Polymer Translocation Through Narrow Pores: Role of Semiflexibility and Hydrodynamics

A thesis
submitted by

Rajneesh Kumar

in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

Indian Institute of Science Education and Research (IISER), Mohali

July 2018

Certificate of Examination

This is to certify that the dissertation titled *Polymer translocation through narrow pores: role of semiflexibility and hydrodynamics* submitted by **Mr. Rajneesh Kumar** (Reg. No. Ph11092) for the partial fulfillment of Doctor of Philosophy programme of the Institute, has been examined by the thesis committee duly appointed by the Institute.The committee finds the work done by the candidate satisfactory and recommends that the report be accepted.

Dr. Dipanjan Chakraborty Dr. Abhishek Chaudhuri Dr. Rajeev Kapri
(Supervisor)

Declaration

The work presented in this dissertation has been carried out by me under the guidance of Dr. Rajeev Kapri at the Indian Institute of Science Education and Research Mohali.

This work has not been submitted in part or in full for a degree, a diploma, or a fellowship to any other university or institute. Whenever contributions of others are involved, every effort is made to indicate that clearly, with due acknowledgement of collaborative research and discussions. This thesis is a bonafide record of original work done by me and all sources listed within have been detailed in the bibliography.

Rajneesh Kumar
(Candidate)
20 July 2018

In my capacity as the supervisor of the candidates doctoral thesis, I certify that the above statements by the candidate are true to the best of my knowledge.

Dr. Rajeev Kapri
(Supervisor)

Acknowledgements

There are no proper words to convey my deep gratitude and respect for my Ph.D advisor, Dr. Rajeev Kapri, for his constant support and encouragement throughout the time of degree. At the same time, I acknowledge my debt to him for introducing me to a dynamical areas of statistical physics and enhancing my computational skills. His qualities of teachings, temperament in research and discussions has made me learn many things on scientific ground that brings positive attitude and enthusiasm in both research and life. Because of his enormous support, I could maintain a balance between my professional and personal life towards end of the programme. I will always be grateful to him for this.

I most gratefully thank Dr. Abhishek Chaudhuri for his contributions in all three problems I did during my degree. I would also like to thank him for being very supportive and enhancing my knowledge about biological systems. His curiosity and positive approach towards every problems has been a great source of motivation for me.

I appreciate contributions of time and support from my doctoral committee members, Dr. Abhishek Chaudhuri and Sanjeev Kumar. Their suggestions towards research has been valuable at times. I especially thank Dr. Debasish Chaudhuri for his valuable suggestions and comments on my work and allowing me to work with him for some time. I gratefully thank Prof. Gerhard Gompper and his post-doc Anoop Varghese (ICS Forschungszentrum Jülich) for their initial help in understanding the multi-particle collision dynamics algorithm. I also take this chance to thank respectfully, Prof. N. Sathyamurthy (Director), Dr. Ramandeep S. Johal (Dean Academic) and Prof. Jasjeet Bagla (Head of the Physical Sciences) at IISER Mohali for their kind support in administrative aspects whenever I needed them.

I also gratefully acknowledge support from IISER Mohali for providing a platform for doing research with high performance computing facility and also for research fellowship.

My heartiest and special thanks to my friends Swagatam Nayak, Satyam Shrivastva and Arun with whom I had shared professional and personal life and has been a constant support for having a wonderful time. I also thank Arash, Sandeep, Anshul, Vivek, Nisha, Aman, Sameep, Ramu, Shubhendu, Gopal, Goarge, Varinder, Satnam, Preetinder, Pooja, Ankit, Chandrakala, Karishma, Archana, Ayushi, Pranay, Suman Kalyan, Sumanlata, Deepraj, Nitish for their moral support and delightful discussions. I sincerely thank Prof. P. K. Ahluwalia for their support, teaching and introduce me to the world of computation.

I should mention the ceaseless support of my family, my mother, father and younger brothers and sister. Their unconditional support, love, and patience gave me the strength to go through the most difficult times of research programme. I shall be forever grateful to my younger brother, Satish, for taking care of almost everything back at home so that I could concentrate on my research. I would also like to thank Medha Karnik for lovely friendship and be there always with me in every phase of my life, especially in bad times.
I also offer my humble thanks to many others who have helped me directly or indirectly throughout my research.

Rajneesh Kumar

List of Publications

1. **Rajneesh Kumar**, Abhishek Chaudhuri and Rajeev Kapri, *Sequencing of semiflexible polymers of varying bending rigidity using patterned pores*, J. Chem. Phys. **148**, 164901 (2018).

2. **Rajneesh Kumar**, Abhishek Chaudhuri and Rajeev Kapri, *Driven translocation of a semiflexible polymer through an interacting conical pore* (Manuscript in preparation).

3. **Rajneesh Kumar**, Rajeev Kapri and Abhishek Chaudhuri, *Fluid flow driven translocation of semiflexible polymer chain through nanopore* (Manuscript in preparation).

List of acronyms

DNA	Deoxyribonucleic Acid
ssDNA	Single-stranded Deoxyribonucleic Acid
dsDNA	Double-stranded Deoxyribonucleic Acid
RNA	Ribonucleic Acid
mRNA	Messenger Ribonucleic Acid
FJC	Freely-Jointed Chain
CM	Center of Mass
WLC	Worm Like Chain
MspA	Mycobacterium Smegmatis Porin A
TEM	Transmission Electron Microscopy
TP	Tension Propagation
FENE	Finitely Extensible Non-linear Elastic
BD	Brownian Dynamics
LJ	Lennard-Jones
BDTP	Brownian Dynamics Tension Propagation
IFTP	Iso-Flux Tension Propagation
LGA	Lattice Gas Automaton
LB	Lattice Boltzmann
DPD	Dissipative Particle Dynamics
MPCD	Multi-Particle Collision Dynamics
SRD	Stochastic Rotation Dynamics
MBS	Maxwell Boltzmann Scaling
VACF	Velocity Auto-Correlation Function

Abstract

Polymer translocation is relevant to various biological processes, like passage of mRNA from nucleus to cytoplasm through nuclear pores after transcription, horizontal gene transfer in bacterial conjugation, transport of proteins and viral injection of DNA into the host cells. Polymer translocation also finds application in gene therapy, controlled drug delivery, and rapid DNA sequencing. Due to these technological applications, polymer translocation has gained considerable attention in the last two decades both theoretically and experimentally. Experiments have demonstrated that single-stranded DNA and RNA molecules can be electrophoretically driven through biological and synthetic nanopores. By using *S. aureus* α-hemolysin to form a single channel across a lipid bilayer separating two buffer-filled compartments, it was found that the ionic current through pore depends strongly on the polynucleotide sequence passing through the pore and could be used for sequencing of DNA and RNA molecules. Currently, three types of nanopores are used for sequencing purpose: biological, synthetic and hybrid nanopores. However, developing such a sequencing devices is still a challenge mainly because of remarkably fast translocation rate of ssDNA molecule through the nanopore ($\sim$ 1 nucleotide/μs). Current research has been focused on to the slowing down translocation of DNA molecule inside the pore. One possible solution can be the introduction of pore-polymer interactions. Also, these biopolymers and proteins are semiflexible in nature. However, most theoretical studies on polymer translocation assume completely flexible polymers. In this thesis, we study theoretically the driven translocation of a *semiflexible* polymer through narrow pores. The goal of this thesis is to study the effect of varying pore-polymer interactions, pore geometry and hydrodynamics on the translocation time statistics of semiflexible polymer to gain a deeper understanding of their fundamental role in polymer translocation and devise better sequencing strategies.

In the first problem, we study the *sequencing of semiflexible polymers of varying bending rigidity using patterned pores*. We first establish interplay between bending rigidity and pore polymer interactions for the translocation of a homogeneously semiflexible polymer. We then consider a heteropolymer made up of alternate stiff and flexible segments. We find that statistical fluctuations in the translocation time could be utilised for efficient sequencing of heteropolymers with varying bending rigidity, by suitably engineering pore-polymer interactions and combining readouts from multiple pores.

In the second problem, we focus on the role of pore geometry in the translocation

process. Experiments on translocation of a single stranded DNA through a protein channel MspA, which has a nearly conical geometry, indicate that such a pore is a promising candidate for nanopore DNA sequencing and other nanosensor applications. Here, we consider the *driven translocation of a semiflexible polymer through an interacting conical pore*. We study the effect of (i) the apex angle of the pore, (ii) the rigidity of the polymer, (iii) the stickiness of the pore, and (iv) the driving force, on the translocation time. We show that the translocation time shows interesting non-monotonic behavior as the pore geometry is altered by changing the apex angle of the pore.

In the third problem, we study flow driven translocation of a polymer through a narrow channel. Experimental and theoretical studies have indicated the importance of hydrodynamic interactions for driven polymer translocation through pores. Specifically, flow driven translocation involve the crossing of a free energy barrier set by the competition of the hydrodynamic drag and the entropic pressure due to the confinement of the polymer inside the pore. Theoretical and numerical studies indicate that the critical flow rate to overcome this barrier is independent of the length of the polymer and the pore geometry. We study *fluid flow driven translocation of semiflexible polymer chain through nanopore*. To incorporate hydrodynamics, we implement hybrid molecular dynamics-multiparticle collision dynamics algorithm. We show that the critical flow rate shows a striking dependence on the bending rigidity of the polymer as well as on the width of the pore.

We believe that in future, these studies on the dynamics of semiflexible polymer, will be useful for designing nanopore based devices for sequencing purpose and also understanding the physical aspects of biomolecular transport in different pore geometries.

Contents

List of Figures

Chapter 1

Introduction

Polymers play a very important role in life. The building blocks of life such as proteins, nucleic acids (DNA and RNA), actin filaments and microtubules are biopolymers. In many biological processes such as the passage of mRNA through nuclear pores [4], transfer of DNA across bacterial membranes during bacteriophage infection [5], translocation of proteins through sub cellular membranes [6], and viral injection of DNA into the host cells [4, 7–9], the biopolymer translocates from one region to another through a very narrow pore whose diameter is of the order of a few nanometers (10^{-9} m). This process is called polymer translocation.

The DNA contains genetic information and the knowledge of the genome, as given by the DNA sequence, is needed to have an understanding of the fundamental mechanisms of life occurring at the cellular level. The main difficulties in getting this information are the size of the genome, the speed and the cost of sequencing used by the current technology. The potential of polymer translocation as a low cost rapid DNA sequencing technique [10] has revolutionized the field. Apart from DNA sequencing, polymer translocation also finds applications in gene therapy and controlled drug delivery [11], and making devices that uses micro- and/or nano-fluidics for cheap medical investigations [12]. Due to these reasons, polymer translocation has been studied extensively in the last two decades.

Let us first discuss important concepts from polymer physics that will be needed to understand the physics of polymer translocation.

1.1 Polymer Physics

Polymers are large molecules made up of repeating structural units, called monomers, connected by covalent bonds [13]. The monomers contain groups of atoms which

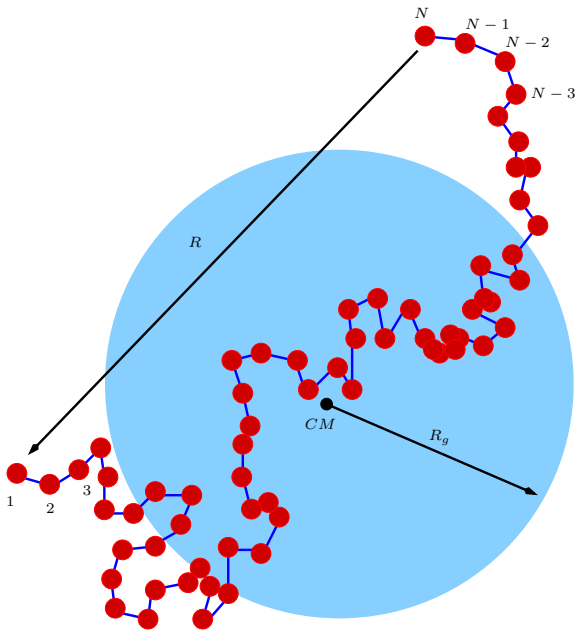

Figure 1.1: Schematic representation of a linear polymer chain with N monomer units, $\boldsymbol{R}$ is the end-to-end distance and $\boldsymbol{R}_g$ is radius of gyration. CM denotes the center of mass of the chain.

can be either identical (as in polyethylene) or chemically different (as in a protein or DNA). A polymer can take different conformations such as globular, coil-like, and rod-like, which depends on the chemical properties, nature of solvent, and the number of monomers present in the chain.

To describe the configuration of a polymer, one has to know the location of each monomer in space. The simplest model of a polymer chain is *Freely Jointed Chain* (FJC) model (Fig. 1.1), in which, the angular constraints between bonds and the excluded volume interactions are ignored. In this model, the chain consisting of N monomers and hence $(N-1)$ bonds, each of length b, is able to orient in any direction independent of each other. The conformation of the FJC is given by the set of N position vectors $\{\boldsymbol{R}_1, \ldots, \boldsymbol{R}_N\}$, or equivalently by the set of $(N-1)$ bond vectors, $\{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_{N-1}\}$, where

$$\boldsymbol{r}_i = \boldsymbol{R}_{i+1} - \boldsymbol{R}_i, \qquad i = 1, 2, \ldots, N-1.$$

The end-to-end vector of the chain is given by

$$\boldsymbol{R} = \boldsymbol{R}_N - \boldsymbol{R}_1 = \sum_{i=1}^{N-1} \boldsymbol{r}_i. \tag{1.1}$$

There are two important quantities which characterize the size of the polymer:

1. *Mean square end-to-end distance*: The mean square end-to-end distance is given as

$$\langle R^2 \rangle = \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \langle \boldsymbol{r}_i \cdot \boldsymbol{r}_j \rangle, \tag{1.2}$$

where, $\langle \cdot \rangle$ denotes average over an ensemble of chains with all possible conformations. For FJC model,

$$\langle R^2 \rangle = (N-1)b^2. \tag{1.3}$$

2. *Radius of gyration*: The square radius of gyration is defined as the average square distance between monomers in a given conformation and the center of mass of the polymer, $\boldsymbol{R}_{CM} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{R}_i$. It is obtained as:

$$R_g^2 = \frac{1}{N}\sum_{i=1}^{N} (\boldsymbol{R}_i - \boldsymbol{R}_{CM})^2 = \frac{1}{2N^2}\sum_{i=1}^{N}\sum_{j=1}^{N} (\boldsymbol{R}_i - \boldsymbol{R}_j)^2. \tag{1.4}$$

For FJC model, there exist a relation between the mean square end-to-end distance and the mean square radius of gyration in the limit of large N:

$$\langle R_g^2 \rangle \simeq \frac{\langle R^2 \rangle}{6}. \tag{1.5}$$

- If we ignore the constant of proportionality in the above equations, and use the generic symbol R, we can write

$$R \sim N^{\nu},$$

where R is like the average size of the polymer and ν is the size exponent. For FJC model, $\nu = 1/2$. For real polymers, due to the excluded volume interactions, two monomers cannot occupy the same space. In general, $\nu = 3/(d+2)$ is the size exponent in d dimensions, and is obtained by Flory within a mean field approach [13]. In one, two and three dimensions, the values of ν are 1, 3/4, and 3/5 respectively, which are very close to the experimentally [14, 15] observed values and numerical simulations [16]. In Fig. 1.2 we have plotted the end-to-end distance, R, and the radius of gyration, R_g, as a function of chain length N for a flexible polymer in three dimensions. A straight line fitting to the data gives us the size exponent as $\nu = 0.58 \pm 0.01$ which is close to the value predicted as 3/5. The simulation details are described later.

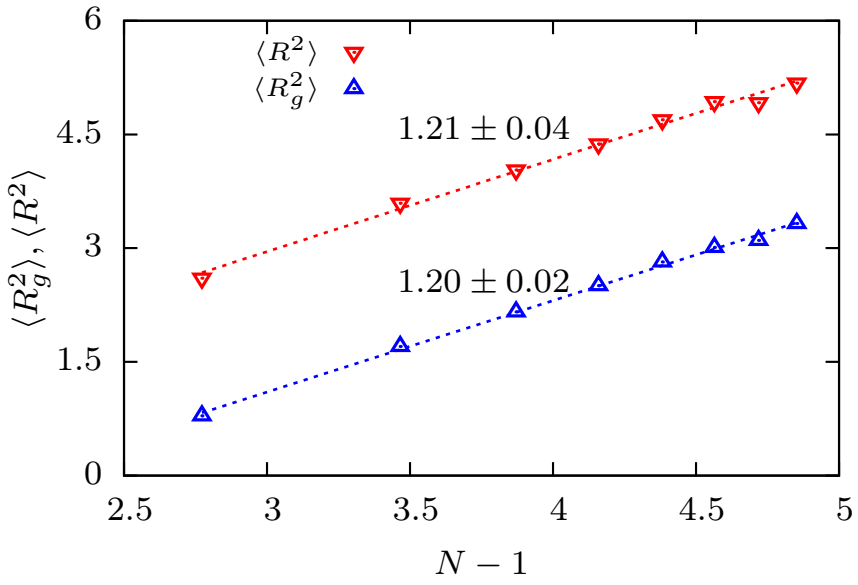

Figure 1.2: Left Panel: $\langle R_g^2 \rangle$ and $\langle R^2 \rangle$ as a function of chain length N.

- The total number of conformations, $\mathcal{Z}$, of a polymer chain with N monomers are given by

$$\mathcal{Z}(N) = \bar{z}^N N^{\gamma-1}, \tag{1.6}$$

where $\bar{z}$ is the effective coordination number for the orientation of the adjacent bonds, and γ is a critical exponent, which depends on the nature of the polymer, the background fluid, and any spatial restriction imposed on the polymer. The Helmholtz free energy F_N of the chain is then given by

$$\frac{F_N}{k_B T} = -\ln \mathcal{Z}(N) = -N \ln \bar{z} - (\gamma - 1) \ln N. \tag{1.7}$$

The second term in the above expression plays a crucial role in constructing the free energy landscape for polymer translocation.

- The FJC model fails to describe the behavior of a *semiflexible polymer* because of relaxation of the angular constraints. A semiflexible polymer is generally characterized by a large bending stiffness. This implies that there is a competition between bending energetics, which favor a straight conformation, and the entropic propensity of the polymer to crumple. To describe the physical properties of semiflexible polymers, the *worm-like chain* (WLC), also known as Kratky-Porod model, is used. The Hamiltonian for the semiflexible chain is

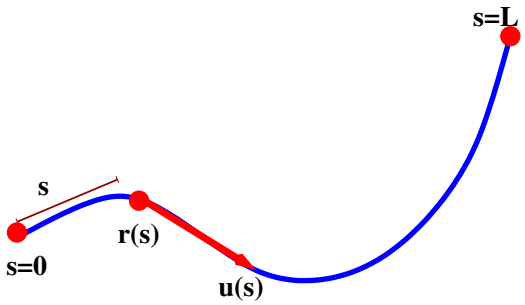

Figure 1.3: Space curve representation: Conformation of a worm like chain is specified by $\boldsymbol{r}(s)$, and $\boldsymbol{u}(s)$ is the local tangential vector.

given by [17]

$$H = \frac{1}{2}\kappa_b \int_0^L \left(\frac{\partial \boldsymbol{u}(s)}{\partial s}\right)^2 ds, \tag{1.8}$$

where, κ_b is the bending rigidity, L is the contour length, and s is measured along the contour from one of the chain ends ($0 < s < L$) (See Fig. 1.3). The tangential unit vector $\boldsymbol{u}(s)$ at s is given by

$$\boldsymbol{u}(s) = \frac{d\boldsymbol{r}(s)}{ds}, \tag{1.9}$$

which represents the local orientation of the segment in the semiflexible chain. The correlation between tangent vectors $\boldsymbol{u}(s)$ and $\boldsymbol{u}(s')$ of two segments at s and s', respectively, decreases to zero as:

$$\langle \boldsymbol{u}(s) \cdot \boldsymbol{u}(s') \rangle = \exp\left(-\frac{|s - s'|}{\ell_p}\right), \tag{1.10}$$

where ℓ_p is known as the persistence length of the semiflexible polymer. The worm like chain is described by two parameters, the persistence length, ℓ_p, and the contour length L. The persistence length can also be expressed as

$$\ell_p = \frac{2\kappa_b}{(d-1)k_B T}, \tag{1.11}$$

where T is the absolute temperature, d is the dimension and k_B is the Boltzmann constant.

The mean square end-to-end distance for a semiflexible polymer is given by

$$\langle R^2 \rangle = 2\ell_p L \left[1 - \frac{\ell_p}{L}\left(1 - e^{-L/\ell_p}\right)\right]. \tag{1.12}$$

The limiting cases of large persistence length with respect to contour length,

and vice versa, follows from the above formula

$$\langle R^2 \rangle = \begin{cases} 2\ell_p L \equiv bL, & \ell_p \ll L \\ L^2, & \ell_p \gg L, \end{cases} \tag{1.13}$$

where $b = 2\ell_p$ is defined as *Kuhn segment length* which is twice the persistence length of the polymer. Hence, it is also a measure of the polymer chain stiffness.

Therefore, the WLC chain reduces to a rod-like polymer in the limit $\ell_p \gg L$, and a FJC for $\ell_p \ll L$. Similar to Eq. 1.13, the limiting cases for the radius of gyration of a WLC are given by

$$\langle R_g^2 \rangle = \begin{cases} bL/6, & \ell_p \ll L \\ L^2/12, & \ell_p \gg L. \end{cases} \tag{1.14}$$

Therefore, for a semiflexible polymer, depending on the ratio of the persistence length to the chain length, the same polymer behaves differently and have different size exponents.

1.2 Physics of polymer Translocation

We are now in a position to see what happens when a polymer situated in the left compartment (*cis* side) tries to go to the right compartment (*trans* side) through a narrow pore that connects the two compartments (See Fig. 1.4). The chain on the *cis* side can have a large number of conformations $\mathcal{N}$ and therefore the chain has large entropy. The free energy of the system is given by $F = E - TS = E - k_B T \ln \mathcal{N}$, where E is the energy of interaction between monomers and surrounding molecules. When such a chain is placed in a restricted environment, such as inside a narrow pore, the number of possible conformations are reduced. Hence, the chain entropy decreases and the free energy of the chain increases [1]. This introduces a free-energy barrier in the translocation process. The polymer must overcome this barrier for a successful translocation to the *trans* side (right compartment). Let F_L, F_R, and F_P represent the free energies of the polymer in the left compartment, right compartment, and inside the pore, respectively. Generally, $F_R \neq F_L$ because the left and right compartments may not have similar environments. The polymer has to overcome a barrier of height $(F_P - F_L)$ to successfully complete the translocation process. The barrier height can be affected by introducing pore-polymer interaction. An external driving force is also useful to overcome the barrier.

For the polymer which is placed in an ionic solution to complete the translocation

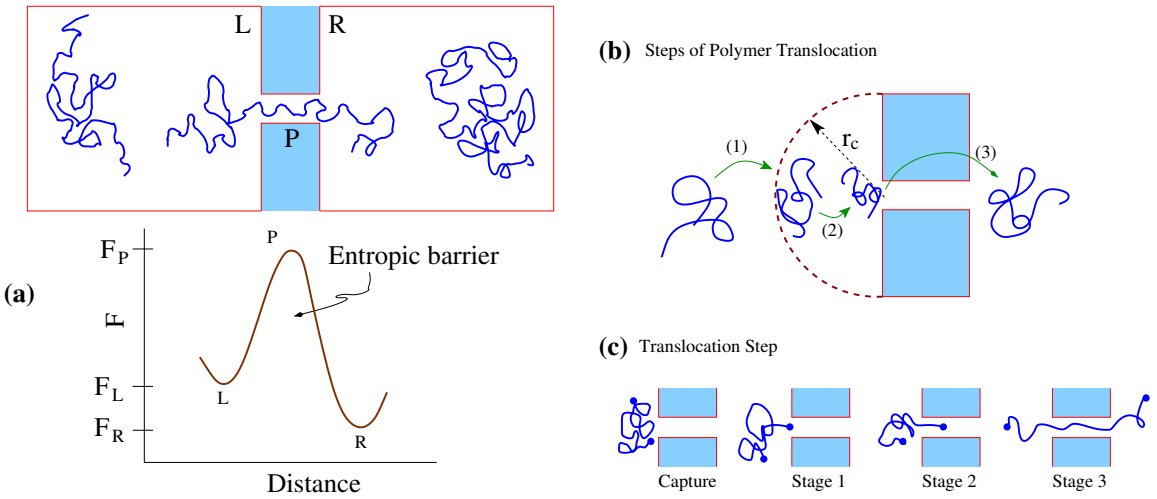

Figure 1.4: (a) Idea of the entropic barrier in polymer translocation. (b) Various steps of polymer translocation process: (1) drift-diffusion, (2) capture, and (3) translocation. (c) Three stages of the translocation step. Adapted from Ref [1]

process from the left compartment to the right across the nanopore (Fig. 1.4(b)), has to go through three stages [1]: (1) drift-diffusion, (2) capture, and (3) translocation. In the first stage, the polymer which is away from the pore entrance diffuses in the solvent and could also drift depending on the presence of external fields. In the second stage, the polymer which has drift-diffused close to the pore entrance, faces an effective sucking force arising from potential gradients at the pore entrance. Once the polymer is captured, it still has to come out of the pore exit, which involves overcoming the entropic barrier. This third stage (Fig. 1.4(c)) is usually facilitated by pore-polymer interactions, pore geometry and external forces in the form of field gradients. In this thesis, we will ignore the first stage of drift-diffusion of the polymer. In all the cases we study, one end of the polymer will always be placed at the pore entrance, so that the capture and subsequent translocation stages are enforced.

1.3 Polymer Translocation through nanopore: A brief survey

In the last two decades, the polymer translocation has attracted considerable attention both theoretically and experimentally. In this section, we give a very brief account of relevant experimental and theoretical work.

1.3.1 Experimental studies

The first experimental study of translocation was done by Wallace Coulter in 1940 to count blood cells by passing them through a hole of size roughly 10 μm, slightly larger than the size of blood cells [18]. His apparatus consisted of two compartments separated by a hole. The compartments were filled with an electrolyte solution and a voltage difference was applied. As the cells passed through the hole, the ionic current through the hole drops which is proportional to the volume of cells. Experimental studies on polymer translocation have largely used three different types of pore [19]: (i) biological pores, (ii) solid state nanopores, and (iii) hybrid of the two.

1.3.1.1 Biological nanopores

Biological nanopores are formed by pore forming proteins. The most frequently used biological pores are α-haemolysin and Mycobacterium Smegmatis porin A (MspA) [20]. In α-haemolysin pore, the channel is made up of a 3.6 nm diameter vestibule or cap, which is connected to a 2.6 nm wide and 5 nm long transmembrane barrel like structure called the β barrel. The diameter of the region where vestibule and the β barrel join is just 1.4 nm. Therefore, only a single stranded DNA molecule can pass through this pore. The side and top view of α-haemolysin pore is shown in Fig. 1.5(A) (Left panel).

In 1996, Kasianowicz, Brandin, Branton and Deamer [21] used a transmembrane protein, an $\alpha-$hemolysin channel, as a pore and translocated a homopolymer RNA molecule. They used the same principle of Coulter and showed that the sequence of nucleotide in RNA could be identified by observing the ionic current across the pore. Using the same technique, they were able to show that it is possible to distinguish between two different types of block copolymers $A_{30}C_{70}$ by monitoring both the amplitude and duration of the current signal. The translocation rate of DNA molecule through $\alpha-$hemolysin pore is remarkably fast: the time spent by a nucleotide inside the pore is $\sim 1\mu s$, which means that very few ions (around 100) are passing through the pore. Therefore, the difference in the ionic current of different bases are overwhelmed by the statistical fluctuations. Most of the work reported on this pore is to regulate the ssDNA transport through the nanopore [22, 23].

One drawback of α-haemolysin is that it cannot distinguish between 10–15 nucleotides which are present inside the pore. As a result, the ionic current statistics of each nucleotide becomes diluted and hence reducing the overall signal-to-noise ratio in sequencing applications. Another example of a biological nanopore is the Mycobacterium Smegmatis porin A (MspA) [24] which has a conical shape. MspA

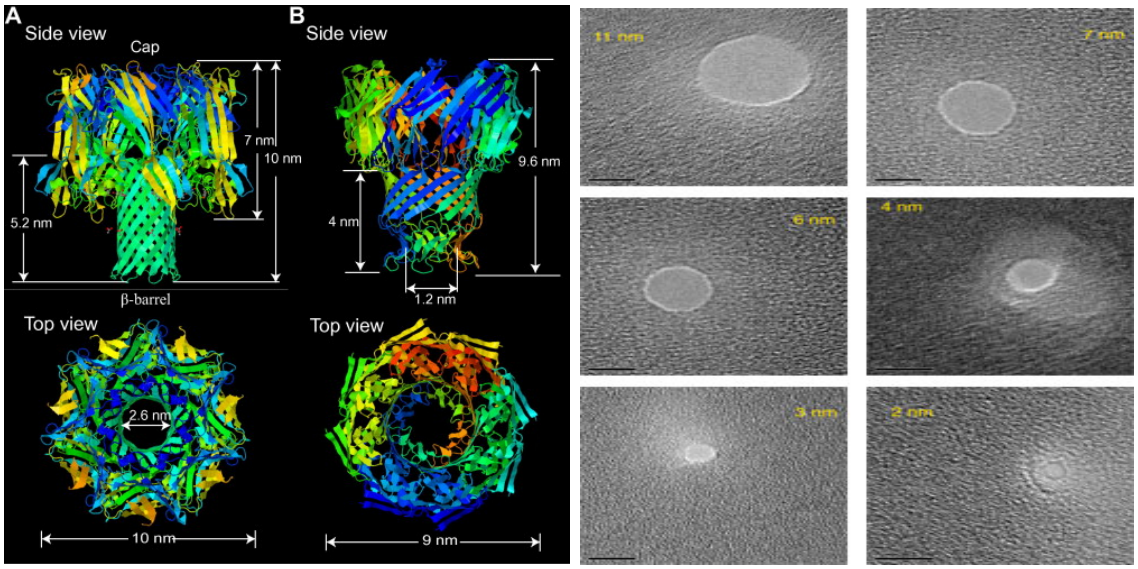

Figure 1.5: Left Panel: Ribbon representations of the $\alpha-$hemolysin protein and MspA channel. (**A**) Side and top view for $\alpha-$hemolysin channel. (**B**) Side and top view for MspA channel [2]. **Right Panel:** Different pore diameter formation on SiN membranes by using high intensity electron beam of a field-emission Transmission Electron Microscope (TEM) [3].

has constriction diameter $\sim$ 1.2nm (see Fig. 1.5(B) Left Panel). In MspA, only 3 nucleotides which are close to the constriction, are simultaneously affected by the current and hence current statistics of the four bases are sufficiently wider. This pore also faces the same challenges, like very fast translocation of ssDNA and the difficulty to sequence in real time. Another important example for biopores is the phi29 viral packaging motor, which was used to transporting dsDNAs [25].

The biological nanopores have both advantages and disadvantages:

- The main advantages of using biological nanopores are the atomic precision of their assembled structure and that they can be tuned easily through genetic modifications [26]. Another advantage of biological pores are that they rarely exceed 2nm in diameter and allow translocation of single-stranded DNA or RNA as well as unfolded protein chains.

- The main disadvantage of biological nanopores are the fixed size of the pore, loss of stability when experimental conditions such as pH, temperature or salt concentration are changed, and difficulty in integrating them into large-scale arrays.

1.3.1.2 Synthetic nanopores

Synthetic nanopores were designed to address some of the difficulties posed by biological nanpores. The disadvantages of biological pores paved the path for engineering of synthetic nanopores. These pores are more stable, the pore diameter (1-20 nm) can be changed with nanometer precision, have tailored surface properties, and can be very easily integrated to large-scale arrays. The solid state nanopores are generally fabricated by making a very narrow hole in thin silicon compound (such as SiN) membranes, by ion beam sculpting and electron beam. In Fig. 1.5(Right Panel) the solid state nanopores of various diameters (ranging from 1-20 nm) are shown [3].

In 2001, Li *et al.* [27] were able to fabricate a solid-state nanopore having diameter of 1.8 nm using ion-beam sculpting technique which they used to detect the passage of DNA molecules. In 2003, Storm *et al.* [28] gave a technique to fabricate solid state nanopores by drilling hole through a silicon membrane using electron beam lithography. The translocation time of long DNA molecules (up to 97000 base pairs) through a solid state nanopore was measured by Dekker's group [29] in 2005. More recently, graphene has been used as a material for solid state nanopores [30] due to its remarkable mechanical, electrical and thermal properties. The translocation speed of polymer through nanopores is quite large and needs to be controlled. To achieve this, different groups have used different techniques. For example, Keyser *et al.* [31] used optical tweezers, while Movileanu's group [32] used electrostatic trapping.

1.3.1.3 Hybrid nanopores

A major drawback of solid state nanopore is the lack of chemical specificity – it can not differentiate between chemical species of nearly same size. This drawback is overcome by a recent technique in which a protein pore is attached to solid state nanopore to create a hybrid nanopores. For example, a genetically engineered α-haemolysin protein has been attached to the nanopore in a SiN membrane [33].

1.3.2 Theoretical studies

The experimental study of Kasianowicz *et al.* [21] inspired several theorists to develop analytical and numerical approaches to the problem. In this section, we briefly discuss some fundamental theoretical studies and focus on the ones relevant to this thesis.

Theoretical studies on polymer translocation mostly focuses on the statistics of translocation time, τ, i.e. the time taken by the polymer to successfully move from the *cis* to the *trans* side of the pore. Sung and Park [34] studied the translocation

of a Gaussian polymer through a pore in the absence of any external force in the equilibrium limit. Their approach was to treat translocation dynamics as a diffusion process across a free energy barrier. It was found that the translocation time τ scales with the chain length N as $\tau \sim N^3$ and $\tau \sim N^{2.5}$ for chain obeying Rouse and Zimm dynamics, respectively. Muthukumar [35,36] also considered translocation as a diffusion process across a free energy barrier and used nucleation theory arguments to solve the mean first-passage time problem for the chain through the channel. He found that the translocation time for a self avoiding chain scales as $\tau \sim N^2$. It was argued by Chuang *et al.* [37] that the translocation time for unbiased translocation cannot be smaller than the Rouse equilibration time ($\tau_R \sim N^{1+2\nu}$), where ν is the Flory exponent ($\nu_{3D} = 0.58$ and $\nu_{2D} = 0.75$), and showed using Monte Carlo simulations that τ scales as $\tau \sim N^{1+2\nu}$.

In the case of a driven translocation, an externally applied force F is also a parameter in the scaling relation along with the chain length N. In majority of experiments, the source of driving force is an applied electric field. However, there are some experiments that considers the bias due to the interactions between the attractive particles present in the *trans* side and the translocating polymer. In most theoretical studies of driven translocation a constant force at the pore is put by hand irrespective of its origin.

The scaling laws for the forced translocation of very long polymer chains were derived by Storm *et al.* [29]. They argued that depending on the polymer length, there are different dominating frictional regimes. For short chains, the friction of the chain is dominated by the friction inside the pore, and the translocation time scales linearly with N. For longer chains, the hydrodynamic drag of the subchains outside the channel dominates over the friction inside the channel, and the translocation time scales as $\tau \sim R_g^2 N^{2\nu}$ or $\tau \sim N^{1+\nu}$ for Rouse or Zimm dynamics, respectively. This was consistent with their experimental translocation exponent of 1.27 measured for long dsDNA.

Luo, Ala-Nissila and collaborators [38–40] present results from Monte Carlo and Langevin dynamics simulations of biased and unbiased polymer translocation. For the unbiased translocation, the chain is initially placed halfway through the channel and can escape on either side of the pore. The translocation time for both the models scales as $\tau \sim N^{1+2\nu}$ in the absence of a driving force. For an unbiased translocation from a very long pore as compared to the chain length, it was found that $\tau \sim N$. In the presence of an external driving force they observed two scaling laws with both models: (i) $\tau \sim N^{2\nu}$ for relatively short chains, and (ii) $\tau \sim N^{1+\nu}$ for longer chains. It was also observed that the biased translocation time is inversely proportional to the field

strength. There have been a plethora of later studies predicting different exponents using arguments like dynamical scaling [37, 41], mass and energy conservations [42] and tension propagation (TP) along the length of the polymer [43–45]. The TP theory, introduced originally by Sakaue [43] for an infinite chain and subsequently modified by Ikonen *et. al.* [46, 47] and Dubbeldam *et. al.* [48] to finite chains have proved to be successful in explaining the non-equilibrium facets of driven translocation. Ikonen *et. al.* [46] introduced the Brownian dynamics tension propagation theory (BDTP), which correctly accounted for the role of pore friction and thermal fluctuations due to the solvent and their effects on the scaling exponent. Further, it explained the various values of the scattering exponents observed in previous studies, thereby providing a unifying picture of polymer translocation. The BDTP theory was recently modified with a constant monomer *iso-flux* approximation by Sarabadani *et. al.* [49, 50], which leads to self consistent theory for polymer translocation with effective pore friction as the only free parameter. Bhattacharya and others [51–54] studied translocation dynamics of a semiflexible polymer and showed the dependence of mean translocation time on the stiffness of the polymer. They showed from their simulations that the underlying TP picture along the polymer backbone is valid for semiflexible chains as well.

A large number of the results discussed above were for pores where the interactions of the polymer with the pore were negligible. The translocation process is inherently stochastic in nature and poses a challenge to devise efficient sequencing strategies without compromising on the speed of the translocation process. Solid state nanopores with tailored surface properties [28, 55] make it possible to regulate the interactions of the polymer with the pore as well as reducing noise [56–60]. Theoretically, this has been achieved by typically looking at extended channels where the interaction of the pore with the channel is tuned. Luo *et. al.* [61–63] showed that the mean translocation time of a polymer across an attractive channel increases with the strength of attraction. This suggests a possibility to separate polymers with varying interactions with the pore. Furthermore, the translocation dynamics of a heterogeneous polymer show a strong dependence on the sequence. The heterogeneity has been introduced in a variety of ways. Luo *et. al.* [64] considered heteropolymers which are distinguished by the driving force they experience inside the pore. The residence time of each bead inside the pore was found to be a strong function of the sequence. Mirigian *et. al.* [65] considered polymers with differing frictional interaction with the pore and charge. The mean translocation time of the multiblock polymers depends on the fraction as well as the arrangement of the blocks. At a certain optimum length of the charged block, the mean translocation rate is the slowest.

de Haan *et al.* [54] considered the chain heterogeneity based on their bending rigidity (rod-coil polymers). They observed distinct plateaus for the rod-coil polymer in the progression towards complete translocation. They showed that, in the quasi-static limit where pore-friction dominated translocation (low viscosity), the stiff segment translocates faster than the flexible ones. Later, Adhikari *et al.* [53] studied the effect of these rod-coil segments on the total translocation time and found that at high viscosity the stiffer segments translocate slower than the flexible segments.

Recent theoretical studies [66–68] considered channels with varying pore-polymer interactions along its length. The translocation time distributions showed significant variations across the differently decorated channels. Katkar and Muthukumar [68] showed that translocation time across a nanpore of alternate charged and uncharged sections, depends non-monotonically on the length of the charged section. In the studies by Cohen *et. al.* [66, 67], it was shown that the statistical fluctuations in the translocation time could be utilised for efficient sequencing of heteropolymers. We extend the idea in the context of semiflexible polymers with varying bending rigidity along the polymer backbone in Chapter 2 of the thesis.

Recently, Nikoofard *et al.* [69,70] has studied the translocation of a flexible polymer through a cone shaped channel in the absence of external driving. They start with an initial configuration of the chain in which the monomers are arranged along the axis of the conical channel in such a way that the two parts of the chain, which are outside the pore, are of equal lengths. It was found that due to the asymmetry of the channel, an effective driving force of entropic origin acts on the polymer such that the polymer always translocates from the end with larger diameter. The authors used local equilibrium assumptions to obtain the equilibrium free energy for the polymer inside the channel. They found the translocation time is nonmonotonic function of the apex angle of the channel. However, translocation processes mostly occur in the presence of an external drive and with significant pore-polymer interactions. We present a study of such a situation through a conical nanopore in Chapter 3 of this thesis.

Recent theoretical works of Sakaue and collaborators [71–73] have reported the translocation of the polymer through narrow pores in the presence of fluid flow. It was found that the translocation is controlled by the competition between the entropic and the hydrodynamic forces. In 2005, Sakaue *et. al.* [71] derived a theoretical expression for the threshold velocity flux of the fluid flow based on the de Gennes polymer blob model [74]. Below the threshold flux the probability of translocation of the polymer is zero. Furthermore, it was claimed that the threshold velocity flux for a linear polymer chain is independent of the length of the chain, and also of

the dimensions of the pore, but depends only on the temperature of the system and the viscosity of the fluid. In recent years, several mesoscopic techniques have been developed to simulate the dynamics of fluid particles such as, lattice gas automata (LGA) [75], lattice Boltzmann (LB) [76], dissipative particle dynamics (DPD) [77], multiparticle collision dynamics (MPCD) [78]. In 2009, Yeomans *et al.* [73] obtained the threshold velocity flux using lattice Boltzmann simulation, which which were in very good agreement to the theoretical prediction of Sakaue. In 2010, Nikoubashman and Likos [79] showed that this expression is also valid for the branched polymer. In 2012, Ledesma-Aguilar *et al.* [80] showed that this expression is valid only for polymers with longer chain lengths. In the case of smaller chain lengths, it was found that the threshold velocity flux depends on the polymer chain length. All the studies reported above are for flexible polymer chain. It is of interest to know how the threshold velocity flux behaves if the underlying polymer in the fluid flow is semiflexible. Therefore, we attempt to study the effect of semiflexiblity on the threshold velocity flux in the presence of fluid flow using MPCD, which we presented in Chapter 4 of this thesis.

1.4 Simulation techniques

We used the bead-spring model proposed by Grest and Kremer [81] in our simulations. In this model, each bead on the polymer chain represents a group of molecules, such as nucleotides or nucleic acid. The consecutive beads are connected by a spring which mimics the bond between the respective groups. In addition, there is a bending potential to account for semiflexiblity of the chain.

1.4.1 Interaction potentials

The beads of the polymer experience an excluded volume interaction modeled by the Weeks-Chandler-Andersen (WCA) potential of the form

$$U_{\text{bead}}(r) = \begin{cases} 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] + \epsilon & \text{for } r \leq r_{min} \\ 0 & \text{for } r > r_{min}, \end{cases} \tag{1.15}$$

where, ϵ is the strength of the potential and σ is the diameter of the polymer bead. The cutoff distance, $r_{min} = 2^{1/6}\sigma$, is set at the minimum of the potential. Consecutive monomers in the chain interact via the finite extension nonlinear elastic (FENE)

potential of the form

$$U_{\text{bond}}(r) = -\frac{1}{2}kR_m^2 \ln\left(1 - \frac{r^2}{R_m^2}\right), \tag{1.16}$$

where k is the spring constant which controls the strength of the bond, and R_m is the maximum allowed separation between the consecutive monomers of the chain.

The bending potential is obtained by discretizing the WLC Hamiltonian

$$U_{\text{bend}} = \frac{\kappa_b}{2\sigma}\sum_{i=1}^{N-2}[\mathbf{t}_{i+1} - \mathbf{t}_i]^2, \tag{1.17}$$

where κ_b is the bending rigidity of the polymer, σ is the equilibrium bond length and $\mathbf{t}_i = [\mathbf{r}_{i+1} - \mathbf{r}_i]/b_i$ is the local tangent. Here, $b_i = |\mathbf{r}_{i+1} - \mathbf{r}_i|$ is the instantaneous bond length. κ_b represents the stiffness of the polymer, and in two dimensions it is related to the persistence length as $\kappa_b/k_BT = l_p/2$, where k_B is the Boltzmann's constant and T is the temperature.

Attractive interaction between the polymer monomers and the pore beads are modeled by the standard LJ potential of the form:

$$U_{\text{pore}}(r) = \begin{cases} 4\epsilon_{\text{pore}}\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] & \text{for } r \le r_c \\ 0 & \text{for } r > r_c, \end{cases} \tag{1.18}$$

where ϵ_{pore} denotes the potential depth and $r_c = 2.5\sigma$ is the cutoff distance. Repulsive interaction between pore beads and the polymer ($U_{\text{pore}}(r)$) is the same as the WCA interaction (U_{bead}). The interaction between the wall beads and of the polymer (U_{wall}), is also the same as the WCA interaction (U_{bead}).

1.4.2 Langevin Dynamics

To integrate the equation of motion for the monomers of the polymer, we use Langevin dynamics. The Langevin equation is a stochastic differential equation to describe the erratic motion of a Brownian particle colliding with molecules in the fluid medium in which it moves [82]. Two additional forces are introduced in the Newton's equations. The first is a viscous force arising from the friction experienced by the particles colliding with the fluid particles and is proportional to the velocity of the particle. The other force is the random force due to the collision whose time average is zero. The equation of motion for an individual particle obeying Langevin dynamics is then

written as:

$$m_i \frac{d^2 \boldsymbol{r}_i}{dt^2} = -\boldsymbol{\nabla} U_i - \zeta \boldsymbol{v}_i + \boldsymbol{f}_{ext} + \boldsymbol{\eta}_i = \boldsymbol{F}_i. \tag{1.19}$$

In our system, $\boldsymbol{r}_i$ and $\boldsymbol{v}_i$ the position and velocity vectors, respectively, of the ith monomer with mass m_i. The first term on the right hand side of Eq. 1.19 is the total force caused by different potentials U_i experienced by the ith monomer,

$$U_i = U_{\text{bend}} + U_{\text{bond}} + U_{\text{bead}} + U_{\text{wall}} + U_{\text{pore}}. \tag{1.20}$$

The second term is the frictional force experienced by the monomer with ζ as friction coefficient, the third term $\boldsymbol{f}_{ext}$ is the externally applied force, and the last term $\boldsymbol{\eta}_i$ is the random force which satisfy the fluctuation-dissipation theorem

$$\langle \boldsymbol{\eta}(t) \rangle = 0 \qquad \text{and} \qquad \langle \boldsymbol{\eta}_i(t) \cdot \boldsymbol{\eta}_j(t') \rangle = 2\zeta k_B T \delta_{ij} \delta(t - t'), \tag{1.21}$$

where, T is the temperature of the fluid bath.

There are many methods to numerically solve Eq. 1.19. In the absence of friction and noise terms in the force field of Eq. 1.19, one of the most popular method to obtain monomers trajectories $\{\boldsymbol{r}_i(t), \boldsymbol{v}_i(t)\}$ over time, is velocity Verlet algorithm. According to this algorithm, given the positions and velocities at time t, the position and velocity for the next time step $t + \Delta t$ are updated as:

$$\boldsymbol{r}_i(t + \Delta t) = \boldsymbol{r}_i(t) + \Delta t \boldsymbol{v}_i(t) + \frac{1}{2m} \Delta t^2 \boldsymbol{F}_i(t), \tag{1.22}$$

and

$$\boldsymbol{v}_i(t + \Delta t) = \boldsymbol{v}_i(t) + \frac{1}{2m} \Delta t [\boldsymbol{F}_i(t) + \boldsymbol{F}_i(t + \Delta t)], \tag{1.23}$$

where $\boldsymbol{F}_i$ denotes the total force term in the R.H.S. of Eq 1.19.

To solve the Langevin equation of motion in presence of friction and noise term numerically, we used Ermak's algorithm [83]. Ermak's algorithm treats both systematic dynamics and stochastic elements of the Langevin dynamics (Eq. 1.19) properly. In this approach, updates of the particle positions and velocities are given as

$$\mathbf{r}_i(t + \delta t) = \mathbf{r}_i(t) + c_1 \delta t \mathbf{v}_i(t) + c_2 \delta t^2 \mathbf{a}_i(t) + \delta \mathbf{r}_i^G, \tag{1.24}$$

$$\mathbf{v}_i(t + \delta t) = c_0 \mathbf{v}_i(t) + (c_1 - c_2) \delta t \mathbf{a}_i(t) + c_2 \delta t \mathbf{a}_i(t + \delta t) + \delta \mathbf{v}_i^G, \tag{1.25}$$

where $\mathbf{r_i}, \mathbf{v_i}, \mathbf{a_i}$ are the position, velocity and acceleration, respectively, of particle i and the numerical coefficients are given as $c_0 = e^{-\zeta \delta t}$, $c_1 = (\zeta \delta t)^{-1}(1 - c_0)$, $c_2 =$

$(\zeta\delta t)^{-1}(1-c_1)$, and $\delta\mathbf{r}_i^G, \delta\mathbf{v}_i^G$ are random variables. Each pair of vectorial components of $\delta\mathbf{r}^G, \delta\mathbf{v}^G$ is drawn from a bivariate Gaussian distribution with zero mean and variances σ_r and σ_v as follows.

$$\sigma_r{}^2 = \frac{mk_BT}{\zeta^2}[2\zeta t/m - 3 + 4e^{-\zeta t/m} - e^{-2\zeta t/m}], \tag{1.26}$$

$$\sigma_v{}^2 = \frac{k_BT}{m}(1 - e^{-2\zeta t/m}). \tag{1.27}$$

The correlation coefficient c_{rv} is given by

$$c_{rv}\sigma_r\sigma_v = \frac{k_BT}{\zeta}(1 - e^{-2\zeta t})^2. \tag{1.28}$$

Once $\delta\mathbf{r}_i^G, \delta\mathbf{v}_i^G$ are drawn, we calculate $\mathbf{r}_i(t + \delta t)$ and $\mathbf{v}_i(t + \delta t)$ and the algorithm proceeds from one time step to the next.

A polymer in the *cis* side compartment would diffuse and take a significant time to reach the pore entrance. As explained before, we will not be interested in this stage of the translocation process. Rather, we start with a chain configuration with its first bead placed at the entrance of the pore. This bead is fixed while the remaining beads of the chain are allowed to fluctuate to obtain the initial equilibrium chain conformation. The first bead is then released and the potential gradients at the pore entrance coming from external forces or pore-polymer interactions sucks the polymer inside the pore. Then the translocation of the polymer continues.

1.4.3 Multi-Particle Collision Dynamics

In Langevin dynamics, although the effect of the fluid bath is present in the form of the frictional and random forces, the solvent particles are not modeled explicitly. In the study of polymer translocation occurring in the presence of fluid flows, the solvent needs to be incorporated explicitly. This is done in this thesis using Multi-Particle Collision Dynamics (MPCD), also known as Stochastic Rotation Dynamics (SRD). MPCD is a mesoscopic simulation technique to solve linearized Navier-Stokes equations for fluid particles conserving local mass, energy and momentum, which satisfies hydrodynamic equations [78,84]. It is a particle based method in which both positions and velocities are continuous variables, but the time evolves in discrete steps. It is important to mention that particle here do not necessarily mean the actual particles of fluid but, can be interpreted as energy and momentum carriers. In MPCD, the solvent particles are modeled by N point-like particles, each of mass m, which move in space and time.

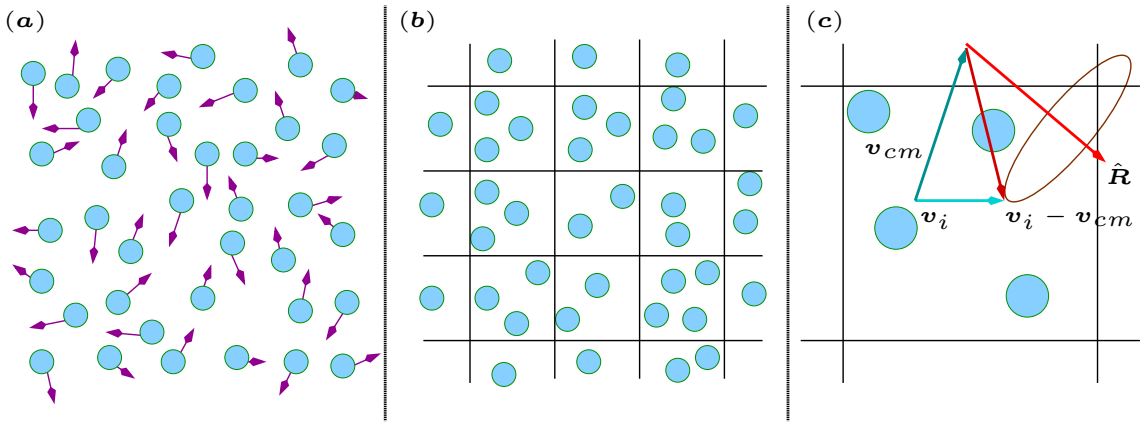

Figure 1.6: Schematic diagram of MPCD algorithm. **(a)** Streaming Step: Particles moving ballistically. **(b)** Collision Step: Sorting of particles inside the cells. **(c)** Rotation of relative velocity vector about random axis by an angle α

The MPCD algorithm consists of two steps: (1) Streaming, and (2) Collision. In the streaming step, the particles move ballistically (Fig. 1.6 (a)), and in time step h, their positions change according to

$$\boldsymbol{r}_i(t+h) = \boldsymbol{r}_i(t) + h\boldsymbol{v}_i(t). \tag{1.29}$$

Therefore, in this step, the particles move without interacting with each other. During the streaming step, additional external forces can be applied to the particles if needed. If there is gravitational force, then an extra term $(1/2)\boldsymbol{g}h^2$, where $\boldsymbol{g}$ is the gravitational force, will be added to Eq. 1.29 to simulate a gravity-driven flow.

The streaming step is followed by the collision step. In this step, the particles are first sorted into cubic cells of side length a ($a = 1$) (Fig. 1.6 (b)). Then the velocity $\boldsymbol{v}_i$ of each particle relative to the center of mass velocity $\boldsymbol{v}_{c.m.}$, i.e., $(\boldsymbol{v}_i - \boldsymbol{v}_{c.m.})$, is rotated about a random axis chosen independently in each cell (Fig. 1.6 (c)). The velocity of ith particle in the cell is given by,

$$\boldsymbol{v}_i(t+h) = \boldsymbol{v}_{c.m.}(t) + \mathcal{R}(\alpha)(\boldsymbol{v}_i(t) - \boldsymbol{v}_{c.m.}(t)), \tag{1.30}$$

where $\mathcal{R}(\alpha)$ is a stochastic rotation matrix with rotation angle α, which is fixed throughout the simulation. In this model, the local momentum and kinetic energy is conserved. In two-dimensions, the rotation of relative velocities are carried by an angle $\pm\alpha$, where the sign is chosen randomly for each cell. In three-dimensions, there are various schemes for performing rotation [85, 86]. The one we use in this thesis consists of choosing a random direction in space for each cell, around which the relative velocities are rotated by an angle α. This scheme is shown in Fig. 1.7(a). If a randomly chosen unit axis is denoted by $\hat{R} = [\hat{R}_x, \hat{R}_y, \hat{R}_z]$, then the stochastic

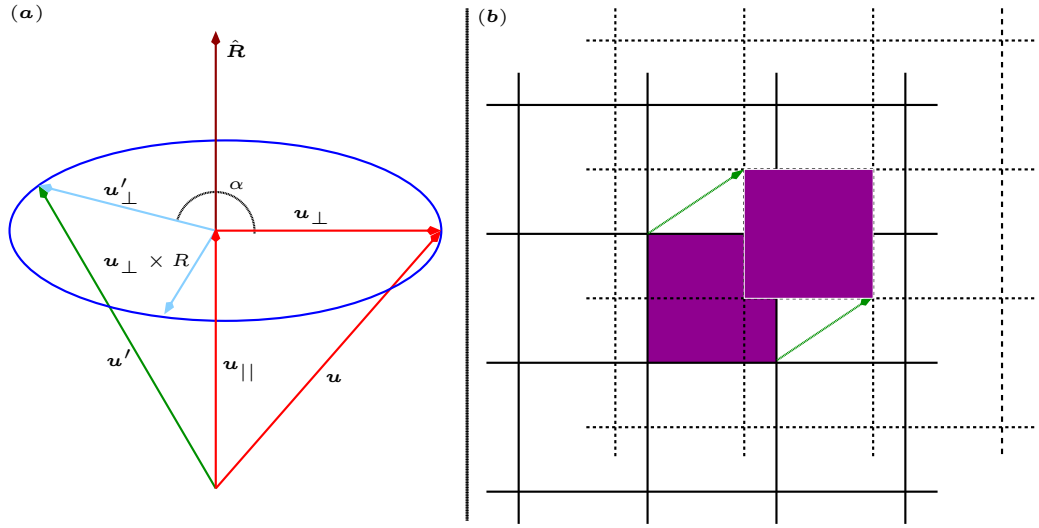

Figure 1.7: **(a)** Rotation of vector $\boldsymbol{u} = (\boldsymbol{v} - \boldsymbol{v}_{c.m.})$ around the unit vector axis $\hat{R}$ by an angle α and get new vector after rotation as $\boldsymbol{u}' = \mathcal{R}(\alpha)\boldsymbol{u}$. **(b)** Shifting of cell randomly by $[-a/2 : a/2]$.

rotation matrix $\mathcal{R}(\alpha)$ is given by

$$\mathcal{R}(\alpha) = \begin{bmatrix} \hat{R}_x^2 + (1-\hat{R}_x^2)\cos\alpha & \hat{R}_x\hat{R}_y(1-\cos\alpha) - \hat{R}_z\sin\alpha & \hat{R}_x\hat{R}_z(1-\cos\alpha) + \hat{R}_y\sin\alpha \\ \hat{R}_x\hat{R}_y(1-\cos\alpha) + \hat{R}_z\sin\alpha & \hat{R}_y^2 + (1-\hat{R}_y^2)\cos\alpha & \hat{R}_y\hat{R}_z(1-\cos\alpha) - \hat{R}_x\sin\alpha \\ \hat{R}_x\hat{R}_z(1-\cos\alpha) - \hat{R}_y\sin\alpha & \hat{R}_y\hat{R}_z(1-\cos\alpha) + \hat{R}_x\sin\alpha & \hat{R}_z^2 + (1-\hat{R}_z^2)\cos\alpha \end{bmatrix}. \tag{1.31}$$

For the algorithm to satisfy Galilean invariance, the grid is randomly shifted in each time step [87, 88]. The shifting of grid redistributes particles in the neighboring cells and ensures that the particles do not have memory of previous collisions and are uncorrelated in space and time. Such memory effects can occur, when the same particles remain in the particular cell for long time and therefore collide among themselves repeatedly. This effect is more prominent when the time step is very small. The range of shifting is $[-a/2 : a/2]$ (Fig. 1.7 (b)). If the mean free path λ is larger than $a/2$, the violation in Galilean invariance is negligible, and grid shift is not necessary.

The following boundary conditions are imposed depending on the system under consideration:

- If no walls are present in the system, the conventional periodic boundary conditions are used.
- If walls are present, the no-slip boundary conditions are used. This is achieved by the simple bounce-back rule.

When the walls are present in the system, a few cells that are near to the walls, remain partially filled after the grid shifting, and the simple bounce-back rule fails

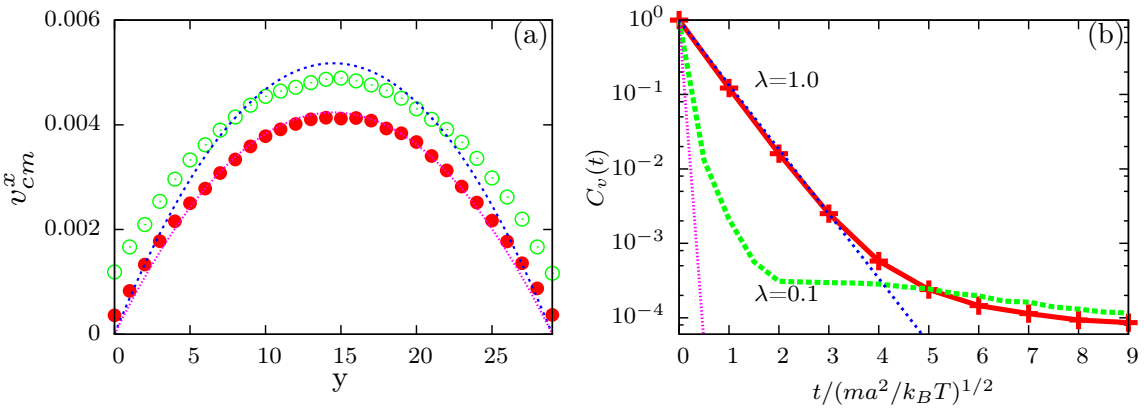

Figure 1.8: **(a)** Poiseuille flow through a channel of size height, H=30 and length L=50, for $\rho = 35$. $k_BT = 0.01275$, corresponding to $\lambda = h(k_BT/m)^{1/2}$, with $h = m = 1$ in 2D. Open circles are for simple bounce back rule and filled circles for generalized bounce-back rule. The dashed and dotted lines are fits to a parabolic flow profile. **(b)** Normalized velocity autocorrelation function as a function of dimensionless time for $\lambda = 1$ and $\lambda = 0.1$. Dashed lines correspond to exponential decay in Eq. 1.34.

to guarantee no-slip boundary conditions. In such a scenario, Lamura et. al. [89] suggested a generalization of the bounce-back rule. According to it, all the partially filled cells are filled with sufficient number of virtual particles (ghost particles) so that the effective density of (real + ghost) particles equals the average density of the bulk fluid. The velocities of the wall particles are taken from the Maxwell-Boltzmann distribution with zero mean, keeping the temperature same as that of the fluid. The collision step is then carried out using the mean velocity of all particles in the cell. The average velocity in Eq. 1.30 rewritten as $\boldsymbol{v}_{c.m} = (1/\rho)\sum_{i=1}^{n}(\boldsymbol{v}_i + \boldsymbol{v_a})$, where n is the number of real particles in the partially filled cells, and $\boldsymbol{v}_a$ is a vector whose components are numbers from a Maxwell Boltzmann distribution with zero mean and variance $(\rho - n)k_BT$, where ρ is the average number of particles inside the cell. Using this procedure, the results obtained for Poiseuille flow with grid shifting, found to be in agreement with the correct parabolic profile as shown in Fig. 1.8(a). Here, the direction of the flow is along x-axis and two walls are placed along $y = 0$ and $y = H$ planes. For the case of simple bounce back rule, the velocity profile does not extrapolate to zero at the walls (i.e., there is a strong slip), whereas there is no slip for the case of generalized bounce-back rule.

The properties of the fluid are characterized by six parameters in MPCD algorithm: (i) the mean free path λ, (ii) the mass of the fluid particle m, (iii) the size of a grid cell a, (iv) the average number of particle per cell ρ, (v) the rotation angle α,

and (vi) the thermal energy of the fluid k_BT. For given values of λ, m and k_BT, the size of the time step h can be obtained as:

$$h = \lambda\sqrt{\frac{m}{k_BT}}. \tag{1.32}$$

The simplicity of MPCD method allows the analytical expressions for many transport coefficients, e.g., the kinematic viscosity of the fluid ν, which is the sum of two contributions, the kinetic viscosity ν_{kin} due to streaming of the fluid particles, and the collisional viscosity ν_{coll} contributed by collision of the fluid particles. In general, $\nu = \nu_{kin} + \nu_{coll}$, whose analytical expressions [85, 87, 90, 91] are as follows:

$$\frac{\nu_{coll}}{\sqrt{k_BTa^2/m}} = \frac{1}{\lambda}\frac{(1-\cos\alpha)}{18}\left(1-\frac{1}{\rho}\right), \tag{1.33a}$$

and

$$\frac{\nu_{kin}}{\sqrt{k_BTa^2/m}} = \lambda\left[\frac{1}{(4-2\cos\alpha-2\cos 2\alpha)}\frac{5\rho}{\rho-1}-\frac{1}{2}\right]. \tag{1.33b}$$

It is evident from Eqs. 1.33a, 1.33b that for small time step h and large angle α, collisional contribution dominates, which corresponds to a fluid-like behavior and for larger h and smaller α, kinetic contribution dominates, which corresponds to a gas-like behavior.

To benchmark our codes, we calculated velocity autocorrelation function (VACF) of the fluid particles (Fig. 1.8(b)) and found exact agreement with previously observed results [92]. The VACF is the time correlation function of velocity of the fluid particles. The time correlation function have been extensively used to describe the dynamic properties and transport processes of the fluid [93]. The expression for VACF is given by [92]

$$C_v(nh) = \frac{\langle \boldsymbol{v}_i(nh)\boldsymbol{v}_i(0)\rangle}{v_i^2(0)} \simeq (1-\gamma)^n, \tag{1.34}$$

where γ is a decorrelation factor defined as

$$\gamma = \frac{2}{3}(1-\cos(\alpha))\left(1-\frac{1}{\rho}\right) \equiv \gamma_\alpha\gamma_\rho. \tag{1.35}$$

As can be seen from fig. 1.8(b), for $\lambda = 1$ the exponential decay predicted in Eq 1.34 is obtained. For $\lambda = 0.1$, exponential decay is followed only in the first collision. For $\lambda = 0.1$, collisional contribution to the total kinematic viscosity is more than the kinetic one and the system behaves like fluid where VACF decay much slower

than the one given by the molecular chaos assumption. So, $\lambda = 0.1$ corresponds to fluid-like behavior and $\lambda = 1.0$ corresponds to gas-like behavior [92] .

1.4.4 Thermostat

To keep the temperature of the fluid constant, we apply Maxwell-Boltzmann scaling (MBS) thermostat [94]. In this method, the distribution of the kinetic energy of the MPCD ideal-gas particles is used to calculate the scale factor Υ. The relative velocities $\boldsymbol{u}_i = (\boldsymbol{v}_i - \boldsymbol{v}_{c.m.})$ of each fluid particles in collision cells is rescaled by Υ, i.e., $\boldsymbol{u}_i' = \Upsilon \boldsymbol{u}_i$ to achieve the constant temperature. Such scaling does not change the total momentum of the collision cell.

The distribution of the kinetic energy of the MPCD ideal-gas particles (Γ distribution) [95] given by,

$$P(E_l) = \frac{1}{E_l \Gamma(f/2)} \left(\frac{E_l}{k_B T}\right)^{f/2} \exp\left(-\frac{E_l}{k_B T}\right), \tag{1.36}$$

where, $f = 3(N_c - 1)$ is the numbers of degree of freedom of the particles in the cell, N_c is the number of particles in each collision cell, and $\Gamma(x)$ is the gamma function. Note that in the limit of very large N_c (i.e., $f \to \infty$), the Γ function becomes a Gaussian with mean $\langle E_l \rangle = f k_B T/2$ and variance $f(k_B T)^2/2$. The scale factor for the velocities in each collision cell can therefore be set to

$$\Upsilon = \sqrt{\hat{E}_l / E_l}.$$

where

$$E_l = \frac{1}{2} \sum_{i=1}^{N_c} m \boldsymbol{u}_i^2. \tag{1.37}$$

is the instantaneous kinetic energy of the collision cell before rescaling the velocities. After rescaling, the average kinetic energy of the particles in the cell is given by

$$\langle E_l \rangle = \langle \frac{1}{2} \sum_{i=1}^{N_c} m \boldsymbol{u}_i'^2 \rangle = \langle \frac{\Upsilon^2}{2} \sum_{i=1}^{N_c} m \boldsymbol{u}_i^2 \rangle = \langle \hat{E}_l \rangle,$$

which means that the average of the kinetic energy of a collision cell is equal to the required mean of the distribution function (Eq. 1.36). This approach to keep the temperature constant gives the correct distribution function of the particle velocities at the level of collision cell [95].

1.4.5 Hybrid MPCD-MD method

The polymer is introduced into the MPCD algorithm by adding N point particles, each of mass M and connected by bonds. The MPCD fluid provides the friction and random force required for the MD to work properly. The time evolution of the chain monomers are determined by Newton's equations of motion, using the velocity Verlet algorithm (Eqs.1.22, and 1.23 with no friction, noise and external force terms in the force field of Eq. 1.19). The time step of MD for polymer, Δt_{MD} is chosen to be smaller than the time step for MPCD fluid particles, h, so that the efficiency of the algorithm gets improved. The monomer-solvent interaction is taken into account by taking monomer mass $M = \rho m$ in the collision step. In this step, the center of mass velocity of the cell is calculated as

$$\boldsymbol{v}_{c.m.} = \frac{\sum\limits_{i \in cell} m\boldsymbol{v}_i + \sum\limits_{p \in cell} M\boldsymbol{V}_p}{\sum\limits_{i \in cell} m + \sum\limits_{p \in cell} M}, \tag{1.38}$$

and velocity of the MPC fluid particles and that of monomers are changed by the same as in Eq. 1.30. The MPCD algorithm is able to correctly reproduce the effect of hydrodynamic in the complex fluids. This method is particularly well suited for:

- Studying phenomena where both thermal fluctuations and hydrodynamics are important.
- Systems with Reynolds numbers and Peclet numbers of order 0.1–10.
- If the exact analytical expressions for the transport coefficients and consistent thermodynamics are needed.

1.5 Organization

The thesis is organized as follows:

In Chapter 2, we study the translocation of a semiflexible polymer through an extended pore. We first explain the dependence of mean waiting time of monomers on the stiffness for a homopolymer using iso-flux tension propagation theory when it is translocated a pore of unit length. We then explain our results for the extended patterned pores. In our study, we have considered three different pores: (1) Pore α, which is an uniformly attractive pore, (2) Pore β, which has an attractive entrance and exit, but a repulsive interaction in between, and (3) Pore γ, which has an attractive entrance but repulsive exit. We show potential energy landscapes for all the

three pores and discuss the effect of pore patterning on the translocation time of the polymer. We then discuss a sequencing strategy which exploits the statistical nature of translocation time. We argue that it could be utilised for an efficient sequencing of heteropolymers, made up of alternate stiff and flexible segments, by suitably engineering pore-polymer interactions and combining readouts from multiple pores.

In Chapter 3, we study the driven translocation of a semiflexible polymer through an interacting conical pore. We derive theoretical expression for the free energy due to confinement of a flexible polymer and its interaction with the pore. Based on this, we discuss results for the driven translocation of a flexible polymer, and explain the effect of driving force, and the pore stickiness on mean translocation time as a function of pore apex angle to explain its non-monotonic behavior. We also discuss the numerical results obtained for the driven translocation of a semiflexible polymer and explain the effect of bending rigidity on the translocation process.

In Chapter 4, we study the flow driven translocation through an extended narrow pore using *hybrid molecular dynamics-multiparticle collision dynamics* (MD-MPCD) simulations. We review the results for a flexible polymer and obtain an expression for the threshold velocity flux using de Gennes blob picture. We present our simulation results for the flow driven translocation of semiflexible polymer and establish the dependence of threshold velocity flux on the persistence length of the polymer and diameter of the pore.

Finally, in Chapter 5 we summarize the main results obtained in this thesis.

Chapter 2

Sequencing of semiflexible polymers of varying bending rigidity using patterned pores

In this chapter, we study the translocation of semiflexible polymers through extended pores with patterned stickiness, using Langevin dynamics simulations. Extending the iso-flux tension propagation theory (IFTP) to homogeneous semiflexible polymers translocating through a pore of unit length, we first explain the dependence of the mean waiting time of monomers on stiffness. For extended patterned pores, we show that the consequence of pore patterning on the translocation time dynamics is dramatic and depends strongly on the interplay of polymer stiffness and pore-polymer interactions. For heterogeneous polymers with periodically varying stiffness along their lengths, we find that variation of the block size of the sequences and the orientation, results in large variations in the translocation time distributions. We show how this fact may be utilized to develop an effective sequencing strategy.

The chapter is organized as follows: In Sec. 2.1 we define our model and the various pore patterns studied. In Sec. 2.2 we discuss results for the driven translocation of semiflexible polymer of homogeneous stiffness through pores of unit length and extended patterned pores. The results for the driven translocation of semiflexible polymer consisting of alternate blocks of stiff and flexible segments and the sequencing method are discussed in Sec. 2.3. Finally we draw our conclusions in Sec. 2.4.

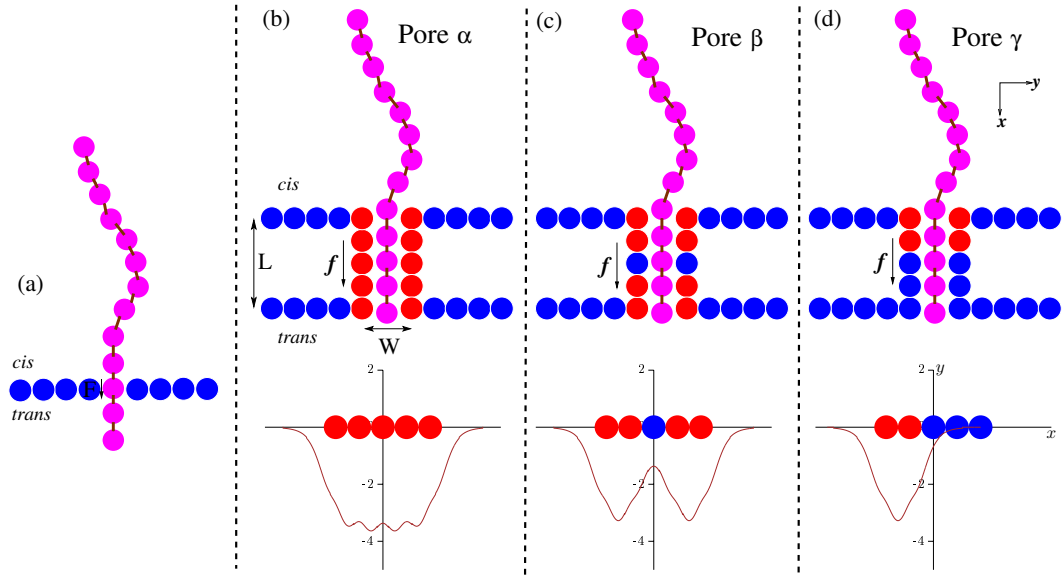

Figure 2.1: A schematic diagram of a semiflexible polymer with uniform bending rigidity κ_b translocating from the *cis* to the *trans* side through (a) Pore of unit length ($L = \sigma$) and three extended patterned pores (b) Pore α (c) Pore β, and (d) Pore γ of length L and width W. The driving force $\boldsymbol{f}_{\text{ext}} = f\hat{\boldsymbol{x}}$ acts on every monomer inside the pore. The potential energy landscape in the center ($y = 0$) along the length of the pore for various pore types is also shown.

2.1 Model and Simulation Details

2.1.1 Homopolymer model

The polymer is modeled as a self-avoiding semiflexible polymer by using beads and springs in two dimensions (Fig. 2.1). The excluded volume interaction between any two monomers is given by Weeks-Chandler-Andersen (WCA) potential $U_{\text{bead}}(r)$ (Eq. 1.15). The consecutive monomers in the chain interact via the finite extension nonlinear elastic (FENE) potential $U_{\text{bond}}(r)$ (Eq. 1.16). The semiflexibility of the polymer is introduced by the bending potential $U_{\text{bend}}(r)$ (Eq. 1.17). The strengths of FENE, and bending interactions are characterized by the spring constant k, and the bending rigidity κ_b, respectively. The length of the polymer is given by $(N-1)\sigma$, where N is the number of beads in the chain. We chose a dimensionless parameter $\lambda = \ell_p/\ell$ to characterize the stiffness of the polymer, where ℓ is the average contour length of the polymer.

2.1.2 Heteropolymer model

The heteropolymer is modeled similarly by using beads and springs with the polymer segment representing n monomers each of stiff (S) and flexible (F) beads arranged

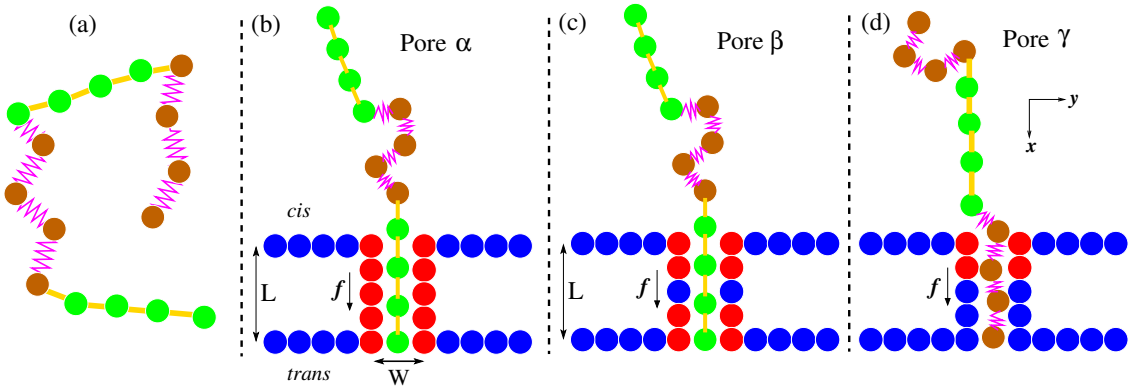

Figure 2.2: (a) Schematic diagram of a polymer with alternate blocks of stiff (S) and flexible (F) segments S_nF_n each having $n = 4$ bonds. The stiff (S) and flexible (F) bonds are shown by straight and zig-zag lines, respectively. (b) and (c) Polymer S_4F_4 translocating through Pores α and β, respectively, with the stiff end entering the pore first. (d) Polymer F_4S_4 translocating trough Pore γ with the flexible end entering the pore first.

in symmetric blocks S_nF_n. A schematic diagram of such a polymer with $n = 4$ is shown in Fig. 2.2(a). As an example, for a polymer with $N = 128$, the minimum value of $n = 1$ is for $(S_1F_1)_{64}$, i.e., 64 repeat units of S_1F_1, and the maximum value of $n = N/2 = 64$ is for a single unit of $S_{64}F_{64}$. For a heteropolymer, it makes a difference whether a flexible or a stiff end enters the pore first (Figs. 2(b), 2(c), 2(d)).

2.1.3 Pore model

The pore and the wall are made from stationary monomers separated by a distance of σ from each other. The pore is made up of two rows of monomers symmetric about the x-axis. The length of the pore is taken to be L with a diameter W. For a *pore of unit length*, $L = \sigma$ (Fig. 2.1(a)). In this case, the interaction of the polymer with the pore is given by the WCA potential with ϵ replaced by $\epsilon_{\rm pore} = 1.2\epsilon$. For *patterned pores*, we choose an extended pore of length $L = 5\sigma$ with three different pore patterns:

(1) Pore α is an attractive pore. All the monomers of the pore interact with the polymer by the LJ potential $U_{\rm pore}(r)$ (Eq. 1.18).

(2) Pore β has an attractive entrance and exit. The first two and the last two monomers of the pore interact with the polymer by the LJ potential, and the middle monomer by WCA potential as in the pore of unit length.

(3) Pore γ has an attractive entrance and repulsive exit. The first two monomers of the pore interact with the polymer by the LJ potential and the last three

monomers of the pore by WCA potential as above.

We have used repulsive walls to avoid the penetration of polymer through it. Therefore, the interaction between the wall beads and of the polymer (U_{wall}), is taken same as the intra-monomer interaction ($U_{\text{bead}}(r)$).

To facilitate the transfer of polymer from the *cis* to the *trans* side of the pore, the polymer experiences a driving force, $\boldsymbol{f}_{\text{ext}} = f\hat{\boldsymbol{x}}$ directed along the pore axis with magnitude f. This mimics the electrophoretic driving of biopolymers through nanopores. Due to the larger entropic cost involved in confining the polymer in extended pores, the pore entrance in such cases are chosen to be attractive to initiate the translocation successfully. A schematic diagram of semiflexible polymers translocating from the *cis* to the *trans* side through the pore of unit length and pores α, β, and γ are shown in Fig. 2.1(b)-(d), respectively.

To integrate the equation of motion for the monomers of the chain we use Langevin dynamics (Eq. 1.19) with Ermak's algorithm (Eqs. 1.24, 1.25).

The unit of energy, length and mass are set by the familiar LJ units ϵ, σ and m respectively. This sets the unit of time as $\sqrt{m\sigma^2/\epsilon}$. Following Luo et al. [63], we assume the bead size in our coarse-grained polymer model as $\sigma = 1.5$ nm. This is equal to the Kuhn length of a single-stranded DNA, which is approximately three nucleotide bases. Hence the mass of the bead is $m \approx 936$ amu (given that the mass of a base in DNA is ≈ 312 amu) and charge of the bead $q \approx 0.3$ e (each base having a charge of 0.1 e effectively [63]). To allow comparison with known results, we set $\zeta = 0.7$ and $k_BT = 1.2$. At $T = 295$ K, the interaction strength is given by $\epsilon = k_BT/1.2 \approx 3.4 \times 10^{-21}$ J, which gives a time scale of $(m\sigma^2/\epsilon)^{1/2} \approx 30$ ps and force scale of $\epsilon/\sigma \approx 2.4$ pN.

In these units, we choose $N = 128$, $L = 5$, $W = 2.25$, $\epsilon_{\text{pore}} = 1.2$ (homopolymer) and 2 (heteropolymer), $r_c = 2.5$, $\zeta = 0.7$, $k = 30$, $R = 1.5$ and $k_BT = 1.2$, in our simulations to allow comparison with known results. We have used $f = 5$ and $f = 1$ for analysis with pore of unit length and extended pore, respectively. A time step of $\Delta t = 0.001$ is used in all simulation runs. Our choice of pore width $W = 2.25\sigma$ is to ensure single-file translocation of the polymer and avoid the formation of hairpin configurations inside the pore. To calculate statistical properties, we have considered $1500 - 2000$ successful translocation events.

2.2 Translocation of Homogeneous semiflexible polymer

The role of stiffness of the polymer chain on the translocation dynamics is best addressed by decoupling it from the energetics of the extended pore. We therefore, first considered the translocation of a homogeneous semiflexible polymer with stiffness through a pore of unit length. We first provided a theoretical description of the translocation process by extending the analysis of the tension propagation (TP) theory to account for the stiffness of the polymer for the case of unit length pore, and subsequently provide a qualitative description of the observations for patterned pores by analyzing the effect of pore-polymer interactions.

2.2.1 Case I: Waiting time for unit length pore

2.2.1.1 Iso-flux tension propagation theory

The idea of tension propagation was first proposed by Sakaue [43] in the context of nonequilibrium driven translocation. In TP theory, the translocation process is described in terms of a single variable, the monomer number at the pore, s. The translocation of the polymer in the presence of entropy, driving and pore-polymer interactions can therefore be considered as a barrier crossing problem of the "translocation coordinate" s. Hence, for a pore of unit length, the polymer is considered as translocated once $s = N\sigma$. The corresponding Langevin equation Eq. 1.19 in the overdamped limit reduces to

$$\tilde{\zeta}(\tilde{t})\frac{ds}{d\tilde{t}} = (1-\gamma')\left[\frac{1}{N-s} - \frac{1}{s}\right] + \tilde{f} + \tilde{\eta}(\tilde{t}) \equiv \tilde{f}_{tot}, \tag{2.1}$$

where $\tilde{\zeta}$ is now an effective friction coefficient, $\tilde{f}(\tilde{t})$ includes the various forces, $\tilde{\eta}$ is a Gaussian white noise and $\tilde{f}_{tot}$ is the total force and γ' is the surface exponent (γ' ≈ 0.69, for self-avoiding chains in two dimensions). Note that, in this limit, we use σ, k_BT and ζ to set the units of length, energy and friction. This is different from the LJ units which we described in the previous section, and are used in our simulations. In these new units, $\zeta\sigma^2/k_BT$ sets the unit of time, k_BT/σ the unit of force and $k_BT/(\zeta\sigma^2)$ the unit of flux. To distinguish from the LJ units, all dimensionless parameters except for length (where the unit is the same), are denoted by tilde.

For driven translocation, with sufficiently strong driving forces, the force propagates along the backbone of the polymer, resulting in a non-trivial *time dependent* drag. This forms the basis of the argument of tension propagation. According to

this, the part of the polymer on the *cis*-side is divided into two distinct domains. The external driving force that acts inside the pore, pulls the monomers nearer to the pore and sets them in motion. The remaining monomers that are farther away from the pore, do not experience the pull and on average remain at rest. As the polymer gets sucked inside, more and more monomers on the *cis* side start responding to the force, with a tension front separating the two domains propagating along the length of the polymer. At time $\tilde{t}$, the drag experienced by a monomer inside the pore, can be written as the sum of the friction due to the length of the chain in the *cis* side ($\tilde{\zeta}_c$) up to which the tension has propagated, that due to the *trans* side segment ($\tilde{\zeta}_t$) and the pore friction ($\tilde{\zeta}_p$) as $\tilde{\zeta}(\tilde{t}) = \tilde{\zeta}_c + \tilde{\zeta}_t + \tilde{\zeta}_p$. It is then easy to see that $\tilde{\zeta}(\tilde{t})$ increases as the tension front propagates and more number of monomers on the *cis* side get involved.

To derive the equation of motion for semiflexible polymer, we have taken the assumption of constant monomer flux [42]. Here, the flux of monomers, $\tilde{\phi} = ds/d\tilde{t}$, through the pore and the mobile domain in the *cis* side is assumed constant in space, but evolving in time. Defining the pore location as $\tilde{x} = 0$, the tension front is at a distance $\tilde{x} = -\tilde{R}(\tilde{t})$ from the pore. For sufficiently strong external forces, a small portion of the chain close to the pore is straightened while the portion behind is coiled, giving rise to the *stem-flower* regime.The magnitude of the tension force at distance $\tilde{x}$ can be obtained as $\tilde{f}(\tilde{x}) = \tilde{f}_0 - \tilde{\phi}(\tilde{t})\tilde{x}$ where $\tilde{f}_0 = \tilde{f}_{tot} - \tilde{\zeta}_p\tilde{\phi}(\tilde{t})$ is the force at pore entrance (see Appendix A for details). Integrating the force balance equation over the mobile domain in the *cis* side, we get

$$\tilde{\phi}(\tilde{t}) = \frac{\tilde{f}_{tot}}{\tilde{\zeta}_t + \tilde{\zeta}_p + \tilde{R}(\tilde{t})}. \tag{2.2}$$

From Eqs. 2.1 and 2.2 and since $\tilde{\phi} = ds/d\tilde{t}$, we get the effective friction as

$$\tilde{\zeta}(\tilde{t}) = \tilde{R}(\tilde{t}) + \tilde{\zeta}_t + \tilde{\zeta}_p. \tag{2.3}$$

To get the full solution, we need to find the equation of motion for the tension front, $\tilde{R}(\tilde{t})$, which is done in two stages. In the TP stage, we assume that in time $\tilde{t}$, the tension in the chain has propagated to M monomers of the chain. Then, $\tilde{R}$ can be evaluated as the root-mean-square of the end-to-end distance of a semiflexible chain of length M. The end-to-end distance of a semiflexible self-avoiding chain in the limit $M/l_p \gg 1$, scales as [50], $\tilde{R} = A_\nu \tilde{\ell}_p^{\nu_p} M^\nu = A_\nu \tilde{\ell}_p^{\nu_p}(l+s)^\nu$, where l is the number of mobile monomers in the *cis* side, and in two dimensions [96], $\nu_p = 1/4$ and $\nu = 3/4$. The equation of motion satisfied by the tension front in the propagation

stage is obtained by taking the derivative of $\tilde{R}(\tilde{t})$ and is given by (see Appendix A for derivation)

$$\dot{\tilde{R}}(\tilde{t}) = \frac{\mathcal{A}\tilde{R}(\tilde{t})^{\frac{\nu-1}{\nu}}\left[(\mathcal{B}_{SF}+\mathcal{C}_{SF})\times\dot{\tilde{f}}_{tot}(\tilde{t})+\tilde{\phi}(\tilde{t})\right]}{1+\mathcal{A}\tilde{R}(\tilde{t})^{\frac{\nu-1}{\nu}}\mathcal{B}_{SF}\times\tilde{\phi}(\tilde{t})}, \tag{2.4}$$

where,

$$\mathcal{A} = \begin{cases} \nu A_\nu^{1/\nu}, & \text{for} \quad \lambda = 0 \\ \nu A_\nu^{1/\nu}\tilde{\ell}_p^{\nu_p/\nu}, & \text{for} \quad \lambda \neq 0, \end{cases}$$

$\dot{\tilde{f}}_{tot}$ is the time derivative of $\tilde{f}_{tot}$ and $\mathcal{B}_{SF}$ and $\mathcal{C}_{SF}$ are function of $\tilde{\phi}$, $\tilde{\zeta}_p$, $\tilde{\zeta}_t$, and ν and given by

$$\mathcal{B}_{SF} = -\frac{1}{\tilde{\phi}(\tilde{t})} + \frac{\nu-1}{(2\nu-1)[\tilde{\zeta}_p+\tilde{\zeta}_t+\tilde{R}(\tilde{t})]\tilde{\phi}(\tilde{t})^2}, \tag{2.5}$$

and

$$\mathcal{C}_{SF} = \frac{1}{\tilde{\phi}(\tilde{t})}. \tag{2.6}$$

In the post propagation stage, the tension front has reached the last monomer of the chain in the *cis* side, i.e., $M = N$. We therefore have $s + l = N$, which gives the equation of motion for the tension front as (see Appendix A for details)

$$\dot{\tilde{R}}(\tilde{t}) = \frac{(\mathcal{B}_{SF}+\mathcal{C}_{SF})\times\dot{\tilde{f}}_{tot}(\tilde{t})+\tilde{\phi}(\tilde{t})}{\mathcal{B}_{SF}\times\tilde{\phi}(\tilde{t})}. \tag{2.7}$$

We finally solve for the IFTP equations ignoring noise. Further, in our solutions, we have assumed that the friction from the *trans* side of the chain is negligible and can be absorbed into the constant pore friction $\tilde{\zeta}_p$. Although earlier numerical results [49] suggest that this is a good approximation for a flexible pore, it is not so for a semiflexible chain [50]. However, for our analysis, we have considered the fairly flexible limit $N/l_p \gg 1$, and our approximation is meaningful. Thus, in our solutions for the IFTP equations, $\tilde{\zeta}_p$ is the only free parameter which we fixed by fitting to the result for $\lambda = 0$.

2.2.1.2 Dependence of the mean waiting time on the stiffness of polymer

In Fig. 2.3, we have shown the mean waiting time as a function of translocating coordinate s for different values of λ, calculated from both simulation and theory. Mean waiting times, $w(s)$, defined as the amount of time a monomer s spends on average inside the pore. The results from the simulation of a homopolymer passing through a pore of unit length, shows an initial increase with s, implying that subse-

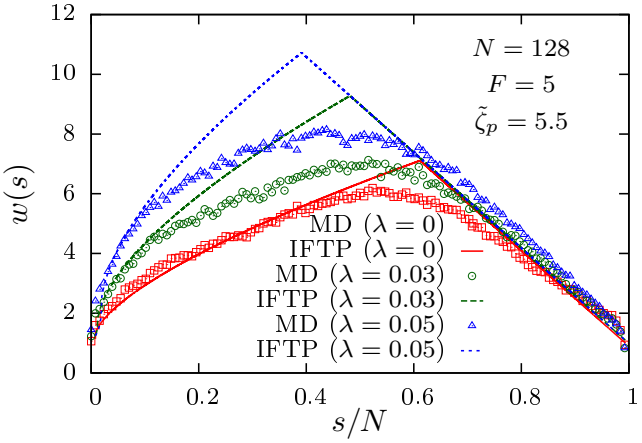

Figure 2.3: The mean waiting times, $w(s)$, for monomers of a semiflexible polymer of length $N = 128$ for various stiffness, λ, translocating through a pore of unit length with an external force of magnitude $f = 5$. The points show the MD simulation data for $\lambda = 0$ ($\square$), $\lambda = 0.03$ ($\bigcirc$), and $\lambda = 0.05$ ($\triangle$), and the lines are the results from IFTP theory with pore friction $\tilde{\zeta}_p = 5.5$ and in the absence of noise.

quent monomers spends more time inside the pore. This increase in waiting time is because of the additional drag $\tilde{\zeta}(\tilde{t})$ from the moving monomers on the *cis*-side. As we have discussed above, $\tilde{\zeta}(\tilde{t})$ increases as more monomers translocated through the pore which also results in the increase of the waiting times. This continues until $\tilde{\zeta}(\tilde{t})$ becomes maximum when the tension front reaches the last monomer. The time at which this happens, is called the tension propagation time ($\tilde{t}_{tp}$). At $\tilde{t}_{tp}$, maximum number of monomers at the *cis*-side participate in the translocation process and the monomer s, which is inside the pore at that instant, has maximum waiting time $w(s)$. For $\tilde{t} > \tilde{t}_{tp}$, the system enters the tail retraction stage or the post propagation (PP) stage, where the monomers on the *cis* side starts decreasing and therefore $\tilde{\zeta}(\tilde{t})$ decreases, and so does the waiting time w. As the stiffness of the polymer increases, the waiting times for every monomer increases indicating that a stiffer chain takes longer to translocate. Also the peak of the waiting time distribution, and hence $\tilde{t}_{tp}$, shifts to the left indicating that tension propagates faster with increasing chain stiffness. We got the same behavior from our theoretical analysis which is shown by solid lines in Fig. 2.3. The simulation data matches well with the results obtained for the deterministic case from the IFTP theory. It correctly reproduces the dependence on the stiffness (λ) of the chain, observed in the simulations, as well as the gradual shift of the peak of the distribution towards lower s as λ increases.

2.2.2 Case II: Mean waiting time for extended patterned pores

For extended patterned pores, the assumption of a constant pore friction is no longer valid. Moreover, the friction from the *trans* side is also a complicated function of the pore patterning, bending rigidity and the external force. It is difficult to analytically incorporate these effects within the IFTP theory. In what follows, we gave a qualitative description of the effects of pore patterning on the mean waiting times and the translocation time distributions, based on the surface energetics of the pores.

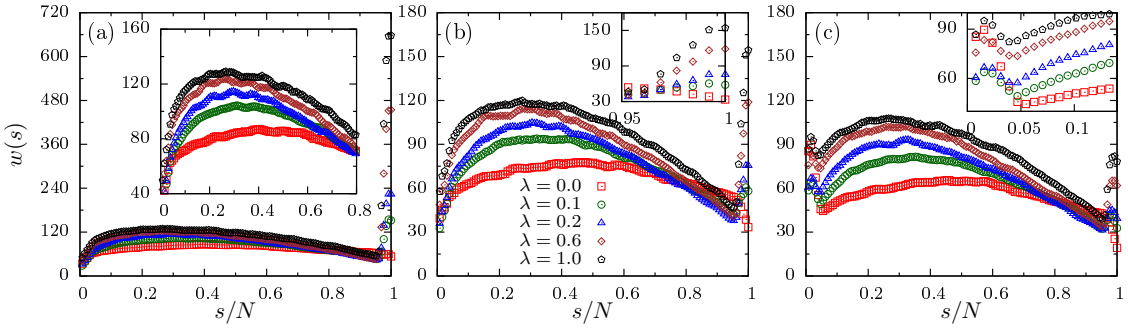

Figure 2.4: Mean waiting times $w(s)$ for monomers of a semiflexible polymer of various stiffness λ for (a) Pore α, (b) Pore β, and (c) Pore γ. The inset in (a) represents $w(s)$ for Pore α excluding the end monomers. Inset (b) shows the end monomers region for Pore β while inset (c) shows the behavior for the initial monomers entering the pore for Pore γ. Note that error bars are smaller than the point size and are not shown here.

For the extended patterned pores in our simulations, we calculated the mean waiting time of a monomer as it translocates from the *cis* to the *trans* side, for small external forces. We observed that the gross features, like a peak in the waiting times and the dependence on chain stiffness, observed for the pore of unit length, are reproduced. Further, there are additional features near $s = 1$ and $s = N$, which can be attributed to the pore polymer interactions. For Pore α, $w(s)$ shows a sharp rise in the large s limit (Fig. 2.4(a)). This feature persists for Pores β and γ as well, although it is less pronounced. Pore γ shows an initial dip in $w(s)$(Fig. 2.4(c)). For monomers in the bulk of the polymer, the non-monotonic variation of $w(s)$ as predicted from the TP theory persists. In order to understand these features, we focus on the surface energetics of the various patterned pores.

The potential energy landscape in the center ($y = 0$) of the pore and along its length for various pore types are shown in the bottom panel of Fig. 2.1. It is obtained by placing stationary test point particles inside the pore along ($y = 0$) which are separated by a small distance Δx ($\Delta x \ll 1$). We then calculated the potential

Pore	λ	s	t_{tp}	$\langle\tau\rangle$	$\frac{t_{tp}}{\langle\tau\rangle}$
Pore α	0.0	54	845.9	1938.2	0.43
	0.1	43	782.3	2212.5	0.35
	0.2	39	768.2	2347.4	0.32
	0.6	37	815.7	2729.1	0.29
	1.0	34	791.9	3079.2	0.25
Pore β	0.0	56	780.6	1717.7	0.45
	0.1	43	710.2	1935.9	0.36
	0.2	39	698.3	2030.4	0.34
	0.6	36	732.0	2270.8	0.32
	1.0	34	741.4	2484.5	0.29
Pore γ	0.0	55	666.4	1445.3	0.46
	0.1	44	637.6	1644.9	0.38
	0.2	39	638.1	1778.3	0.35
	0.6	36	680.1	2004.0	0.33
	1.0	33	668.9	2194.6	0.30

Table 2.1: The tension propagation time t_{tp}, the translocation coordinate s at which the tension front reaches the last monomer, mean translocation time $\langle\tau\rangle$, and ratio $t_{tp}/\langle\tau\rangle$, for semiflexible polymer of length $N = 128$ with various stiffness λ for different pore types.

experienced by these particle from the wall beads which has been shown along y-axis in Fig. 2.1. Pore α is a uniformly attractive pore. Although this makes it easier to pull the polymer inside the pore, the attractive interactions makes it difficult to exit the pore from the *trans* side for small external forces. Pore β has a shallower potential well compared to Pore α. Therefore, the waiting times for the end monomers are significantly less for this pore. For Pore γ, which has a repulsive exit, this effect is the least. However, since Pore γ has a large potential barrier beyond the *cis* side, the waiting times for the first few monomers as they enter the pore shows a dip. Naively, one would expect that the Pore γ will have the longest filling time, which is indeed correct. However, once the first monomer crosses the energy barrier, the repulsive interaction immediately ejects it out of the pore. It drags a few monomers following it, resulting in their ejection and a waiting time scale less than the first one. This is the cause of the dip in the waiting times for the first few monomers. Once there are monomers on the *trans* side, the usual tension propagation argument ensures that the waiting time start increasing until the post propagation stage.

In Table 2.1 we have shown the tension propagation time t_{tp}, the translocation coordinate at which the tension front reaches the last monomer, the mean translocation time $\langle\tau\rangle$, and the ratio $t_{tp}/\langle\tau\rangle$ for a chain of length $N = 128$ for various chain stiffness λ and the three different pore types. This table clearly shows that, for all pore types, the ratio $t_{tp}/\langle\tau\rangle$ decreases as the stiffness λ of the polymer increases, which indicates that the tension propagates relatively faster for the stiff chain.

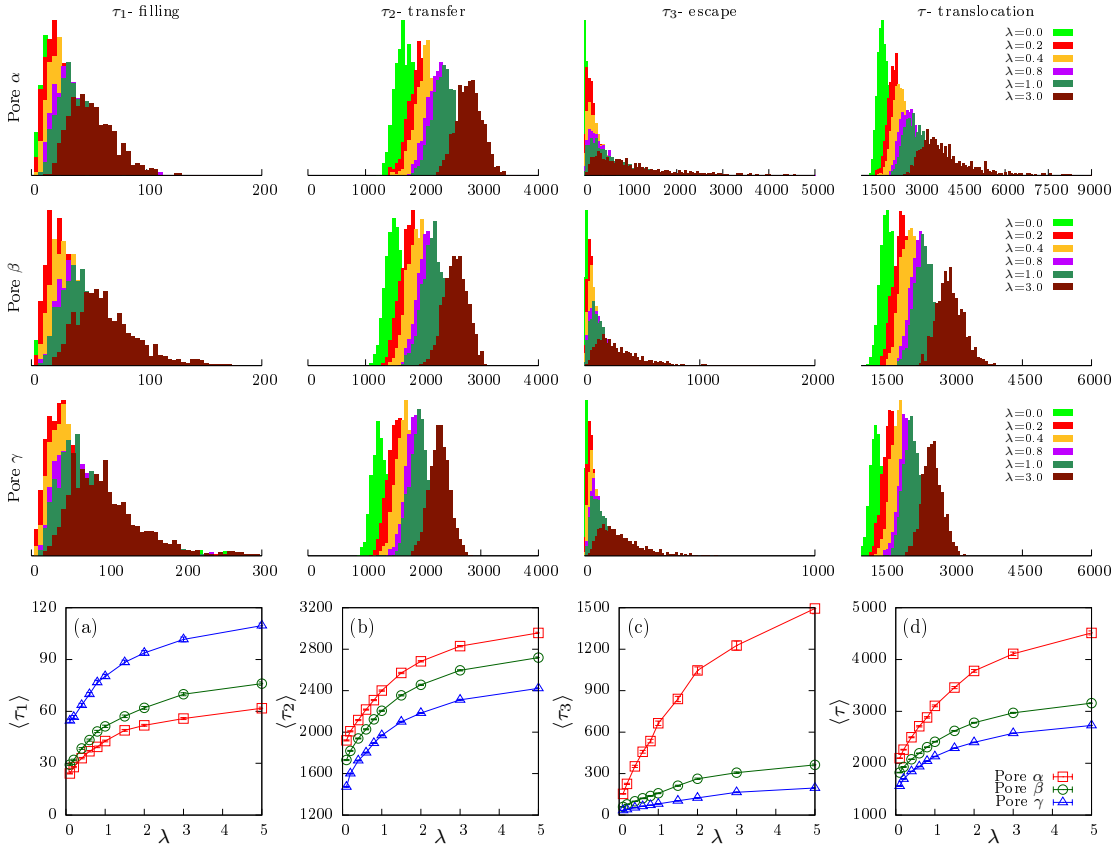

Figure 2.5: Translocation time statistics for semiflexible polymers with homogeneous bending rigidity. First, second and third rows : Translocation time distributions for pores α, β and γ respectively, for $f = 1.0$, as λ is varied. Not only are the the three distributions different in their moments across the three pores, they also vary with varying λ. Fourth row : Average (a) filling time, $\langle\tau_1\rangle$ (b) transfer time, $\langle\tau_2\rangle$, (c) escape time, $\langle\tau_3\rangle$ and (d) mean translocation time $\langle\tau\rangle$ as a function of λ. This corroborates the observation from the distributions.

2.2.3 Translocation times and their distributions

We have divided the total translocation time in three stages as $\tau = \tau_1 + \tau_2 + \tau_3$ where (i) τ_1 is the initial *filling* time, the time taken by the first monomer of the polymer

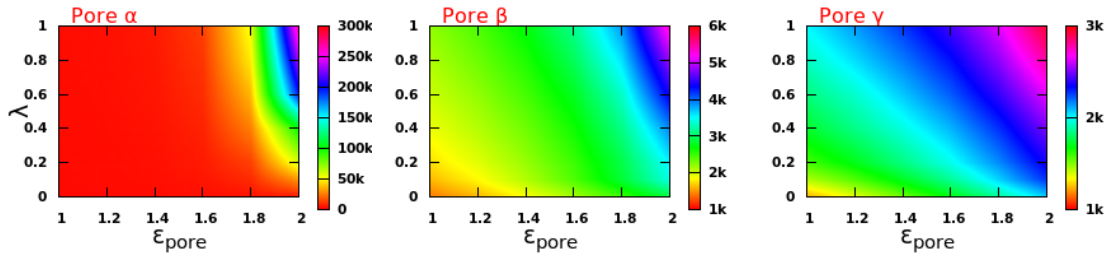

Figure 2.6: Density plots of average translocation time as a function of strength of pore polymer interaction, ϵ_{pore}, and bending rigidity $\lambda = \ell_p/\ell$. (k in colorbox stands for thousand (1k=1000)).

to reach the exit without returning to the pore, (ii) τ_2, the transfer time, the time taken from the exit of the first monomer into the *trans*-side to the entry of the last monomer from the *cis*-side and (iii) τ_3, the escape time, the time between the entry of the last monomer in the pore and its escape to the *trans*-side. We compare the average time scales for filling, transfer and escape for the three different pore patterns to investigate the effect of changing pore-polymer interactions. In first three rows of (Fig. 2.5), we have shown the distributions for the various λ values along three pores. These distributions are quite distinct from each others which indicated the possibility of detecting bending rigidity dependent sequences of heteropolymer.

In the fourth row of (Fig. 2.5), we have shown various times as a function of the stiffness λ of the polymer. We observed that all time scales show a monotonic increase with increasing stiffness. This behavior is expected from the discussion of waiting times which increases with increasing λ. The attractive nature of Pore α ensures that its filling time, $\langle\tau_1\rangle$ is the minimum of the three pores. Pore γ which has a repulsive exit has the largest $\langle\tau_1\rangle$ due to the large potential barrier encountered by the monomers entering the pore. The distributions of $\langle\tau_1\rangle$ also reveals these characteristics, with longer tails observed for Pore γ (Fig. 2.5). The transfer time $\langle\tau_2\rangle$ shows more regular behavior, the repulsive exit of Pore γ ensuring that monomers leave the pore to the *trans* side faster than the other pores. Evidently, Pore α has the longest $\langle\tau_2\rangle$. The escape time, $\langle\tau_3\rangle$ is determined by the pore-polymer interaction at the exit. The strong attractive interactions near the exit for Pores α and β ensures longer times compared to that of Pore γ. The total translocation time shows clearly that the transfer and escape times dominate with Pore α taking the longest time to translocate followed by pores β and γ (Fig. 2.5(d)).

2.2.4 Translocation time for different pore polymer interaction

We have also explored the dependence of translocation time on ϵ_{pore} and λ, keeping the driving force same. In Fig. 2.6, we show the density plots of average translocation time as a function of the strength of pore-polymer interaction ϵ_{pore} and the bending rigidity $\lambda = \ell_p/\ell$ for the driving force $f = 1.0$ for Pores α, β and γ. From Fig. 2.6, we can easily see that the average translocation time increases (irrespective of pore) as

1. the strength of pore-polymer interaction ϵ_{pore} is increased for a translocating polymer of constant stiffness λ.

2. the stiffness λ of the translocating polymer increases keeping the pore-polymer interaction ϵ_{pore} constant.

We will utilize this fact to detect a heteropolymer sequence consisting of repeat units of stiff and flexible segments by translocating through multiple pores.

2.3 Translocation of heterogeneous polymer

As elaborated earlier, we have introduced heterogeneity by varying the stiffness of the polymer along the chain backbone. The heterogeneity introduced in our polymer model is periodic with alternate flexible and stiff segments. We first discuss the effect of heterogeneity and orientation on mean waiting times and translocation time dynamics and then discuss about our sequencing strategy. We have chosen $\lambda = 0.5$ as the rigidity of the stiff segments and $\epsilon_{pore} = 2.0$ as the strength of pore polymer interactions so that the translocation time of flexible and stiff segments are fairly separated and can be distinguished easily.

2.3.1 Dependence of Waiting times

In Fig. 2.7, we have shown the waiting times $w(s)$ of individual monomers with index s translocating through Pores α, β, and γ for heteropolymers S_8F_8 and F_8S_8. We observed oscillations in the waiting times in all the three cases, except near the end when the pore synergetics becomes dominant. From our analysis of waiting times of monomers of a homogeneous polymer we know that (i) tension propagates faster for chains with increasing stiffness and hence (ii) leads to larger waiting times. In the case of heterogeneous polymers, tension propagates intermittently through blocks of

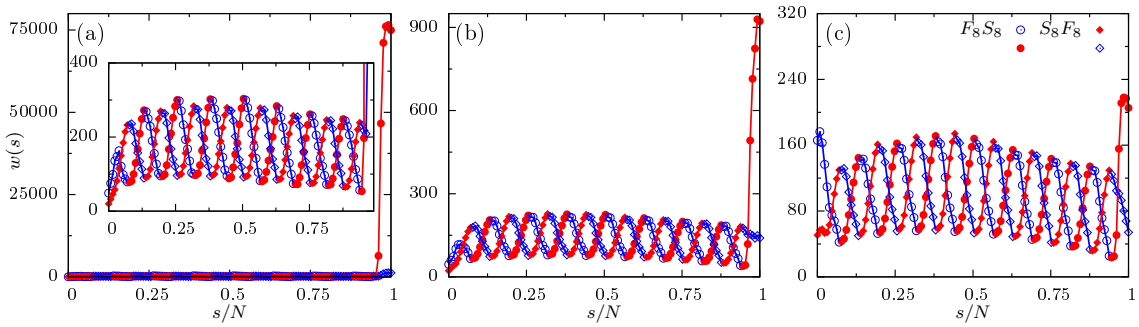

Figure 2.7: Waiting times for heteropolymers, S_8F_8 and F_8S_8, translocating through (a) Pore α, (b) Pore β, and (c) Pore γ. The open and filled symbols represent flexible (F) and stiff (S) segments, respectively. The results for heteropolymers, F_8S_8 and S_8F_8, are shown by circles and diamonds, respectively.

stiff and flexible segments leading to the oscillations in the waiting time distribution. A stiff block has a larger waiting time, followed by a flexible block with lower waiting time and so on. When the orientation of the chain is reversed, the oscillations for S_8F_8 and F_8S_8 are exactly out of phase as expected. The waiting times for the end monomers of the chain however show distinct features for different orientations of the chain.

In sync with its homopolymer counterpart, the end chain dynamics of heteropolymers is strongly influenced by the pore-polymer interactions. For Pores α and β, the attractive interactions near the *trans* side of the pore dominate, leading to large waiting times. Evidently, the waiting times for the end monomers of the chain are significantly larger for pore α compared to pore β. This effect is significantly less for Pore γ which has a repulsive exit.

The end chain dynamics for the reversed conformation S_8F_8 is not significantly affected by these interactions. Pore α due to the large potential barrier does make it difficult for the end monomers to exit the pore leading to larger waiting times. However, the waiting times are considerably less compared to F_8S_8. Pore β and γ are largely unaffected. This is expected from our earlier analysis of larger waiting times for stiffer chains. Polymer in the conformation S_8F_8 enters the pore with the stiffer block entering first followed by a flexible block. This implies that a flexible block exits the pore last in this conformation. In contrast, in the conformation F_8S_8, it is a stiff block which exits the pore last from the *trans* side in the translocation process leading to much larger waiting times. Note that in both conformations, in the final stages of the tail retraction stage of the translocation process, pore polymer interactions start playing a significant role.

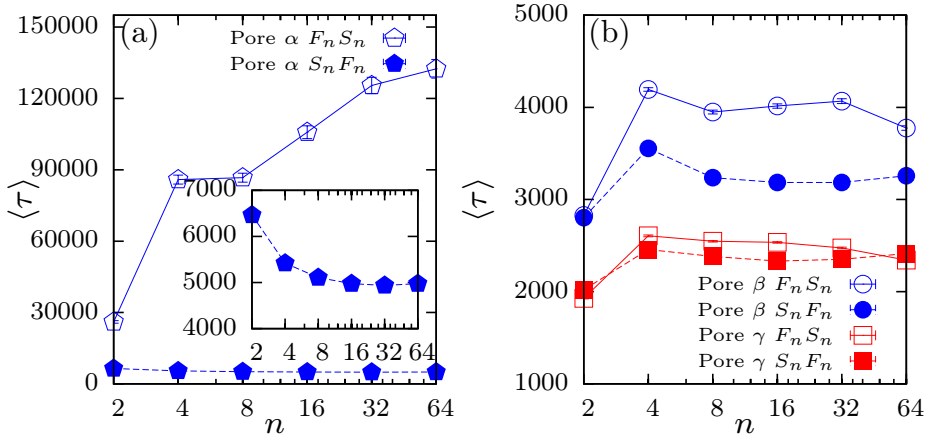

Figure 2.8: Average translocation time $\langle\tau\rangle$ as a function of block length n (in logscale) for a polymer of length $N = 128$ having alternate stiff (S) and flexible (F) segments. (a) Polymer translocating through Pore α, and (b) Pores β (circles) and γ (squares). Results for polymer entering the pore through the stiff (flexible) end, represented by S_nF_n (F_nS_n), are shown by filled (open) symbols. The inset in (a) zooms the data for the case where the polymer enters the Pore α through the stiff end.

2.3.2 Dependence of the mean translocation time on the segment length

In Fig. 2.8, we have shown the mean translocation time $\langle\tau\rangle$ as a function of segment length n for heteropolymers translocating through Pores α, β and γ. For the case of polymer entering the pore from the flexible end (i.e.,F_nS_n), it is the stiff segments that leave the pore at the end (or vice versa). We have seen earlier (Fig. 2.7) that the waiting times of the last few monomers for the polymer entering the pore from the flexible end (F_8S_8) are usually larger than the waiting time for the polymer entering the pore from the stiff end (S_8F_8).

The dependence of the mean translocation times on segment length n for the heteropolymers as they pass through the patterned pores can be understood by looking at the mean waiting times of individual monomers of the heteropolymer. For Pore α, the difference in the mean translocation times for the two different orientations (i.e., S_nF_n and F_nS_n) is significant and increases with increasing block length n (Fig. 2.8). For pore β, the effect of the longer waiting times for end monomers on the total translocation time is less significant while for pore γ, it is effectively the same for both orientations of the polymer during the translocation.

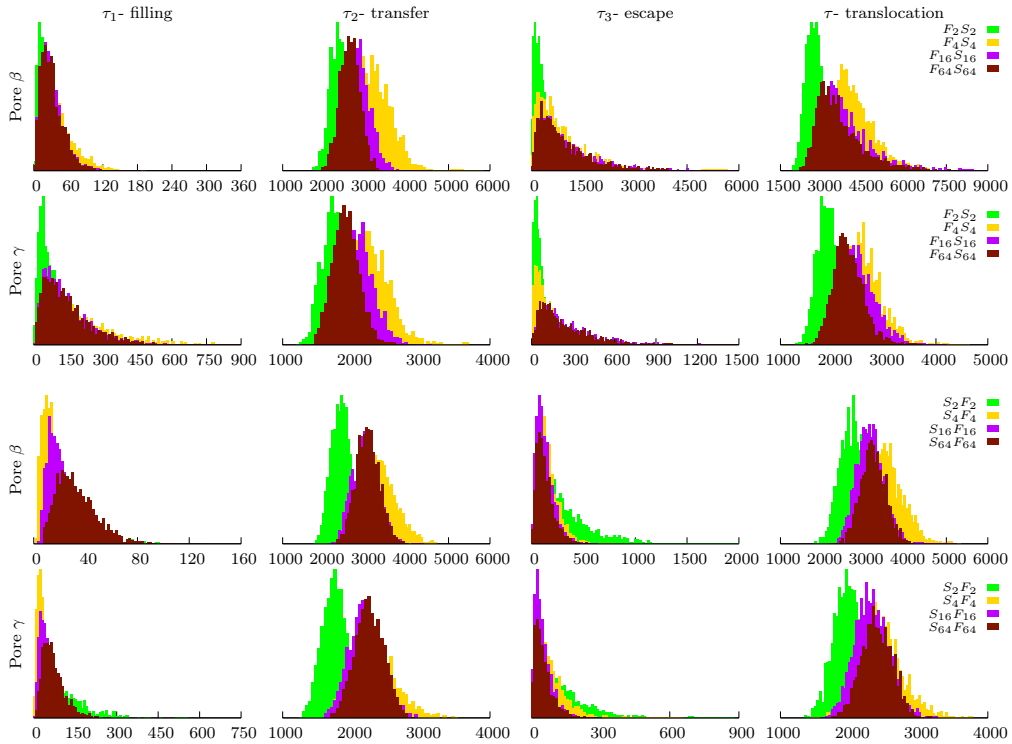

Figure 2.9: Translocation time distribution for heteropolymers. Comparison of filling, transfer and escape time distribution for Pores β and γ for four different sequences of the heteropolymer. **First two rows** show the translocation time distribution for Pores β and γ when polymer enters the pore from the flexible end (F_nS_n). **Last two rows** show the translocation time distribution for Pores β and γ when polymer enters from the stiff end (S_nF_n).

2.3.3 Sequencing of polynucleotide with varying bending rigidity

The sensitivity of the translocation dynamics on the varying bending rigidity of heteropolymers and the patterning of pores opens up the possibility of sequencing heteropolymers based on their unique translocation time statistical properties. For example, Fig. 2.9 shows the translocation time distributions for four different sequences of types S_nF_n and F_nS_n translocating through Pores β and γ. These translocation time distributions exhibit distinct features corresponding to the variation in the block lengths of the heteropolymer. Also, the translocation time distributions of the same sequence, but entering the pore from different orientation (S_nF_n/F_nS_n), show clear distinction, as the length of the segment increases (Fig. 2.10). We calculate the mean translocation time ($\langle\tau\rangle$) and the standard deviation ($\sqrt{\langle\tau^2\rangle - \langle\tau\rangle^2}$) from these distributions, and construct scatter plots by plotting mean translocation time along the x-axis and standard deviation along y-axis, as shown in Fig. 2.11. The scatter plots

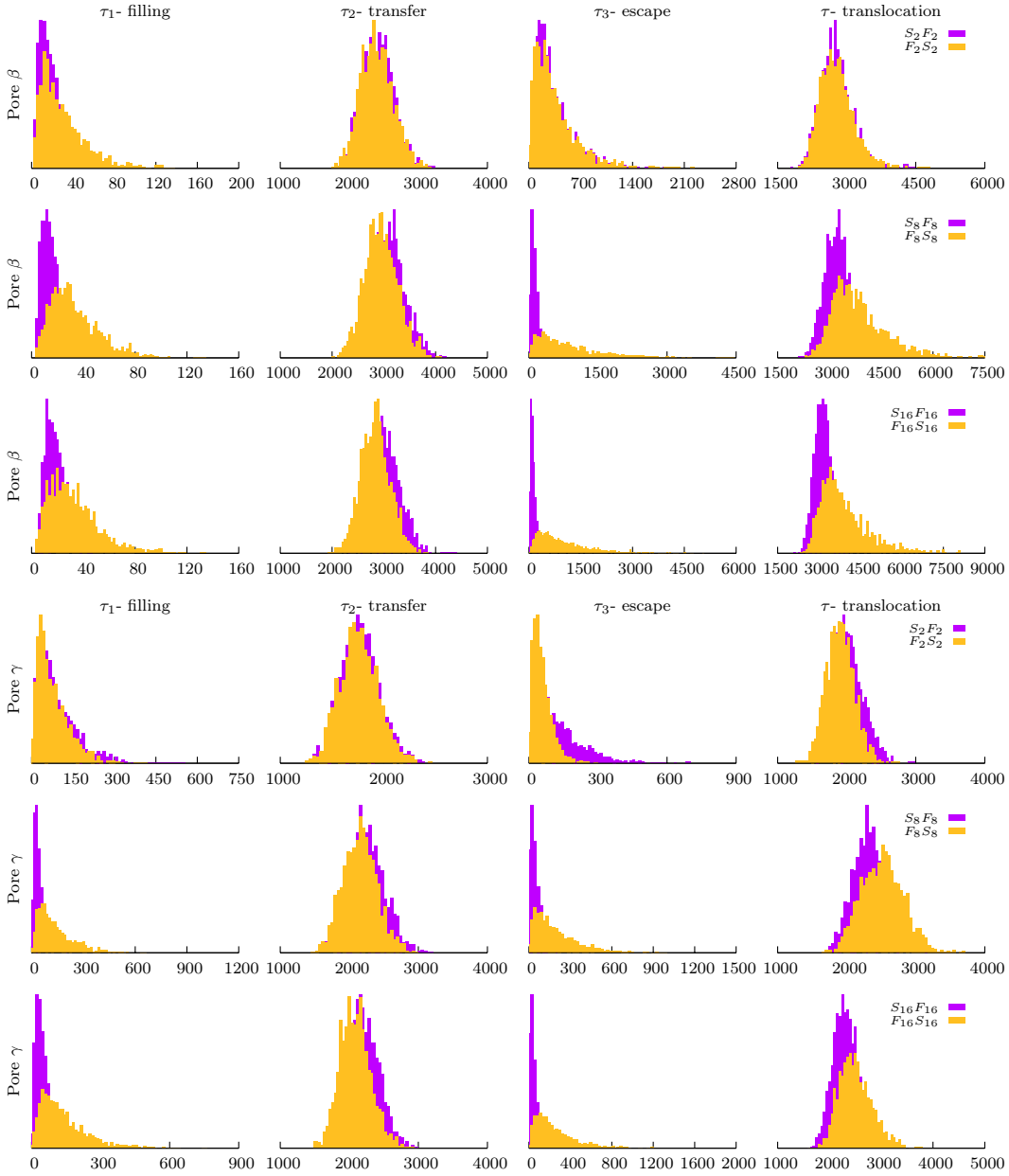

Figure 2.10: Translocation time distribution for heteropolymers for three segment lengths, n=2,8 and 16. Comparison of filling, transfer and escape time distribution for different orientation of the heteropolymer for Pores β (first three rows) and Pore γ (last three rows). Violet/gold color represent flexible(F_nS_n)/stiff (S_nF_n) end entering the pore first. As the segment length increases, one can see a clear distinction in the distributions for different orientation

reveal several interesting features. For example, Pore γ cannot distinguish between $(S_{32}F_{32})_2$ and $(S_{64}F_{64})_1$, but Pore β can. Similarly, Pore β cannot distinguish between $(S_2F_2)_{32}$ and $(F_2S_2)_{32}$, but Pore γ can. These differences in the scatter plots for various pores, clearly show that the combination of translocation time measurements

across different pores could be utilized to differentiate, and thus identify, sequences which would otherwise be difficult to distinguish using measurements from a single pore.

Our simulation methodology for sequence detection is as follows. We choose a sequence of the heteropolymer (say S_kF_k, with a specific orientation) from the set of all available sequences (defined as the training set) used to plot Figs. 2.11(a-c) and call it an "unknown" sequence. This sequence is then passed through a single pore of α, β or γ type. For each pore type, the heteropolymer is passed multiple times and successful translocation events are registered (say m). For every attempted translocation, the chain configuration of the heteropolymer is chosen from the equilibrated configurations obtained in accordance with the simulation strategy discussed before (Sec. 2.1). Having registered the successful translocation events across each pore, the mean translocation time and standard deviation are calculated for each pore type, which correspond to respective points in the scatter plots. These numbers are then compared with those of the training set for the corresponding pore using a "distance" metric. The larger the distance between the point corresponding to the "unknown" sequence is from a particular known sequence in the scatter plot, the greater is the relative error for that sequence. The total error, which is the sum of the distance from a particular known sequence in all the plots, is minimized to predict the "unknown" sequence. If the predicted sequence matches the sequence we started with, then this marks a successful detection.

The ratio of the number of times a sequence is correctly detected to the total number of attempts, gives the accuracy of the measurement process (see Fig. 2.11(d)-(f)). For full algorithm, see Appendix B. The samples per pore type merely indicate the number of registered successful translocation events chosen for the unknown sequence across each pore type. Evidently, if we use a very large number of samples for a given pore type, the sequence detection would be accurate. However, this scheme suggests that a combination of different pore types gives very high accuracy of prediction for a relatively small number of copies of each pore. In Fig. 2.11, we have used the statistical data for only two pore types, Pores β and γ, to test our hypothesis. Employing the above scheme we found, for example, that the accuracy of detection for the sequence, $F_{64}S_{64}$ (Fig. 2.11(d)), reaches 100%, even for ~ 130 samples per pore type. It is important to note that in our method of sequence detection, we have used only the first two moments of the probability distribution of translocation times. As observed from the results of the distributions, this is far from accurate. Inclusion of higher moments would most definitely improve the accuracy of the scheme and lead to a more rapid convergence. Further, our method has considered only a few possible pore types. It

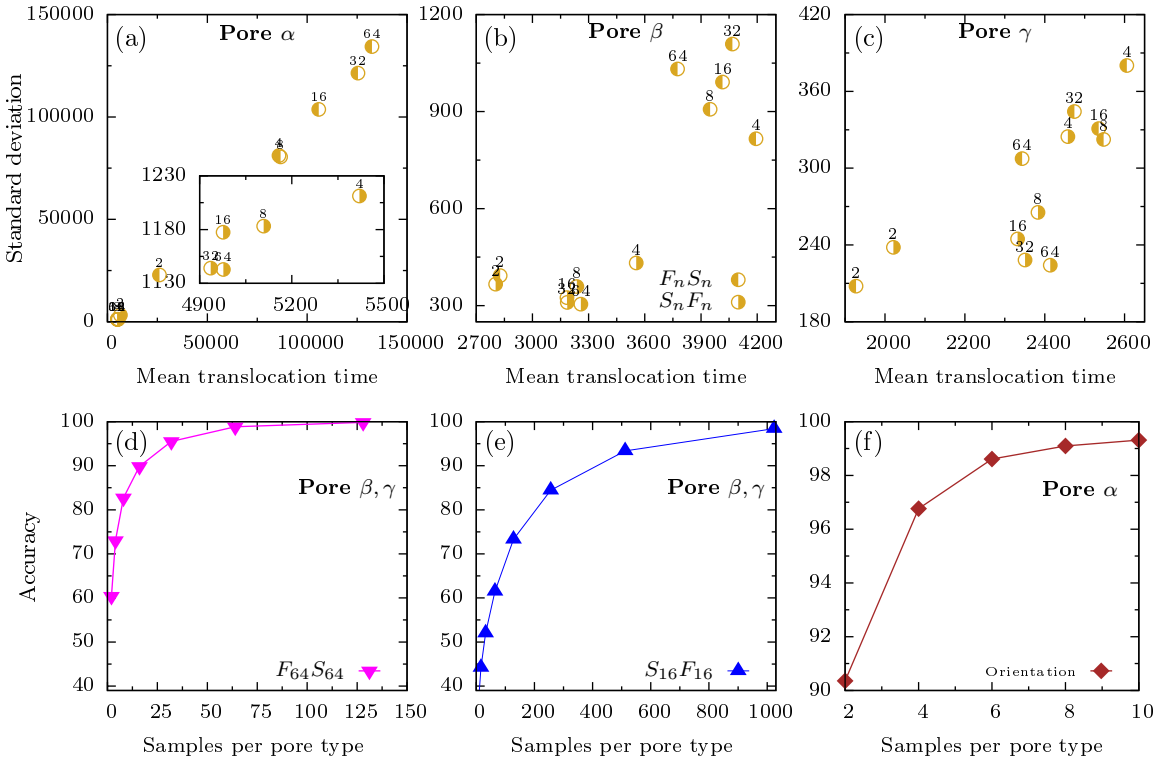

Figure 2.11: [(a) - (c)] Scatter plots showing mean and standard deviation of translocation times for Pores α, β and γ. The inset in (a) shows the zoomed bottom left portion for Pore α. In each of these plots, the polymer entering the pore from the stiff end (S_nF_n) is shown by symbol half yellow filled circle (right), while the polymer entering from the flexible end (F_nS_n) is shown by half yellow filled circle (left). The accuracy of detecting heteropolymers (d) $F_{64}S_{64}$ (pink inverted triangles) and (e) $S_{16}F_{16}$ (blue triangles) through Pores β and γ. (f) Accuracy of detection by including Pore α to distinguish orientation of the heteropolymer $S_{16}F_{16}$.

would be interesting to design pores leading to even more distinguishable translocation time statistics, which when used in conjunction with varying semiflexibility across the polymer backbone, would lead to enhanced sequence detection.

From the set of pores chosen for this study, it is evident that for Pore α, there is an order of magnitude difference in the translocation times of stiff and flexible segments. Therefore, this pore is an ideal candidate to detect the difference in orientation between S_kF_k and F_kS_k and make the detection process even more precise. Indeed, we find that the accuracy of detecting the correct orientation is almost 100 percent even for a small number of copies of Pore α (Fig. 2.11(f)). We would like to stress that it is not necessary to distinguish the orientation of the polymer before passing them through the pores. Our statistical analysis simply suggest that it requires far less samples per pore type if we manage to do so.

Our result needs to be compared with the case where heterogeneous segments were distinguished by their relative interactions with the pore [66]. It turns out that in such a scenario, the translocation time distributions have sharper and more distinguished features, leading to better sequencing accuracy. Our analysis shows that using different patterned pores can lead to efficient sequencing strategies for such cases.

2.4 Conclusions

In this chapter, we studied translocation of homogeneous semiflexible polymer and heteropolymer through single-slit and extended nanopore. We extended IFTP theory for homogeneous semiflexible polymer for single-slit nanopore which correctly captured the dependence of mean waiting time of monomers on the polymer stiffness. We have demonstrated that for extended pore, not only waiting times increases and peak shifts toward left, as stiffness increases, but last few monomers take longer time to translocate through the pore and show a sharp increase in the mean waiting time of monomers towards the end of the chain which was missing in single-slit nanopore and this increase in the waiting time for last few monomers is attributed to the pore polymer interaction which is shown in escape time of the polymer. As we increase the strength of pore polymer interaction, last few monomer will take longer and longer time to translocate. Thus for extended pores, the breakup of the total translocation time into the filling, transfer and escape times proves useful and in this context reveal interesting features for semiflexible polymer translocation hitherto unobserved for single-slit nanopores.

It has been found in experimental studies that the sequence dependent bending rigidity is important for DNA-protein interaction and nucleosome positioning [97,98]. Such a dependence is confirmed from cyclization studies of short DNA fragments, which allows accurate measurement of persistence length [99]. Other examples of polymers with varying bending rigidity includes partially melted DNA and proteins which exhibit stiff and flexible segments along the polymer backbone [100–103].

The strong dependence of translocation time on the bending rigidity of the polymer and the distinguishable translocation time statistics generated by translocation through pores with different patterned stickiness, allows us to efficiently detect polymers with varying bending rigidity.

The sequencing theme proposed in this chapter, could be experimentally realized using fabricated nanofluidic channels with surface decoration. Arrays of nanochannels interfaced with microfludic loading channels have been shown to be a highly

parallel platform for the restriction mapping of DNA [104–106]. The first task is to construct the set of translocation time distributions for known sequences. This requires passing sequences with a particular orientation multiple times through these functionally modified nanofluidic channels. Solid state nanopores are other highly plausible candidates to achieve the same. With the training set characterised, the sequencing of heteropolymers with "unknown" sequences can be efficiently achieved in limited time by passing them through these channels using our purely statistical analysis. The detection of the orientation of the polymer could be achieved using a fluorescent dye on either the stiff or flexible end of the polymer [107].

Chapter 3

Driven translocation of a semiflexible polymer through an interacting conical pore

In this chapter, we study the driven translocation of a semiflexible polymer through an interacting conical pore. There has been experimental studies of polymer translocation through conical pores, which report better distinguishability of ionic current signatures from DNA sequences, making them promising candidates for sequencing applications [20]. A few theoretical studies on translocation through such pores have focused on the effective driving achieved due to pore asymmetry in the absence of an external drive [69, 70]. However, for all practical applications, the translocation process is driven and pore-polymer interactions are significant. In this study, we focus on the effects of pore-polymer interactions and flexibility of the polymer in a *driven* translocation process.

The chapter is organized as follows. In Sec. 3.1 we define our model and the the geometry of pore studied. In Sec. 3.2, we derive theoretical expression for the free energy due to confinement of the polymer and its interaction with the pore. In Sec. 3.3 we discuss results for the driven translocation of flexible polymer, where we explain the effect of driving force, and the pore stickiness on the translocation process. We study the mean waiting times of monomers for different values of the driving force, and for different pore apex angles. In Sec. 3.4 we discuss the numerical results obtained for the driven translocation of a semiflexible polymer and explain the effect of bending rigidity on the translocation process. Finally, in Sec. 3.5, we draw our conclusions.

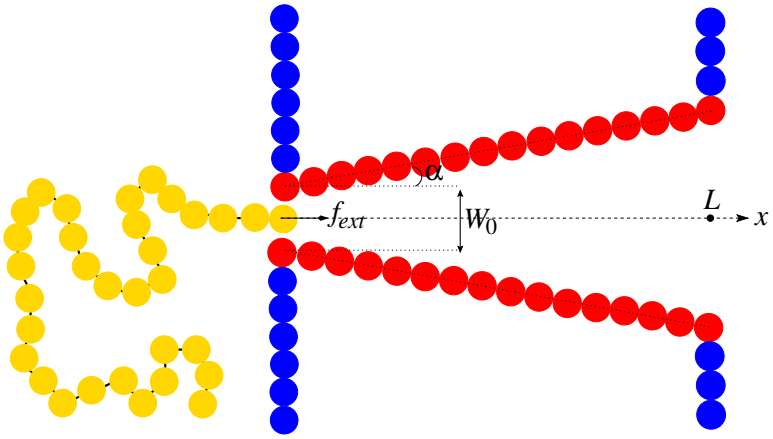

Figure 3.1: Schematic diagram of a semiflexible polymer translocating through a conical pore with an apex angle $\theta = 2\alpha$. The width (at the apex) and the length of the pore is $W_0 = 2.25\sigma$ and $L = 16\sigma$, respectively.

3.1 Model

As before, our model for the polymer is a self-avoiding semiflexible chain made up of beads and springs in two dimensions. We have used WCA potential $U_{\text{bead}}(r)$ (Eq. 1.15), FENE spring potential $U_{\text{bond}}(r)$ (Eq. 1.16), and a bending potential $U_{\text{bend}}(r)$ (Eq. 1.17) to mimic the excluded volume, the bond stretching between two consecutive monomers, and, the stiffness of the polymer chain, respectively. The stiffness of the polymer is changed by changing the bending rigidity parameter κ_b in Eq. 1.17. The pore and walls are made from stationary monomers separated by a distance of σ from each other. The conical pore is made up of two rows of monomers symmetric about the x-axis with an apex angle $\theta = 2\alpha$. Henceforth, we will refer to α as the apex angle. The length of the pore is taken to be L with a width W_0 at the apex. This pore width allows only single-file translocation of the polymer and avoid the formation of hairpin configurations at the apex opening. The interaction of the pore with the polymer is chosen to be the standard LJ potential $U_{\text{pore}}(r)$ (Eq. 1.18).

In addition, the polymer experiences a driving force, $\boldsymbol{f}_{\text{ext}} = f\hat{\boldsymbol{x}}$ directed along the pore axis with magnitude f, which mimics the electrophoretic driving of biopolymers through nanopores. A schematic diagram of semiflexible polymers translocating from the *cis* to the *trans* side of a conical pore is shown in Fig. 3.1.

The translocation time τ is defined as the time elapsed between the entrance of the first bead of the polymer and the exit of the last bead from the pore. All failed translocation events are discarded. We use Langevin dynamics (Eq. 1.19) with Ermak's algorithm (Eq. 1.24, 1.25) to integrate the equation of motion for the monomers.

3.2 Theoretical Results for a flexible polymer ($\kappa_b = 0$)

Consider a linear polymer chain made up of N monomers of diameter σ. In equilibrium, the size of the free chain in is given by $R \simeq \sigma N^{\nu}$, where ν is the Flory exponent. In two-dimensions $\nu = 3/4$. It is well known that the size of the polymer chain changes when it is confined in a channel of width less than the size of the free chain R.

To understand the process of translocation, we need to estimate the free-energy change due to the confinement, F_c, of the polymer inside the conical channel. Let us consider a partly confined chain in the channel in the presence of an external driving force $\boldsymbol{f}_{ext}$. The confinement of the polymer inside the pore costs entropy. We assume that the part of the polymer chain that is inside the pore breaks up into blobs of size $\xi(x) = W_0 \cos\alpha + 2x \sin\alpha$ that are tangent to the pore walls as shown in Fig 2.2. Then, the entropic penalty due to the confinement of chain in the conical channel is of the order of k_BT per blob. If $N_b(y)$ represents the number of blobs that penetrate a distance y into the channel, we have $F_c(y) \sim k_BTN_b(y)$. To count $N_b(y)$, we use the scheme given by Nikoofard et al. [69]. As can be seen from Fig. 3.2, the two consecutive blobs inside the conical pore at positions x_j and x_{j+1} with diameters $\xi(x_j)$ and $\xi(x_{j+1})$, respectively, satisfy the relation

$$x_{j+1} = x_j + \frac{1}{2}\left[\xi(x_j) + \xi(x_{j+1})\right]. \tag{3.1}$$

Using this recursion relation, one can easily obtain an expression between the position of the blob x_j and its number j along the pore as

$$x_j = \mathcal{Q}^{j-1}\left(x_1 + \frac{W_0}{2\tan\alpha}\right) - \frac{W_0}{2\tan\alpha}, \tag{3.2}$$

where $\mathcal{Q} = (1+\sin\alpha)/(1-\sin\alpha)$. The first blob is tangent to the beginning of the pore and its location along the axis is given by $x_1 = \xi(x_1)/2$. The last (say nth) blob is tangent to the pore at a distance y from the pore entrance, hence its location x_n along the axis is given by $x_n = y - \xi(x_n)/2$. The blob number n, and hence $N_b(y)$, can be obtained by Using Eq. 3.2 along with x_1 and x_n, and hence the free-energy cost due to the confinement as a function of distance y along the pore axis is given by

$$\frac{F_c(y)}{k_BT} \sim N_b(y) \sim \frac{\log\mathcal{P}}{\log\mathcal{Q}}, \tag{3.3}$$

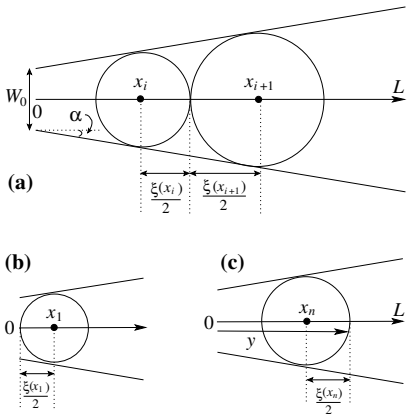

Figure 3.2: (a) Two consecutive blobs inside the conical pore. The diameter $\xi(x_i)$ of a blob depends on the position x_i along the pore axis. This figure also shows the relation satisfied by two consecutive blobs. (b) The location of the first blob is tangent to the begnining of the conical pore. (c) The location of the nth blob is at a distance y from the entrance of the conical pore.

where $\mathcal{P} = 1 + 2y \tan\alpha / W_0$.

The other contribution to the free-energy is due to the constant external force, f, experienced by the segment of the polymer which is inside the pore, denoted by F_f. The value of F_f changes with the number of monomers which are present inside the pore. The free energy change corresponds to the work done to displace the chain by a distance y inside the pore and is given by

$$\begin{aligned} \frac{F_f(y)}{k_B T} &\sim -\int_0^y f N_b(x) dx \\ &\sim -\frac{f}{2\log\mathcal{Q}} \left[2y - \left(2y + \frac{W_0}{\tan\alpha} \right) \log\mathcal{P} \right], \end{aligned} \tag{3.4}$$

where f is the magnitude of the external force $\boldsymbol{f}_{ext}$ acting inside the pore towards the trans direction.

There is one more contribution to the free-energy due to the attractive interaction with the walls of the channel. To find this, we need to determine the number of monomers in the blobs that are in contact with two walls of the conical channel. In 2D, the number of monomers in jth blob is given by $m(x_j) = (\xi(x_j)/\sigma)^{4/3}$. The total number of monomers N_n in n blobs that are penetrated up to distance y inside the channel can then be obtained from the constraint $N_n(y) = \sum_{j=1}^{n} m(x_j)$. Substituting x_j from the recursion relation Eq. 3.2, and number of blobs $N_b(y)$, we get

$$N_n(y) = \sigma^{-\frac{4}{3}} \frac{(2y\sin\alpha + W_0\cos\alpha)^{\frac{4}{3}} - (W_0\cos\alpha)^{\frac{4}{3}}}{(1+\sin\alpha)^{\frac{4}{3}} - (1-\sin\alpha)^{\frac{4}{3}}}. \tag{3.5}$$

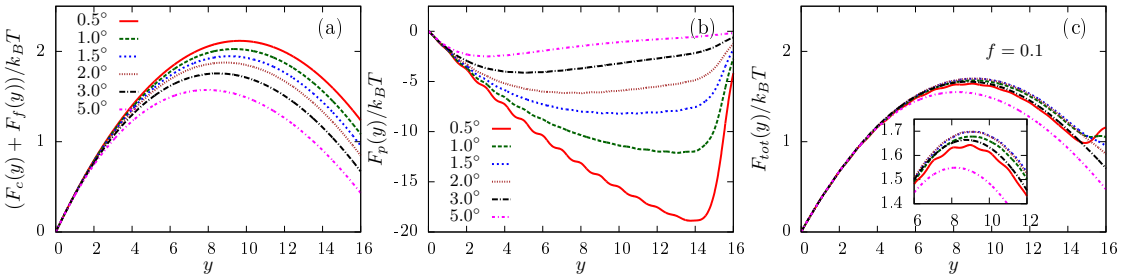

Figure 3.3: (Color online) (a) Sum of free energy contributions due to confinement and external force $(F_c+F_f)/k_BT$, (b) Free energy due to pore interaction F_p/k_BT, (c) Total free energy F_{tot}/k_BT, as a function of distance y from the pore apex along the pore axis for different half-apex angle α for a polymer of length $N = 64$ with an external force $f = 0.1$. The inset shows the data near the peak region.

The fraction of monomers that are in contact with a wall of the channel is then given by $N_n(y)(\sigma/\xi(y))$. If $V(y)$ denotes the interaction energy, which is due to the pore-polymer interaction and given by the LJ potential (Eq. 1.18), the free energy contribution due to the attractive interaction for a polymer that has entered the pore up to a distance y is given by

$$\frac{F_p(y)}{k_BT} \sim V(y)N_n(y)\left(\frac{\sigma}{\xi(y)}\right). \tag{3.6}$$

The total free energy is then given by

$$\frac{F_{tot}(y)}{k_BT} = \frac{1}{k_BT}\left(F_f(y) + AF_p(y) + CF_c(y)\right), \tag{3.7}$$

where A and C are undetermined factors.

In Fig. 3.3, we have plotted the total free energy F_{tot}/k_BT, given by Eq (3.7), for parameter values $A = 0.035$ and $C = 1$, as a function of distance y along the pore axis for different half-apex angle α ranging from $\alpha = 0.5°$ to $\alpha = 5°$ for a polymer of length $N = 64$ with an external force $f = 0.1$ and pore polymer interaction energy strength $\varepsilon_p = 1$. The free energy exhibits a barrier that needs to be overcome by the polymer to translocate towards the trans side successfully. As α increases, the barrier height also increases indicating that it is relatively more difficult for the polymer to translocate towards the trans end. The free energy barrier attains a maximum height at $\alpha = 2°$ and on increasing α further, the barrier starts decreasing showing that polymer can easily translocate from the conical pore for larger apex angle. This clearly explains qualitatively the hump in the average translocation time plotted in Fig.3.4.

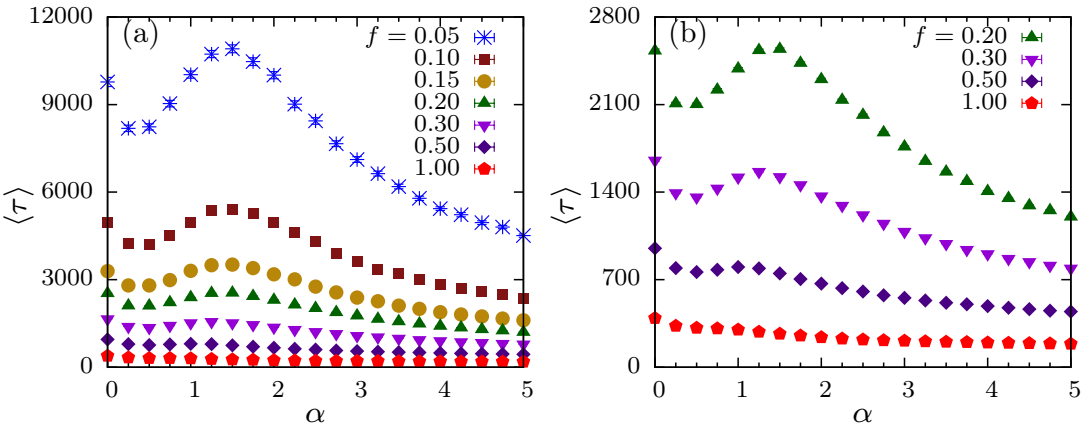

Figure 3.4: (a) Average translocation time $\langle\tau\rangle$ as a function of pore apex angle α for a fixed pore-monomer strength $\epsilon_{\text{pore}} = 1$. (b) Same as (a) for higher values of driving forces.

3.3 Case I: Numerical results for a flexible polymer

We next consider the simulation results for the translocation of a flexible polymer. In this case, $\kappa_b = 0$, hence $U_{\text{bend}} = 0$. We discuss the effect of driving force and the pore-monomer interaction strength ϵ_{pore} on translocation time of the polymer. We also discuss the mean waiting time of monomers at different pore apex angles and driving forces.

3.3.1 Translocation time of flexible polymer

3.3.1.1 Effect of driving force

In Fig. 3.4, we have plotted the mean translocation time $\langle\tau\rangle$ for a flexible polymer as a function of apex angle α, for a fixed ϵ_{pore} and for different driving forces, f. For low driving forces, we observe a non-monotonicity in the variation of $\langle\tau\rangle$ as the apex angle is changed. For small apex angles, $\langle\tau\rangle$ initially falls and then rises as α increases. For large α, there is a monotonous decrease. For example, for $f = 0.05$, as the pore apex angle is increased from 0°, the filling time first decreases and shows a minimum at 0.25°, after that it increases up to a maximum value at 1.50° and then decreases monotonically as pore angle is further increased. This non-monotonicity gradually disappears with increasing driving force.

We attempt to give a qualitative explanation of this behavior from our theoretical analysis of the free energies of confinement and pore-polymer interaction. As observed in Fig 3.3, the confinement free energy of the polymer exhibits a barrier. This barrier

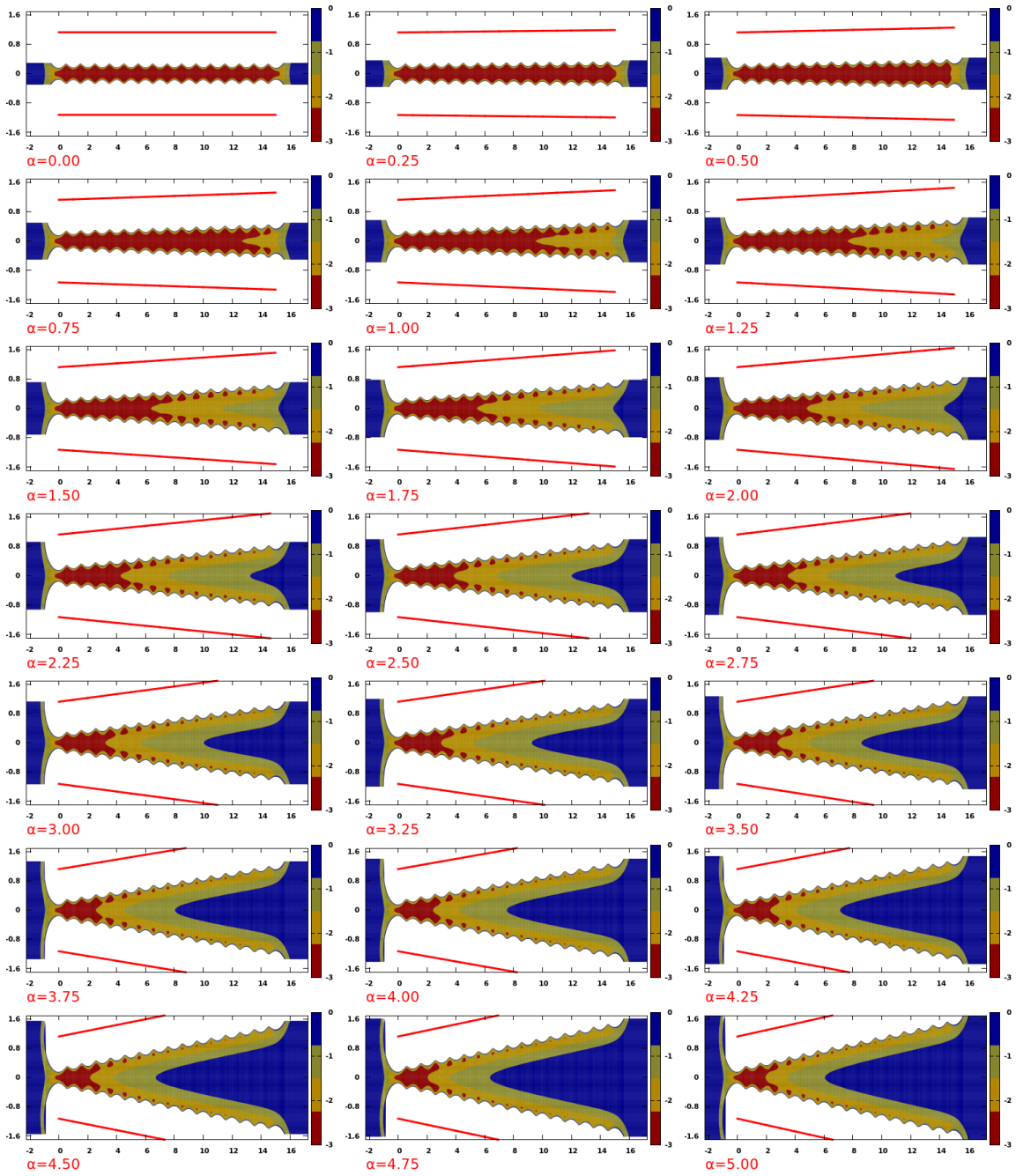

Figure 3.5: Potential energy landscape for various pore apex angles α used in our simulations. Potential depth increases from blue to red.

is closer to the pore exit for small apex angles and shifts towards the pore entrance as α increases. The height of the barrier also decreases with increasing α. Therefore, one would expect that the time taken to translocate would fall as α increases. However, the free energy due to pore-polymer interactions indicate that free energy becomes shallower near the entrance as α increases. The free energy barrier at the exit also decreases with increasing α as expected. Therefore, the pull that the polymer experi-

ences near the *cis* side of the pore due to the strong attractive interactions, decreases as the apex angle (and hence the asymmetry of the pore) is increased. This would then indicate that it will take the polymer longer to complete the translocation process. It is this competition between the force due to confinement and external drive and the force due to pore-polymer interactions for small apex angles, that sets an angle α^* where the translocation time is a minimum. Beyond α^*, the translocation time again increases until up to α^{**}. For $\alpha > \alpha^{**}$, the effect of confinement is again negligible. The polymer needs only to overcome the barrier posed near the exit which reduces with increasing α. Hence the mean translocation time keeps on decreasing beyond α^{**}.

The effect of pore-polymer interactions can be captured from the potential energy landscape for an attractive conical pore. Fig. 3.5 shows the potential energy landscape for various apex angles shown with the color gradient. Blue to red represents an increase in the potential strength (depth of the potential well). For very small values of apex angle, the position of the potential barrier is nearer to the trans side of pore, which corroborates the free energy picture. So for small apex angles, the translocation process is an interplay between confinement effects and interactions of the polymer with the pore. But as the pore apex angle increases, position of the potential barrier start moving towards the apex of the pore as can be seen from the red region moving close to the pore entry.

As the driving force is increased, the translocation is faster. The polymer will take less time to overcome the barrier as compared to the smaller values of driving forces. At sufficiently strong driving force, effect of pore-polymer interaction becomes negligible and translocation time decreases monotonically with the apex angle of the pore (Fig. 3.4(a)).

3.3.1.2 Effect of pore stickiness

In Fig. 3.6, we plot the variation of mean translocation time $\langle\tau\rangle$ as a function of α for different strengths of pore-polymer interaction, ϵ_{pore}, at a fixed value of the external force f. As is evident from the free energy argument, F_p dominates the translocation process as ϵ_{pore} increases. Therefore, the non-monotonicity in the mean translocation times with varying α, gradually disappears at higher values of ϵ_{pore}. Here, the free energy barrier near the exit of the pore is high and the polymer spends longer time inside the pore. As α increases, this barrier falls drastically and the translocation times are significantly less. For low ϵ_{pore}, the competition between confinement and interaction is significant and the non-monotonicity persists.

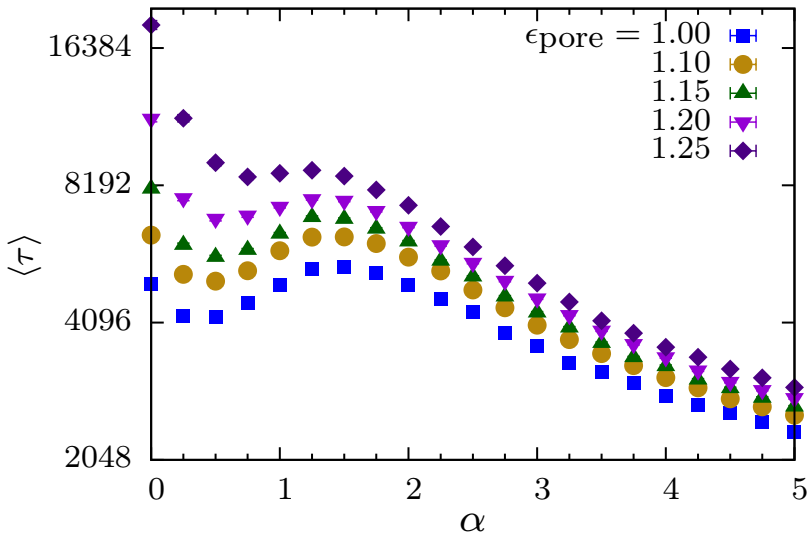

Figure 3.6: Average translocation time $\langle\tau\rangle$ as a function of pore apex angle α for various strength of pore-polymer interaction ϵ_{pore}. The external driving force is $f = 0.1$.

3.3.2 Waiting times of flexible polymer for fixed pore-polymer interaction strength

As argued before, in the context of driven translocation, the mean waiting time, $w(s)$ plays an important role in understanding the translocation process. As defined in Chapter 2, mean waiting time, $w(s)$, of a monomer s is the total time spent by it inside the pore while translocating from the *cis* to the *trans* side of the pore. In Fig. 3.7, we have shown the mean waiting time $w(s)$ for a flexible polymer as function of monomer index s when it is translocated through conical pores of different apex angles α, and for various external driving forces f. The pore-polymer interaction strength, ϵ_{pore} is kept fixed. We first consider the case of a small directed drive (Figs. 3.7(a-b)). For $\alpha = 0$, i.e. for a flat pore, $w(s)$ shows a hump near the pore entrance. This hump can be explained on the basis of tension propagation theory. According to this, the part of the polymer on the *cis*-side is divided into two distinct domains. The external driving force and the attractive interactions of the pore with the polymer, pulls on the monomers nearer to the pore and sets them in motion. The remaining monomers that are farther away from the pore, do not experience the pull and on average remain at rest. As the polymer gets sucked inside, more and more monomers on the *cis* side start responding to the force, with a tension front separating the two domains propagating along the length of the polymer. The time dependent drag experienced by a monomer $\zeta(t)$ increases as the tension front propagates and more number of monomers on the *cis*

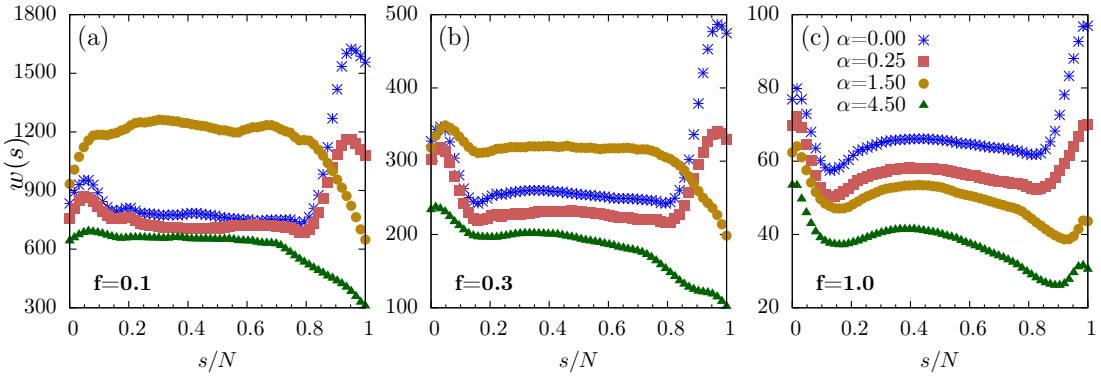

Figure 3.7: The mean waiting time $\langle w(s) \rangle$, for a flexible polymer of length $N = 64$ as a function of s/N for various pore apex angles α for external driving forces (a) $f = 0.1$, (b) $f = 0.3$, and (c) $f = 1.0$. The strength of pore-monomer interaction is $\epsilon_{\text{pore}} = 1$.

side get involved. This increase in the effective friction is manifested in the waiting times which show an initial increase with s, implying that subsequent monomers spends more time inside the pore. This continues until $\zeta(t)$ becomes maximum when the tension front reaches the last monomer, maximum number of monomers at the *cis*-side participating in the translocation process. Beyond this, the system enters the tail retraction stage where the monomers on the *cis* side starts decreasing and therefore $\zeta(t)$ decreases and so does the waiting time $w(s)$. For a short pore, this hump spreads across the polymer length. In this situation, since the pore is long, the tension propagation effect is seen upto $N \approx 16$, beyond which $w(s)$ saturates . The effect of pore-polymer interactions is again observed near the pore exit, when the attractive nature of the pore makes it difficult to pull the polymer out of it.

As we start increasing α, we observe several interesting effects which can only be attributed to the conical nature of the pore. For $\alpha = 0.25$, we note that the characteristic features observed for $\alpha = 0$ are preserved. However, $w(s)$ falls sharply for the end monomers of the chain. The marginal entropic gain in increasing the pore width near the exit, facilitates translocation. As we increase α further, for example at $\alpha = 1.5$, there is a dramatic change. The increasing width of the pore implies that the effect of the interactions of the polymer with the pore walls reduces sharply. Therefore, the effective pull on the polymer is substantially less than that for smaller α. It takes longer time for the tension to propagate to the end monomers of the chain to set them in motion. Eventually, once the tail retraction stage is reached, $w(s)$ starts to drop. Note that, for these pore angles, the large width near the *trans* side means that the attractive interactions which drags the end monomers has no effect. On increasing the pore angle further, for example $\alpha = 4.5$, the entropic gain is large

enough to ensure smooth polymer translocation with smaller $w(s)$ for all s.

For large forces, for example $f = 1.0$, this subtle competition between entropy and pore synergetics is absent. At these forces, tension propagates almost instantaneously to the end monomers of the chain and therefore we do not observe the initial hump. The waiting times of the monomer mimics the potential energy profile of the pore with an initial decrease for small s, a predominantly flat region afterwards and finally the increase near the exit. The entropic gain with increasing α is also applicable in a global sense with the decreasing $w(s)$ for all s (Fig. 3.7(c)).

3.4 Case II: Numerical results for a semiflexible polymer

In this section, we discuss the effect of bending rigidity, κ_b, of the polymer on the translocation dynamics across the conical pore. To facilitate the description, we divide translocation time τ into two parts :

1. The *filling-time*, denoted by τ_f, is the time elapsed between the entrance of the first and the last monomers of the polymer chain in the conical pore from the *cis* side.

2. The *escape-time*, denoted by τ_e is the time taken between the entrance of the last monomer of the polymer in the pore, and the exit of all the monomers from the *trans* end of the conical pore.

With these definitions, we clearly have $\tau = \tau_f + \tau_e$. Note that the transfer time discussed previously in Chapter 2, is included in the expression for the pore filling time in this approach.

3.4.1 Translocation time as a function of pore apex angle

From our previous study of driven translocation through a pore of unit length and extended interactive pores, we have seen that the translocation time increases with increasing bending rigidity, i.e. increasing stiffness. The mean translocation time $\langle\tau\rangle$ as a function of α for different polymer rigidities show this expected behavior (Fig. 3.8(a)). The translocation time $\langle\tau\rangle$ for a given α increases with increasing κ_b. However, it shows the opposite behavior in the low bending rigidity regime. For small bending rigidity, filling time dominates the total translocation process. Therefore, translocation time will decrease as bending rigidity increases initially. Note that

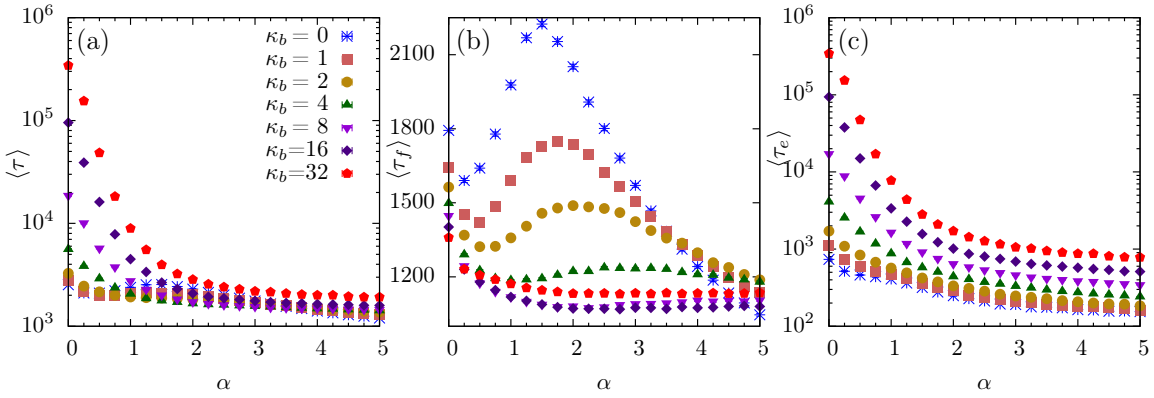

Figure 3.8: (a) Mean translocation time, $\langle\tau\rangle$ (b) Mean filling-time, $\langle\tau_f\rangle$, and (c) Mean escape-time $\langle\tau_e\rangle$, as a function of pore apex angle α for a semiflexible polymer of various bending stiffness κ_b. The external driving force is $f = 0.2$ and the strength of pore-polymer interaction is $\epsilon_{\text{pore}} = 1$.

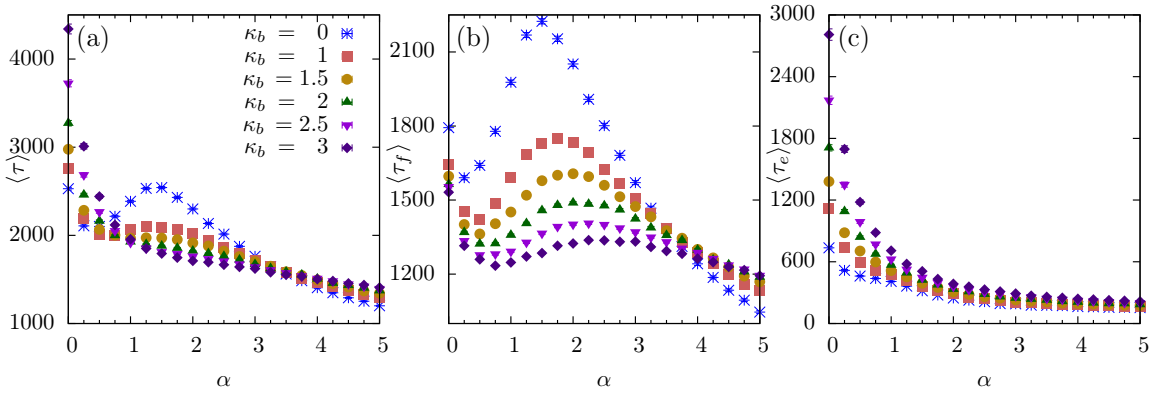

Figure 3.9: Same as Fig. 3.8 but for small bending stiffness κ_b.

this effect vanishes for $\alpha \approx 0$, which is consistent with our earlier study reported in Chapter 2. For further increase in the bending rigidity, escape time starts dominating the translocation process and hence, polymer with higher bending rigidity will take higher time to translocate. The non-monotonic behavior in the translocation times as α is varied, observed for the flexible polymer, is absent for very stiff polymers. The break up of the mean translocation time into the filling and escape times provides useful information. As seen in Fig. 3.8(b), the filling time does exhibit the non-monotonic feature as α is varied. The feature gradually disappears as the stiffness increases with decreasing effect of confinement for stiffer polymers. Moreover, as Fig. 3.8(b) suggests, $\langle\tau_f\rangle$ actually decreases with increasing κ_b for a given α. This suggests that the filling time statistics for a semiflexible polymer is still dependent on the same competition between confinement effects and pore-polymer interaction as observed for a flexible polymer. However, as Fig. 3.8(c) indicates, the translocation

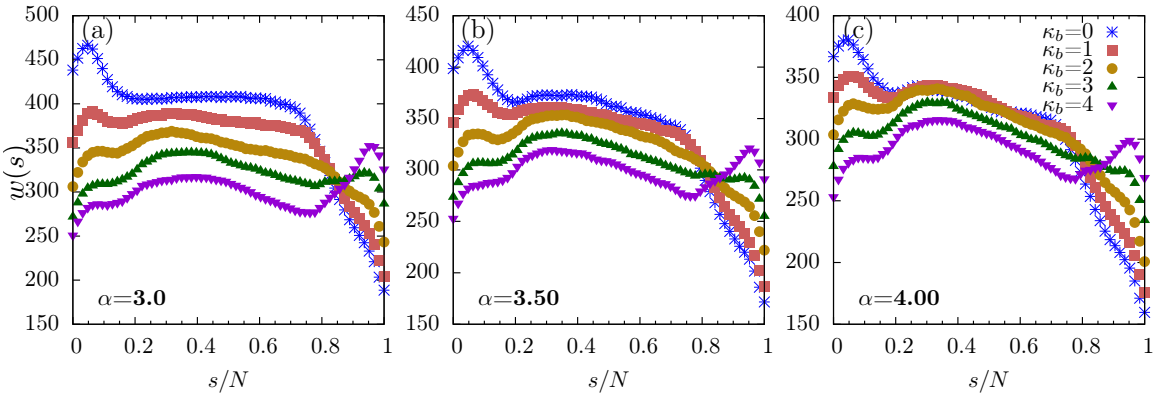

Figure 3.10: Mean waiting time $w(s)$ as a function of s/N for semiflexible polymer of various stiffness κ_b for three different pore apex angles α.

time for the semiflexible polymer is dominated by the escape time $\langle\tau_f\rangle$ which exhibits a monotonically decreasing behavior as α is increased. A blow-up of the small κ_b region indicates that the non-monotonic feature in the translocation times persists, the behavior at small κ_b resembling the situation for a flexible polymer (Fig. 3.9(a-c)).

3.4.2 Waiting time for semiflexible polymer

In Fig. 3.10, we plot the waiting time $w(s)$ distribution for a monomer inside the pore for varying polymer stiffness and for three different apex angles. Monomers of the polymer, except for those at the end, exhibit a rise of waiting time as κ_b is increased. This is consistent with the behavior observed for the mean filling times $\langle\tau_f\rangle$. It takes lesser time to fill as the stiffness is increased. The end monomers of the polymer show a completely different behavior with a decrease in mean waiting times as the stiffness of the polymer is increased. Escape is difficult for stiffer polymers. For very large stiffness (not shown here), the mean waiting times for monomers shows a monotonic increase with increasing κ_b and the chain takes longer to translocate.

3.4.3 Translocation of semiflexible polymer at low and high driving force

We also studied the effect of driving force on the $\tau - \alpha$ curve at small and large values of κ_b. In Fig. 3.11, we have shown various translocation time as a function of α for small stiffness and for three different driving forces. At small force, $\tau - \alpha$ curve shows non-monotonic dependence and this non-monotonicity vanishes as the strength of driving force is increased (Fig. 3.11(c)).

In Fig. 3.12, we have shown various translocation time as a function of α for high

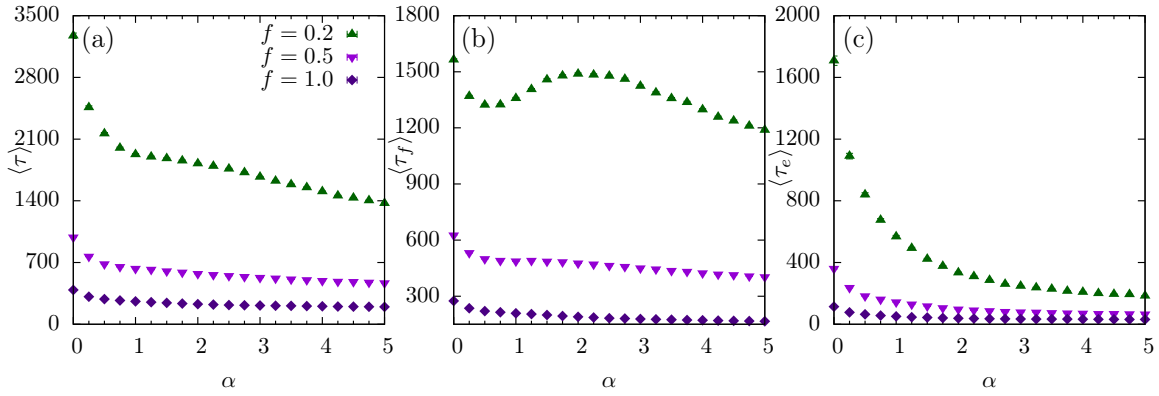

Figure 3.11: (a) Mean translocation time, $\langle\tau\rangle$ (b) Mean filling-time, $\langle\tau_f\rangle$, and (c) Mean escape-time $\langle\tau_e\rangle$, as a function of pore apex angle α for a semiflexible polymer of bending stiffness $\kappa_b = 2$ for various values of driving force f. The strength of pore-polymer interaction is $\epsilon_{\text{pore}} = 1$.

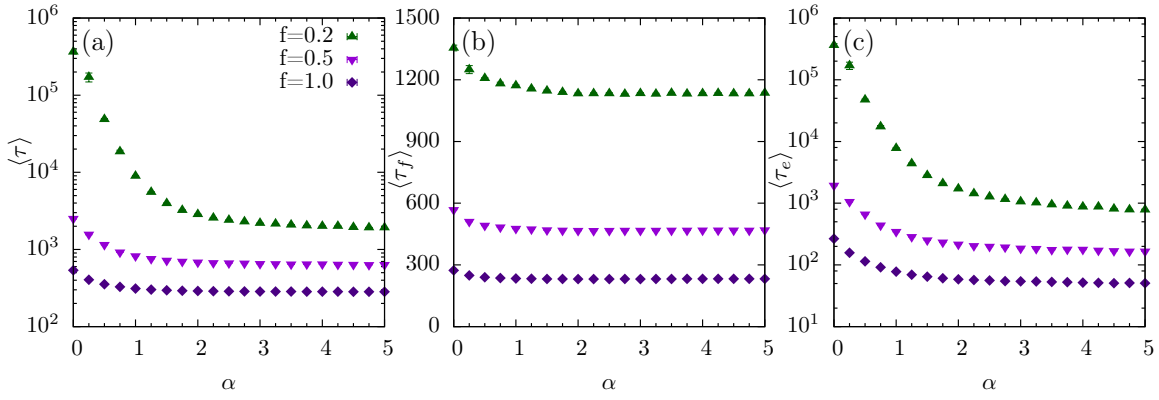

Figure 3.12: Same as Fig. 3.11 but for a semiflexible polymer of bending stiffness $\kappa_b = 32$.

stiffness and for three different driving forces. For very stiff polymers, the translocation times at a given α shows a continuous decrease with increasing driving force. The behavior across $\langle\tau_f\rangle$, $\langle\tau_e\rangle$ and $\langle\tau\rangle$ remains the same.

3.5 Conclusions

We numerically studied the driven translocation of a flexible and semiflexible polymer through an interacting cone shaped pore. We calculated the free energy of (confinement + external force) and that due to pore-polymer interactions and showed that a competition between the two can give rise to non-monotonic features in the translocation time distributions as the pore apex angle α is changed. From our earlier studies on semiflexible polymers translocating through parallel pores, we expect that stiffer polymers will take longer time to translocate. While this is in general true for translo-

cation dynamics across a conical shaped pore as well, there are additional interesting features observed at small κ_b. Here, the translocation time actually decreases as κ_b is increased, a feature which requires a detailed theoretical investigation.

Chapter 4

Flow driven translocation through a narrow channel

In this chapter, we present results for the fluid flow induced translocation of a semiflexible polymer through an extended narrow pore using *multi-particle collision dynamics* (MPCD) simulation. Sakaue *et al.* [71] used the blob picture by de Gennes, to predict the threshold velocity flux to push a polymer inside a narrow channel as

$$j_v^{th} = k_B T/\eta, \tag{4.1}$$

where k_B is the Boltzmann's constant, T is the temperature and η is the viscosity of the solvent. This suggests that the velocity flux is independent of the length of the polymer or the dimensions of the channel. Markesteijn *et. al.* [73] verified this prediction using mesoscale simulations where the fluid was modeled using Lattice Boltzmann. In this study we investigate the effect of stiffness of the polymer on this threshold flux.

In Sec. 4.1, we review the calculation of Sakaue *et. al.* for threshold velocity flux for a flexible polymer using de Gennes blob model. In Sec. 4.2 we define our model for both polymer and fluid particles and give simulation details. In Sec. 4.3 we discuss numerical results for the flow driven translocation of semiflexible polymer through a narrow pore where we establish the dependence of threshold velocity flux on the persistence length of the polymer and diameter of the pore. Finally we draw our conclusions in Sec. 4.4.

4.1 Theoretical results for a flexible polymer

We briefly review the theoretical analysis to calculate the threshold flux in flow driven translocation for a flexible polymer. Here we follow the analysis by Ledesma-Aguilar *et. al.* [80]. In the blob picture of de Gennes, a chain of length L confined inside a channel of width W breaks up into N_b blobs of size ξ. From the Flory theory it follows that the size of the blob is comparable to the pore size i.e.,

$$\xi \approx W. \tag{4.2}$$

Let us assume that in the presence of a weak driving, the polymer has entered the channel partially upto a length y. The confinement of the polymer inside the channel costs entropic penalty of the order k_BT per blob. The number of blobs, given that the chain has penetrated a distance y inside the channel, is given as $N_b \sim yW^2/\xi^3$. Then the energy cost is

$$F_C \sim k_BT\frac{yW^2}{\xi^3} \sim C_1k_BT\left(\frac{y}{W}\right), \tag{4.3}$$

where we have used Eq. 4.2. Here C_1 is a numerical prefactor.

The fluid flow drags the polymer inside the channel giving rise to hydrodynamic drag per blob as $\eta u\xi$, where $u = j_v/W^2$ is a typical velocity inside the channel, j being the velocity flux. Therefore, for N_b blobs, the free energy contribution is given as the work done by the fluid to displace the polymer a distance y:

$$F_H \sim -\frac{C_2\eta j_v}{2}\left(\frac{y}{W}\right)^2, \tag{4.4}$$

where C_2 is another constant prefactor. We can see that the two free energy terms compete with each other setting a length scale y^* for which the difference

$$\Delta F = C_1k_BT\left(\frac{y}{W}\right) - \frac{C_2\eta j_v}{2}\left(\frac{y}{W}\right)^2, \tag{4.5}$$

is a minimum. That length is

$$y^* = \left(\frac{C_1}{C_2}\right)\frac{k_BTW}{\eta j_v}. \tag{4.6}$$

This gives

$$\Delta F^* = \frac{C_3}{2}\frac{(k_BT)^2}{\eta j_v}, \tag{4.7}$$

where $C_3 = C_1^2/C_2$. Sakaue *et al.* [71] assumed that the strength of the barrier is of the order of thermal fluctuations, i.e. $\Delta F^* \sim k_B T$. This gives $y^* \sim W$ and hence the critical flow rate $j_v^{th} \sim k_B T/\eta$.

4.2 Model and Simulation Details

In this section we introduce our model. We simulated the fluid particles by MPCD method and polymer chain by conventional Molecular dynamics (MD) simulations.

We have used WCA potential $U_{\text{bead}}(r)$ (Eq. 1.15), FENE spring potential $U_{\text{bond}}(r)$ (Eq. 1.16) and a three body bond bending potential $U_{\text{bend}}(r)$ (Eq. 1.17) to mimic excluded volume, bond stretching between two consecutive monomers, and stiffness of the chain, respectively. The persistence length of a semiflexible polymer is changed by changing the bending rigidity parameter κ_b in Eq. 1.17.

The solvent is modeled explicitly in this study. The fluid particles evolve in time according to the MPCD algorithm that we have discussed in Chapter 1. To induce flow in the system, we employed gravitationally driven flow in which each fluid particle experiences a force $\boldsymbol{g} = g\hat{\boldsymbol{x}}$, acting along x-axis. In order to incorporate flow in the system, the streaming step given by Eqs. 1.29 can be rewritten as:

$$\boldsymbol{r}_i(t+h) = \boldsymbol{r}_i(t) + h\boldsymbol{v}_i(t) + \frac{h^2}{2} g\hat{\boldsymbol{x}}, \tag{4.8}$$

$$\boldsymbol{v}_i(t+h) = \boldsymbol{v}_i(t) + hg\hat{\boldsymbol{x}}. \tag{4.9}$$

Periodic boundary conditions are applied along the $\boldsymbol{x}$- and $\boldsymbol{z}$-axes, and no-slip boundary conditions in the $\boldsymbol{y}$-direction. To simulate such conditions, conventional bounce-back rule is employed in which the velocities of each particle that collide with the walls are inverted after the collision. For planar walls coinciding with the boundaries, such conditions are conveniently simulated by employing a bounce-back rule, i.e., the velocities of particles that hit the walls are inverted after the collision. But due to cell shifting, some of the cells remain partially filled near the wall and simple bounce-back rule fails to guarantee no-slip boundary conditions. So we use a generalization of the bounce-back rule which has been detailed in the introduction.

The polymer is introduced into the solvent by adding N_p point particles, each of mass M and connected by bonds. The time evolution of the chain monomers, namely positions $\boldsymbol{R}_p$ and velocities $\boldsymbol{V}_p$, are determined by solving Newton's equations of motion using the velocity Verlet update (Eq. 1.22, 1.23). The time step of Molecular dynamics for polymer, Δt_{MD} is chosen to be smaller than the time step for MPCD

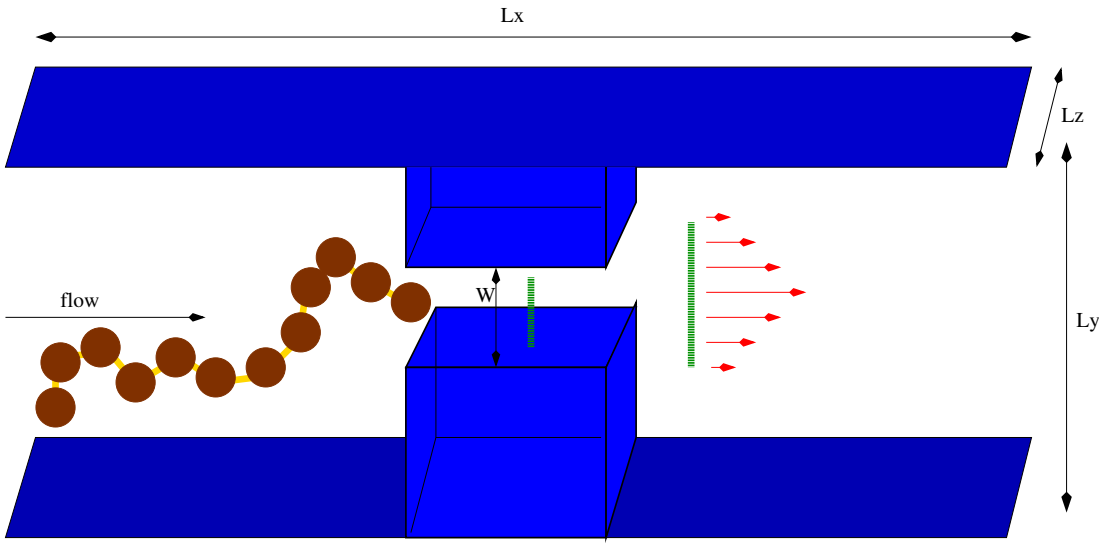

Figure 4.1: The geometry used in simulations of flow driven translocation of a semi-flexible polymer through a narrow pore. The green dotted line indicates the velocity flow profile inside the pore.

fluid particles, h. The monomer-fluid interaction is taken into account by taking monomer mass $M = \rho m$, where ρ is the number density of fluid particles in each cell. In the collision step, the center of mass velocity of the cell is calculated as

$$\boldsymbol{v}_{c.m.} = \frac{\sum\limits_{i \in cell} m\boldsymbol{v}_i + \sum\limits_{p \in cell} M\boldsymbol{V}_p}{\sum\limits_{i \in cell} m + \sum\limits_{p \in cell} M}, \tag{4.10}$$

and velocities of fluid particles and that of monomers are given by

$$\boldsymbol{v}_i(t+h) = \boldsymbol{v}_{c.m.}(t) + \mathcal{R}(\alpha)(\boldsymbol{v}_i(t) - \boldsymbol{v}_{c.m.}(t)), \tag{4.11a}$$

$$\boldsymbol{V}_p(t+h) = \boldsymbol{v}_{c.m.}(t) + \mathcal{R}(\alpha)(\boldsymbol{V}_p(t) - \boldsymbol{v}_{c.m.}(t)). \tag{4.11b}$$

where $\mathcal{R}(\alpha)$ is the stochastic rotation matrix introduced in Chapter 1.

4.2.1 Simulation details

The cell length, a, and the energy k_BT, sets the length and energy scales respectively. The unit of time is set by $\tau = h$, i.e., the time step of MPCD step.

The dimensions of the simulation box is chosen as: $L_x \times L_y \times L_z$. In this geometry, a narrow pore of size $(L_x^p \times W \times L_z^p)$. The schematic diagram of the geometry used in the simulation is shown in Fig. 4.1. Periodic boundary conditions are imposed along x-, and z-axes, and the no-slip boundary conditions along the y-axis and other wall

surfaces. For a Poiseuille flow within the pore, the mean velocity flux per unit width of the pore can be calculated by solving Navior-Stokes equation with the boundary conditions, which is given as:

$$j_v^0 = \frac{W^3 g}{12\eta}. \tag{4.12}$$

By putting the parameters values, for W, η and g, we get the reference value for the velocity flux j_v^0.

We start with a chain configuration with the first monomer placed at the entrance of the pore. This monomer is kept fixed and the remaining monomers of the chain are allowed to fluctuate to relax the polymer to its equilibrium configuration. The equilibration time for the polymer chain is taken to be 5×10^4 steps in MPCD units h (equal to 2.5×10^6 steps in MD units Δt_{MD}). A virtual wall is place inside the pore to prevent the monomers to enter the pore during equilibration. The wall is then removed and the first monomer is released from its fixed position, and a gravitationally driven flow is applied. The translocation of the polymer across the pore is than monitored. For a successful translocation, all the monomers of the polymer chain must translocate to the other side of the pore. To calculate the probability of translocation, we have taken 100 equilibrated samples.

4.3 Numerical results

4.3.1 Dependence on persistence length of the polymer

Fig. 4.2 shows the probability of translocation as a function of the applied velocity flux, for different values of the persistence length l_p. All other parameters are kept fixed.

The simulation data is in units of the reference value for the velocity flux j_v^0. The points are from our simulations and the solid lines are from fitting the data to a functional form:

$$P(j_v) = \frac{1}{2}\left[1 + \mathrm{erf}\left\{\frac{j_v - j_v(1/2)}{B}\right\}\right], \tag{4.13}$$

where $j_v(P_t)$ is the value of applied velocity flux for which threshold fraction P_t of the polymer has been translocated, B is a fitting parameter and $\mathrm{erf}\{\cdot\}$ is an error function. On increasing the velocity flux j_v, the drag increases, so the probability of translocation of polymer chain also increases which is reflected in the transition from a non-translocation (i.e., $P = 0$) to a successful translocation ($P = 1$) of the polymer as velocity flux is increased (Fig 4.2). This transition is similar to the results obtained by Markesteijn *et. al.* [73] for a flexible chain using lattice Boltzmann method. While

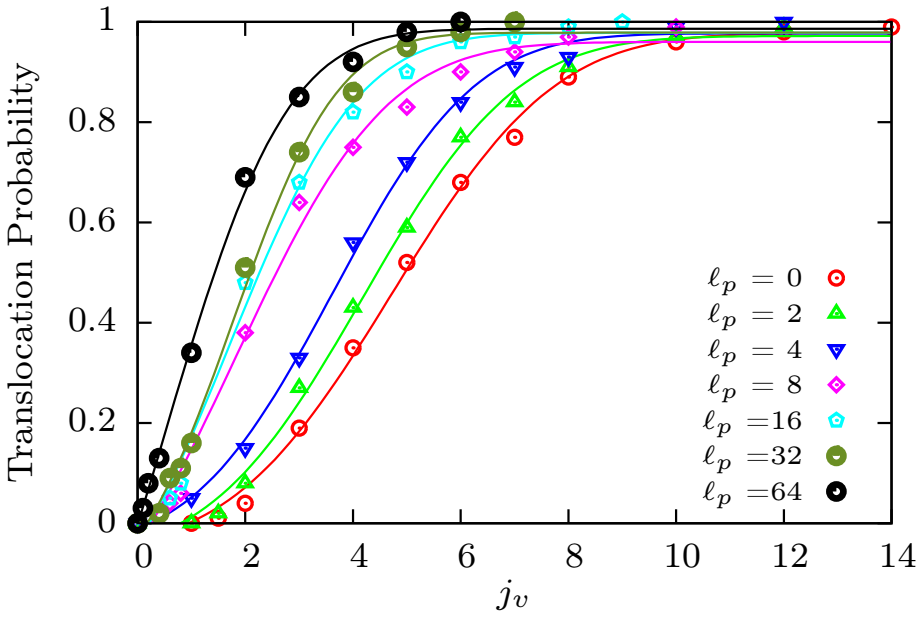

Figure 4.2: Simulation results of the translocation probability as a function of applied velocity flux for polymers of different persistence lengths ℓ_p. The points are results from simulations and the solid lines are the best fit to the data.

this smooth transition from $P = 0$ to $P = 1$ as a function of applied velocity flux j_v holds for all polymer chains with different persistence length considered, we found that the curves shift towards left as ℓ_p is increased. This indicates that the threshold velocity flux needed to start the translocation of the chain gets lower as persistence length of the polymer increases. To define this behavior more quantitatively, we define a threshold velocity flux as j_v^{th} as $j_v(P_t = 0.5)$ (which means that value of j_v^{th} for which 50% of polymer chains have been translocated) and calculate the corresponding velocity flux using Eq. 4.13. The measured values of the threshold velocity flux for different ℓ_p of the polymer is tabulated in Table 4.1. The threshold velocity flux, j_v^{th}

Persistence length (ℓ_p)	Threshold velocity flux (j_v^{th})
0	4.92
2	4.52
4	3.68
8	2.61
16	2.27
32	1.98
64	1.46

Table 4.1: Threshold velocity flux j_v^{th} for semiflexible polymer of different persistence length ℓ_p.

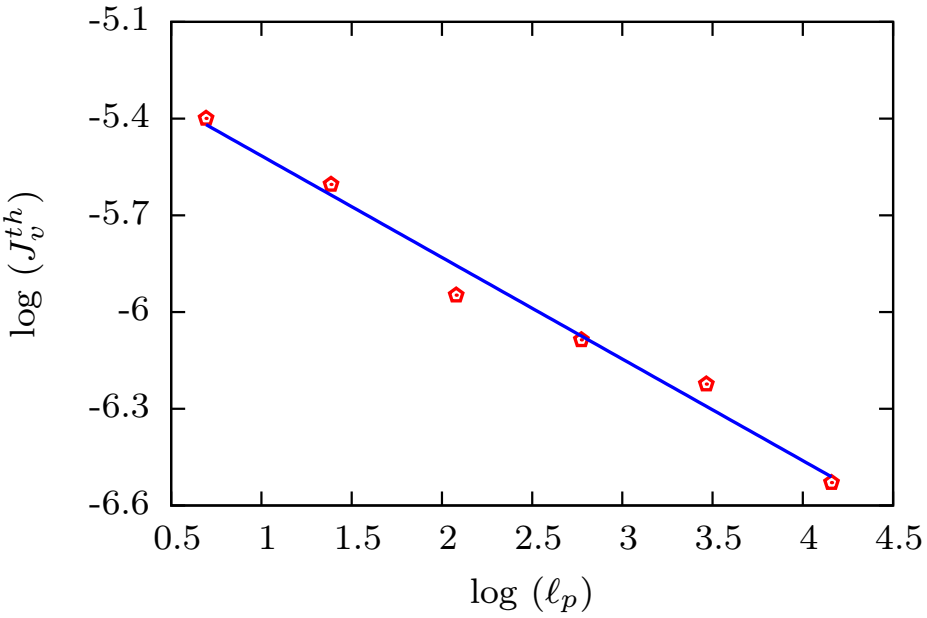

Figure 4.3: Log-Log plot of threshold velocity flux j_v^{th} as a function of persistence length ℓ_p of the polymer.

as a function of ℓ_p when plotted in a log-log scale shows a straight line (see Fig. 4.3) indicating a power law dependence between them:

$$\frac{j_v^{th}}{j_v^{th}(\ell_p = 0)} \sim L_s \frac{1}{\ell_p^{\beta}}, \tag{4.14}$$

where L_s is some length scale to identified. Note that requirement of L_s to be a length scale follows from dimensional analysis. On fitting the data with a straight line, we obtain the exponent $\beta = 0.315 \pm 0.025 \sim 1/3$.

4.3.2 Dependence on diameter of the pore

In Fig 4.4, we have shown the probability, P, of translocation of polymer chain as a function of applied velocity flux, j_v, for different pore diameter and two different values of the persistence length. To check the dependence of threshold velocity flux on the diameter of pore, we choose respective values of W and ℓ_p in such a way, so that the ratio W/ℓ_p remains constant. In Table 4.2, we have shown values of threshold velocity flux for various pore diameter and two different persistence length corresponding to each pore diameter.

We found that for the choosen values of W and ℓ_p, scaled threshold velocity flux,$j_v^{th}(\ell_p)/j_v^{th}(\ell_p = 0)$ remains the same which indicates that j_v^{th} is increasing function of W with the same exponent as we measured with the persistence length,

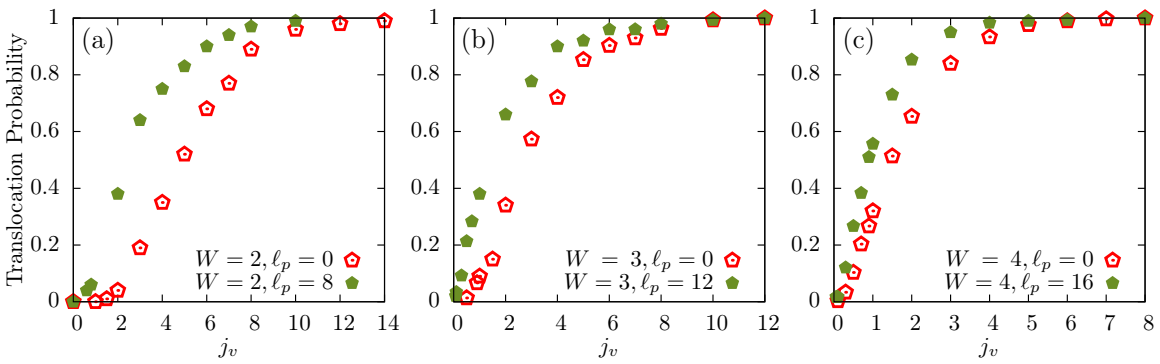

Figure 4.4: Simulation results for the translocation probability as a function of applied velocity flux for different pore widths and different persistence lengths ℓ_p of the polymer.

W	ℓ_p	$\frac{W}{\ell_p}$	$j_v^{th}(\ell_p = 0)$	$j_v^{th}(\ell_p)$	$\frac{j_v^{th}(\ell_p)}{j_v^{th}(\ell_p = 0)}$
2	8	0.25	4.92	2.61	0.528
3	12	0.25	2.89	1.51	0.522
4	16	0.25	1.53	0.80	0.523

Table 4.2: Threshold velocity flux j_v^{th} for semiflexible polymer of different persistence lengths ℓ_p translocating through pores of different widths.

$\beta = 0.315$. This suggests that for semiflexible polymer, threshold velocity flux is given by the following relation,

$$\frac{j_v^{th}}{j_v^{th}(\ell_p = 0)} \simeq \left(\frac{W}{\ell_p}\right)^{\frac{1}{3}}. \tag{4.15}$$

4.4 Conclusions

In this chapter, we studied flow driven translocation of semiflexible polymer through a narrow pore where we are looking for the translocation probability of polymer as a function of applied velocity flux for different values of persistence lengths. We have showed numerically that the threshold velocity flux of semiflexible polymer not only depends on the temperature and viscosity of the fluid, but also on the persistence length of the polymer and diameter of the pore, which sets a scale for confinement of the polymer.

Chapter 5

Summary

In this thesis, we have studied theoretically the driven translocation of a semiflexible polymer through narrow pores having different pore patterning and pore geometries. We have also studied the flow induced driven translocation.

In Chapter 2, we studied the translocation of homogeneous and heterogeneous semiflexible polymer through a slit (i.e., pore of unit length) and extended patterned pores. For translocation through a slit, we found that the waiting time first increases, reaches a maximum, and then decreases with the increase in monomer index. We showed using IFTP theory, that as the stiffness of the semiflexible polymer increases, the peak in the waiting time also increases, but shifts towards lower monomer index. For the case of an extended pore, we have taken three different types of pores with different patterned stickiness. We found that the pore-polymer interaction has a significant effect on the total translocation time of homogeneous semiflexible polymer. As the stiffness of the polymer increases, it is found that the waiting time increases and the peak shifts towards the lower monomer index, same as reported for translocation through a slit, but we also observed a sharp increase in the waiting time of last few monomers. This increase in the waiting time is attributed to the pore polymer interaction which slows down the escape of the polymer from the pore. The division of total translocation time in three parts: filling, transfer and escape times, is found to be useful which reveal many interesting features for semiflexible polymer translocation through extended pores. We studied the translocation of heteropolymer which consists of periodic blocks of stiff and flexible segments. Since the stiffer segments took longer time to translocate through the pore than the flexible segments, we obtained periodic patterns in the mean waiting times for heteropolymers. We proposed a strategy for efficient sequencing of heteropolymers with varying bending rigidity along the polymer backbone, which could be verified experimentally. We argue that the statistical fluctuations in the translocation time could be utilized for efficient se-

quencing of heteropolymers by suitably engineering pore-polymer interactions, and combining the readouts from multiple pores.

In Chapter 3, we studied the driven translocation of a semiflexible polymer through an interacting pore with cone shaped geometry. We obtained an expression for the free energy due to partial confinement of a flexible polymer inside the pore and found that the competition between the sum of free energy of confinement and external force, and the free energy due to pore-polymer interactions can give rise to non-monotonic features in the translocation time distributions as the pore apex angle α is changed. From our studies on translocation of semiflexible polymers through extended pore (Chapter 2), we expect that a stiffer polymer will take longer time to translocate. While this is generally true for the translocation dynamics across a conical shaped pore also, we observed additional interesting features at small stiffness κ_b. In this region, we found that the translocation time actually decreases as κ_b is increased.

In Chapter 4, we studied flow driven translocation of a semiflexible polymer through a narrow pore using hybrid MD-MPCD algorithm. We monitored the translocation probability of semiflexible polymer as a function of applied velocity flux for polymers of various stiffness, and the diameter of the pore. Earlier studies, on flexible polymer, have indicated that the threshold velocity flux depends only on the temperature and viscosity of the fluid, but independent of polymer properties. We found that the translocation probability curves shift towards lower values of applied velocity flux as the persistence length of polymer (i.e., stiffness) increases. We found that the threshold velocity flux for a semiflexible polymer not only depends on the temperature and viscosity of the fluid, but also on the pore diameter and persistence length of the polymer as $j_v^{th} \sim (W/\ell_p)^{\beta}$, with exponent $\beta \approx 1/3$.

In this thesis, we have found that the pore-polymer interactions affect the translocation time significantly. This could be exploited in designing extended nanopores with suitable pore patterning for rapid DNA sequencing in future.

Appendix A

Equation of motion for the tension front

In this appendix, we give mathematical derivation of the equations of motion for the tension front, i.e., Eqs. 2.4, 2.7, used in Chapter 2. In Sec. A.1, we calculate the force experienced by monomers on the *cis* side as a function of distance from the pore and in Sec. A.2, we derive the equations of motion for the tension front in the stem-flower regime.

A.1 Force experienced by monomers on *cis* side

In order to calculate the force experienced by monomers on the *cis* side, we use iso-flux assumption [42], i.e., the flux of mobile monomers ($\tilde{\phi} = ds/d\tilde{t}$) on the *cis* side and passing through the pore is constant. The driving force acting on monomers inside the pore, also pulls few monomers that are nearer to the pore entrance and set them in motion. In our study, we have taken an external driving force $f = 5$. This force is strong enough to straighten a small portion of the chain near the pore entrance, but weak for the remaining monomers in the *cis* side up to which the tension has propagated and hence the chain is still in coiled shape. The later is called a flower and the straighten part is called a stem, and this regime is called stem-flower regime (Fig A.1).

The estimation of the distance, $\tilde{x}$, to which tension has propagated is given by $\tilde{R}(\tilde{t})$, i.e., $\tilde{x} = -\tilde{R}(\tilde{t})$. At $\tilde{x} = \tilde{r}(\tilde{t})$, which is the boundary between these two regions, the tension force has the value of unity. Thus, by integrating the force balance relation $d\tilde{f}(\tilde{x}') = -\tilde{\phi}(\tilde{t})d\tilde{x}'$, for a differential element for the stem region and solve for $\tilde{r}(\tilde{t})$,

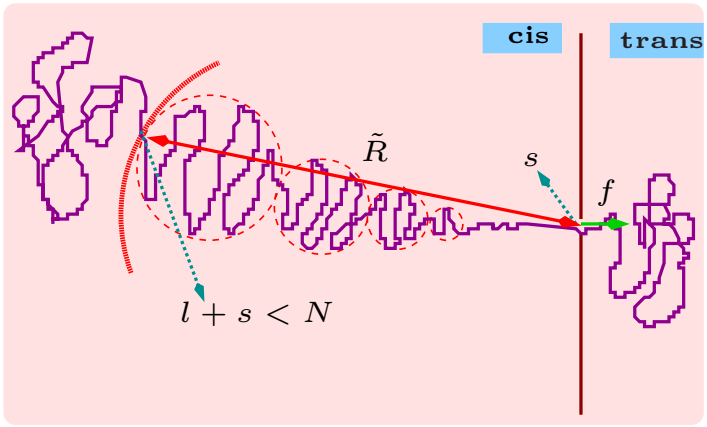

Figure A.1: Schematic diagram of driven translocation during tension propagation stage in stem-flower regime. The driving force f acts on the monomer that is inside the pore. The total number of monomers in the polymer chain is N and s denotes the number of monomers that are translocated in the *trans* side, and l represents the number of mobile monomers in the *cis* side. The location of the last monomer from the pore entrance up to which the tension is propagated is $\tilde{R}$.

one gets;

$$\tilde{r}(\tilde{t}) = \frac{\tilde{f}_0 - 1}{\tilde{\phi}(\tilde{t})}. \tag{A.1}$$

By integrating the force balance equation over the distance between $\tilde{r}$ and $\tilde{x}$, and putting $\tilde{f}(\tilde{r}) = 1$, in the flower regime, one obtains the following relation:

$$\tilde{f}(\tilde{x}) = 1 - \tilde{\phi}(\tilde{t})(\tilde{x} - \tilde{r}). \tag{A.2}$$

Combining Eqs. A.1 and A.2, one gets the expression for the force as a function of distance as

$$\tilde{f}(\tilde{x}) = \tilde{f}_0 - \tilde{\phi}(\tilde{t})\tilde{x}, \tag{A.3}$$

where $\tilde{f}_0 = \tilde{f}_{tot} - \tilde{\zeta}_p\tilde{\phi}(\tilde{t})$ is the force at the entrance of the pore.

A.2 Equations of motion

The end-to-end distance of a semiflexible self-avoiding chain scales as [50],

$$\tilde{R} = A_\nu \tilde{\ell}_p^{\nu_p} M^\nu. \tag{A.4}$$

In two dimensions, $\nu_p = 1/4$ and $\nu = 3/4$ [96]. Here, M is the sum of the number of mobile monomers in the *cis* side l and the number of translocated monomers s. So,

Eq. A.4 can be rewritten as,

$$\tilde{R} = A_\nu \tilde{\ell}_p^{\nu_p} (l+s)^\nu. \tag{A.5}$$

To calculate the equation of motion of the tension front located at $\tilde{R}$, one needs to calculate l, s and then take the time derivative of $\tilde{R}$ which is given as

$$\dot{\tilde{R}}(\tilde{t}) = \nu A_\nu^{1/\nu} \tilde{\ell}_p^{\nu_p/\nu} \tilde{R}(\tilde{t})^{\frac{\nu-1}{\nu}} [\dot{l}(\tilde{t}) + \dot{s}(\tilde{t})], \tag{A.6}$$

where by definition,

$$\frac{ds(\tilde{t})}{d\tilde{t}} = \dot{s}(\tilde{t}) = \tilde{\phi}(\tilde{t}). \tag{A.7}$$

The number of mobile *cis* monomers, i.e., $l(\tilde{t})$, is calculated by integrating the linear monomer number density, $\tilde{\sigma}(\tilde{t})$, over the distance between 0 and $\tilde{R}$. Here, we have used the blob theory to calculate $\tilde{\sigma}(\tilde{t})$. When a blob is formed by applying the tension force on the backbone of the polymer chain, the blob size, $\tilde{\xi}(\tilde{x})$, is given by $\tilde{\xi}(\tilde{x}) = 1/|\tilde{f}(\tilde{x})|$. The number of monomers inside the each blob, $g(\tilde{x})$, also depends on the position of the blob and is given by size scales as $g(\tilde{x}) = \tilde{\xi}(\tilde{x})^{1/\nu}$. Thus the monomer number density is given by $\tilde{\sigma}(\tilde{x},\tilde{t}) = g/\tilde{\xi} = \tilde{\xi}^{-1+1/\nu} = |\tilde{f}(\tilde{x})|^{1-1/\nu}$. Therefore, the number of mobile monomers in the *cis* side can be derived as

$$\begin{aligned}
l_{SF}(\tilde{t}) &= \int_0^{\tilde{R}(\tilde{t})} \tilde{\sigma}(\tilde{x},\tilde{t}) d\tilde{x}, \\
&= \int_0^{\tilde{r}(\tilde{t})} \tilde{\sigma}(\tilde{x},\tilde{t}) d\tilde{x} + \int_{\tilde{r}(\tilde{t})}^{\tilde{R}(\tilde{t})} \tilde{\sigma}(\tilde{x},\tilde{t}) d\tilde{x}, \\
&= \tilde{r}(\tilde{t}) + \int_{\tilde{r}(\tilde{t})}^{\tilde{R}(\tilde{t})} |\tilde{f}(\tilde{x})|^{(\nu-1)/\nu} d\tilde{x}, \\
&= \frac{\tilde{\phi}(\tilde{t})\tilde{R}(\tilde{t}) - 1}{\tilde{\phi}(\tilde{t})} + \int_{\tilde{r}(\tilde{t})}^{\tilde{R}(\tilde{t})} |\tilde{\phi}(\tilde{t})\tilde{R}(\tilde{t}) - \tilde{\phi}(\tilde{t}\tilde{x})|^{(\nu-1)/\nu} d\tilde{x}.
\end{aligned} \tag{A.8}$$

After solving the integral, $l_{SF}(\tilde{t})$ becomes,

$$l_{SF}(\tilde{t}) = \tilde{R}(\tilde{t}) + \frac{1-\nu}{(2\nu-1)\tilde{\phi}(\tilde{t})}, \tag{A.9}$$

where the subscript SF denotes the stem-flower regime. Taking time derivative of Eq. A.9, the time derivative of the number of mobile beads, $\dot{l}(\tilde{t})$ is given by

$$\dot{l}_{SF}(\tilde{t}) = \mathcal{B}_{SF} \times [\dot{\tilde{f}}_{tot}(\tilde{t}) - \tilde{\phi}(\tilde{t})\dot{\tilde{R}}(\tilde{t})] + \mathcal{C}_{SF} \times \dot{\tilde{f}}_{tot}(\tilde{t}), \tag{A.10}$$

where,

$$\mathcal{B}_{SF} = -\frac{1}{\tilde{\phi}(\tilde{t})} + \frac{\nu - 1}{(2\nu - 1)[\tilde{\zeta}_p + \tilde{\zeta}_t + \tilde{R}(\tilde{t})]\tilde{\phi}(\tilde{t})^2}, \tag{A.11}$$

and

$$\mathcal{C}_{SF} = \frac{1}{\tilde{\phi}(\tilde{t})}. \tag{A.12}$$

From Eqs. A.6, A.7 and A.10, the equation of motion of the tension front can be written as

$$\dot{\tilde{R}}(\tilde{t}) = \frac{\mathcal{A}\tilde{R}(\tilde{t})^{\frac{\nu-1}{\nu}} \left[(\mathcal{B}_{SF} + \mathcal{C}_{SF}) \times \dot{\tilde{f}}_{tot}(\tilde{t}) + \tilde{\phi}(\tilde{t})\right]}{1 + \mathcal{A}\tilde{R}(\tilde{t})^{\frac{\nu-1}{\nu}} \mathcal{B}_{SF} \times \tilde{\phi}(\tilde{t})}, \tag{A.13}$$

where,

$$\mathcal{A} = \begin{cases} \nu A_\nu^{1/\nu}, & \text{for} \quad \lambda = 0 \\ \nu A_\nu^{1/\nu} \tilde{\ell}_p^{\nu_p/\nu}, & \text{for} \quad \lambda \neq 0, \end{cases}$$

In the post propagation stage, the tension front has reached the last monomer of the chain in the *cis* side. We therefore have $s + l = N$ and the time derivative of this closure relation is given by

$$\dot{l} + \dot{s} = 0 \tag{A.14}$$

From Eqs.. A.7, A.10 and A.14 the equation of motion for the tension front in the post propagation stage is given by

$$\dot{\tilde{R}}(\tilde{t}) = \frac{(\mathcal{B}_{SF} + \mathcal{C}_{SF}) \times \dot{\tilde{f}}_{tot}(\tilde{t}) + \tilde{\phi}(\tilde{t})}{\mathcal{B}_{SF} \times \tilde{\phi}(\tilde{t})}. \tag{A.15}$$

Appendix B

Algorithm for the sequence detection of hetereopolymer

In this appendix, we give the details of the algorithm used in Chapter 2 for the detection of heteropolymer sequence having alternate blocks of stiff and flexible segments. The sequence is passed through Pores β and γ multiple times, and the statistical nature of readouts are used to predict the sequence. The orientation of the polymer sequence can be detected by translocating it through Pore α.

The detailed algorithm of sequence detection is as follows:

Algorithm to find Accuracy of Detection

Step 1. Take m number of copies of Pore β and Pore γ.

Step 2. Choose a random sequence r from the set

$$r \in \{(S_2F_2)_{32}, (S_4F_4)_{16}, (S_8F_8)_8, (S_{16}F_{16})_4, (S_{32}F_{32})_2, (S_{64}F_{64})\}.$$

Step 3. Pass the polymer sequence r, m times from Pores β and γ and record the translocation time through each pore independently. Thus, we will have m values each for Pores β and γ.

Step 4. Let $x_r^{\beta,\gamma}$ and $y_r^{\beta,\gamma}$ represent the mean and the standard deviation, respectively, calculated from the m values of the translocation times obtained for Pores β and γ separately. Therefore, $(x_r^{\beta}, y_r^{\beta})$ and $(x_r^{\gamma}, y_r^{\gamma})$ are the coordinates of a point in the scatter plot for Pore β (Chapter 2 Fig. 9(b)) and Pore γ (Chapter 2 Fig. 9(c)), respectively.

Step 5. Calculate the Euclidean distance of these points from the tabulated values $x_\mu^{\beta,\gamma}$ and $y_\mu^{\beta,\gamma}$ $(\mu = 2, 4, 8, 16, 32, 64)$ of mean and standard deviation of the translocation times of the above sequences through Pore β and Pore γ

$$(dS_\mu^{\beta,\gamma})^2 = \frac{(x_r^{\beta,\gamma} - x_\mu^{\beta,\gamma})^2 + (y_r^{\beta,\gamma} - y_\mu^{\beta,\gamma})^2}{(x_r^{\beta,\gamma})^2 + (y_r^{\beta,\gamma})^2}.$$

The $(dS_\mu^{\beta,\gamma})^2$ is a 1×6 row matrix.

Step 6. The error matrix $R_{\mu,\nu}$ is obtained using

$$(dS_\mu)^2 = R_{\mu,\nu}(dS_\mu^\beta)^2(dS_\mu^\gamma)^2.$$

The error matrix is diagonal as we are looking for a particular sequence r, in both the pores. To get the sequence we need to minimize the above expression

$$(dS_r)^2 = min[R_{\mu,\nu}(dS_\mu^\beta)^2(dS_\mu^\gamma)^2].$$

If our unknown sequence matches with the detected sequence, we mark it as a successful attempt. The accuracy of detection is thus is calculated as

$$\text{Accuracy} = \frac{\text{Number of successful attempts}}{\text{Total numbers of attempts}} \times 100.$$

A similar algorithm, with multiple pores of type "Pore α", is used to predict the orientation of the polymer sequence.

Bibliography

[1] M. Muthukumar, *Polymer Translocation* (Taylor & Francis, Boca Raton, 2011).

[2] Y. Feng *et al.*, Genomics Proteomics Bioinformatics **15**, 4 (2015).

[3] A. Meller, Accessed: 2017-08-27, http://meller.bm.technion.ac.il/research_Nanopore_fab.html.

[4] B. Alberts *et al.*, *Molecular biology of the cell* (2nd ed. Garland Publishing, New York, 1989).

[5] B. Dreiseikelmann, Microbiol. Rev. **58**, 293 (1994).

[6] W. Wickner and R. Schekman, Science **310**, 1452 (2005).

[7] M. Bukrinsky, Mol. Med. **10**, 1 (2004).

[8] J. Frank and R. L. Gonzalez Jr., Ann. Rev. Biochem. **79**, 381 (2010).

[9] H. Salman *et al.*, Proc. Natl. Acad. Sci. USA **98**, 7247 (2001).

[10] D. Branton *et al.*, Nature Biotechnol. **26**, 1146 (2008).

[11] J. M. Tsutsui, F. Xie, and R. T. Porter, Cardiovasc Ultrasound **2**, 23 (2004).

[12] S. K. Sia and G. M. Whitesides, Electrophoresis **24**, 3563 (2003).

[13] P. Flory, *Principles of Polymer Chemistry* (Cornell University Press, Ithaca, N. Y., 1971).

[14] B. Maier and J. O. Rdler, Phys. Rev. Lett. **82**, 1911 (1999).

[15] J. J. Prentis and D. R. Sisan, Phys. Rev. E **65**, 031306 (2002).

[16] A. Berretti and A. D. Sokal, J. Stat. Phys. **40**, 483 (1985).

[17] M. Rubinstein and R. H. Colby, *Polymer Physics* (Oxford University Press, 2003).

[18] M. Muthukumar, C. Plesa, and C. Dekker, Physics Today **68**, 40 (2015).

[19] B. M. Venkatesan and R. Bashir, Nat. Nanotechnol. **6**, 615 (2011).

[20] T. M. Butler *et al.*, Proc. Nat. Acad. Sci. U S A **52**, 20647 (2008).

[21] J. J. Kasianowicz, E. Brandin, D. Branton, and D. W. Deamer, Proc. Natl. Acad. Sci. U.S.A. **93**, 13770 (1996).

[22] S. Benner *et al.*, Nature Nanotechnol. **2**, 718 (2007).

[23] S. L. Cockroft*et al.*, J. Am. Chem. Soc. **130**, 818 (2008).

[24] I. M. Derrington *et al.*, Proc. Natl. Acad. Sci. USA **107**, 1606016065 (2010).

[25] D. Wendell *et al.*, Nat. Nanotechnol. **4**, 765772 (2009).

[26] U. F. Keyser and J. R. Soc, Interface **8**, 1369 (2011).

[27] J. Li *et al.*, Nat. Materials **2**, 611 (2003).

[28] A. J. Storm *et al.*, Nature Mater. **2**, 537 (2003).

[29] A. J. Storm *et al.*, Nano Lett. **5**, 11931197 (2005).

[30] M. D. Fischbein and M. Drndic, Appl. Phys. Lett. **93**, 113107 (2008).

[31] U. F. Keyser *et al.*, Nature **2**, 473 (2006).

[32] M. M. Mohammad, S. Prakash, A. Matouschek, and L. Movileanu, J. Ame. Chem. Soc. **130**, 4081 (2008).

[33] A. R. Hall *et al.*, Nature Nanotechnol. **5**, 874 (2010).

[34] W. Sung and P. J. Park, Phys. Rev. Lett. **77**, 783 (1996).

[35] M. Muthukumar, J. Chem. Phys. **111**, 10371 (1999).

[36] M. Muthukumar, Phys. Rev. Lett. **86**, 3188 (2001).

[37] J. Chuang, Y. Kantor, and M. Kardar, Phys. Rev. E **65**, 011802 (2001).

[38] K. Luo, T. Ala-Nissila, and S. C. Ying, J. Chem. Phys. **124**, 034714 (2006).

[39] K. Luo, I. Huopaniemi, T. Ala-Nissila, and S. C. Ying, J. Chem. Phys. **124**, 114704 (2006).

[40] I. Huopaniemi, K. Luo, T. Ala-Nissila, and S. C. Ying, J. Chem. Phys. **125**, 124901 (2006).

[41] Y. Kantor and M. Kardar, Phys. Rev. E **69**, 021806 (2004).

[42] P. Rowghanian and A. Y. Grosberg, J. Phys. Chem. B **115**, 14127 (2011).

[43] T. Sakaue, Phys. Rev. E **76**, 021803 (2007).

[44] T. Sakaue, Phys. Rev. E **81**, 041808 (2010).

[45] T. Saito and T. Sakaue, Eur. Phys. J. E **34**, 135 (2011).

[46] T. Ikonen, A. Bhattacharya, T. Ala-Nissila, and W. Sung, Phys. Rev. E **85**, 051803 (2012).

[47] T. Ikonen, A. Bhattacharya, T. Ala-Nissila, and W. Sung, J. Chem. Phys. **137**, 085101 (2012).

[48] J. L. A. Dubbeldam, V. G. Rostiashvili, A. Milchev, and T. A. Vilgis, Phys. Rev. E **85**, 041801 (2012).

[49] J. Sarabadani, T. Ikonen, and T. Ala-Nissila, J. Chem. Phys. **141**, 214907 (2014).

[50] J. Sarabadani *et al.*, Scientific Reports **7**, 7423 (2017).

[51] A. Bhattacharya, Polymer Science, Ser. C **55**, 60 (2013).

[52] R. Adhikari and A. Bhattacharya, J. Chem. Phys. **138**, 204909 (2013).

[53] R. Adhikari and A. Bhattacharya, Europhys. Lett. **109**, 38001 (2015).

[54] H. W. de Haan and G. W. Slater, Phys. Rev. Lett. **110**, 048101 (2013).

[55] M. J. Kim, M. Wanunu, D. C. Bell, and A. Meller, Adv. Mater. **18**, 3149 (2006).

[56] T. Ohshiro and Y. Umezawa, Proc. Natl. Acad. Sci. USA **103**, 1014 (2006).

[57] M. D. Iqbal, S. M. andAkin and R. Bashir, Nature Nanotechnol. **2**, 243 (2007).

[58] M. Wanunu and A. Meller, Nano Lett. **7**, 1580 (2007).

[59] P. Chen *et al.*, Nano Lett. **4**, 1333 (2004).

[60] V. Tabard-Cossa *et al.*, Nanotechnology **18**, 305505 (2007).

[61] K. Luo, T. Ala-Nissila, S. C. Ying, and A. Bhattacharya, Phys. Rev. Lett. **99**, 148102 (2007).

[62] K. Luo, T. Ala-Nissila, S. C. Ying, and A. Bhattacharya, Phys. Rev. E **78**, 061918 (2008).

[63] K. Luo, T. Ala-Nissila, S. C. Ying, and A. Bhattacharya, Phys. Rev. Lett. **100**, 058101 (2008).

[64] K. Luo, T. Ala-Nissila, S. C. Ying, and A. Bhattacharya, J. Chem. Phys. **126**, 145101 (2007).

[65] S. Mirigian, Y. Wang, and M. Muthukumar, J. Chem. Phy **137**, 064904 (2012).

[66] J. A. Cohen, A. Chaudhuri, and R. Golestanian, Phys. Rev. X **2**, 021002 (2012).

[67] J. A. Cohen, A. Chaudhuri, and R. Golestanian, J. Chem. Phys. **137**, 204911 (2012).

[68] H. H. Katkar and M. Muthukumar, J. Chem. Phys. **140**, 135102 (2014).

[69] N. Nikoofard, H. Khalilian, and H. Fazli, J. Chem. Phys. **139**, 074901 (2013).

[70] N. Nikoofard and H. Fazli, Soft Matter **11**, 4879 (2015).

[71] T. Sakaue *et al.*, Europhys. Lett. **72**, 83 (2005).

[72] T. Sakaue and E. Raphäel, Macromolecules **39**, 2621 (2006).

[73] A. P. Markesteijn *et al.*, Soft Matter **5**, 4575 (2009).

[74] P. G. de Gennes, *Scaling concepts in polymer physics* (Cornell University press, Ithaca, 1979).

[75] U. Frisch, B. Hasslacher, and Y. Pomeau, Phys. Rev. Lett. **56**, 1505 (1986).

[76] G. R. McNamara and G. Zanetti, Phys. Rev. Lett. **61**, 2332 (1988).

[77] P. J. Hoogerbrugge and J. M. V. A. Koelman, Europhys. Lett. **19**, 155 (1992).

[78] A. Malevanets and R. Kapral, J. Chem. Phys. **110**, 8605 (1999).

[79] A. Nikoubashman and C. N. Likos, J. Chem. Phys. **133**, 074901 (2010).

[80] R. Ledesma-Aguilar *et al.*, Soft Matter **8**, 1884 (2012).

[81] G. Grest and K. Kremer, Phys. Rev. A **33**, 3628 (1986).

[82] M. Doi and S. F. Edwards, *The theory of polymer dynamics* (Oxford University Press, 1988).

[83] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, 1987).

[84] A. Malevanets and R. Kapral, J. Chem. Phys. **112**, 7260 (2000).

[85] E. Tuzel, M. Stauss, T. Ihle, and D. M. Kroll, Phys. Rev. E **68**, 036701 (2003).

[86] E. Allahyarov and G. Gompper, Phys. Rev. E **66**, 036702 (2002).

[87] T. Ihle and D. M. Kroll, Phys. Rev. E **67**, 066705 (2003).

[88] T. Ihle and D. M. Kroll, Phys. Rev. E **63**, 020201 (2001).

[89] A. Lamura, G. Gompper, T. Ihle, and D. M. Kroll, Europhys. Lett. **56**, 319 (2001).

[90] T. Ihle and D. M. Kroll, Phys. Rev. E **67**, 066706 (2003).

[91] N. Kikuchi, C. M. Polley, J. F. Ryder, and J. M. Yeomans, J. Chem. Phys. **119**, 6388 (2003).

[92] M. Ripoll, K. Mussawisade, R. G. Winkler, and G. Gompper, Phys. Rev. E **72**, 016701 (2005).

[93] R. Zwanzig, Ann. Rev. Phys. Chem. **16**, 67 (1965).

[94] C. C. Haung, A. Varghese, G. Gompper, and R. G. Winkler, Phys. Rev. E **91**, 013310 (2015).

[95] C. C. Haung *et al.*, J. Comput. Phys. **229**, 168 (2010).

[96] N. Nakanishi, J. Physique **48**, 979 (1987).

[97] J. Widom, Q Rev Biophys **34**, 269 (2001).

[98] A. A. Travers, Philos Trans Roy Soc A **362**, 1423 (2004).

[99] S. Geggier and A. Vologodskii, Proc. Natl. Acad. Sci USA **107**, 15421 (2010).

[100] C. Branden and J. Tooze, *Introduction of Protein Structure* (Garland Publishing, New York, 1998).

[101] M. Lee, B. K. Cho, and W. C. Zin, Chem. Rev. **101**, 3869 (2001).

[102] E. Stellwagen, Y. Lu, and N. Stellwagen, Biochemistry **42(11)**, 3275 (2003).

[103] S. W. Kowalczyk, A. R. Hall, and C. Dekker, Nano Lett. **10**, 324 (2010).

[104] F. Persson and J. O. Tegenfeldt, Chem. Soc. Rev **39**, 985 (2010).

[105] W. Reisner *et al.*, Proc. Natl. Acad. Sci. USA **107**, 13294 (2010).

[106] R. B. Schoch, J. Han, and P. Renaud, Rev. Mod. Phys. **80**, 839 (2008).

[107] J. Ju *et al.*, Proc. Natl. Acad. Sci. USA **92**, 4347 (1995).