less than three (except for solid-like conformations), $\rho(r)$ decreases algebraically with the radial distance (Figure 2.4a). This result is contrasted with the corresponding result for a solid object in Figure 2.4b. Therefore, the topological correlation arising from chain connectivity leads directly to long-ranged correlation between monomer densities for polymer chains that assume noncompact and ramified conformations.

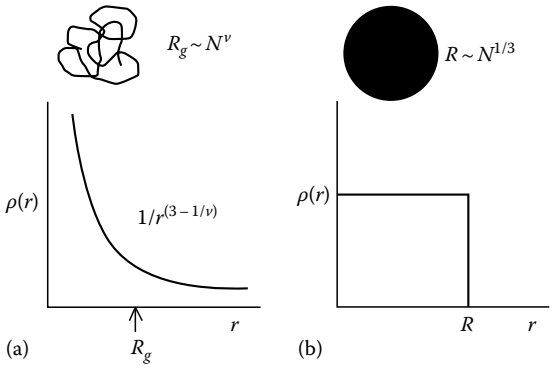


Figure 2.4 (a) Algebraic decay of monomer density correlation with radial distance. The determining factor is the fractal dimension (reciprocal of the size exponent ν). (b) The density profile is a step function for solid objects.