

Electronic Thesis and Dissertation Repository

4-22-2020 10:00 AM

The journey of a single polymer chain to a nanopore

Navid Afrasiabian, *The University of Western Ontario*

Supervisor: Denniston Colin, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Applied Mathematics

© Navid Afrasiabian 2020

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Statistical, Nonlinear, and Soft Matter Physics Commons](#)

Recommended Citation

Afrasiabian, Navid, "The journey of a single polymer chain to a nanopore" (2020). *Electronic Thesis and Dissertation Repository*. 6966.

<https://ir.lib.uwo.ca/etd/6966>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

Abstract

Single chain translocation has been eagerly studied for more than two decades due to its importance in biological processes and also, providing a better understanding of polymer dynamics. Polymer translocation can be divided into three stages of reaching the pore, entering the pore and passing through it. We study the delivery of the chain from the bulk to the entrance, which is called the capture process, for a single chain driven by hydrodynamic flow. Our molecular dynamics-lattice Boltzmann simulations show that the converging flow around the nanopore not only facilitates the process of finding the pore but also deforms the chain in a way that the insertion happens mostly by one of the ends. We explain why single-file capture happens most often despite the formation of folded shapes or hairpins.

Keywords: Polymer translocation, polymer capture, capture radius, pulley effect, molecular dynamics, lattice-Boltzmann

Summary for Lay Audience

One of the fundamental processes involved in the survival and functioning of cells is the ability of the cell to exchange matter with its environment. This exchange consists of the motion of different molecules into or out of the cell through very narrow pores, namely nanopores. An important group of molecules taking part in this process are polymers, long chain-like molecules which look like spaghetti noodles. In the present work, we investigate the effect of fluid flow on polymer capture, the process of finding the entrance of a nanopore, using computer programs. We observed that the non-uniform fluid flow approaching the nanopore stretches the polymer chain which can result in the unravelling of folded shapes (hairpins) and promoting single-file capture.

Co-Authorship Statement

The work presented in this thesis was done in collaboration with my supervisor, Prof. Colin Denniston. Chapter 3 will form the basis for a publication that will be co-authored with my supervisor.

Acknowledgements

I wish to express my deepest gratitude to my supervisor, Professor Colin Denniston, without whom this work would not come together. His kind feedback and comments were the best guide throughout the journey and I learnt many great things from him not only as an academic supervisor but as a person. I would like to thank Sean Snider for fruitful discussions and occasional help with simulations. I would like to pay my special regards to my parents, Farideh and Omid, who like always have been supportive, although from miles away, and my sister, Nasim, who always knows how to cheer me up and give me hope even at the toughest times.

Contents

Abstract	i
Co-Authorship Statement	ii
Acknowledgements	iii
List of Figures	v
List of Tables	viii
1 Introduction	1
1.1 Polymer shape and size	2
1.2 Coarse-grained models of polymers	4
1.3 Polymer Dynamics	5
1.3.1 Rouse Model	6
1.3.2 Zimm Model	8
1.4 Polymer Capture and Translocation	8
1.5 Motivation	14
1.6 Outline	14
2 Methodology	15
2.1 Fluid model	16
2.1.1 Kinetic Theory and the Boltzmann equation	16
2.2 Fluid simulations with Lattice-Boltzmann	19
2.2.1 Our system: a polymer chain in LB-FLUID	23
3 Results	30
3.1 Arrival time	30
3.2 Capture radius and chain extension	34
3.3 Pulley effect: Unravelling of folded conformation	37
3.4 summary	42
4 Conclusion	43
Bibliography	45
Curriculum Vitae	50

List of Figures

1.1	A polymer chain can take different conformations. When the intra-chain attractions overcome the solvent-monomer interactions, the polymer compacts to a globule (a). In (b), a rod-like conformation is shown. This conformation appears as a result of mechanisms causing the polymer and solvent to increase their contact area. (c) shows a polymer coil.	2
1.2	A polymer chain as a sequence of random vectors.	3
1.3	The graph shows the potential of a van der Waals force between non-neighbour monomers i and j as a function of their distance.	4
1.4	The blob model split the chain to blob which consist of random chains. The size of the confinement dictates the radius of the blobs.	9
1.5	The polymer chain threads through the nanopore and translocates from the cis side, where it is originally located, to the trans side.	10
1.6	A hairpin structure threading through the nanopore. (a) shows the lateral view and (b) shows the back view of the chain. The polymer is shown with gradient colouring to distinguish the ends from each other and from the middle segments.	13
2.1	A triangular mesh was used in HLG model to improve the isotropy of the LGCA model. The collisions in the HLG model are head-on collisions. The particles from occupied sites move along the links, collide at an empty site, and leave the site in a perpendicular direction to their original path. The single and double arrows show the propagation at time t and $t + 1$	19
2.2	In D1Q3 model, the fictitious particle can travel to one of the adjacent sites via $\mathbf{e}_1 = 1$ or $\mathbf{e}_2 = -1$ in lattice units, or it can stay at the current site.	20
2.3	In D2Q5, the streaming happens only vertically or horizontally while in D2Q9, the digonal motion is also possible.	21
2.4	Illustration of common 3D lattice arrangements for LB method. (a) shows a model with 15 neighbouring sites and (b) shows the 19 neighbour model.	22
2.5	An immersed surface in a fluid mesh is shown. In trilinear stencil, the mapping from surface nodes to lattice sites and vice versa are done with help of weights that are defined based on the area between the surface node and the mesh sites.	24

2.6	(a) shows the fluid flow in terms of streamlines. The narrow red area is inside the hole where the flow is stronger. (b) shows the same flow but with the Peskin stencil. Some of streamlines go through the walls.	25
2.7	(a) shows a composite particle. The pink particles are the shell atoms and the jade particle is the central atom which altogether make a rigid sphere. (b) is an illustration of our polymer chain.	26
2.8	A snapshot of the system. The simulation box is a box of size $80\text{nm} \times 52\text{nm} \times 52\text{nm}$ with periodic boundaries in all directions. The size of the box is selected to allow extended chains without crossing the boundaries or interacting with itself. The length of the nanopore is chosen to be of the order of the radius of gyration of the chain.	28
3.1	Within a certain radius from the nanopore, the polymer feels a stronger pull from the flow (a). In (b), an end reaches the nanopore and the translocation begins. In (c), a part of the chain travelled to the trans side while in (d), the translocation has successfully happened and the whole polymer is on the trans side.	31
3.2	The velocity of the fluid is illustrated as a function of distance from the pore. The red curve is obtained from a system with thermal noise while the black one shows the velocity for a system without noise. The red curve follows the same pattern as the black one within error bounds. The relaxation velocity v_R is shown with a horizontal line (blue dash-dotted). The intersection of the flow velocity and the relaxation velocity is marked with a black circle. For velocities faster than v_R the chain does not have enough time to relax. The inset shows the velocity of the flow approaching the pore in a logarithmic scale. An inverse-square relation for the approaching flow is observed.	32
3.3	The journey of a chain which threads into the nanopore in hairpin conformation is illustrated through series of snapshots at different times.	33
3.4	The arrival time of the COM of the chain is normally distributed from which we can conclude that the motion of the chain is a mix of diffusion and drift. The average arrival time is 107ns and the standard deviation equals 36ns	35
3.5	The position of the front monomer as a function of time for a chain with 32 monomers (a) and 64 monomers (b). The smaller gap between the curve of the front and tail monomer in the case of hairpin conformations (dotted lines) stems from the fact that the tension propagates faster along the chain because the strands are shorter than the whole backbone of the chain which is the length that tension must spread along for the single-file conformation.	35
3.6	(a)The chain in the bulk experiences 2 stages to arrive at the nanopore. In the first stage, the chain gets the opportunity to diffuse and equilibrate while the stronger flow deforms the chain in the second stage. The radius of gyration declines to a value less than equilibrium value as it gets out of the hole. (b)The graph shows the relative distance of the front monomer from the main body of the chain. The distance between the front and centre of mass increases and reaches a peak when the front monomer leaves the nanopore.	36
3.7	The most probable conformation of insertion is single-file. Due to chain unravelling and statistical uncertainty, the bin widths of less than 4 monomers would not demonstrate the essential information.	38

3.8	The faster motion of shorter strand of a hairpin conformation is shown in a sequence of snapshots. For illustration purposes, the wall (not shown) was wider in this figure. (a) shows the chain entering a pore of length 40 nm. The leading monomer is the twentieth bead. (b) shows another snapshot of the system when the chain is half-way through. The leading monomer is nineteenth bead. In (c), the chain is shown as it leaves the nanopore. The leading monomer is bead number sixteen.	39
3.9	(a) shows the probability of the position of the hairpin vertex along the polymer backbone for a 32-mer chain and (b) shows the same property for a 64-mer chain. The x axis is scaled by the chain length. The medium-sized hairpins open up within the capture radius due to the pulley effect. The unravelling continues inside the nanopore. In fact, the stronger force in the pore speeds up the pulley effect. This is more obvious for longer nanopores.	40
3.10	(a) shows the velocity of polymer segments while travelling toward the nanopore for a single-file capture averaged over 32 realizations and (b) shows the same quantity for a hairpin capture averaged over 23 realizations. The polymer motion deviates from a perfect inverse-square relation due to the chain constraints. Because of the averaging steps included in obtaining both (a) and (b), we consider a cut-off distance (8nm from the origin) that beyond which the curves are not accurate enough.	41
3.11	In the (a), the velocity of different segments of a chain that enters the nanopore in a single-file conformation has been shown while in the (b), the same feature has been shown for a chain with hairpin conformation. The tension simultaneously propagates along the both strands of the chain in the case of hairpin conformation and consequently, the velocity graphs for the first segments look alike.	42

List of Tables

2.1	The table shows the weight factors for different models.	23
2.2	The moments of equilibrium distribution M_{eq}^a and of the continuum random process $\bar{\chi}$ for model D3Q15	27

Chapter 1

Introduction

It has been a century since Staudinger introduced the concept of polymerization [1]. Throughout these hundred years, polymers have been the subject of many studies. We have studied their structure, motion, reactions, interactions with other molecules, etc. and yet, at this milestone which is marked as “Year of Polymers”, there are many remaining questions, about polymers and processes that they are involved in, that are in need of answers.

The question of how a polymer chain moves in confined spaces has been troubling scientists for decades. One particular process of this kind which has been under the microscope for more than two decades is polymer translocation, the delivery of a polymer chain across a barrier from one side to the other through a narrow channel. This picture may remind the reader of transport of matter across membranes which is to a great extent what happens in polymer translocation. However, due to the special structure of macromolecules, this process is more difficult to investigate compared to small molecules or ion translocation. The huge interest in understanding this process stems from not only the fascinating nature of it but also its essential role in biological systems and processes. For example, synthesized mRNA transcripts in the nucleus need to move through the nuclear pore complex to get to the cytoplasm where they can be translated to proteins [2].

Other examples of polymer translocation in biology are insertion of DNA by a virus into the host, translocation of proteins through the cell membrane, and drug delivery [3]. In addition to these processes, polymer translocation is a crucial step of fast sequencing techniques, and one can claim that designing better and cheaper sequencing methods is the primary motivation for studying polymer translocation [4, 5].

Before a macromolecule can thread through a nanopore, the chain must be captured by the pore, i.e. it must first find the nanopore. After arriving at the entrance, the insertion happens and then the chain translocates from the cis side, where the chain starts its journey, to the trans side, the destination. The dynamical behaviour of the chain may differ depending on whether the process is driven or unbiased. To understand the features of polymer translocation and capture, we must first learn about measures of polymer shape and theories of polymer dynamics. In this chapter, we first describe some measures of polymer conformation. In section 1.2, the motion of a polymer in dilute solutions is discussed and fundamental theories for such systems are pointed out. The mathematical formulation of the theories is shown and derived wherever necessary. However, we try to avoid any cumbersome mathematics. In section 1.3, we give a brief history of experimental and theoretical works in the field of polymer translocation,

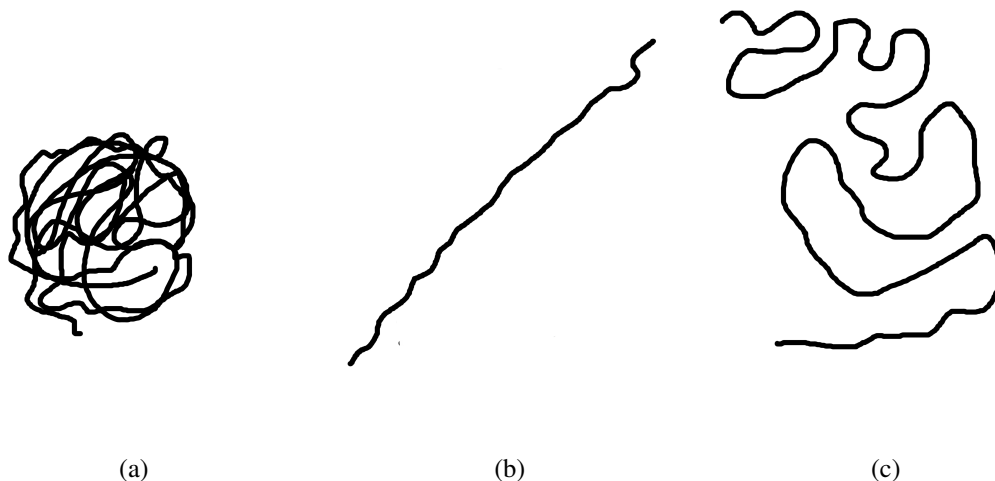


Figure 1.1: A polymer chain can take different conformations. When the intra-chain attractions overcome the solvent-monomer interactions, the polymer compacts to a globule (a). In (b), a rod-like conformation is shown. This conformation appears as a result of mechanisms causing the polymer and solvent to increase their contact area. (c) shows a polymer coil.

describe the polymer translocation more thoroughly, and review the models proposed for both stages, capture and translocation.

1.1 Polymer shape and size

A polymer chain consists of many monomers which are connected by covalent bonds. This allows the chain to form various shapes or conformations in space. However, there are constraints on possible conformations a chain can take stemming from the interactions different parts of the chain have with each other or their environment. For a polymer in a fluid, these interactions are divided into three types: monomer-monomer, solvent-solvent, and solvent-monomer. Depending on the relative strength of these interactions, the spatial density of the chain may vary. For example, the chain shrinks to a globule when the solvent-monomer interactions are not as favoured as monomer-monomer and solvent-solvent interactions, Fig. 1.1a, while the chain swells to a rod as a result of strong attraction between the solvent molecules and the polymer chain, Fig. 1.1b. Between these two extremes, the chain adopts a coil conformation, Fig. 1.1c.

For scientific purposes, rough qualitative measures are not always enough and one may seek quantities by which the conformation of a chain in a specific environment can be specified and even more importantly, be compared with a chain under different circumstances and among various experiments. One of these quantities is the end-to-end distance of the chain that is defined as the difference between the position vector of one of the end monomers and the other one, as shown in 1.2:

$$\sqrt{\langle R_E^2 \rangle} = \sqrt{\langle (\mathbf{r}_1 - \mathbf{r}_N)^2 \rangle} \quad (1.1)$$

where \mathbf{r}_1 and \mathbf{r}_N are the position vectors of the ends for a polymer with N monomers and the

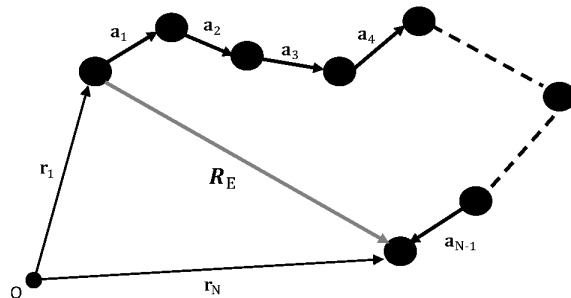


Figure 1.2: A polymer chain as a sequence of random vectors.

angle brackets show the average over all possible conformations.

Although the end-to-end distance gives a general picture of how stretched or compact the chain is, it merely depends on the end monomers. In other words, this quantity does not describe the spatial distribution of the monomers or allow comparison between different topologies. In addition, it is difficult to measure this quantity in experiments. Therefore, the radius of gyration of a chain (R_G) which can be determined from scattering experiments [6] and provides information about how a polymer is distributed around its centre of mass comes into the picture,

$$R_G = \sqrt{\frac{1}{M} \sum_{i=1}^N m_i (\mathbf{r}_i - \mathbf{r}_{com})^2} \quad (1.2)$$

where M is the total mass of the chain, m_i is the mass of a monomer, \mathbf{r}_i is the position vector of the i th monomer, and \mathbf{r}_{com} is the position vector of the chain's centre of mass.

By describing a polymer chain as a series of beads (monomers) and links (bonds), and after working out the math, both the radius of gyration and the end-to-end distance will be only a function of the number of monomers and the bond length which leads to the conclusion that the size of a macromolecule must scale with the number of monomers,

$$R \sim aN^\nu \quad (1.3)$$

where R is a general notation for polymer radius, and ν is a size exponent, namely the Flory exponent [7]. The interaction between monomers is not limited to chemical bonds between neighbours. The monomers that are far from each other along the chain typically interact by van der Waals-type force. Since the van der Waals interactions are highly repulsive in short distances, as seen in Fig. 1.3, the non-neighbour monomers cannot get closer than a certain distance in space which results in an excluded-volume. In the following sections, the relation between the Flory exponent and the excluded-volume will be discussed in greater depth.

The experimental work on polymers has illustrated that the radius of gyration of various macromolecules is related to the number of monomers with a similar excluded-volume exponent despite their chemical details. This suggests that the relation must be universal and independent of the local constraints of the chain. In the next section, we will discuss the coarse-grained models of polymers and explain how the observed universal behaviour of macromolecules helped in developing these models.

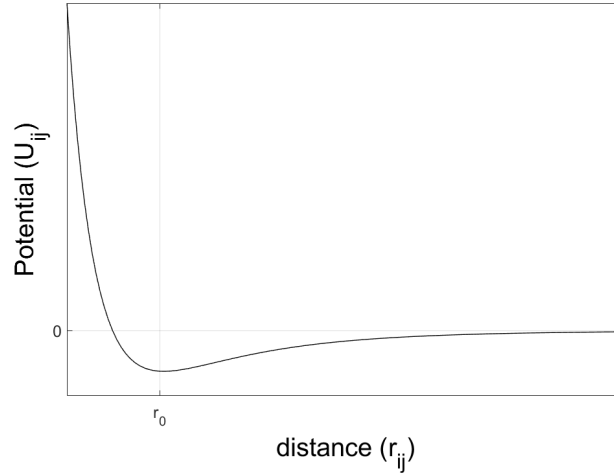


Figure 1.3: The graph shows the potential of a van der Waals force between non-neighbour monomers i and j as a function of their distance.

1.2 Coarse-grained models of polymers

Due to many possible conformations that a polymer chain can possess in equilibrium, modelling their shape can effectively be studied as a statistical problem. In the simplest model, a polymer chain with N monomers which are connected by bonds of length a can be considered as a random walk with a total number of steps N and step size a [8]. Then, each step can be presented as a vector which takes any direction independent of other steps, as shown in Fig. 1.2. Therefore, the end-to-end vector of the walk is

$$\mathbf{R}_E = \sum_{i=1}^{N-1} \mathbf{a}_i \quad (1.4)$$

Since the vectors are random and independent of each other, for $N \rightarrow \infty$, the average value of the end-to-end distance $\langle \mathbf{R} \rangle = 0$ and the variance can be found as follows:

$$\langle R_E^2 \rangle = \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \langle \mathbf{a}_i \cdot \mathbf{a}_j \rangle = \sum_{i=1}^{N-1} \langle \mathbf{a}_i^2 \rangle + 2 \sum_{i>j} \langle \mathbf{a}_i \cdot \mathbf{a}_j \rangle = Na^2 \quad (1.5)$$

This model is called a freely jointed chain. Despite the unrealistic assumptions of this model, the scaling behaviour between the polymer radius and the number of monomers obtained from this model is close to scaling observed in experiments. This generated the idea of the existence of a length scale beyond which the polymer segments are connected in a freely jointed manner. This length is called the Kuhn length (l_k) and the number of Kuhn segments that creates the chain is labelled as N_k and is referred to as the number of Kuhn steps [7].

If a polymer molecule is long enough that it consists of many Kuhn segments then, the universal features of such a chain can be captured by replacing it with a chain with N_k beads and bonds of length l_k . Such a coarse-grained model is called a Gaussian chain. As mentioned for freely jointed chain, the end-to-end distance of a Gaussian chain is $\langle R_E^2 \rangle \sim N_k l_k^2$ which

means $\langle R_E^2 \rangle^{1/2} \sim N_k^{1/2}$ and $\nu = 1/2$. To define a potential for the harmonic bonds, we first find the probability of finding the ends at a fixed distance \mathbf{R} at temperature T by applying central limit theorem [8]:

$$P(\mathbf{R}) = \left(\frac{3}{2\pi \langle R_E^2 \rangle} \right)^{3/2} \exp \left(-\frac{3\mathbf{R}^2}{2 \langle R_E^2 \rangle} \right) \quad (1.6)$$

Where $\langle R_E^2 \rangle = g(k, T) N_k l_k^2$ and $g(k, T)$ is the flexibility coefficient which depends on the bending rigidity (σ_b) and temperature. The coefficient approaches 1 in the high-temperature limit, and the chain becomes fully flexible. The probability distribution $P(\mathbf{R})$ can be used to find the entropy of the chain, and as a result, the free energy of the chain constrained to end-to-end distance \mathbf{R} is [9],

$$F(\mathbf{R}) = E - TS = F_0 + \left(\frac{3k_B T}{2 \langle R_E^2 \rangle} \right) R^2 \quad (1.7)$$

where k_B is Boltzmann constant, T is temperature, and F_0 is the free energy of a chain with no constraint. As illustrated in Eq.1.7, the free energy of the bond is a quadratic function of the distance like a Hookean spring.

Although the Gaussian chain explains much of the observed elasticity of polymeric materials, there are two shortcomings in the model. Firstly, bonds of the Gaussian chain are infinitely extensible which is unrealistic. Secondly, the Flory exponent found in experiments of a dilute solution of polymers is different from the $\nu = 1/2$ obtained from this model.

When it comes to simulation, the infinite extensibility of Gaussian bonds makes bond-crossing possible and generates nonphysical results. To overcome this problem, a different bond model was proposed, the Finitely Extensible Non-linear Elastic (FENE) bonds [10]. For our molecular model, we use a similar type of bonds with the following potential [11]:

$$U = -\frac{1}{2} k R_0^2 \ln \left(1 - \left(\frac{r}{R_0} \right)^2 \right) + 4\epsilon \left(\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 + \frac{1}{4} \right) H \left(2^{1/6} - \frac{r}{\sigma} \right) \quad (1.8)$$

Where K is the elastic constant, R_0 is the maximum extension of the bonds, and r is the distance between adjacent beads. The second term on the left is a repulsive Lennard-Jones (LJ) potential, and ϵ is the energy scale and σ is the length scale of the LJ-potential.

Beads in a Gaussian chain are allowed to move in all directions without constraints. However, in reality, a bead cannot be at a spot already occupied by other beads. With this logic, it was suggested that a self-avoiding walk would be a more realistic model for polymer chains. The $\nu \approx 0.6$ obtained from this model is much closer to the values measured for polymers in dilute solution in experiments. To have such an effect present in our system, we added a truncated-shifted LJ force between the beads. The molecular model and the exact values of the parameters will be discussed in the next chapter to provide the reader with a better picture of our molecular model.

1.3 Polymer Dynamics

There are different ways to describe the dynamics of a rigid sphere (a Brownian particle) moving inside a fluid, as well described by Doi and Edwards [12]. One of the commonly used

expressions for modelling this motion, known as Brownian motion, is the Langevin equation,

$$m \frac{d\mathbf{v}(t)}{dt} = -\zeta \mathbf{v}(t) + \boldsymbol{\xi}(t) + \mathbf{F}_{ext} \quad (1.9)$$

where m and v are the mass and velocity of the particle, respectively, ζ is the friction coefficient, $\boldsymbol{\xi}$ is the random force due to the thermal agitation, and F_{ext} stands for any forces other than fluid friction and random forces acting on the particle. The thermal agitation stems from the collision between fluid molecules and the Brownian particle. These collisions are also responsible for the systematic resistance against the inertial drift of the particle. Therefore, one may expect a relation between random and viscous forces. The relation is made in the form of the ‘‘fluctuation-dissipation theorem’’ which can be represented in terms of the correlation function of the random forces,

$$\langle \boldsymbol{\xi}(t) \cdot \boldsymbol{\xi}(t') \rangle = 2\zeta k_B T \delta(t - t') \quad (1.10)$$

where δ is the Dirac delta function [12]. Although studying the motion of the big blob of the polymer as a Brownian particle partly explains the dynamics of the chain, the effect of chemical bonds and internal dynamics are missing. The chain connectivity impacts the collective motion of the segments and it is important to be considered, especially where chain deformation occurs.

In the next section, we discuss the Rouse model which models a Gaussian chain with no excluded-volume effect. The Rouse model provides a helpful insight into polymer dynamics, even though it is not successful in quantitatively modelling real-world examples. The Zimm model which obtains more realistic results will be described in section 1.3.2.

1.3.1 Rouse Model

The Rouse model is one of the earliest models that counts for connectivity forces in a polymer chain. Although this model is simple, it helps to get a better idea of the dynamics of a macromolecule. In this section, we give a brief description of the model and more detailed discussions can be found in any polymer physics textbook [12].

Newton’s second law for the i th segment of a polymer with N Kuhn segments of length l in the absence of hydrodynamic interactions is,

$$m \frac{d\mathbf{v}_i}{dt} = (\mathbf{F}_{bond})_i + (\mathbf{F}_{fric})_i + \boldsymbol{\xi}_i + (\mathbf{F}_{ext})_i \quad (1.11)$$

where \mathbf{v}_i is the velocity of the segment (bead), \mathbf{F}_{bond} is the connectivity forces, \mathbf{F}_{fric} is the resistance of the fluid against the bead’s motion, $\boldsymbol{\xi}$ is the thermal fluctuation force, \mathbf{F}_{ext} is any force other than the ones mentioned, and i indicates the segment number.

The friction forces spring from the collision between the solvent molecules and i th bead. Therefore,

$$(\mathbf{F}_{fric})_i = -\zeta_b \left(\frac{d\mathbf{r}_i}{dt} - \mathbf{u}_{bg} \right) \quad (1.12)$$

where \mathbf{r}_i is the position vector of the i th bead, ζ_b is the friction coefficient of the bead, and \mathbf{u}_{bg} is the background velocity in the absence of the polymer. The thermal fluctuation of the solvent

has a Gaussian nature with the first and second moments as follows:

$$\langle \xi_i(t) \rangle = 0 \quad (1.13)$$

$$\langle \xi_{i\alpha}(t) \cdot \xi_{j\beta}(t') \rangle = 2\zeta_b k_B T \delta_{ij} \delta_{\alpha\beta} \delta(t - t'). \quad (1.14)$$

With no hydrodynamics present, the Gaussian chain is simply a series of beads connected by springs. Hence, the chain is like N coupled harmonic oscillators with potential energy U_H ,

$$U_H = \frac{k_h}{2} \sum_{i=2}^N (r_{i,i-1} - r_0)^2 \quad (1.15)$$

where $r_{i,i-1} = |r_i - r_{i-1}|$, and r_0 is the equilibrium length of the bond. The bond forces \mathbf{F}_{bond} on the i th bead from its adjacent neighbours can be derived from the derivative of the potential, giving

$$(\mathbf{F}_{bond})_i = -k_h \sum_{i=1}^{N-1} (2\mathbf{r}_i - \mathbf{r}_{i+1} - \mathbf{r}_{i-1}) \quad (1.16)$$

where $k_h = \frac{3k_B T}{r^2}$ is the spring constant of the harmonic oscillator [12] and as a result, the equation of motion,

$$\zeta_b \left(\frac{d\mathbf{r}_i}{dt} - \mathbf{u}_{bg} \right) + k_h \sum_{i=1}^{N-1} (2\mathbf{r}_i - \mathbf{r}_{i+1} - \mathbf{r}_{i-1}) = \xi_i. \quad (1.17)$$

A standard way of solving the problem of coupled harmonic oscillators is to introduce a set of normal coordinates which transform the system of equations to an uncoupled one [13]. This method is well-known and the reader can find the steps of the derivation in any fundamental polymer physics or classical mechanics book [13, 12]. Here, we focus on the physical meaning of the normal coordinates. \mathbf{X}_p is the normal coordinate related to the p th mode of motion. The coordinate \mathbf{X}_0 is the position of COM of the chain. Therefore, the diffusion coefficient of the centre of mass of the chain can shown to be,

$$D = \frac{k_B T}{N\zeta_b} \quad (1.18)$$

which means that the the diffusion constant of the chain is inversely proportional to the length of the chain [12]. The parameter with the most physical significance is the longest relaxation time of the chain which corresponds to the Rouse mode $p = 1$, known as Rouse time, and it is the time that it takes for the whole chain to equilibrate after a deformation. One can show that [12],

$$\tau_{Rouse} = \frac{\zeta_b N^2}{12\pi^2 k_B T} \sim N^2. \quad (1.19)$$

The predicted scaling relation between the chain length and the relaxation time fails to match with results found in experiments and therefore it is not the best model for polymer in dilute solutions. The shortcomings of the Rouse model originates from neglecting hydrodynamic interactions.

1.3.2 Zimm Model

As shown in the previous section, ignoring the long-range hydrodynamic interactions in modelling the motion of a flexible polymer chain results in unrealistic scaling of the relaxation time and viscosity. In 1948, Kirkwood and Riseman proposed a model in which the hydrodynamic forces were considered [14],

$$\bar{\bar{H}}_{ij} = \frac{1}{8\pi\eta_s|\mathbf{R}_{ij}|} \left(\frac{\mathbf{R}_{ij}\mathbf{R}_{ij}}{|\mathbf{R}_{ij}|^2} + \bar{\bar{I}} \right) \quad (1.20)$$

where $\bar{\bar{H}}_{ij}$ is the mobility tensor, as defined in Ref. [12], η_s is the solvent viscosity, $\mathbf{R}_{ij} = \mathbf{r}_i - \mathbf{r}_j$, and $\bar{\bar{I}}$ is the identity tensor. Implementing this mobility tensor in the Langevin equation of the polymer chain, Zimm created a model for polymer dynamics in dilute solution [15],

$$\left(\frac{d\mathbf{r}_i}{dt} - \mathbf{v}_{field} \right) = \sum \bar{\bar{H}}_{ij} \cdot \left(K \frac{\partial^2 r_j}{\partial j^2} + \boldsymbol{\xi}_i \right) \quad (1.21)$$

where K is the bond strength constant and $\frac{\partial^2 r_j}{\partial j^2}$ is the continuous format of $\mathbf{r}_{j+1} + \mathbf{r}_{j-1} - 2\mathbf{r}_j$. Mathematically, there are conditions to be met in order to transform the equations from discrete to continuous format which for the current purposes, they are considered to be satisfied [12, 15]. Unlike the Rouse model, the mobility tensor of the Zimm model is a nonlinear function of R_{ij} .

In order to solve this equation analytically, Zimm replaced $\bar{\bar{H}}_{ij}$ with its average (pre-averaging approximation) and used the normal coordinates method to solve the approximate problem [15]. The scaling relation that he obtained for the relaxation time is in good agreement with experimental data,

$$\tau_{Zimm} \sim R_G^3 \sim N^{3\nu} \quad (1.22)$$

where R_G is the radius of gyration of the chain and ν is the Flory exponent.

1.4 Polymer Capture and Translocation

The impact of geometric confinements on polymer conformation and motion has been studied for more than four decades. In one of the early attempts, de Gennes described the size and the energy of a real chain squeezed into a tube employing scaling arguments [9], resulting in

$$R_{\parallel} \approx Na \left(\frac{a}{d_t} \right)^{\frac{1}{\nu}-1} \quad (1.23)$$

where R_{\parallel} is the length of the tube occupied by the chain, N is the number of monomers, a is the bond length, and d_t is the diameter of the tube which is considered to be much smaller than the equilibrium radius of the chain (R_F). This relation can be obtained using the so-called blob model. Since the tube sets a boundary for the chain's shape, for the lengths longer than the tube size the conformation is restricted and biased from a random path, whereas for length scales smaller than the tube size, the chain's local conformation is random. As a result, the chain can be modelled as a succession of rigid blobs that each consists of a random chain,

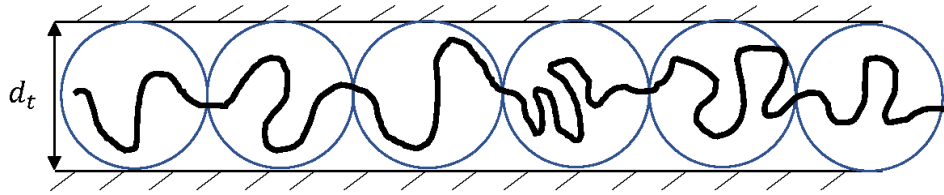


Figure 1.4: The blob model split the chain to blob which consist of random chains. The size of the confinement dictates the radius of the blobs.

$$R_{\parallel} = \left(\frac{N}{g}\right) d_t \quad (1.24)$$

where g is the number of monomers inside each blob (see Fig. 1.4), and it is related to the size of the blob by the Flory exponent,

$$d_t = ag^{\nu} \quad (1.25)$$

which can be substituted into Eq. 1.24 to obtain Eq. 1.23. As can be seen, the confinement results in $R_{\parallel} > R_F$ which is expected. de Gennes also found the confinement free energy of the chain inside a tube,

$$F_{conf} = k_B T N \left(\frac{a}{d_t}\right)^{\frac{1}{\nu}} \quad (1.26)$$

where k_B is the Boltzmann constant, and T is the temperature.

In a few chapters later in his book [9], de Gennes worked out a scaling relation for the dynamics of a polymer coil going through a tube by means of the chain mobility,

$$\mu_{chain} = v_p / f_{tot} \approx \frac{g}{N \eta d_t} \sim \frac{1}{\eta_s R_{\parallel}} \quad (1.27)$$

where μ_{chain} is the chain mobility, v_p is the polymer velocity, f_{tot} is the total force on the coil, and η_s stands for the viscosity of the solvent. Although this type of theoretical work provided great insight into the behaviour of confined polymer coils [16, 17], their importance only came to light almost two decades later with studies on polymer translocation and DNA sequencing.

Polymer translocation is the process by which a polymer chain travels from the chamber in which it is originally located to the destination chamber through a nano-scale pore (nanopore or nanochannel). Fig. 1.5 shows an imaginary translocation process. The nanopore can have either a biological basis, for example, a nanochannel through a lipid bilayer [18, 19], or a solid-state basis, for instance, fabricated nanopores in silicon nitride [4, 20]. In 1996, Kasianowicz et al. showed that single-stranded DNA in an ionic solvent can be driven through a nanochannel by means of an electric field [21]. They found that the blockade in the ionic current is the result of the molecule's presence in the nanochannel and the duration of the blockade is proportional to the polymer length. In the same year, Sung and Park conducted a theoretical study on the threading of a generic flexible polymer through a nanopore [22]. Due to these two works, 1996 is one of the important milestones in the field of polymer translocation.

Sung et al. [22] simplified the dynamics of translocation to Brownian motion of a single particle, namely, the reaction or translocation coordinate (s) over a free energy barrier. The term

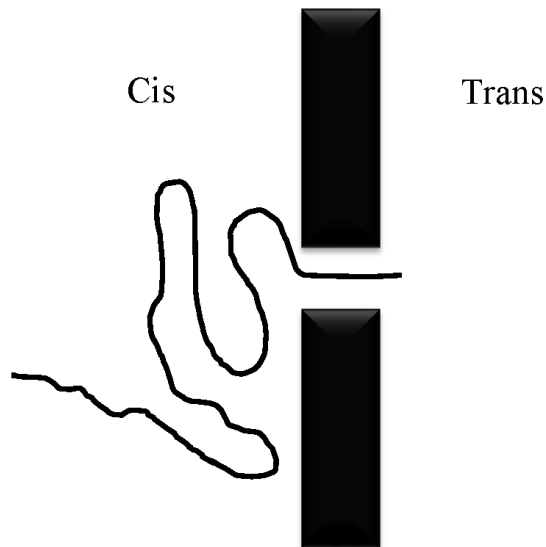


Figure 1.5: The polymer chain threads through the nanopore and translocates from the cis side, where it is originally located, to the trans side.

reaction coordinate is borrowed from the phase transition and nucleation theory [23, 24]. In their model, translocation was considered as diffusion of the translocation coordinate which can be described by a Fokker-Planck equation of the probability density of the reaction coordinate s at time t ,

$$\frac{\partial}{\partial t} P(s, t) = \left(\frac{k_B T}{l} \right)^2 \left[\left(\frac{1}{k_B T} \frac{\partial D(s)}{\partial s} - D(s) \frac{\partial F(s)}{\partial s} \right) \frac{\partial F(s)}{\partial s} \right] P(s, t) \quad (1.28)$$

where l is the Kuhn length, k_B is the Boltzmann constant, T is the temperature, $D(s) \sim N^{-\nu}$ is the diffusion coefficient which is related to the chain length by the Flory exponent ν , and $F(s)$ stands the free energy. Consequently, the translocation time (τ) could be estimated by mean first passage time. The scaling relation obtained from these calculations was in form of $\tau \sim N^{2+\nu}$.

In the following years, various groups showed interest in the polymer translocation and since a number of competing predictions have been offered.

Using a similar formulation to that of Sung et al., Muthukumar [25] found that the translocation time scales with the square of polymer length, $\tau \sim N^2$. In this case, he assumed that the diffusion coefficient and the friction force are solely depending on the details of the pore. Both studies focused on the translocation of a polymer chain in the absence of any driving force.

According to the above phase transition analogies, each phase must be in a quasi-equilibrium state during the nucleation. In other words, the translocating part of the chain on the cis side must be in equilibrium as well as the translocated part on the trans side. Chuang, Kantor and Kardar found such a formulation counter-intuitive [26]. Chuang et al. argued that the chain translocates through the nanopore one segment (bead) at a time which restricts the motion of the chain. Therefore, the motion of the chain must be slower than a chain travelling with no constraints. The distance that the chain translocates through a thin pore is virtually equal to its radius of gyration. The time that it takes for the chain to move a radius of gyration under no restriction is the longest Rouse time which scales with the polymer length as $\tau_{Rouse} \sim N^{1+2\nu}$.

Chuang et al. pointed out that the Rouse time must set the lower bound of the translocation time.

To support their idea, they performed two dimensional Monte Carlo (MC) simulations using a fluctuating-bond model (FB). In this model, each bead is represented by a four-site lattice block. The bonds connecting the beads have a length between $4^{1/2}$ and $16^{1/2}$ which prevents bond cutting [27]. In each realization, they placed an end at the hole and applied a condition on the first bead to only move forward so that the time measured is exactly the translocation time and not the time required for the chain to find the pore. The hole thickness and width were set to three and two lattice units, respectively. Their results indicated a scaling of the form $\tau \sim N^{1+2\nu}$ with $\nu \approx 0.6$ which gives a scaling exponent greater than what was obtained from Muthukumar's model. The mean-squared displacement of s measured from their simulation scales with translocation time as follows:

$$\langle \Delta s^2(t) \rangle \sim t^{2\beta} \quad (1.29)$$

Where $\beta = 1/(1 + 2\nu)$ resulting in $\beta \approx 0.92$ for 3 dimensions and without hydrodynamic interactions. Hence, they arrive at the conclusion that polymer translocation must have anomalous dynamics. Luo et al. also used a FB model with a slightly different setting to determine the scaling behaviour of translocation. In their simulation, the polymer was initially placed halfway through the nanopore and instead of translocation time, the escape time τ_e was measured [28]. This way, there was no need for adding any artificial constraints on the first bead. The escape time is defined as the time required for the chain to exit the pore to either side. In the case of short pores ($L_p \ll R_G$), they found $\tau_e \sim N^{1+2\nu}$ which was in agreement with the prediction of Chuang et al. while for long pores ($L_p \gg R_G$), the escape time scaled with the polymer's length linearly, $\tau_e \sim N$.

The fact that there is no real notion of time in Monte Carlo simulations may make the reader hesitate at the reliability of the time scaling relations obtained from MC simulations and this is probably one of the main reasons for the popularity of molecular dynamics and mesoscale methods over Monte Carlo method [29, 30, 31, 32, 33, 34, 35, 36, 37]. For example, Guillouzic and Slater performed molecular dynamics with an explicit fluid, by adding fluid particles to the system, to investigate the effect of hydrodynamic interactions on the scaling behaviour of the translocation. They observed a scaling relation of form $\tau \sim N^{2.27 \pm 0.04}$ [29]. In another study, Lehtola et al. [33] accounted for hydrodynamic effects through stochastic rotation dynamics (SRD) simulation. They found scaling exponent 2.55 ± 0.05 for 2D and 2.33 ± 0.05 for 3D. In both aforementioned studies, the middle of the chain was initially placed in the nanopore. Most of the results obtained are reasonably consistent with $\alpha = 1 + 2\nu$ exponent for the scaling of translocation time (or escape time) and the chain length.

In spite of the great deal of information that studying unbiased translocation provides on polymer dynamics in confined spaces and the nature of polymer-surface interactions, the omnipresence of forced translocation in many biological systems has made forced translocation more relevant.

The external force applied to the system can be of any nature, for instance, electric or hydrodynamic. Despite the nature of the force, we expect the force, from the cis side toward the trans side, to speed up the process, and therefore the translocation time scales as,

$$\tau \sim N^\alpha / f^\gamma \quad (1.30)$$

where f is the force amplitude, and α and γ are the scaling exponents.

Like unforced translocation, many approaches to the problem of driven translocation have been tried and different estimates of the exponents have been made [38, 18, 39, 40, 31, 28, 41, 42, 43, 44, 45]. In an experimental work, Meller et al. studied voltage-driven translocation of a single-stranded DNA molecule through a biological membrane [18]. They found that the translocation time is proportional to the chain length for molecules longer than the length of the pore, $\tau \sim N$, while for shorter polymers, a steeper dependency was observed. In a theoretical work, Kantor et al. challenged the scaling relations obtained from quasi-equilibrium approaches based on a similar line of reasoning they provided for unbiased translocation [39]. Comparing the relaxation time of the translocating polymer and the translocation time obtained from quasi-equilibrium approach, Kantor et al. observed that the translocation time is shorter than the relaxation time and as a result, the chain segments on either side do not get the chance to relax. They addressed this problem by Monte Carlo simulations and obtained a scaling behaviour of $\tau \sim N^{1+\nu}/f$ which led them to the conclusion that driven translocation has anomalous dynamics.

Although they observed that polymer translocation is not an equilibrium process from Brownian dynamics (BD) simulations, Tian and Smith showed that their one-dimensional diffusion model obtains results that are in qualitative agreement with BD simulations for the dynamics of driven translocation [40]. In their simulation which was governed by Langevin dynamics, the membrane potential was incorporated by applying an external field across the nanopore. The results of their LD simulation showed a linear scaling between the translocation time and the chain length, chemical potential difference, and friction coefficient, $\tau \sim N\zeta/\Delta\mu$ where ζ is friction coefficient and $\Delta\mu$ is the chemical potential gradient. In their study, the nature of a second peak in the translocation time distribution which was observed in experiments [21] and simulations [46] was also studied. Formation of trapped conformations on the trans side, for example, a loop-like shape was found to be responsible for the second peak exhibition. They also pursued the effect of the existence of attractive polymer-pore interactions. Tian et al. observed that the capture process is facilitated by an absorbing channel and the translocation time decreased in the presence of attractive polymer-pore interactions.

In another work based on Langevin dynamics, Luo et al. found a weak-to-strong driving force crossover. Their results showed a scaling relation where $\alpha \approx 1.37 \pm 0.02$ for fast translocation, caused by stronger force or lower viscosity, and $\alpha \approx 1.52 \pm 0.01$ for slow translocation [47]. The exponents found are in good agreement with $\alpha = 1 + \nu$ reported in previous works [39, 48].

In all the studies of forced translocation mentioned above, the hydrodynamic interactions have been neglected primarily because of the complicated nature of systems involving this kind of force as well as the computational cost of including hydrodynamic forces in simulations. In 2009, Lehtola et al. carried out SRD simulations to investigate the effect of hydrodynamics on forced translocation [43]. Their study showed that the translocation time drops considerably in the presence of hydrodynamic interactions. They found that the scaling exponent α depends on both the pore force and friction, and concluded that no universal scaling of τ with N exists.

For any translocation to happen, the chain must first be delivered to the nanopore and enter the pore. This process is called polymer capture; the process of finding the entrance of a nanopore. As we mentioned before, most of the computational and numerical studies on polymer translocation start with the chain being already captured by an end or placed in the

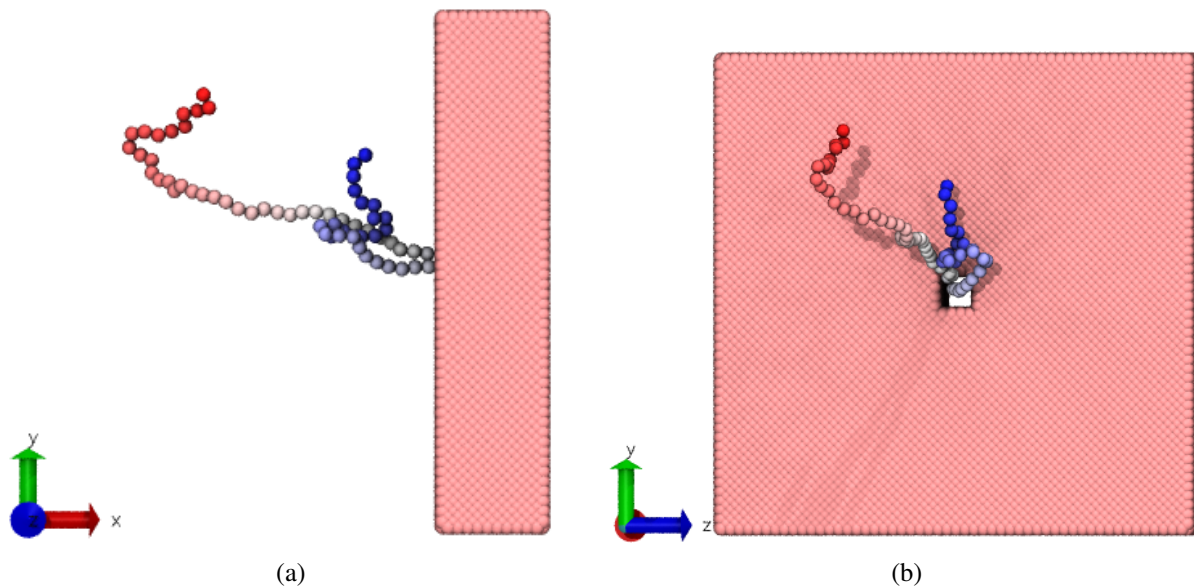


Figure 1.6: A hairpin structure threading through the nanopore. (a) shows the lateral view and (b) shows the back view of the chain. The polymer is shown with gradient colouring to distinguish the ends from each other and from the middle segments.

middle of the nanopore. This is mainly because reproducing the capture process has been found to be costly, both time-wise and resource-wise. However, some experimental works showed that the assumption that the chain is always captured by its ends fails to draw a complete picture and the capture can actually have an impact on translocation conformation[20]. The formation of folded shapes was later observed in Langevin dynamics (LD) simulations as well [49]. The most common folded conformation observed in experiments is a single-folded shape, which is called a hairpin. Fig .1.6 shows a hairpin capture from one of our simulations.

The dynamics of hairpin translocation has recently been studied [50]. In this study, Ghosh et al. describe the motion of a hairpin conformation through a nanopore using a generalized form of iso-flux tension propagation (IFTP) theory which accounts for translocation of two strands of the chain at the same time. The theory of tension propagation in translocation was originally introduced and derived by Sakaue [44] and has been modified since then [51, 52].

The hairpin translocation of DNA creates secondary current drops which are similar to the effect of the presence of a target molecule or a specific base sequence in the structure. This poses a challenge in analyzing the results obtained from single-molecule sequencing methods. In order to solve this problem, a reasonable amount of effort has been put into understanding the capture process, hairpin formation, and dynamics of hairpin unravelling and single-file capture promotion.

In one of the early works on polymer capture, Wong and Muthukumar theoretically showed that an electro-osmotic flow can cause extension in the polymer chain within a certain distance from the nanopore. They called this distance the capture radius (r_c) and proposed a quantitative prediction for this quantity [53]. However, most of the studies since have focused on the translocation process. In the studies on the capture process, the system in the absence of hydrodynamics, or capture to a very narrow nanopore which does not allow for a hairpin insertion was examined. In the present study, we investigate the effect of hydrodynamic flow on polymer

capture using multi-scale molecular dynamics-lattice Boltzmann simulations. The size of the nanopore is set to allow for hairpin passing.

1.5 Motivation

Polymer translocation is a process that appears in many biological systems. In this process, a polymer chain travels from one side of a barrier to the other side through a nanopore. However, before this process takes place, the chain must find the pore first. This process is called polymer capture. Recent studies on the polymer capture showed that the common approach to capture, which assumes that the chain arrives at the entrance as a random coil and after several attempts, an end gets captured, does not draw an accurate picture of forced capture [54, 55]. Mihovilovic et al. conducted an experimental study on electrophoretic capture of DNA. In Fig. 2b of Ref. [54], it is shown that despite the formation of hairpins, DNA molecules have a strong tendency to thread into the nanopore near the ends. In another study, Farahpour et al. showed that chain extensions occur during the capture of a DNA molecule in an electric field [55]. As a part of our goal to provide a better insight into the polymer capture, we will shed light on why the capture happens often by the ends, and if there is any relation between the extensions observed and the insertion conformation.

1.6 Outline

In the following chapter, we provide the details regarding our fluid and molecular model as well as some background on the models. we present the results obtained from our simulation in chapter 3. The fluid in our system is driven by a pressure jump at the boundary which results in two flow regions, a weakly-driven bulk and a converging area near the nanopore. The definition of the different regions and the impact of having this geometry is discussed in 3.1. In this section, we show that a hydrodynamic flow can eliminate the possibility of unsuccessful capture even when it is weakly-driven in the bulk. In section 3.2, we discuss the possible deformations in polymer conformation during the capture process. The dependency of the intensity of a pressure-driven flow on the cross-section area creates a flow with two regimes. The stronger regime close to the nanopore induces extensions on the chain. Section 3.3 is dedicated to introducing the concept of the pulley effect which causes the unravelling of the polymer. Pulley effect is the result of a balance between elastic recoiling forces of the two strands of a hairpin and the drag force of the flow. Finally, in chapter 4, we summarize with conclusions and possible future work.

Chapter 2

Methodology

In the previous chapter, we discussed polymer molecules, their measures of shape and dynamics, and finally, introduced polymer capture and translocation. In order to successfully reproduce the behaviour of polymer capture under hydrodynamic flow, we need a fluid model that accounts for such interactions. A fluid system consists of an enormous number of molecules on the microscopic scale. This creates an obstacle for simulating such systems using molecular dynamics. Molecular dynamics (MD) is a method that solves Newton's equation for all the microscopic particles in the system and determines the molecular state of the system at each time step [56]. The molecular dynamics method incorporates the thermodynamic features of the system using statistical dynamics calculations. With this definition, one can see that the obstacle mentioned before is the time and resources required to conduct molecular dynamics for fluid systems.

The other reason that MD methods are not the most popular way to simulate the behaviour of fluids is that the macroscopic quantities of interest do not depend on the detail of the motion of every particle and their interactions, but they are the result of the behaviour of a collection of particles. The continuum approaches directly deal with macroscopic variables and does not suffer from this problem. The Navier-Stokes equations can be solved for a certain boundary and initial conditions, and the dynamics of the fluid and its flow can be predicted. Since solving such equations are difficult due to non-linearity, a group of discretizing schemes, for example, finite volume method, has been used to convert the Navier-Stokes equations to a set of algebraic equations [57]. The drawbacks of using these methods are the truncation error rising from the discretization and the potential for numerical instability.

The mesoscale approaches are something between the microscopic and macroscopic methods. Since the molecular level details are not of interest, these models simulate the dynamics of the system by following the motion of groups of molecules and the evolution of their representative distribution function. One of these methods is the lattice-Boltzmann (LB) method. In the next section, we first touch on the fundamental physical theory behind the LB method. Then we describe the general scheme of LB algorithms and its features, and elaborate on the fluctuating lattice-Boltzmann package, which we used to reproduce the fluid in our system. In the last section of this chapter, we discuss our system setup and the advantages of applying hybrid molecular-dynamics lattice-Boltzmann to understand the dynamics of polymer capture and translocation.

2.1 Fluid model

2.1.1 Kinetic Theory and the Boltzmann equation

Imagine a box consisting of a total number of N molecules with various positions \mathbf{r} and velocities \mathbf{v} . Since we are not interested in the motion of each molecule in detail, we focus on the probability distribution function $f(\mathbf{r}, \mathbf{v}, t)$ of molecules at time t in the volume element $d^3\mathbf{r}$ with velocities that lie within a $d^3\mathbf{v}$ element of the velocity space, and number density n ,

$$n = f(\mathbf{r}, \mathbf{v}, t)d^3\mathbf{v}. \quad (2.1)$$

In all our discussion of this chapter, when we mention a volume element, we mean a very small element from a macroscopic viewpoint. That is, the volume element may consist of a large number of molecules. Therefore, the distribution function shows the distribution of position and momenta among all the members of the small finite volume element.

Since each particle is at a certain position and has a certain velocity, we can construct a phase space of positions and velocities. A specific configuration of a system with a total number of molecules N is represented with N points in the phase space. This phase space can be divided into elements with a volume of $d^3\mathbf{r}d^3\mathbf{v}$. Therefore, the total number of molecules of the system can be found from the summation of the total number of molecules in each element at time t ,

$$N = \sum n(\mathbf{r}, \mathbf{v}, t)d^3\mathbf{r}. \quad (2.2)$$

Considering the fact the volume elements consist of a large number of molecules, we can assume that the density of the phase space points does not change drastically from one element to the neighbours and hence, the distribution function can be regarded as a continuous function and

$$N = \int n(\mathbf{r}, \mathbf{v}, t)d^3\mathbf{r} = \int f(\mathbf{r}, \mathbf{v}, t)d^3\mathbf{r}d^3\mathbf{v}. \quad (2.3)$$

In order to understand the dynamic behaviour of the system, we need to find the form of the distribution function and how it evolves in time. This is essentially the main goal of kinetic theory [58].

The first step in finding the distribution function is to find the equation of motion of this function. The change of the distribution function over time is due to the exchange of molecules between different volume elements. In the case of no collisions between the moving molecules, a molecule with coordinates (\mathbf{r}, \mathbf{v}) at time t will move to $\mathbf{r} + \mathbf{v}\delta t$ at time $t + \delta t$, assuming δt is small. The new velocity of the molecule will be $(\mathbf{v} + \frac{\mathbf{F}}{m}\delta t)$ where \mathbf{F} is the external force acting on a single molecule and m is the mass of the molecule. Therefore, we can find all the molecules that were in $d^3\mathbf{r}d^3\mathbf{v}$ at time t in $d^3\mathbf{r}'d^3\mathbf{v}'$ at time $t + \delta t$,

$$f(\mathbf{r} + \mathbf{v}\delta t, \mathbf{v} + \frac{\mathbf{F}}{m}\delta t, t + \delta t)d^3\mathbf{r}d^3\mathbf{v} = f(\mathbf{r}', \mathbf{v}', t)d^3\mathbf{r}'d^3\mathbf{v}'. \quad (2.4)$$

In equilibrium, the probability distribution or the equivalent here, the number of particles with a given velocity \mathbf{v} , is conserved. Therefore, we must have $f(\mathbf{r}, \mathbf{v}, t) = f(\mathbf{r}', \mathbf{v}', t)$ and

$$f(\mathbf{r} + \mathbf{v}\delta t, \mathbf{v} + \frac{\mathbf{F}}{m}\delta t, t + \delta t) = f(\mathbf{r}, \mathbf{v}, t). \quad (2.5)$$

Up to this point, the collisions between molecules were ignored. This effect can be added as the rate of change of the distribution function to Eq. 2.5,

$$\frac{Df}{Dt}\delta t = f(\mathbf{r} + \mathbf{v}\delta t, \mathbf{v} + \frac{\mathbf{F}}{m}\delta t, t + \delta t) - f(\mathbf{r}, \mathbf{v}, t) = \Delta(f) \quad (2.6)$$

where Δ is the collision operator and D is the material derivative defined as,

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + (\mathbf{v} \cdot \nabla_{\mathbf{r}}) + \left(\frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}}\right) \quad (2.7)$$

where $\nabla_{\mathbf{r}}$ and $\nabla_{\mathbf{v}}$ are the gradients with respect to position and velocity, respectively. Eq.2.6 is the Boltzmann transport equation.

To be able to characterize the distribution function, it is necessary to know the nature of the collision operator. This can be difficult due to the complexity of molecular interactions and collisions. However, it has been shown that the first-order approximation in an expansion about equilibrium can provide reasonable results for most practical problems. In this model, the collision function is written as a linear function of the distribution function,

$$\frac{Df}{Dt} = -\frac{1}{\tau_b}(f - f^{eq}) \quad (2.8)$$

where τ is the collision time and f^{eq} is the equilibrium distribution function.

The best way of testing a model is to find the connection between the model and measurable real-world quantities. This is how experiments confirm mathematical models. By applying the conservation theorem relevant to the Boltzmann transport equation and defining mass density,

$$\rho(\mathbf{r}, t) = mn(\mathbf{r}, t) \quad (2.9)$$

we can derive the conservation of mass, momentum, and energy equations:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \quad (2.10)$$

$$\rho \left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla \right) \mathbf{u} = \frac{\rho}{m} \mathbf{F} - \nabla \cdot \bar{\bar{P}} \quad (2.11)$$

$$\rho \left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla \right) \theta = -\frac{2}{3} \nabla \cdot \mathbf{q} - \frac{2}{3} \bar{\bar{P}} \cdot \bar{\bar{\Gamma}} \quad (2.12)$$

where $\mathbf{u}(\mathbf{r}, t) \equiv \langle \mathbf{v} \rangle$. The average value is defined as,

$$\langle \mathbf{v} \rangle \equiv \frac{\int d^3v f \mathbf{v}}{\int d^3v f} = \frac{1}{n} \int d^3v f \mathbf{v} \quad (2.13)$$

where d^3v is a velocity volume element around \mathbf{v} . The other quantities are defined as follows:

$$\rho(\mathbf{r}, t) \equiv m \int d^3v f(\mathbf{r}, \mathbf{v}, t) \quad (\text{mass density}) \quad (2.14)$$

$$\mathbf{u}(\mathbf{r}, t) = \langle \mathbf{v} \rangle \quad (\text{average velocity}) \quad (2.15)$$

$$\theta(\mathbf{r}, t) \equiv \frac{1}{3} m \langle |\mathbf{v} - \mathbf{u}|^2 \rangle \quad (\text{temperature}) \quad (2.16)$$

$$\mathbf{q}(\mathbf{r}, t) \equiv \frac{1}{2} m \rho \langle (\mathbf{v} - \mathbf{u}) |\mathbf{v} - \mathbf{u}|^2 \rangle \quad (\text{heat flux vector}) \quad (2.17)$$

$$P_{ij} \equiv \rho \langle (v_i - u_i)(v_j - u_j) \rangle \quad (\text{Pressure tensor}) \quad (2.18)$$

$$\Gamma_{ij} \equiv \frac{1}{2} m \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \quad (2.19)$$

The physical interpretation of the equations above and the names assigned to the quantities become more clear when the first-order approximation is applied and the equations are rearranged. We skip over the details of the derivation, which can be found in any standard kinetic theory and statistical mechanics book [58, 59], and focus on the few important steps that help us to recover the equations of hydrodynamics. The Maxwell-Boltzmann distribution is a possible equilibrium distribution which obtains good result in the case of a dilute gas and is used as an example here [58].

We define $g(\mathbf{r}, \mathbf{v}, t) = f - f^{eq}$ where

$$f^{eq} = n \left(\frac{m}{2\pi\theta} \right)^{3/2} \exp \left(-\frac{m}{2\theta} (\mathbf{v} - \mathbf{u})^2 \right). \quad (2.20)$$

Eq. 2.6 in the limit of $\delta t \rightarrow 0$ and with first-order collision operator can be rewritten as,

$$\frac{f(\mathbf{r} + \mathbf{v}\delta t, \mathbf{v} + \frac{\mathbf{F}}{m}\delta t, t + \delta t) - f(\mathbf{r}, \mathbf{v}, t)}{\delta t} = \left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{r}} + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} \right) f \approx -\frac{f - f^{eq}}{\tau_b}. \quad (2.21)$$

By plugging $g(\mathbf{r}, \mathbf{v}, t)$ in Eq. 2.21 assuming that $g \ll f^{eq}$ (which is reasonable for small Knudsen number $Kn = \lambda/L$ where λ is the free mean path and L is the characteristic length),

$$-\tau_b \left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla_{\mathbf{r}} + \frac{\mathbf{F}}{m} \cdot \nabla_{\mathbf{v}} \right) f^{eq} \approx g \quad (2.22)$$

which can be substituted into Eq. 2.17 and 2.18 to obtain \mathbf{q} and P_{ij} ,

$$\mathbf{q} = \frac{m\rho}{2n} \int d^3v (\mathbf{v} - \mathbf{u}) |\mathbf{v} - \mathbf{u}|^2 g = -K \nabla \theta \quad (2.23)$$

$$P_{ij} = \delta_{ij} P - \frac{2\eta}{m} \left(\Gamma_{ij} - \frac{m}{3} \delta_{ij} \nabla \cdot \mathbf{u} \right) \quad (2.24)$$

where $K = 5/2\tau_b\theta n$ is the thermal conductivity and $\eta = \tau_b n \theta$ is the viscosity. By substituting Eq. 2.23 and 2.24 into the conservation equations and neglecting the higher derivatives of ρ , θ ,

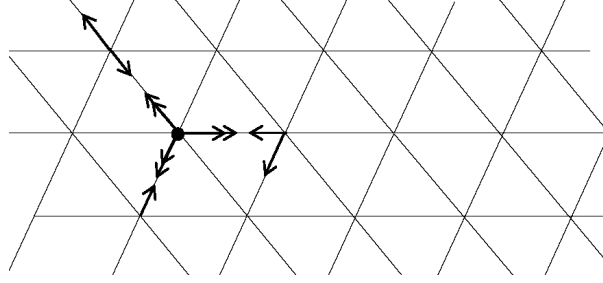


Figure 2.1: A triangular mesh was used in HLG model to improve the isotropy of the LGCA model. The collisions in the HLG model are head-on collisions. The particles from occupied sites move along the links, collide at an empty site, and leave the site in a perpendicular direction to their original path. The single and double arrows show the propagation at time t and $t + 1$.

and \mathbf{u} ,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \quad (2.25)$$

$$\left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla \right) \mathbf{u} = \frac{\mathbf{F}}{m} - \frac{1}{\rho} \nabla \left(P - \frac{\eta}{3} \nabla \cdot \mathbf{u} \right) + \frac{\eta}{\rho} \nabla^2 \mathbf{u} \quad (2.26)$$

$$\left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla \right) \theta = -\frac{1}{c_V} (\nabla \cdot \mathbf{u}) \theta + \frac{K}{\rho c_V} \nabla^2 \theta \quad (2.27)$$

where $\eta = \sqrt{mk_B T}/a^2$ is the viscosity coefficient, a is the molecular diameter, and $c_V = 3/2k_B\theta$ is heat capacity. The first equation is the continuity equation, the second one is the Navier-Stokes equation (NSE), and the third equation is the heat equation. The fact that Navier-Stokes equations can be derived from the first-order approximation of the Boltzmann transport equation raises the hope that the behaviour of fluids can be simulated using this approach through a reasonable algorithm.

2.2 Fluid simulations with Lattice-Boltzmann

The conventional computational fluid dynamics (CFD) methods attempt to solve Eq. 2.26 (NSE) by discretizing them in space and time. The $\mathbf{u} \cdot \nabla \mathbf{u}$ term on the left-hand side of Eq. 2.26 brings non-linearity to this system of second-order partial differential equations. Key requirements for accurate simulations of fluids are conservation of mass, momentum, and energy. The truncation errors arising from the discretizing NSE makes meeting these requirements difficult and the need for implicit iterative algorithms to obtain a converged solution makes simulating fluid behaviour based on NSE even more challenging. A great deal of effort has been dedicated to developing more efficient and stable techniques to handle Navier-Stokes equations. On the other hand, finding alternative formulations for fluid dynamics has been pursued by a group of scientists. The "lattice-gas automata" is one of the old methods which was used to reproduce hydrodynamics [60].

The original lattice-gas cellular automata (LGCA) was proposed by Hardy et al. in 1973 and is known as the HPP model after the three authors Hardy, de Pazzis, and Pomeau. The HPP model is set up on a regular, square, 2D lattice with nodes that are connected by unit

length links [61]. The lack of isotropy and Galilean invariance of the HPP model encouraged Frisch et al. to come up with a different scheme that accounts for the isotropy [62]. They found out that a triangular mesh brings better isotropy into the model. Since each node is surrounded by 6 neighbour sites, the model was called the hexagonal lattice gas (HLG), as shown in Fig. 2.1. Although the collision rules and mesh structure were modified from HPP to HLG, and as a result, the lack of isotropy was overcome, the HLG model still suffered from some other fundamental drawbacks, such as spurious invariants and noise [63]. The LGCA methods inherit their noisy nature from the Boolean variables used in the formulation. The problem of noise was solved by switching from Boolean variables to particle distribution functions which brought a new branch of hydrodynamic simulations to life, namely, the lattice-Boltzmann (LB) method [64].

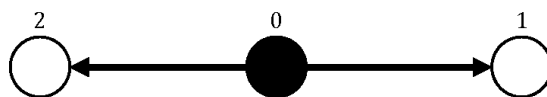


Figure 2.2: In D1Q3 model, the fictitious particle can travel to one of the adjacent sites via $\mathbf{e}_1 = 1$ or $\mathbf{e}_2 = -1$ in lattice units, or it can stay at the current site.

The LB method is based on solving a discretized, both in time and space, version of Boltzmann's transport equation on a lattice,

$$f_i(x_\kappa + \Delta x, t + \Delta t) - f_i(x_\kappa, t) = \Delta_i \quad (2.28)$$

where f_i is a partial distribution function, x_κ is the coordinate in the direction κ ($\kappa = x, y, z$ in Cartesian coordinates), Δx is the lattice spacing, Δt is the time step, and Δ_i is a discretized collision operator.

Each site of the lattice is occupied with a fictitious group of particles representing the distribution function at that site. In a typical LB algorithm, particles move from their current site to the neighbouring sites in the streaming step, and then the distribution function is updated according to collision rules in the collision step. The number of components of the partial distribution function and the velocity vector may vary depending on the lattice arrangement.

In one dimension, the most common arrangement is *D1Q3*, which stands for a one-dimensional model with 3 neighbouring sites, as illustrated in Fig. 2.2. In general, the notation *DdQn* labels the LB model where d is the spatial dimension and n is the number of neighbouring sites including the current site. The velocities by which the particle moves to the other sites are vectors \mathbf{e}_i where $i = 0, 1, 2, \dots, n - 1$. The 2D and 3D models are shown in Fig. 2.3 and Fig. 2.4. The velocities for different models can be obtained based on their definition, for example, the velocities of D3Q15 are,

$$\begin{aligned} \mathbf{e}_0 &= (0, 0, 0)c_l, \quad \mathbf{e}_1 = (1, 0, 0)c_l, \quad \mathbf{e}_2 = (0, 1, 0)c_l \\ \mathbf{e}_3 &= (-1, 0, 0)c_l, \quad \mathbf{e}_4 = (0, -1, 0)c_l, \quad \mathbf{e}_5 = (0, 0, 1)c_l \\ \mathbf{e}_6 &= (0, 0, -1)c_l, \quad \mathbf{e}_7 = (1, 1, 1)c_l, \quad \mathbf{e}_8 = (-1, 1, 1)c_l \\ \mathbf{e}_9 &= (-1, 1, -1)c_l, \quad \mathbf{e}_{10} = (1, 1, -1)c_l, \quad \mathbf{e}_{11} = (1, -1, 1)c_l \\ \mathbf{e}_{12} &= (-1, -1, 1)c_l, \quad \mathbf{e}_{13} = (-1, -1, -1)c_l, \quad \mathbf{e}_{14} = (1, -1, -1)c_l. \end{aligned}$$

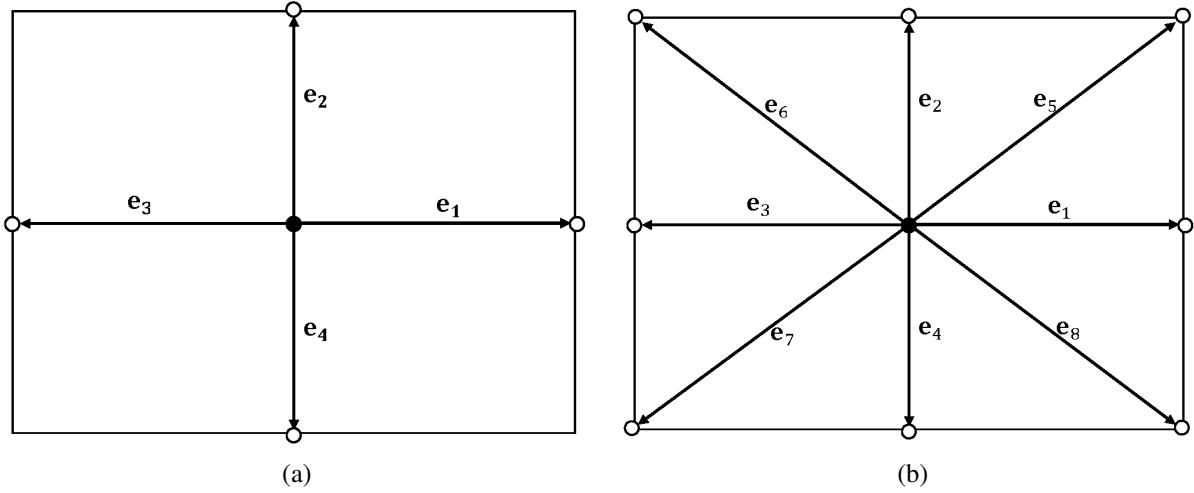


Figure 2.3: In D2Q5, the streaming happens only vertically or horizontally while in D2Q9, the digonal motion is also possible.

where $c_l = \delta x / \delta t$ is the lattice velocity. As mentioned before, the evolution of the distribution function depends on the collision operator and being able to solve the LB equation and simulate fluids hinges on knowing this operator. A very common and broadly-used LB model is the single-relaxation time model. This model is known as the BGK model after Bhatnagar, Gross, and Krook who introduced this model in 1954 [65]. Substituting the first-order discrete collision term into Eq. 2.28 and rewriting the equation in differential format, we obtain the equation for the evolution of f_i

$$D_i f_i \equiv (\partial_t + e_{ik} \partial_k) f_i = -\frac{1}{\tau_b} (f_i - f_i^{eq}). \quad (2.29)$$

where D_i is defined as the material derivative in direction \mathbf{e}_i , and τ_b is the relaxation time. This is the simplest form of LB equation without any external forcing or thermal fluctuations term. The local mass and momentum are defined as,

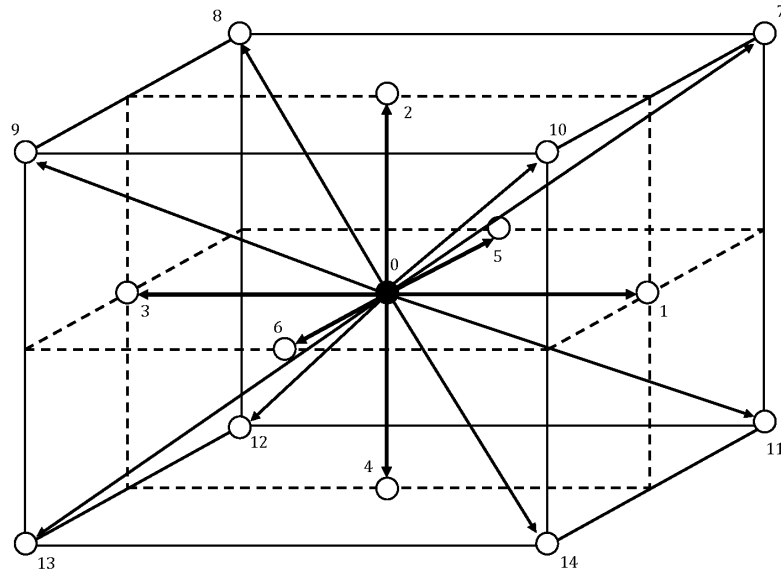
$$\rho(\mathbf{x}, t) = \sum_i f_i(\mathbf{x}, t) \quad (2.30)$$

$$\rho(\mathbf{x}, t) \mathbf{u}(\mathbf{x}, t) = \sum_i f_i(\mathbf{x}, t) \mathbf{e}_i \quad (2.31)$$

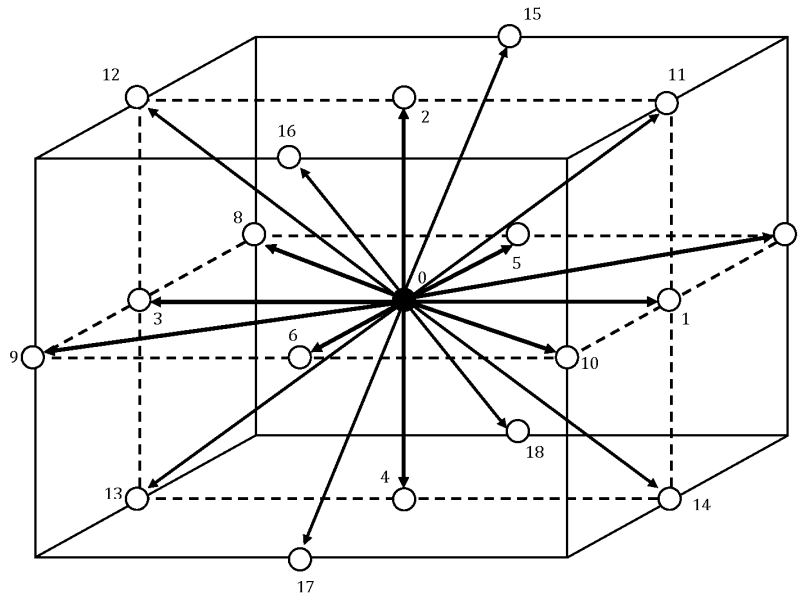
where $\mathbf{x} = (i, j, k) \Delta x$, $\rho(\mathbf{x}, t)$ is the mass density and \mathbf{e}_i is the velocity vector in the i th direction and the conservation laws are reinforced by the following constraints,

$$\sum_i f_i^{eq}(\mathbf{x}, t) = \rho(\mathbf{x}, t) \quad (2.32)$$

$$\sum_i f_i^{eq}(\mathbf{x}, t) \mathbf{e}_i = \rho(\mathbf{x}, t) \mathbf{u}(\mathbf{x}, t) \quad (2.33)$$



(a)



(b)

Figure 2.4: Illustration of common 3D lattice arrangements for LB method. (a) shows a model with 15 neighbouring sites and (b) shows the 19 neighbour model.

The relation between the terms in Eq. 2.29 and 2.26 can be figured out through a Chapman-Enskog expansion [60] of the equilibrium distribution function,

$$f_i^{eq} = \rho \left[a + b \mathbf{e}_i \cdot \mathbf{u} + c (\mathbf{e}_i \cdot \mathbf{u})^2 + d (\mathbf{u} \cdot \mathbf{u}) \right] \quad (2.34)$$

where a , b , c , and d are constants. To find the constants we consider the conditions 2.32 and

model	ω_0	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8	ω_9	ω_{10}	ω_{11}	ω_{12}	ω_{13}	ω_{14}
D1Q3	$\frac{4}{6}$	$\frac{1}{6}$	$\frac{1}{6}$												
D2Q5	$\frac{2}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$										
D2Q9	$\frac{4}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$						
D3Q15	$\frac{16}{72}$	$\frac{8}{72}$	$\frac{8}{72}$	$\frac{8}{72}$	$\frac{8}{72}$	$\frac{8}{72}$	$\frac{8}{72}$	$\frac{1}{72}$	$\frac{1}{72}$	$\frac{1}{72}$	$\frac{1}{72}$	$\frac{1}{72}$	$\frac{1}{72}$	$\frac{1}{72}$	$\frac{1}{72}$

Table 2.1: The table shows the weight factors for different models.

2.33 for the equilibrium distribution and, therefore,

$$f_i^{eq} = \rho \omega_i \left[1 + 3\mathbf{e}_i \cdot \mathbf{u} + \frac{9}{2}(\mathbf{e}_i \cdot \mathbf{u})^2 - \frac{3}{2}(\mathbf{u} \cdot \mathbf{u}) \right] \quad (2.35)$$

where ω_i is the weight associated with \mathbf{e}_i . Table 2.1 shows the weights for different LB models. Now that the equilibrium distribution is specified, one can derive other continuum quantities by working out the higher velocity moments of the equilibrium distribution function.

As shown above, the LB method is capable of reproducing the hydrodynamics equations. However, in systems like the one in our study, there are other parameters to be considered, for instance, the interactions between the polymer molecule immersed in the fluid and the thermal motion of the polymer caused by the collisions between the macromolecule and the fluid molecules. A significant number of studies have been done on implementing fluid-structure interactions and thermal fluctuations into the LB algorithm [66, 67, 68]. Different programs and packages have been built based on proposed models among which we use the LAMMPS (Large-scale atomistic/molecular massively parallel simulator) [69] standard package LB-FLUID introduced by Mackay et al. in 2013 [70].

2.2.1 Our system: a polymer chain in LB-FLUID

Many biological phenomena and industrial processes involve motion of suspended particles in a fluid, for example, colloidal suspensions, cell motion in the bloodstream, polymer processing, etc. Polymer capture also belongs to this group of processes. Therefore, To successfully simulate the polymer capture, a tool is required that can handle the motion of embedded particles in a fluid. The LB method is a mesoscale method which is flexible with different boundary conditions and hence, very well capable of simulating fluid-structure systems. In our study, the structure is made up of molecular dynamics particles.

In addition, since polymer capture and translocation is a nano-scale process, the thermal agitation becomes crucial for describing the motion of the polymer correctly. Therefore, we do not apply a basic LB algorithm but a fluctuating model that accounts for thermal fluctuations.

The LB-FLUID package of LAMMPS follows the general theme of LB algorithms discussed in the previous section. However, a general forcing term is added to the Boltzmann

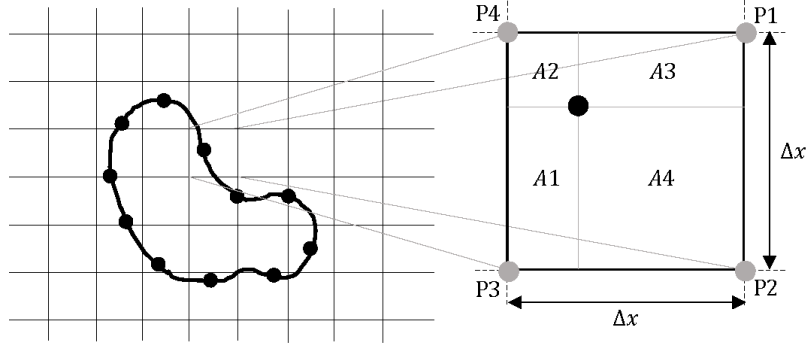


Figure 2.5: An immersed surface in a fluid mesh is shown. In trilinear stencil, the mapping from surface nodes to lattice sites and vice versa are done with help of weights that are defined based on the area between the surface node and the mesh sites.

equation 2.29 through which the thermal fluctuations and fluid-structure interactions are implemented into the method,

$$D_i f_i \equiv (\partial_t + e_{i\kappa} \partial_\kappa) f_i = -\frac{1}{\tau_b} (f_i - f_i^{eq}) + W_i \quad (2.36)$$

where W_i is the general forcing term [68]. This term encompasses two parts, the external forces p_i and thermal noise ξ_i ,

$$W_i = p_i + \frac{1}{\tau_b} \bar{\xi}_i \quad (2.37)$$

where $1/\tau_b$ is added to make the definition of $\bar{\xi}_i$ and f_i^{eq} consistent. The external forcing term is related to the applied external force (electric, gravity, etc.) through its moments,

$$\sum p_i = 0 \quad (2.38)$$

$$\sum p_i e_{i\kappa} = F_\kappa \quad (2.39)$$

$$\sum p_i e_{i\kappa} e_{i\beta} = u_\kappa F_\kappa + u_\beta F_\beta \quad (2.40)$$

where F_κ is the external force density. To calculate the force exerted on the fluid by any embedded surface, the surface must be mapped onto the fluid mesh. The first step is to discretize the surface to a set of nodes (surface nodes). Then, the surface nodes are distributed among surrounding lattice sites. The distribution weights can be chosen using different schemes. For example, in 2D, a bilinear stencil gives weights of the form $a_{ij}^{(T)} = A_{ij}/\Delta x^2$ where A_{ij} is the opposing area to the area between mesh site i and surface node j . An illustration of this method is shown in Fig. 2.5. This method is general and can be used in 3D as well where instead of relative areas, the weights are found from the ratio of the volumes to the total volume. There is no constraint on the shape of the surface and no extra information, other than the solid's linear and angular velocity, is needed to find its effect on the fluid.

An early, widely used scheme, is the immersed boundary method introduced by Charles Peskin in 1972 [71]. Both the Peskin and trilinear stencil are implemented in LB-FLUID and can be used in simulations. We use a trilinear stencil in our simulations primarily because

of the smearing effects of the Peskin stencil [72]. These effects arise from the spreading of each point across more lattice sites. As we will see, our nano-scale channel is made out of a hole through an MD solid-state wall. With the Peskin stencil, we observed that the fluid leaks through the walls. This problem can be overcome by calibrating the fluid parameters and applying a trilinear stencil. We used both in our case. Fig. 2.6 shows the fluid flow in terms of streamlines. As can be seen, the streamlines penetrate through the wall around the hole when the Peskin stencil was used, but not with the calibrated trilinear stencil.

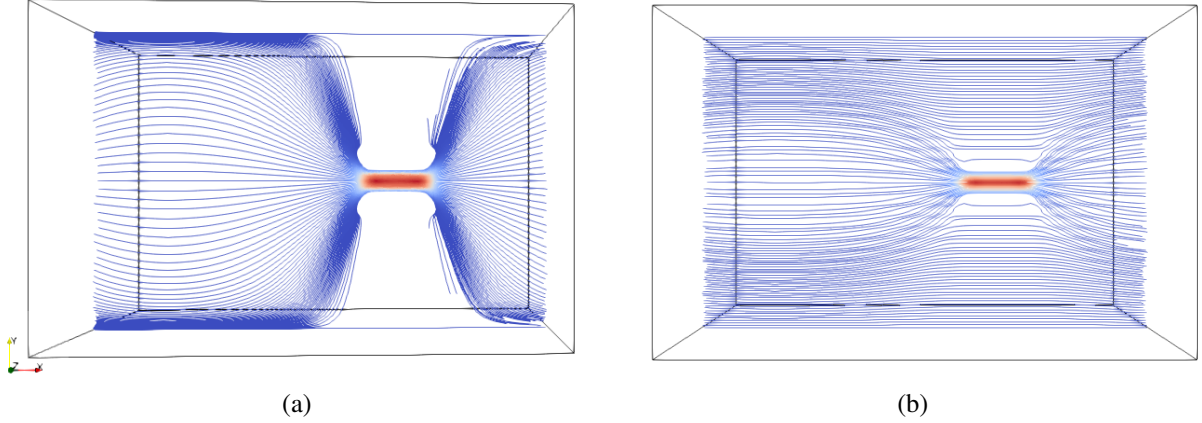


Figure 2.6: (a) shows the fluid flow in terms of streamlines. The narrow red area is inside the hole where the flow is stronger. (b) shows the same flow but with the Peskin stencil. Some of streamlines go through the walls.

While the chain is moving through the fluid, it interacts with the fluid. The LB-FLUID handles these interactions as follows. The location of different nodes of the surface \mathbf{r}_i are calculated with respect to the centre of mass (com) of immersed objects and therefore their velocities are found by $\mathbf{v}_i = \mathbf{v}_{com} + \boldsymbol{\Omega}_j \times \mathbf{r}_i$ where $\boldsymbol{\Omega}_j$ is the angular velocity of the j th monomer. The local forces at each lattice site j imposes a force on the surface node i like,

$$\mathbf{F}_{ij} = (\mathbf{v}_i - \hat{\mathbf{u}}_i) a_{ij} \gamma. \quad (2.41)$$

Since the velocities are known at each mesh site the next step is to determine $\hat{\mathbf{u}}_i$. The simplest way is to use $\hat{\mathbf{u}}_i = \mathbf{u}(x_j)$ where $\mathbf{u}(x_j)$ is the fluid velocity at mesh site j . In contrast with this “noninterpolating” method, the velocity of the fluid at the surface node can be found as following,

$$\hat{\mathbf{u}}_i = \sum_j a_{ij} \mathbf{u}(x_j) \quad (2.42)$$

where for a_{ij} either the trilinear or Peskin weights can be used. A normalization condition is applied on the weights so that the fluid mesh resolution doesn’t have an impact on the calculated area of the immersed body,

$$\sum_j^n a_{ij} = 1 \quad (2.43)$$

where $n = 2d$ for trilinear stencil and $n = 4d$ for Peskin stencil. As shown by Ollila et al. [72], the presence of mesh effects poses a challenge on defining the hydrodynamic radius of a single

particle and as a result, defining the friction coefficient γ . They solved the mesh effect problem using composite particles. For our simulations, we follow the same path.

The LB-FLUID algorithm has a standard LB integrator. The algorithm is second order accurate in Δx and Δt . The stability of the model is guaranteed by $\tau_b/\Delta t \sim O(1)$ and $c_s < c_l$ where $c_s = \sqrt{\partial P/\partial \rho}$ is the speed of sound and P is the normal pressure. The equilibrium distribution of Eq. 2.35 is valid in the low Mach numbers limit which is a good approximation for our system. The model converges to Navier-Stokes equations of hydrodynamics in this limit.

Our polymer molecule is created out of composite particles consist of central and shell atoms, as shown in Fig. 2.7a. The word atom refers to point MD particles. The central atoms (monomers) are connected via FENE bonds as defined in Eq. 1.8,

$$U = -\frac{1}{2}kR_0^2 \ln\left(1 - \left(\frac{r}{R_0}\right)^2\right) + 4\epsilon\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 + \frac{1}{4}\right)H\left(2^{1/6} - \frac{r}{\sigma}\right), \quad (2.44)$$

where r is the distance between adjacent monomers, the first term on the right is the elastic potential with elastic constant $K = 30\epsilon\sigma^{-2}$ in which $\epsilon = k_B T$, where k_B is the Boltzmann constant and T is temperature. The maximum distance $R_0 = 1.5\sigma$ where σ is the Lennard-Jones (LJ) length scale. The second term is the truncated and shifted 12-6 LJ potential multiplied by the Heaviside step function $H\left(2^{1/6} - \frac{r}{\sigma}\right)$ in order to reflect a repulsive behaviour (truncation at the minimum). The same LJ forces are applied between the monomers to reproduce the excluded-volume effect for the chain (self-avoiding chain). Each monomer is surrounded by 31 shell atoms which interact with the fluid. Each group consisting of 1 central atom and 31 shell atoms make a rigid sphere of radius $R_{shell} \approx 0.7\text{nm}$. Hence, all the interactions between the shell atoms and the fluid affect the motion of the monomer as well. The particular coupling method used in this model satisfies the no-slip condition on the shell surface [72].

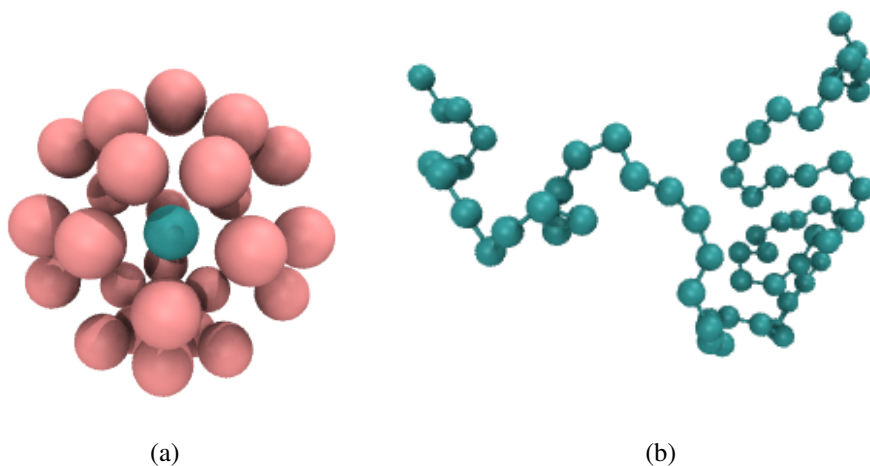


Figure 2.7: (a) shows a composite particle. The pink particles are the shell atoms and the jade particle is the central atom which altogether make a rigid sphere. (b) is an illustration of our polymer chain.

a	0	1	2	3	4	5	6	7	8	9	10-12	13	14
M_{eq}^a	ρ	ρu_x	ρu_y	ρu_z	$P_{xx} + \rho \left(u_x^2 - \frac{c_l^2}{3} \right)$	$P_{yy} + \rho \left(u_y^2 - \frac{c_l^2}{3} \right)$	$P_{zz} + \rho \left(u_z^2 - \frac{c_l^2}{3} \right)$	$P_{xy} + \rho u_x u_y$	$P_{yz} + \rho u_y u_z$	$P_{xz} + \rho u_x u_z$	0	$\rho u_x u_y u_z$	K^{eq}
$\bar{\chi}$	0	0	0	0	s_{xx}	s_{yy}	s_{zz}	s_{xy}	s_{yz}	s_{xz}	$\bar{\chi}^a$	$\bar{\chi}^{13}$	$\bar{\chi}^{14}$

Table 2.2: The moments of equilibrium distribution M_{eq}^a and of the continuum random process $\bar{\chi}$ for model D3Q15.

Up to this point, we discussed the implementation of surface-fluid interaction in the LB-FLUID package and our polymer model which is chosen to match with the LB-FLUID scheme. The other important component in simulating the motion of a polymer chain is the thermal fluctuations. As mentioned before, the LB-FLUID package accounts for this type of dynamics through the general forcing term, via the second term in Eq. 2.37. The continuum random process $\bar{\xi}_i$ can be constructed from the moments of local thermal fluctuations $\bar{\chi}$,

$$\bar{\xi}_i(\mathbf{x}, t) = \omega_i \sum_a m_i^a \bar{\chi}^a(\mathbf{x}, t) N^a \quad (2.45)$$

where m_i^a corresponds to moments of the velocity tensor \mathbf{e}_i , $\bar{\chi}^a$ is the local thermal fluctuations in the corresponding moment, and N^a is a normalization factor. The values of these quantities can be found in Ref. [68]. In the presence of fluctuations the stress tensor can be split into three components,

$$\sigma_{\kappa\beta} = -P_{\kappa\beta} + \eta_{\kappa\beta\alpha\gamma} \partial_\alpha u_\gamma + s_{\kappa\beta} \quad (2.46)$$

where $P_{\kappa\beta}$ is the pressure tensor, $\eta_{\kappa\beta\alpha\gamma}$ is the viscosity tensor, and $s_{\kappa\beta}$ is fluctuating stress tensor. The fluctuating stress tensor is related to viscosity through a fluctuation-dissipation theorem,

$$\langle s_{\kappa\beta}(\mathbf{r}, t) s_{\alpha\gamma}(\mathbf{r}', t') \rangle = 2\eta_{\kappa\beta\alpha\gamma} k_B T \delta(\mathbf{r} - \mathbf{r}') \delta(t - t') \quad (2.47)$$

and to the local thermal fluctuations through the moments of the continuum thermal noise as given in Table. 2.2.

As discussed in the first chapter, the radius of gyration is an appropriate measure of the size of a polymer chain. Therefore, we measured the equilibrium size of the chain by running a free chain simulation of our chain and we found $R_G = R_F \approx 8.6\text{nm}$ for a 64-mer chain. In addition to fluid and polymer molecules, we have a nanochannel in the system. This nanopore is made out of a channel of size $10\text{nm} \times 4\text{nm} \times 4\text{nm}$ with rectangular cross-section through a fixed solid wall. The width of the nanopore is chosen to allow single-folded conformations (hairpins) to thread through while does not allow knots (yarn-like conformations) to pass. Since the research focus is on polymer capture rather than translocation, the shape of the pore would not have an impact on our results. This being said, since rectangular channels are a common geometry used in detection instruments, we go with that. The wall is created by MD particles which have a significantly larger mass than the MD particles used for the polymer chain. The wall interacts with both the fluid and the polymer chain. The wall-polymer interactions are of repulsive LJ form and the no-slip condition is satisfied on the surface of the wall. The wall is of size $10\text{nm} \times 52\text{nm} \times 52\text{nm}$ and blocks the path of the fluid through the system other than

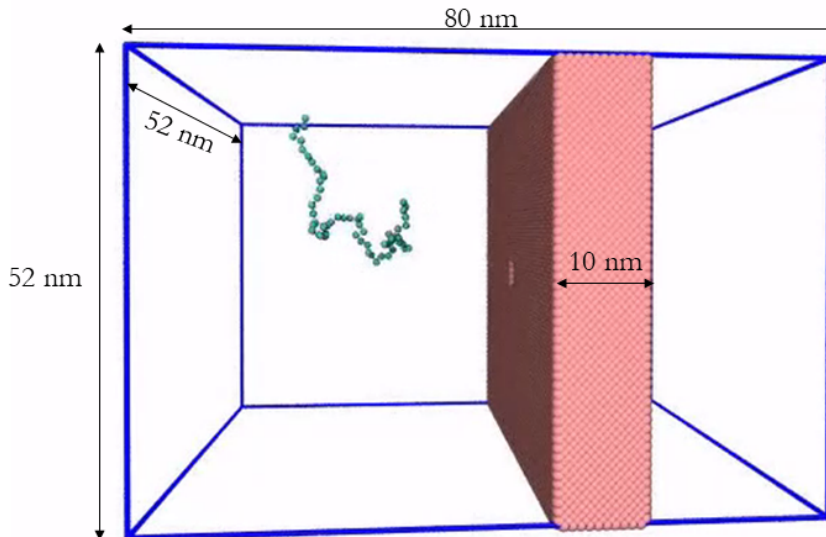


Figure 2.8: A snapshot of the system. The simulation box is a box of size $80\text{nm} \times 52\text{nm} \times 52\text{nm}$ with periodic boundaries in all directions. The size of the box is selected to allow extended chains without crossing the boundaries or interacting with itself. The length of the nanopore is chosen to be of the order of the radius of gyration of the chain.

through the nanopore. The beginning of the wall is placed at 48nm from the origin of the simulation box in the x -direction which is of size $80\text{nm} \times 52\text{nm} \times 52\text{nm}$ 2.8. The box size and the position of the wall are chosen this way mainly because of two reasons. Firstly, we want the chain to start far from the nanopore so that it gets time to explore its different conformations before threading through the pore. Since we fixed the initial position of the centre of mass of the chain to a specific area, namely $\{x = 12\text{nm}, y = 16 - 36\text{nm}, z = 16 - 36\text{nm}\}$, the distance $x_{ent} - x_{COM}^0 \approx 4R_F$ provides enough time for random wandering of the chain. x_{ent} is the position of the entrance of the nanopore in the x -direction, x_{COM}^0 is the initial position of the centre of mass of the chain, and R_F is the equilibrium radius of our chain.

Secondly, the box must be large enough to minimize any periodic boundary (PB) effects in the system. If the box size is comparable to the chain size, some segments might interact with their periodic shadow or the wall. The PB effect can also alter the fluid flow field in the case of small boxes.

The flow in our system is driven by a pressure jump at the x -boundary. The main feature of such a driving force is that it results in an inverse relation between the cross-section area and the flow force. In other words, the flow force increases locally wherever the cross-section decreases. This feature, which is absent in body-force driven flows, allows for a flow with two regimes: uniform weakly-driven flow in the bulk and converging fast flow near the nanochannel. To ensure that the flow is weakly-driven in the bulk, i.e. the motion of the chain is a balanced mix of diffusion and drift, we compared the flow velocity in the bulk with the relaxation time of the polymer chain. We also ran a couple of realizations for systems of lower and higher pressure jumps and compared the results to our main simulation. We take a closer look at this problem in section 3.1.

When it comes to computational studies, it is crucial to choose the methods and parameters

that both are realistic and computationally efficient. For our fluid, we set the density and dynamic viscosity to one-tenth of that of water. This gives similar dynamical behaviour to that of water while the chain can move easier. For our polymer chain, we first tested our hypothesis using a 32-mer chain. Since for longer chains, the size of the box and the wall must be adjusted as well, The simulation time increases almost linearly with the size of the chain. Therefore, a medium length chain is a good place to start. However, we base our analysis on a 64-mer chain which corresponds to a more relevant chain length in the physical examples.

Polymer translocation is a process commonly found in biological systems. Understanding the effect of hydrodynamics on this process and polymer capture can provide great insight into such systems which can lead to important breakthroughs in different fields from medicine to physics. In the next chapter, we show our results which emphasize the effects of hydrodynamic flow on polymer capture.

Chapter 3

Results

Polymer translocation is the process of a single polymer chain going through a biological or solid-state pore, the latter in our case, that is the same size or smaller than the radius of gyration of the chain. The polymer's journey starts in the bulk and far from the hole, as shown in Fig. 3.1a. In this region, the flow field is uniform and weakly-driven, as illustrated in Fig. 2.6a by the uniform streamlines. The change in the system cross-section from the bulk region to the nanopore results in a converging flow with growing strength, due to conservation of mass and momentum, approaching the nanochannel. Fig. 3.2 shows the velocity of the fluid as a function of position for systems with and without thermal noise. The inset is a logarithmic graph of the velocity as the fluid approaches the nanopore. As this graph shows, the velocity increases like $1/|r_{pore}|^2$, where $|r_{pore}|$ is the distance from the pore, as expected based on the fact that all the streamlines from the bulk must converge into the pore as it is the only exit. The converging flow induces extensions in the polymer, as can be seen in Fig. 3.1b. When a segment of the chain arrives at the entrance of the nanopore, Fig. 3.1c, the translocation starts. We will refer to the monomer that threads into the nanopore before any other monomer as the front monomer. The monomer which is ahead of other monomers at any given time is referred to as the leading monomer. The front monomer can be one of the ends (single-file conformation), as in Fig. 3.1c, or any other monomer along the chain (hairpin conformation), as shown in Fig. 3.3c [49, 54]. During the translocation, monomers thread through the nanochannel from the cis side, where the chain originates, to the trans side, where the chain translocates to, as shown in Fig. 3.1d until the whole chain arrives at the trans side, as in Fig. 3.1e. In the rest of this chapter, we discuss how a non-uniform hydrodynamic flow affects the capture of a single polymer chain through the results obtained from MD-LB simulations.

3.1 Arrival time

The pressure jump at the x boundary generates the flow which induces a forced capture and translocation. We are studying the regime in which the thermal motion and flow force are comparable in the bulk (weakly-driven regime [73, 74]). The chain is initially located in the bulk and far from the nanopore and since the motion of the chain is weakly-driven, it gets the chance to diffuse in the fluid. Due to the thermal diffusion, the time required for the chain to reach the converging flow area, the area with significant velocity gradients, may differ between

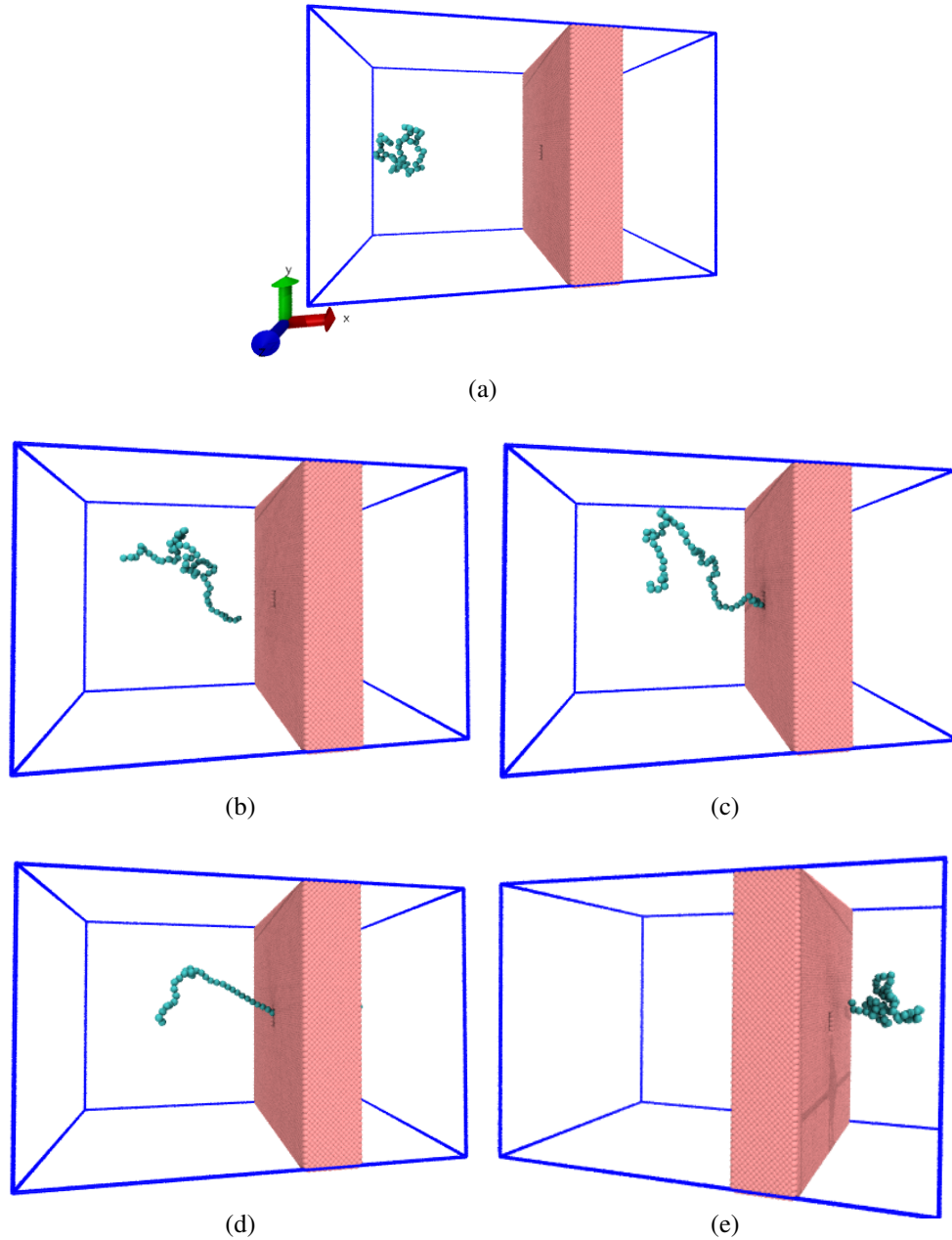


Figure 3.1: Within a certain radius from the nanopore, the polymer feels a stronger pull from the flow (a). In (b), an end reaches the nanopore and the translocation begins. In (c), a part of the chain travelled to the trans side while in (d), the translocation has successfully happened and the whole polymer is on the trans side.

different realizations. As a result, in spite of the drift-controlled regime in the converging area due to the stronger flow, we expect a distribution of centre of mass arrival times (t_a). Fig. 3.4 shows a histogram of arrival times of different realizations and the fitted distribution. The mean arrival time is 107 ns and the standard deviation of the distribution is $w_t = 36 \text{ ns} \pm 3 \text{ ns}$. If the chain were a rigid blob moving with the flow, we would expect a single arrival time equal to

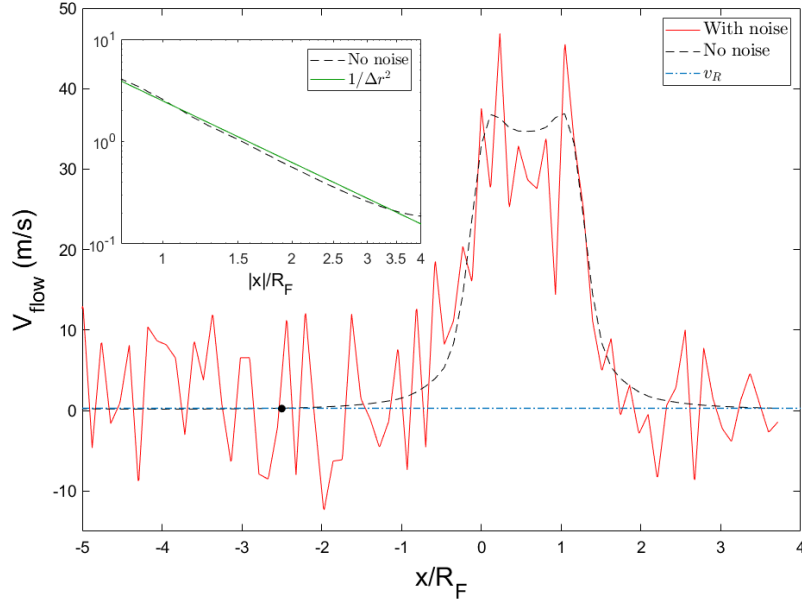


Figure 3.2: The velocity of the fluid is illustrated as a function of distance from the pore. The red curve is obtained from a system with thermal noise while the black one shows the velocity for a system without noise. The red curve follows the same pattern as the black one within error bounds. The relaxation velocity v_R is shown with a horizontal line (blue dash-dotted). The intersection of the flow velocity and the relaxation velocity is marked with a black circle. For velocities faster than v_R the chain does not have enough time to relax. The inset shows the velocity of the flow approaching the pore in a logarithmic scale. An inverse-square relation for the approaching flow is observed.

$$t_a \approx \int_0^{t_a} dt = \int_{x_{int}}^{x_{ent}} \frac{dx}{v(x)} = 94 \text{ ns} \quad (3.1)$$

where $v(x)$ is the flow velocity shown in Fig. 3.2, x_{int} is the initial position of the COM of the chain, and x_{ent} is the position of the entrance of the nanopore in the x-direction. As expected due to diffusion, the arrival time of a single rigid blob is slightly less than the measured mean arrival time obtained from our simulations. To further investigate the significance of diffusion in the system, we define a diffusion velocity based on a one-dimensional diffusion constant (D_0) which is found from the linear part of the measured mean-squared displacement of the chain, giving

$$v_R = \frac{R_F}{t_R} = \frac{2D}{R_F} = 0.21 \text{ nm/ns} \quad (3.2)$$

where R_F is the equilibrium radius of gyration and $t_R = R_F^2/2D$ is the characteristic time of the diffusive motion. When the velocity of the flow increases beyond v_R , the chain does not have time to relax and the motion of the chain becomes more like what we describe as a rigid blob motion. The distance from the nanopore where the flow velocity surpasses the diffusion velocity is shown in Fig. 3.2 by the intersection between the horizontal line showing v_R and the flow velocity curve. This distance is a characteristic quantity of the capture process which we shall define later in the analysis of the radius of gyration of the chain and discuss the impact of the shift in the flow regime on the polymer motion and shape.

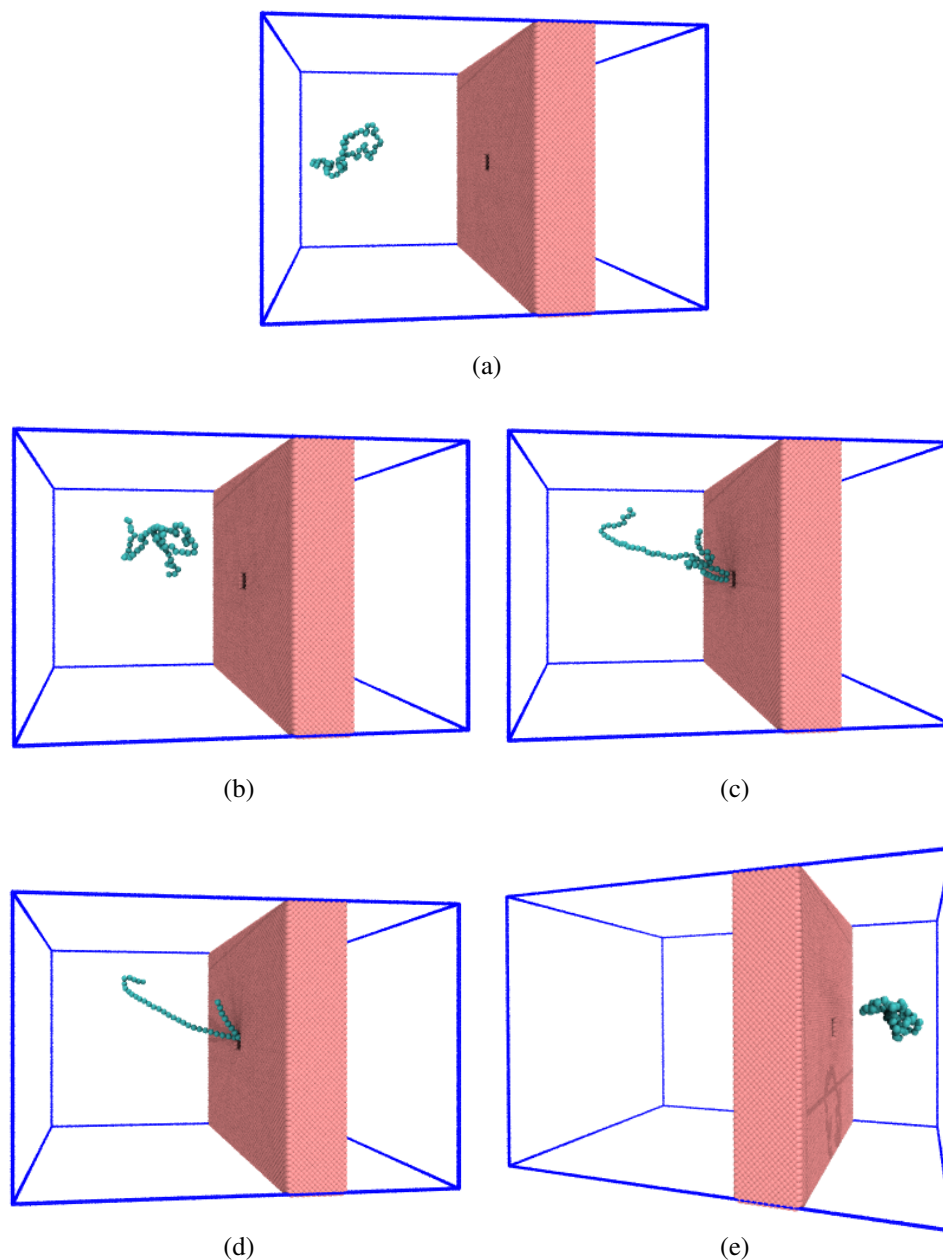


Figure 3.3: The journey of a chain which threads into the nanopore in hairpin conformation is illustrated through series of snapshots at different times.

To ensure that the randomness in the arrival time is not due to unsuccessful capture attempts, we monitored the motion of the chain as a function of time. Fig. 3.5 shows the position of the front monomer and the end monomers as a function of time. The front monomer wanders in the fluid for a reasonable amount of time with an almost constant velocity before it is pulled to the hole and its velocity rises. As mentioned below and shown in Fig. 3.5, the velocity of the chain dramatically increases when the chain arrives at the entrance and the chain is captured without any delay. The translocation stage happens comparably fast. The acceleration of the chain at different stages is obvious from the change in the slope of the graphs.

The extra distance that the chain wanders because of diffusion is denoted as w_x and is related to the standard deviation of the probability distribution of the diffusive motion,

$$P(\Delta x_{com}, t_a) = \frac{1}{(4\pi D_0 t_a)^{1/2}} \exp\left(-\frac{\Delta x_{com}^2}{4D_0 t_a}\right) \quad (3.3)$$

where Δx_{com} is the initial distance of the centre of mass of the chain from the entrance of the nanopore in the x-direction, and D_0 is the diffusion constant which obtained from MSD graph. Consequently, $w_x = \sqrt{2D_0 t_a}$. Equation 3.3 is a standard probability distribution of the motion of a Brownian particle in one dimension [75]. The time that the chain migrates w_x is $w_t = w_x/\bar{v}$ where \bar{v} is the average centre of mass speed and can be estimated by $\Delta x_{com}/\bar{t}_a$. By combining the equations above, we get an approximation for the standard deviation of the arrival time,

$$w_t \sim \frac{w_x \bar{t}_a}{\Delta x_{com}} = \sqrt{2D_0 \bar{t}_a} \left(\frac{\bar{t}_a}{\Delta x_{com}} \right). \quad (3.4)$$

As seen in Fig. 3.4, t_a is normally distributed with mean of 107 ns, and standard deviation of $w_t = 36 \text{ ns} \pm 3 \text{ ns}$. The measured w_t is in reasonable agreement with estimated value from equation 3.4, $w_t = 41 \text{ ns}$.

Comparing estimated diffusion (relaxation) time of the chain and the results from our simulations, one can see that the average arrival time $\bar{t}_a = 107 \text{ ns}$ is much longer than the $t_R = 41 \text{ ns}$ which implies that the chain had enough time to relax. Moreover, the standard deviation of the distribution is very close to the diffusion time which means that the chain wandering time is enough for the chain to explore its possible conformations. Based on this, we conclude that the motion of the chain is not completely drift-controlled and the chain experiences a balanced mix of diffusion and drift in the lead up to translocation.

3.2 Capture radius and chain extension

The discussions up to this point demonstrate that the motion of the chain far from the pore is balanced between diffusion and drift. We also showed that the non-uniform flow around the nanopore facilitates the delivery of the chain to the nanopore. However, the remaining question is how the converging flow helps the chain to find the pore. Seeking an answer for this question, we monitored the relative distance between the front monomer and centre of mass, and radius of gyration of the chain during the process in order to detect any deformations or extensions.

Fig. 3.6a shows how the radius of gyration of the chain changes as a function of the front monomer distance from the entrance of the nanopore (Δr). Both axes are scaled by the equilibrium radius of gyration. As can be seen, the average R_G fluctuates around the equilibrium value when it is far from the nanopore, $R_G/R_F \approx 1$.

Although the drag force is comparably weak in the bulk, it still creates a drift in the x-direction which slowly pushes the chain toward the pore. The drift-diffuse dynamics governs the motion until a part of the chain meets the high intensity flow near the pore ($\langle v_{flow} \rangle > v_R$). The flow in the converging area is stronger which pulls the leading segments out of the equilibrium blob that is mostly still in the uniform flow (see Fig. 3.1b and 3.3b). The chain, which is trying to go back to equilibrium, pulls the extended segments back toward the bulk

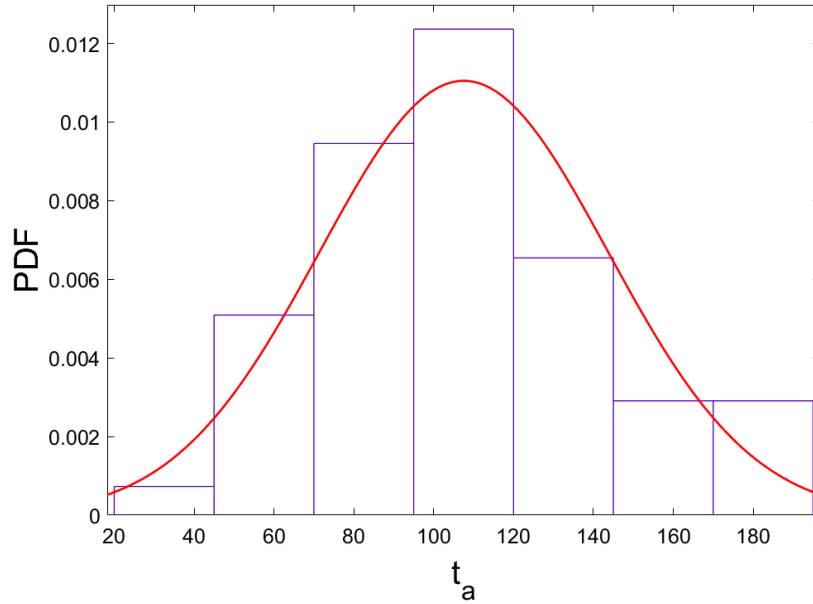


Figure 3.4: The arrival time of the COM of the chain is normally distributed from which we can conclude that the motion of the chain is a mix of diffusion and drift. The average arrival time is 107ns and the standard deviation equals 36ns .

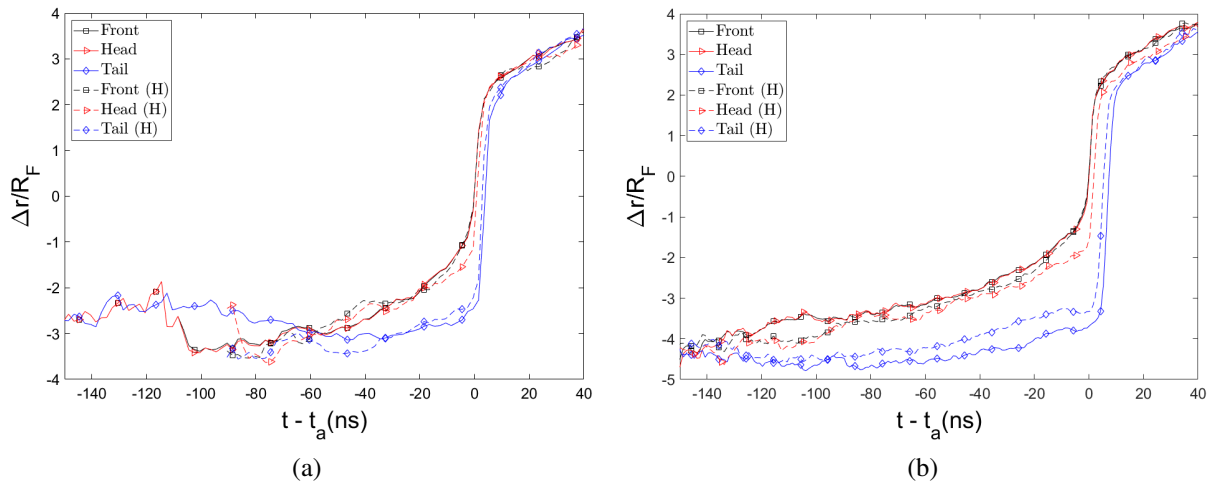
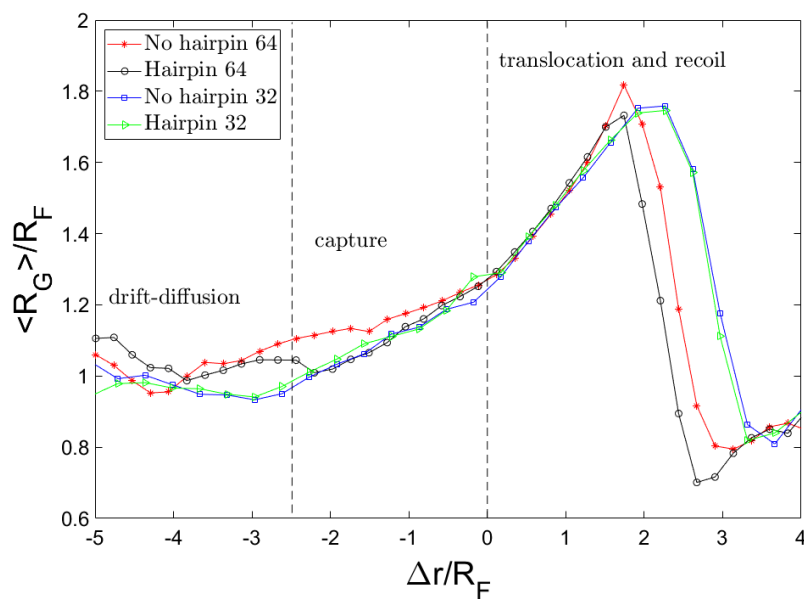
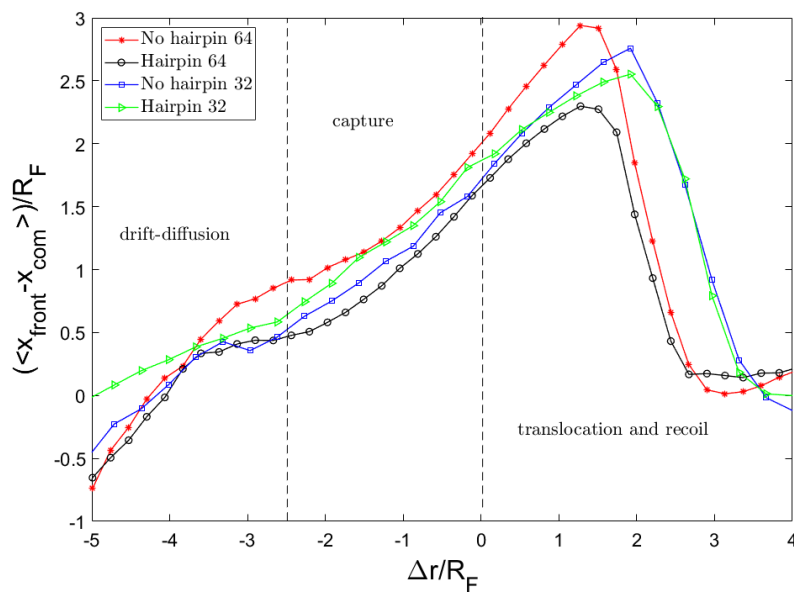


Figure 3.5: The position of the front monomer as a function of time for a chain with 32 monomers (a) and 64 monomers (b). The smaller gap between the curve of the front and tail monomer in the case of hairpin conformations (dotted lines) stems from the fact that the tension propagates faster along the chain because the strands are shorter than the whole backbone of the chain which is the length that tension must spread along for the single-file conformation.

while the flow drags the chain toward the nanopore. This competition does not last long before the whole chain stretches out toward the nanochannel. The region where this stretching takes place is labelled as capture region in Fig. 3.6. The weak drift in the flow direction brings the equilibrium blob closer to the strong flow area, and as soon as a segment feels the stronger flow, the recoiling force weakens.



(a)



(b)

Figure 3.6: (a) The chain in the bulk experiences 2 stages to arrive at the nanopore. In the first stage, the chain gets the opportunity to diffuse and equilibrate while the stronger flow deforms the chain in the second stage. The radius of gyration declines to a value less than equilibrium value as it gets out of the hole. (b) The graph shows the relative distance of the front monomer from the main body of the chain. The distance between the front and centre of mass increases and reaches a peak when the front monomer leaves the nanopore.

Fig. 3.6b shows the distance between the front monomer and the centre of mass of the chain as a function of front monomer position. The values on both axes are scaled by R_F . Tracking the distance between the front monomer and the centre of mass in the x-direction, one can see the change in the behaviour of this quantity when the front monomer gets closer than $r_c \approx 2.5R_F$ to the pore. r_c (capture radius) is the distance from the pore where the dynamics of the motion alters. In other words, the capture radius is defined to be the distance from the hole where the front segment accelerates and the chain begins to extend. The extension of the chain within r_c is clearly shown in Fig. 3.6. This coincides with the front segment entering the converging flow area where the velocity is higher than the v_R , as illustrated in Fig. 3.2.

The concept of capture radius was introduced by Muthukumar, and Grosberg and Rabin [73, 76]. Although the possibility of chain extension within the capture radius has been mentioned and mathematically formulated before [53], in most of the available models for the capture process, the whole chain accelerates toward the nanopore within the capture radius and arrives at the entrance as a jammed coil.

In contrast, our simulation shows that the tension created by the converging and accelerating flow deforms the chain as the leading segment enters this area. We expect a mechanism similar to the tension propagation formulated by Sakaue [44] to be responsible for the observed deformation. Since the velocity field points toward the hole, the flow guides the front segment to the entrance of the nanopore where the front monomer experiences the maximum force from the flow and gets sucked into the nanopore. The gradual increase of flow intensity makes it possible to have a no-barrier capture in a weakly-driven flow. Farahpour et al. obtained similar results in the case of an electric voltage-driven system, with this difference that their pore diameter was set so that no hairpin conformation can thread through the nanopore [55]. Both acceleration steps can be clearly observed as the change in the slope of the graphs in Fig. 3.6a. The first acceleration happens when the chain enters the converging flow area (drift-diffusion to capture) and the second one occurs when the chain arrives at the entrance (capture to translocation).

3.3 Pulley effect: Unravelling of folded conformation

Hairpin insertion has been reported as a common feature of driven polymer capture in literature [5, 49, 77]. The presence of a non-uniform force field is a major factor behind formation of hairpins. The hairpin forms when a monomer other than one of the ends becomes the leading monomer and experiences the stronger force of the converging flow first. Fig. 3.7 illustrates the probability of formation of different arrival conformations for polymers of length 32 and 64. The horizontal axis shows the relative position of the vertex of the hairpin ($h = 1$ is a single-file conformation and $h/N = 0.5$ is a polymer folded in half). Although the chain enters the nanopore in a single-file form most of the time, a considerable number of entries happen with hairpin conformations. Because the ends have more freedom compared to other monomers of the backbone, they have a higher chance of entering the high-velocity region before any other part of the chain and consequently, they are dragged to the nanopore as the front monomer. As can be seen in Fig. 3.7, the probability of hairpin capture decreases as the size of the hairpin increases and one might expect that the long hairpins (near 1/2 of chain length) must be very rare. However, as illustrated in Fig. 3.7 for the 64-mer and also captured in experimental

data as well [54], formation of long hairpins happens more often than medium ones. This can be explained by applying a basic force balance argument. When the strands of a hairpin are not of similar length, the time that it takes for the entire strand on the long side to feel the stretching force is longer. This results in a stronger pull-back force (with entropic origin) on the leading monomer in the longer strand's direction. Since the flow field is spatially symmetric, the hairpin unravels and the length of the shorter strand decreases even more. Due to the resemblance of such a motion to that of a rope on a pulley in a gravitational field, we call this phenomenon the "pulley effect". This unravelling continues as the translocation progresses. Fig. 3.8 shows a sequence of snapshot of our polymer chain in a realization in which the pulley effect is illustrated.

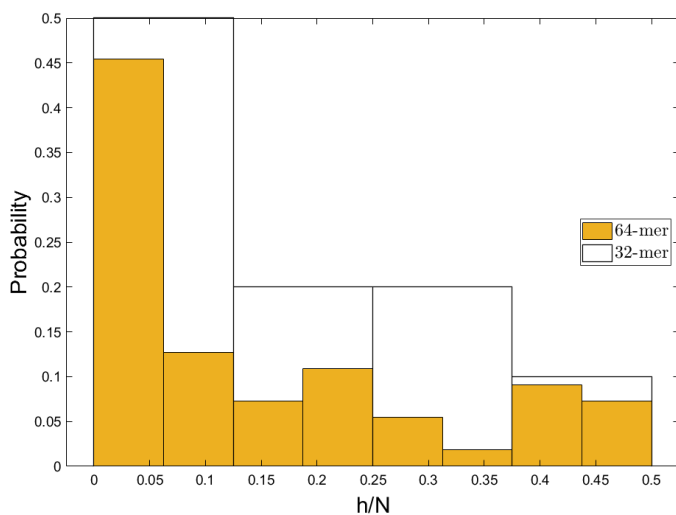


Figure 3.7: The most probable conformation of insertion is single-file. Due to chain unravelling and statistical uncertainty, the bin widths of less than 4 monomers would not demonstrate the essential information.

If the strands create comparable forces on the front monomer, as when the hairpin occurs halfway along the chain the symmetry of the conformation is preserved during capture and translocation. The effect of symmetry and similar tension propagation along symmetric hairpins has been discussed in a recent theoretical paper by Ghosh et al. [50]. Therefore, not only does the freedom of the ends increases the chance of single-file capture, but also the pulley effect promotes it as well.

Fig. 3.9a shows the location of the hairpin vertex along the chain for a chain of size 32 at three different stages. The first stage is when the leading monomer is at the capture radius (blue bars), the second stage is when the leading monomer is at the entrance of the nanopore, and the third stage is when the leading monomer is leaving the pore. We see that the leading monomer location after shifts from the middle of the chain towards the ends within the capture radius. The pulley effect continues to shorten the hairpin even during translocation. Although a similar pattern is observed for the 64-mer chain, as shown in Fig. 3.9b, there are some differences between 32-mer and 64-mer chains. For example, the probability of long hairpins ($0.4N < h$) stays virtually constant for the 64-mer while it drops in the case of 32-mer. In general, a consistent shift from longer to shorter can be observed for the 32-mer. However, for the 64-mer chain, it can be seen that the conformations close to half-chain hairpin tend to

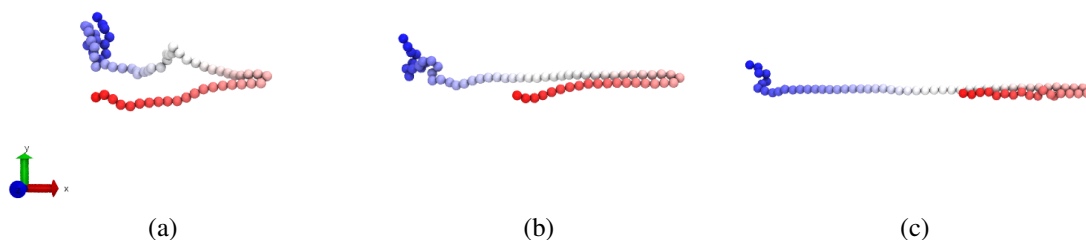
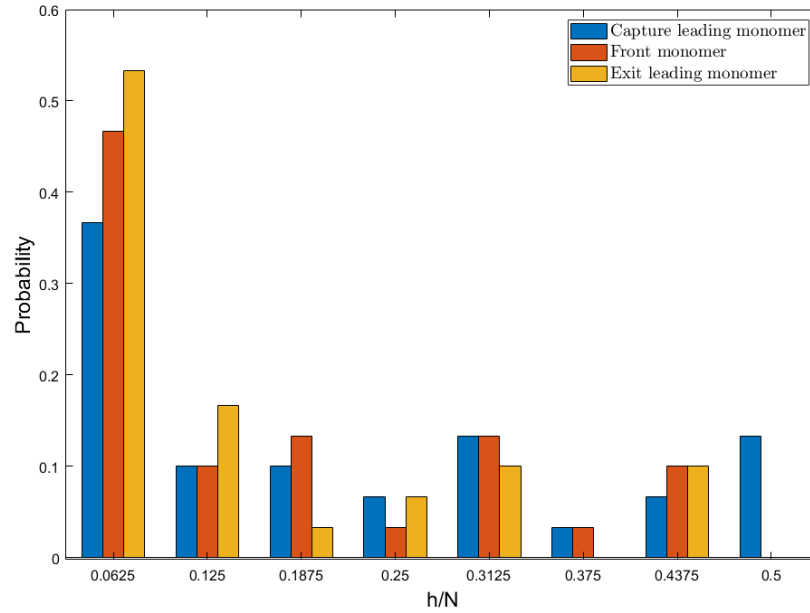


Figure 3.8: The faster motion of shorter strand of a hairpin conformation is shown in a sequence of snapshots. For illustration purposes, the wall (not shown) was wider in this figure. (a) shows the chain entering a pore of length 40 nm. The leading monomer is the twentieth bead. (b) shows another snapshot of the system when the chain is half-way through. The leading monomer is nineteenth bead. In (c), the chain is shown as it leaves the nanopore. The leading monomer is bead number sixteen.

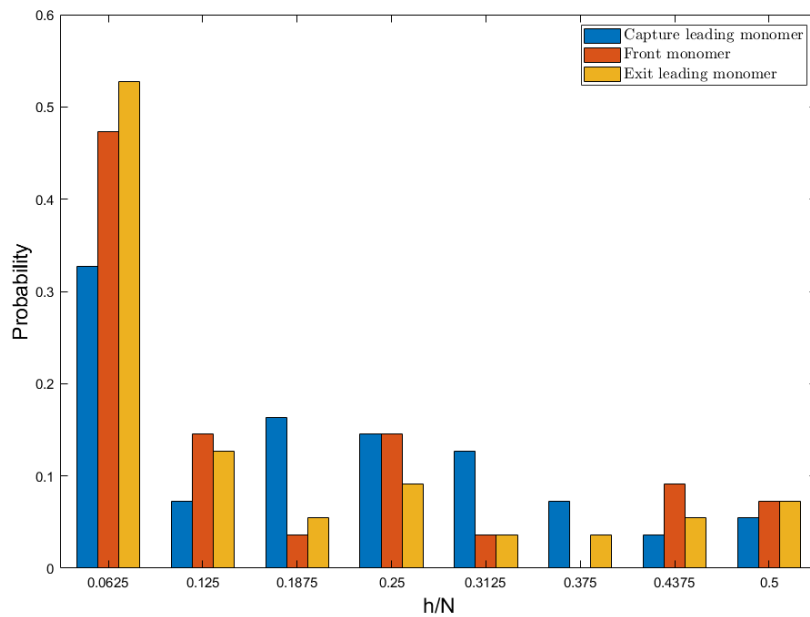
move toward this symmetric conformation. The medium-sized hairpins are still unusual with more than 40% for single-file capture. Since the pulley effect stems from the interplay between flow force exerted on the chain and the entropic forces, we believe that the desired pulley effect for a chain of a certain size can be obtained by modifying the flow field. This adjustment can be done through change of geometry or the strength of the flow field. This being said, more studies must be conducted on the matter to achieve concrete conclusions.

To get a better picture of how the converging flow field affects the dynamics of capture, we take a look at the velocity of the chain. Inspired by the idea behind the blob model [9], the chain is divided to segments and the velocity of different segments are plotted versus their position in Fig. 3.10. The segments are numbered from the end that enters the hole first (head) to the end that arrives the last (tail). Since the converging field has an inverse-square behaviour, Fig. 3.2, the velocity of a rigid sphere in such a field must follow the same rule as well. However, our polymer is not a rigid body and the collective dynamics of the chain results in some deviation from the inverse-square relation. In Fig. 3.10a, although the segments near the tail experience a steeper acceleration, the slope of the curves are fairly similar and close to -2 while in Fig. 3.10b, the difference between the curves is more significant which fits into the picture that ends feel the tension of the field sooner when a hairpin formed. In general, the segments close to the tail move faster within the capture radius because not only are they dragged along the flow but also pulled by the rest of the chain toward the pore. However, with the noise present in the data, it might be hard to draw any quantitative conclusions. It is worth mentioning that in the case of hairpin conformations, the velocity of the first segment increases faster than other first 3 segments which is a confirmation on the occurrence of the pulley effect (the shorter strand moves faster, therefore, the hairpin opens up in the flow direction).

While the front monomer is travelling from the cis side to the trans side, the tail segments (the four segments closer to the tail than head) might be in any conformation. They can still be diffusing in the bulk or feeling the tension on the backbone and moving toward the hole. Therefore, the tail segments can be in various states in different simulations when the front monomer is at a certain distance from the entrance of the nanopore and this is the reason for the lower peaks of the tail segments compared to the head segments in Fig. 3.11. One must keep in mind that the velocity peak in Fig. 3.11a and 3.11b is not corresponding to the maximum velocity of the segment but the most probable velocity of the segment when the front monomer



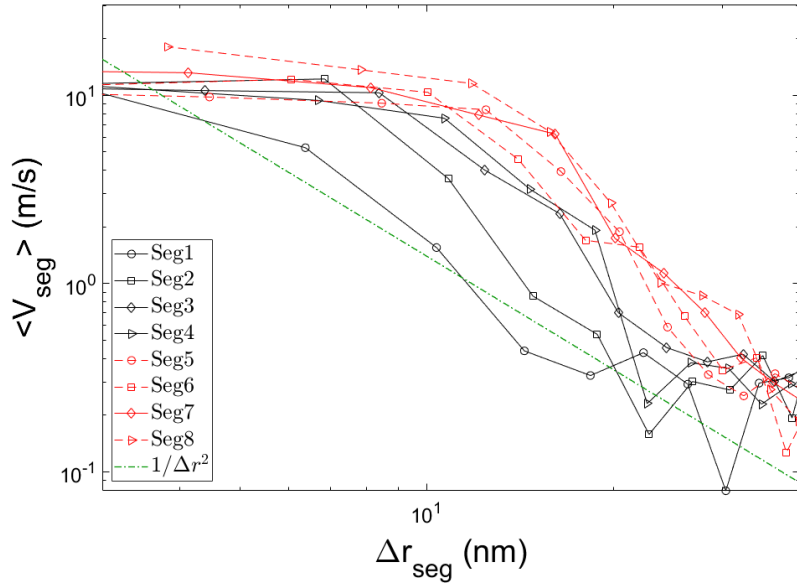
(a)



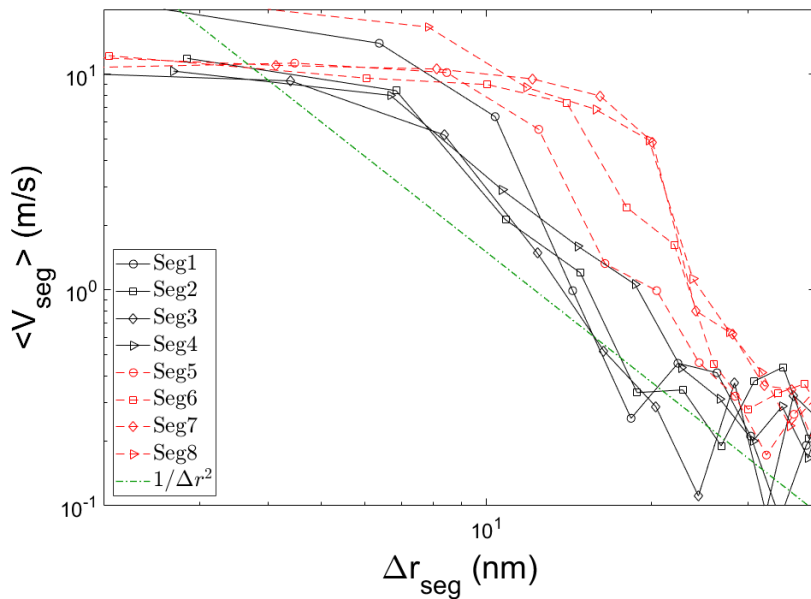
(b)

Figure 3.9: (a) shows the probability of the position of the hairpin vertex along the polymer backbone for a 32-mer chain and (b) shows the same property for a 64-mer chain. The x axis is scaled by the chain length. The medium-sized hairpins open up within the capture radius due to the pulley effect. The unravelling continues inside the nanopore. In fact, the stronger force in the pore speeds up the pulley effect. This is more obvious for longer nanopores.

is at that certain distance from the entrance of the nanochannel. For example, segment 8 can be inside the nanopore, or exiting the nanopore, or already out of the hole and relaxing when the front monomer is at $2R_F$. The velocity of the segment 8 can be totally different in any



(a)



(b)

Figure 3.10: (a) shows the velocity of polymer segments while travelling toward the nanopore for a single-file capture averaged over 32 realizations and (b) shows the same quantity for a hairpin capture averaged over 23 realizations. The polymer motion deviates from a perfect inverse-square relation due to the chain constraints. Because of the averaging steps included in obtaining both (a) and (b), we consider a cut-off distance (8nm from the origin) that beyond which the curves are not accurate enough.

of those states. If the segment is translocating through, its velocity will be high. However, if it is leaving the nanochannel, it probably feels a repulsion from the non-equilibrated jam of translocated monomers and decelerates. The dynamical behaviour would be completely different from the two previous possible states, if the tail has already left the nanopore and is

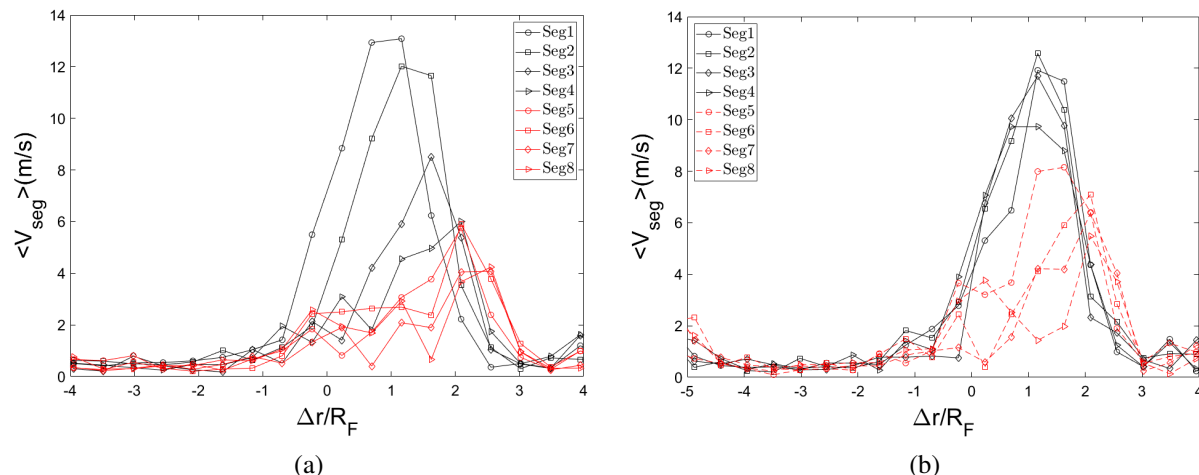


Figure 3.11: In the (a), the velocity of different segments of a chain that enters the nanopore in a single-file conformation has been shown while in the (b), the same feature has been shown for a chain with hairpin conformation. The tension simultaneously propagates along the both strands of the chain in the case of hairpin conformation and consequently, the velocity graphs for the first segments look alike.

equilibrating.

3.4 summary

The polymer capture and translocation in the presence of hydrodynamics is far from well-understood. In this work, we studied the effect of hydrodynamics on the capture process using hybrid molecular dynamics-lattice-Boltzmann simulation. We observed that a pressure-driven hydrodynamic flow guides the chain to the nanopore and makes a no-barrier capture possible. However, a variety of threading conformations were observed. Among those, the single-file capture was the most probable insertion conformation. This can be tied to the freedom of motion of the ends compared to the inner segments. Although this has an impact, it could not explain the other effects observed in simulations and experiments like the comparably high probability of long-hairpin (almost folded in half) capture. We observed an unravelling mechanism, namely the pulley effect, which shortens asymmetric hairpins toward single-file conformation.

Chapter 4

Conclusion

Polymer translocation as a process entangled with many biological phenomena has attracted researchers' interest from all over science and engineering. Despite all the effort made up to now, the capture process, which is a step before translocation, has not been understood well. This is the motivation behind this work. Having the privilege of accessing a multiscale simulation package which takes into account both the hydrodynamic interactions and thermal fluctuations, we observed that a weakly-driven hydrodynamic flow can facilitate the process of finding the pore by the polymer chain and the threading can happen without any necessity for overcoming an entropic barrier.

Comparing the arrival time obtained from our simulations and estimated values from mathematical models, we found that the non-uniform and converging flow near the pore speeds up the motion of the chain despite the fact that the motion of the chain is a balanced mix of diffusion and drift in the bulk. By studying the parameters related to the polymer's shape, we discovered that not only does the velocity of the chain increase as a whole but also the stronger flow causes extensions in the polymer.

Moreover, we investigated the possibility of the formation of hairpins and the effect of the extensions on this process. We observed that the single-file insertion is the most probable insertion conformation, as seen in experiments for DNA capture [54]. This can be associated with the greater freedom of the ends and the fact that they have better chance of entering the high velocity area and be guided to the pore. However, the freedom of the ends couldn't explain the large number of realizations in which long hairpins (almost half of the size of the chain) threaded through the channel. We found out that there is a mechanism, which we call the pulley effect, by which the hairpins with strands of considerably different sizes unravel in the favour of shortening the hairpin, but if the strands have comparable length, the chain keeps its conformation and threads through in hairpin shape. The unravelling due to the pulley effect makes single-file capture more probable which means that hydrodynamic flow can be used to promote single-file threading.

As discussed in section 3.2, a correlation between where the converging flow starts and the extension in the polymer chain was observed. In future work, we intend to study the effect of flow intensity on the capture radius and pulley effect. We also mentioned in section 3.3 that the unravelling of a hairpin due to pulley effect appears to depend highly on the asymmetry of the hairpin shape. Based on this, We suspect that the unravelling can be encouraged by asymmetric nanopores. An investigation on this theory and in general, the relation between geometry and

chain unravelling can be the next step toward obtaining a better picture of the pulley effect and capture process.

Bibliography

- [1] H. Staudinger. Über Polymerisation. *Berichte der deutschen chemischen Gesellschaft (A and B Series)*, 53(6):1073–1085, jun 1920.
- [2] Sean R. Carmody and Susan R. Wentz. mRNA nuclear export at a glance. *Journal of Cell Science*, 122(12):1933–1937, jun 2009.
- [3] Bruce Alberts, Alexander Johnson, Julian Lewis, David Morgan, Martin Raff, Keith Roberts, and Peter Walter. *Molecular Biology of the Cell: Chapter 9 Visualizing Cells*. GARLAND PUB - USA, 2017.
- [4] Arnold J. Storm, Cornelis Storm, Jianghua Chen, Henny Zandbergen, Jean François Joanny, and Cees Dekker. Fast DNA translocation through a solid-state nanopore. *Nano Letters*, 5(7):1193–1197, jul 2005.
- [5] Vladimir V. Palyulin, Tapio Ala-Nissila, and Ralf Metzler. Polymer translocation: The first two decades and the recent diversification. *Soft Matter*, 10(45):9016–9037, 2014.
- [6] M Rubinstein and R H Colby. *Polymer Physics*. OUP Oxford, 2003.
- [7] Murugappan Muthukumar. *Polymer translocation*. CRC Press, 2016.
- [8] M Kardar. *Statistical Physics of Particles*. Cambridge University Press, 2007.
- [9] P. G. de Gennes. *Scaling Concepts in Polymer Physics*. Cornell University Press, 1979.
- [10] Harold R. Warner. Kinetic Theory and Rheology of Dilute Suspensions of Finitely Extensible Dumbbells. *Industrial and Engineering Chemistry Fundamentals*, 11(3):379–387, aug 1972.
- [11] Kurt Kremer and Gary S. Grest. Dynamics of entangled linear polymer melts: A molecular-dynamics simulation. *The Journal of Chemical Physics*, 92(8):5057–5086, 1990.
- [12] M Doi and S F Edwards. *The Theory of Polymer Dynamics*. Comparative Pathobiology - Studies in the Postmodern Theory of Education. Clarendon Press, 1988.
- [13] S T Thornton and J B Marion. *Classical Dynamics of Particles and Systems*. Brooks/Cole, 2004.

- [14] John G. Kirkwood and Jacob Riseman. The intrinsic viscosities and diffusion constants of flexible macromolecules in solution. *The Journal of Chemical Physics*, 16(6):565–573, 1948.
- [15] Bruno H. Zimm. Dynamics of polymer molecules in dilute solution: Viscoelasticity, flow birefringence and dielectric loss. *The Journal of Chemical Physics*, 24(2):269–278, 1956.
- [16] M Daoud and P.G. De Gennes. Statistics of macromolecular solutions trapped in small pores. *Journal de Physique*, 38(1):85–93, 1977.
- [17] F. Brochard and P. G. De Gennes. Dynamics of confined polymer chains. *The Journal of Chemical Physics*, 67(1):52–56, 1977.
- [18] Amit Meller and Daniel Branton. Single molecule measurements of DNA transport through a nanopore. *Electrophoresis*, 23(16):2583–2591, aug 2002.
- [19] Amit Meller. Dynamics of polynucleotide transport through nanometre-scale pores. *Journal of Physics Condensed Matter*, 15(17):R581–R607, may 2003.
- [20] Peng Chen, Jiajun Gu, Eric Brandin, Young Rok Kim, Qiao Wang, and Daniel Branton. Probing single DNA molecule transport using fabricated nanopores. *Nano Letters*, 4(11):2293–2298, nov 2004.
- [21] John J. Kasianowicz, Eric Brandin, Daniel Branton, and David W. Deamer. Characterization of individual polynucleotide molecules using a membrane channel. *Proceedings of the National Academy of Sciences of the United States of America*, 93(24):13770–13773, 1996.
- [22] W. Sung and P. J. Park. Polymer translocation through a pore in a membrane. *Physical Review Letters*, 77(4):783–786, 1996.
- [23] L P Pitaevskii and E M Lifshitz. *Physical Kinetics*. Number v. 10. Elsevier Science, 2012.
- [24] Edmund A. Di Marzio and Arnold J. Mandell. Phase transition behavior of a linear macromolecule threading a membrane. *Journal of Chemical Physics*, 107(14):5510–5514, oct 1997.
- [25] M. Muthukumar. Polymer translocation through a hole. *Journal of Chemical Physics*, 111(22):10371–10374, dec 1999.
- [26] Jeffrey Chuang, Yacov Kantor, and Mehran Kardar. Anomalous dynamics of translocation. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, 65(1), 2002.
- [27] C. Carmesin and Kurt Kremer. The Bond Fluctuation Method: A New Effective Algorithm for the Dynamics of Polymers in All Spatial Dimensions. *Macromolecules*, 21(9):2819–2823, sep 1988.
- [28] Kaifu Luo, T. Ala-Nissila, and See Chen Ying. Polymer translocation through a nanopore: A two-dimensional Monte Carlo study. *Journal of Chemical Physics*, 124(3), 2006.

- [29] Steve Guillouzie and Gary W. Slater. Polymer translocation in the presence of excluded volume and explicit hydrodynamic interactions. *Physics Letters, Section A: General, Atomic and Solid State Physics*, 359(4):261–264, nov 2006.
- [30] M. G. Gauthier and G. W. Slater. Molecular dynamics simulation of a polymer chain translocating through a nanoscopic pore : HHydrodynamic interactions versus pore radius. *European Physical Journal E*, 25(1):17–23, jan 2008.
- [31] Ilkka Huopaniemi, Kaifu Luo, Tapio Ala-Nissila, and See Chen Ying. Langevin dynamics simulations of polymer translocation through nanopores. *Journal of Chemical Physics*, 125(12), 2006.
- [32] Kaifu Luo, Santtu T.T. Ollila, Ilkka Huopaniemi, Tapio Ala-Nissila, Pawel Pomorski, Mikko Karttunen, See Chen Ying, and Aniket Bhattacharya. Dynamical scaling exponents for polymer translocation through a nanopore. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 78(5), nov 2008.
- [33] V. V. Lehtola, R. P. Linna, and K. Kaski. Unforced polymer translocation compared to the forced case. *Physical Review E*, 81(3):031803, mar 2010.
- [34] Dongshan Wei, Wen Yang, Xigao Jin, and Qi Liao. Unforced translocation of a polymer chain through a nanopore: The solvent effect. *Journal of Chemical Physics*, 126(20), 2007.
- [35] Debabrata Panja and Gerard T. Barkema. Simulations of two-dimensional unbiased polymer translocation using the bond fluctuation model. *Journal of Chemical Physics*, 132(1), 2010.
- [36] Felix Kapahnke, Ulrich Schmidt, Dieter W. Heermann, and Matthias Weiss. Polymer translocation through a nanopore: The effect of solvent conditions. *Journal of Chemical Physics*, 132(16), apr 2010.
- [37] Hendrick W. De Haan and Gary W. Slater. Memory effects during the unbiased translocation of a polymer through a nanopore. *Journal of Chemical Physics*, 136(15), apr 2012.
- [38] Shyh Shi Chern, Alfredo E. Cárdenas, and Rob D. Coalson. Three-dimensional dynamic Monte Carlo simulations of driven polymer transport through a hole in a wall. *Journal of Chemical Physics*, 115(16):7772–7782, oct 2001.
- [39] Yacov Kantor and Mehran Kardar. Anomalous dynamics of forced translocation. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 69(2 1), feb 2004.
- [40] Pu Tian and Grant D Smith. Translocation of a polymer chain across a nanopore: A Brownian dynamics simulation study. *Journal of Chemical Physics*, 119(21):11475–11483, 2003.
- [41] Maria Fyta, Jayanta Sircar, Efthimios Kaxiras, Simone Melchionna, Massimo Bernaschi, and Sauro Succi. Parallel multiscale modeling of biopolymer dynamics with hydrodynamic correlations. *International Journal for Multiscale Computational Engineering*, 6(1):25–37, apr 2008.

- [42] Maria Fyta, Simone Melchionna, Sauro Succi, and Efthimios Kaxiras. Hydrodynamic correlations in the translocation of a biopolymer through a nanopore: Theory and multiscale simulations. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 78(3), 2008.
- [43] V. V. Lehtola, R. P. Linna, and K. Kaski. Dynamics of forced biopolymer translocation. *Epl*, 85(5), 2009.
- [44] Takahiro Sakaue. Nonequilibrium dynamics of polymer translocation and straightening. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 76(2), aug 2007.
- [45] Takuya Saito and Takahiro Sakaue. Erratum to: Dynamical diagram and scaling in polymer driven translocation. *The European Physical Journal E*, 35(11):135, 2012.
- [46] Chung Y. Kong and Murugappan Muthukumar. Modeling of polynucleotide translocation through protein pores and nanotubes. *Electrophoresis*, 23(16):2697–2703, aug 2002.
- [47] K. Luo, T. Ala-Nissila, S. C. Ying, and R. Metzler. Driven polymer translocation through nanopores: Slow-vs.-fast dynamics. *Epl*, 88(6), 2009.
- [48] A Yu Grosberg, S Nechaev, M Tamm, and O Vasilyev. How long does it take to pull an ideal polymer into a small hole? *Physical Review Letters*, 96(22), 2006.
- [49] Christopher Forrey and M. Muthukumar. Langevin dynamics simulations of ds-DNA translocation through synthetic nanopores. *Journal of Chemical Physics*, 127(1), 2007.
- [50] Bappa Ghosh, Jalal Sarabadani, Srabanti Chaudhury, and Tapio Ala-Nissila. Pulling a folded polymer through a nanopore. dec 2019.
- [51] Payam Rowghanian and Alexander Y. Grosberg. Force-driven polymer translocation through a nanopore: An old problem revisited. *Journal of Physical Chemistry B*, 115(48):14127–14135, dec 2011.
- [52] T. Ikonen, A. Bhattacharya, T. Ala-Nissila, and W. Sung. Unifying model of driven polymer translocation. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 85(5), may 2012.
- [53] C. T.A. Wong and M. Muthukumar. Polymer capture by electro-osmotic flow of oppositely charged nanopores. *Journal of Chemical Physics*, 126(16), 2007.
- [54] Mirna Mihovilovic, Nicholas Hagerty, and Derek Stein. Statistics of DNA capture by a solid-state nanopore. *Physical Review Letters*, 110(2), jan 2013.
- [55] Farnoush Farahpour, Azadeh Maleknejad, Fathollah Varnik, and Mohammad Reza Eftehadi. Chain deformation in translocation phenomena. *Soft Matter*, 9(9):2750–2759, mar 2013.
- [56] M Tuckerman. *Statistical Mechanics: Theory and Molecular Simulation*. Oxford Graduate Texts. OUP Oxford, 2010.

- [57] A A Mohamad. *Lattice Boltzmann Method: Fundamentals and Engineering Applications with Computer Codes*. SpringerLink : Bücher. Springer London, 2011.
- [58] K Huang. *Statistical mechanics*. Wiley, 1987.
- [59] F Reif. *Fundamentals of Statistical and Thermal Physics*. Waveland Press, 2009.
- [60] D. Arumuga Perumal and Anoop K. Dass. A Review on the development of lattice Boltzmann computation of macro fluid flows and heat transfer, dec 2015.
- [61] J. Hardy, Y. Pomeau, and O. De Pazzis. Time evolution of a two-dimensional classical lattice system. *Physical Review Letters*, 31(5):276–279, 1973.
- [62] U. Frisch, B. Hasslacher, and Y. Pomeau. Lattice-gas automata for the Navier-Stokes equation. *Physical Review Letters*, 56(14):1505–1508, 1986.
- [63] D A Wolf-Gladrow. *Lattice-Gas Cellular Automata and Lattice Boltzmann Models: An Introduction*. Lecture Notes in Mathematics. Springer Berlin Heidelberg, 2000.
- [64] Guy R. McNamara and Gianluigi Zanetti. Use of the boltzmann equation to simulate lattice-gas automata. *Physical Review Letters*, 61(20):2332–2335, 1988.
- [65] P. L. Bhatnagar, E P Gross, and M. Krook. A model for collision processes in gases. I. Small amplitude processes in charged and neutral one-component systems. *Physical Review*, 94(3):511–525, 1954.
- [66] B. Manz, L. F. Gladden, and P. B. Warren. Flow and dispersion in porous media: Lattice-Boltzmann and N M R studies. *AIChE Journal*, 45(9):1845–1854, 1999.
- [67] R. Adhikari, K. Stratford, M. E. Cates, and A. J. Wagner. Fluctuating lattice Boltzmann. *Europhysics Letters*, 71(3):473–479, 2005.
- [68] Santtu T.T. Ollila, Colin Denniston, Mikko Karttunen, and Tapio Ala-Nissila. Fluctuating lattice-Boltzmann model for complex fluids. *Journal of Chemical Physics*, 134(6):64902, 2011.
- [69] Steve Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics*, 117(1):1–19, 1995.
- [70] F. E. Mackay, S. T.T. Ollila, and C. Denniston. Hydrodynamic forces implemented into LAMMPS through a lattice-Boltzmann fluid. *Computer Physics Communications*, 184(8):2021–2031, 2013.
- [71] Charles S. Peskin. Flow patterns around heart valves: A numerical method. *Journal of Computational Physics*, 10(2):252–271, oct 1972.
- [72] Santtu T.T. Ollila, Christopher J. Smith, Tapio Ala-Nissila, and Colin Denniston. The hydrodynamic radius of particles in the hybrid lattice boltzmann-molecular dynamics method. *Multiscale Modeling and Simulation*, 11(1):213–243, 2013.

- [73] M. Muthukumar. Theory of capture rate in polymer translocation. *Journal of Chemical Physics*, 132(19), may 2010.
- [74] Takahiro Sakaue. Dynamics of polymer translocation: A short review with an introduction of weakly-driven regime, 2016.
- [75] Ryogo. Kubo, Morikazu. Toda, and Natsuki. Hashitsume. *Statistical Physics II. Nonequilibrium Statistical Mechanics*, volume 31. Springer Berlin Heidelberg, 1985.
- [76] Alexander Y. Grosberg and Yitzhak Rabin. DNA capture into a nanopore: Interplay of diffusion and electrohydrodynamics. *Journal of Chemical Physics*, 133(16), oct 2010.
- [77] A. Baumgärtner. Insertion and hairpin formation of membrane proteins: A Monte Carlo study. *Biophysical Journal*, 71(3):1248–1255, 1996.

Curriculum Vitae

Name: Navid Afrasiabian

Post-Secondary Education and Degrees: Isfahan University of Technology
Isfahan, Iran
2013 - 2018 BSc

University of Western Ontario
London, ON
2018 - 2020 MSc

Honours and Awards: Western Graduate Research Scholarship
2018 - 2020

Graduate Teaching Assistant Award
London, ON
2018 - 2020

Related Work Experience: Teaching Assistant
The University of Western Ontario
2018 - 2020