Polymer Confinement and Translocation

Chiu Tai Andrew Wong

University of Massachusetts Amherst, ctawong@gmail.com

Follow this and additional works at: https://scholarworks.umass.edu/open_access_dissertations

Part of the Polymer and Organic Materials Commons

Recommended Citation

Wong, Chiu Tai Andrew, "Polymer Confinement and Translocation" (2009). Open Access Dissertations. 50.
https://doi.org/10.7275/bbw5-6j38 https://scholarworks.umass.edu/open_access_dissertations/50

This Open Access Dissertation is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Open Access Dissertations by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
POLYMER CONFINEMENT AND TRANSLOCATION

A Dissertation Presented
by
CHIU TAI ANDREW WONG

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of
DOCTOR OF PHILOSOPHY
February 2009
Polymer Science and Engineering
POLYMER CONFINEMENT AND TRANSLOCATION

A Dissertation Presented

by

CHIU TAI ANDREW WONG

Approved as to style and content by:

______________________________
Murugappan Muthukumar, Chair

______________________________
Gregory N. Tew, Member

______________________________
Jonathan Machta, Member

______________________________
Anthony Dinsmore, Member

______________________________
Narayanan Menon, Member

______________________________
Shaw Ling Hsu, Department Chair
Polymer Science and Engineering
To my Wife (Ka Man Cheung), parents (Yin Kwai Cheung and Kim Chi Wong) and Brother (Chiu Lam Kevin Wong) for their unconditional support. To my Son (Hiu Ching Sage Wong) for the happiness he has brought us. To God for his blessings throughout all these years.
ACKNOWLEDGMENTS

I am most thankful to my thesis advisor, Prof. Murugappan Muthukumar, for the opportunity of graduate study under him and the freedom he had given me in pursuing research problems. I have been very impressed by and have learned a lot from his physical insights, his ability in picking important problems and his taste in solving problems elegantly. I would also like to thank Muthu for providing the best resources to me for conducting research. Under his guidance, I was encouraged to do theoretical calculations, computer simulations and experiments.

Special thanks are due to Prof. Po-Zen Wong, who encouraged me to apply for this graduate program when I was in Hong Kong and knew nothing about polymers. I am very grateful to my wife Mandy, who sacrificed so much and came all the way from Hong Kong to accompany me on my graduate studies. I cannot imagine how I could go through everything in the past years without her.

I am very pleased to thank my committee members, Prof. Gregory N. Tew, Prof. Jonathan Machta, Prof. Anthony Dinsmore and Prof. Narayanan Menon, for their insightful comments and their patience during all the meetings.

I am indebted to Ryan Murphy, who spent a lot of time training me to do the translocation experiments, and Joey Kong, who helped me get started with the Fokker-Planck equation.

I would like to thank the members in Muthu group, Mark, Chris, Rajeev, Jess, Deniz C, Deniz K, Xiangdong, J.P., Jianing, Zhaoyang, Alexander, Arnout, Steve, Nivas, Arindam, Erica, Ajay, Abhijit, Joe, Sonoko, Joey Kong, Joanna, Chilun, Vlad, Peng Bo, Ryan and Jiafang, for fun times and academic discussions.
Finally, I would like to thank the Chinese community in Amherst and the brothers and sisters in the Pioneer Valley Church of Christ. Their friendships and supports throughout the years mean a lot to me.
ABSTRACT

POLYMER CONFINEMENT AND TRANSLOCATION

FEBRUARY 2009

CHIU TAI ANDREW WONG
B.S., CHINESE UNIVERSITY OF HONG KONG
M.Phil., CHINESE UNIVERSITY OF HONG KONG
M.S., UNIVERSITY OF MASSACHUSETTS AMHERST
Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Murugappan Muthukumar

Single polymer passage through geometrically confined regions is ubiquitous in biology. Recent technological advances have made the direct study of its dynamics possible. We studied the capture of DNA molecules by the electroosmotic flow of a nanopore induced by its surface charge under an applied electric field. We showed theoretically that the DNA molecules underwent coil-stretch transitions at a critical radius around the nanopore and the transition assisted the polymer passage through the pore. To understand how a polymer worms through a narrow channel, we investigated the translocation dynamics of a Gaussian chain between two compartments connected with a cylindrical channel. The number of segments inside the channel changed throughout the translocation process according to the overall free energy of the chain. We found a change in the entropic driving force near the end of the process due to the partitioning of the chain end into the channel rather than the initial
compartment. We also developed a theory to account for the electrophoretic mobility of DNA molecules passing through periodic confined regions. We showed that the decrease in the translocation time with the molecular weight was due to the propensity of hairpin entries into the confined regions. To further explore the dynamics of polymer translocation through nanopores, we performed experimental studies of sodium polystyrene sulfonate translocation through α-hemolysin protein nanopores. By changing the polymer-pore interaction using different pH conditions, we identified the physical origins of the three most common event types. We showed that increasing the polymer-pore attraction increased the probability of successful translocation. Motivated by understanding the dynamics of a polymer in a crowded environment, we investigated the dynamics of a chain inside a one dimensional array of periodic cavities. In our theory, the chain occupied different number of cavities according to its confinement free energy which consisted of entropic and excluded volume parts. By assuming that the chain moved cooperatively, the diffusion constant exhibited Rouse dynamics. Finally, we performed computer simulations of a chain inside a spherical cavity. We found that the confinement effect was best described by the hard sphere chain model. We further studied the escape dynamics of the chain out of the cavity through a small hole. The equilibrium condition of the chain during the escape was discussed.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xiii</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1. GENERAL INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2. POLYMER CAPTURE BY ELECTROOSMOTIC FLOW OF OPPOSITELY CHARGED NANOPORES</td>
<td>3</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>3</td>
</tr>
<tr>
<td>2.2 Electroosmotic flow inside a nanopore</td>
<td>6</td>
</tr>
<tr>
<td>2.3 Polymer capture by EOF in the nanopore</td>
<td>8</td>
</tr>
<tr>
<td>2.3.1 Absorption time estimation</td>
<td>8</td>
</tr>
<tr>
<td>2.3.2 Coil-stretch transition</td>
<td>10</td>
</tr>
<tr>
<td>2.3.3 Estimation of the absorption radius $R_c$</td>
<td>10</td>
</tr>
<tr>
<td>2.4 Diffusion limited translocation rate</td>
<td>11</td>
</tr>
<tr>
<td>2.5 Summary and discussions</td>
<td>13</td>
</tr>
<tr>
<td>3. POLYMER TRANSLOCATION THROUGH A CYLINDRICAL CHANNEL</td>
<td>22</td>
</tr>
<tr>
<td>3.1 Theory</td>
<td>23</td>
</tr>
<tr>
<td>3.1.1 Polymer confined in a sphere</td>
<td>24</td>
</tr>
<tr>
<td>3.1.2 Polymer confined in a cylinder</td>
<td>25</td>
</tr>
<tr>
<td>3.1.3 Free energy of confinement</td>
<td>27</td>
</tr>
</tbody>
</table>
5.4 Stochastic model of polymer translocation through αHL .......... 86
   5.4.1 Formalism .............................................. 86
   5.4.2 Results .................................................. 89
5.5 Conclusions and discussions ........................................ 90

6. POLYMER DIFFUSION THROUGH PERIODIC CAVITIES .................... 112
   6.1 Introduction .................................................. 112
   6.2 Confinement free energy ...................................... 113
   6.3 Translocation time ............................................ 116
      6.3.1 1-2-1 translocation ..................................... 116
      6.3.2 2-3-2 translocation ..................................... 119
      6.3.3 m-(m+1)-m translocation ................................ 120
   6.4 Most probable occupation number ................................. 121
   6.5 Most probable translocation time ............................... 121
   6.6 Discussions and conclusion .................................... 122

7. POLYMER CONFINEMENT AND ESCAPE FROM A SPHERICAL CAVITY .......... 131
   7.1 Introduction .................................................. 131
   7.2 Confinement of a hard sphere chain inside a spherical cavity .... 133
   7.3 Pressure exerted by the chain on the cavity .................... 134
   7.4 Polymer escape from a cavity through a small hole ............... 135
   7.5 Escape of a gaussian chain ................................... 136
   7.6 Discussions and conclusions .................................... 137

8. GENERAL CONCLUSION ................................................. 144

APPENDIX: DERIVATIONS OF THE MEAN FIRST PASSAGE TIMES ................. 147
   A.1 General solution of the MFPT equation .......................... 149
   A.2 One absorbing and one reflecting boundaries ..................... 150
   A.3 Two absorbing boundaries ..................................... 151
   A.4 Conditional exit probabilities and mean exit times in two adsorbing boundaries ............................................. 151

BIBLIOGRAPHY .......................................................... 155
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>The reversal voltages ($V_R$) measured in pH 7.5 1 M KCl (cis) and 200 mM KCl with variable pH (trans). The permeability ratio $P_+/P_-$ were calculated using Eq. (5.4).</td>
<td>94</td>
</tr>
</tbody>
</table>
**LIST OF FIGURES**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 A cylindrical nanopore with radius (a) and positive surface potential (\Phi_0).</td>
<td>17</td>
</tr>
<tr>
<td>2.2 EOF in a nanopore. (a) Velocity profile at different monovalent ((z = 1)) salt concentrations. (b) Strength of the electroosmotic flux as a function of salt concentration. See Eqs. ((2.10)) and ((2.11)) for the definition of (\beta). (\Phi_0 = 0.01V, a = 7.5nm, E = 5 \times 10^7V/m, T = 300K) and (\eta = 1cP) are used in both graphs.</td>
<td>18</td>
</tr>
<tr>
<td>2.3 If the surface charge of the nanopore is opposite to the charge of the polymer molecules, EOF is in the same direction as the polymer translocation, assisting the translocation process.</td>
<td>19</td>
</tr>
<tr>
<td>2.4 When the polymer passes the absorbing radius (R_C) given by Eq. ((2.18)), the fluid velocity gradient is larger than the critical velocity gradient. The polymer is unable to relax and coil-stretch transition occurs.</td>
<td>20</td>
</tr>
<tr>
<td>2.5 Concentration profile at different (\alpha) given by Eq. ((2.22)).</td>
<td>21</td>
</tr>
<tr>
<td>3.1 All possible conformations of a polymer of (N) segments during translocation. (p) and (q) are defined as the number of segments inside the channel and the receptor compartment, respectively.</td>
<td>42</td>
</tr>
<tr>
<td>3.2 The free energy profiles of translocation of a chain with length (N = 300) and donor compartment radius (R_1 = 30). The translocation coordinate (q) denotes the number of segments inside the receptor compartment. The various free energy labels (F) in ((a)) are defined in Eqs. ((3.12), (3.13), (3.31), (3.32), (3.20), (3.21)) and ((3.22)). ((a)) (M = 30, R_2 = 30) and (a = 3). ((b)) (M = 30, R_2 = 60) and (a = 3). ((c)) (M = 30, R_2 = 60) and (a = 5). ((d)) (M = 45, R_2 = 60) and (a = 3).</td>
<td>43</td>
</tr>
</tbody>
</table>
3.3 The free energy barrier $F^+$ of channel-filling as a function of different parameters. (a) $M = 30$, $R_1 = R_2 = 30$ and $N = 300$. (b) $a = 5$, $R_1 = R_2 = 30$ and $N = 300$. (c) $a = 5$, $R_2 = 30$, $M = 30$ and $N = 300$. (d) $a = 3$, $R_1 = 10$, $R_2 = 15$ and $M = 6$.

3.4 Total translocation time $\langle \tau_1 \rangle + \langle \tau_2 \rangle$. (a) $a = 3$, $R_1 = 30$, $N = 300$. Solid lines: $R_2 = 30, 40, 50$ and 60. Dotted line: $\langle \tau_1 \rangle$ for $R_2 = 60$. Dashed line: $\langle \tau_2 \rangle$ for $R_2 = 60$. (b) $N = 300$, $a = 5$, $R_1 = 30$, $R_2 = 30, 40, 50$ and 60. (c) $N = 300$, $R_1 = 30$, $R_2 = 50$, $a = 3, 4, 5$ and 6. (d) $R_1 = 10$, $R_2 = 15$, $M = 6$, $a = 3, 4$ and 5.

3.5 Translocation time distribution of $\tau_2$ for $a = 3, 4, 5$, respectively.

3.6 (a) Free energy profiles of translocation involve crossing a free energy barrier $\Delta F$ (step (1)) and an arbitrary profile $G(x)$ (step (2)). The corresponding correction factors to $\langle \tau_1 \rangle$ (in the limits explained in the text) are (b) $C = N$, (c) $C = 1/|\mu|$, (d) $C = \exp(\mu N)/\mu$, (e) $C = \exp(\mu_1 x_m)(1/\mu_1 + 1/|\mu_2|)$ and (f) $C = x_m + 1/|\mu|$, respectively.

4.1 The DNA molecule starts at the beginning of the deep region as shown in (a). Driven by the electric field $E$ (the arrow indicates the direction of the movement of the chain), it takes time $\tau_1$ to migrate to the end of the deep region as shown in (b). The polymer molecule then takes $\tau_2$ to enter the shallow region, as sketched in (c). Once the polymer molecule enters the shallow region successfully, it takes $\tau_3$ to finish passing through the shallow region, as shown in (d).

4.2 Two modes of translocation (from the top view of geometry in Fig. 1). (a) For a linear translocation, the chain enters the shallow region with a chain end, forming $n$ blobs in a linear series, with $m$ segments in the shallow region and a tail of $N - m$ segments in the deep region. (b) For a hairpin translocation, the chain enters the shallow region with a hairpin, with $m$ segments in the shallow region and two tails of $j$ and $N - m - j$ segments in the deep region. In the deep region, there are two linear chains of blobs, each with $n/2$ blobs.

4.3 Typical free energy profile of translocation as a function of the number of segments in the shallow region $m$. $m^*$ is the critical number of segments in the shallow region beyond which the translocation process is favorable.
4.4 A hairpin entry with $m^*_h < m < m^*_h$ and only one tail in the deep region leads to a linear translocation. The chain enters the shallow region as a hairpin but $m$ does not exceed the hairpin critical insertion size $m^*_h$. The chain end forming the hairpin has no entropic penalty to prevent itself from proceeding into the shallow region and gaining electrostatic energy. If $m$ is larger than the linear critical insertion length $m^*_l$, linear translocation occurs.

4.5 (a) Relative probabilities of different translocation modes. The probability of hairpin translocation increases with $N$. $D = 1.5$ and $f = 0.08$ are used. (b) Probability of linear translocations (Eq. (4.24)) decreases with increasing $N$ and $f$. $D = 1.5$, $N = 30$, 60, 90 and 120 respectively. Solid lines are calculated using approximations Eqs. (4.14), (4.18) and (4.23). Dotted lines are calculated using exact formulas. The approximation works well for large $N$.

4.6 (a) Average translocation time $\langle \tau \rangle$ exhibits a crossover from linear to hairpin translocations as $N$ increases for $D = 1.2$, while this feature is missing for wider gaps. Exact formulas with $f = 0.1$ are used. (b) $\langle \tau \rangle$ as a function inverse electric force $1/f$ with $N = 100$. The exact formulas are used.

5.1 The $\alpha$HL protein nanopore were embedded onto a lipid bilayer in a 1M KCl solution. Charged polymers were driven by the applied voltage to pass through the nanopore.


5.3 Two possible models for translocation dynamics through $\alpha$HL, the entropic barrier model and the enthalpic barrier model. Both entropy and enthalpy may be responsible for the trapping of polymer inside the vestibule.

5.4 The ionic current ($I$) is plotted against the applied voltage ($V$) for $\alpha$HL pores in 1M KCl solution with different pH values on the cis and trans sides. Open and solid symbols denote symmetric and asymmetric solution, respectively.

5.5 The equilibrium net charges of the 1.5 nm constriction and the charge ring at the end of the $\beta$-barrel of $\alpha$HL.
5.6 The three most prominent event types (event 1, event 12 and event 2) in NaPSS translocation through αHL. The blockages were produced by 57.5 kg/mol NaPSS in 1M KCl pH 7.5 solution under 150 mV. ......................... 100

5.7 Typical ionic current traces in trans-pH 7.5 and 4.5, respectively, for 57.5 kg/mol NaPSS under 160 mV. The cis-pH were 7.5 in both cases. ................................. 101

5.8 The event diagrams and the percentages of event types for pH 7.5 and 4.5 on the trans side of the pore. The pH on the cis side was always 7.5. (a) and (b): Event diagrams, the normalized current blockage against the event duration, of 16 kg/mol NaPSS under 140 mV in trans-pH 7.5 and 4.5 solution respectively. 1000 events are plotted in both diagrams. The levels 1 and 2 blockages of events 12 are treated separately and denoted as events 12a and 12b, respectively. (c) and (d): Percentages of event types under the same conditions as in (a) and (b), respectively, at different voltages. ................................. 102

5.9 The average event duration is plotted against the molecular weight \(N\) divided by the applied voltage \(V\). Both \(\langle \tau_{12b} \rangle\) and \(\langle \tau_2 \rangle\) obeyed the same linear relation against \(N/V\), while \(\langle \tau_{12a} \rangle\) was insensitive to the changes in \(N\) and \(V\). ......................... 103

5.10 The threshold voltage of successful translocation for 16 kg/mol NaPSS in cis-pH 7.5 decreased as trans-pH deceased. The error bars are smaller than the symbols. ................................. 104

5.11 The average blockage times of 16 kg/mol NaPSS at 140 mV changed with trans-pH. The durations of level 1 blockages, \(\langle \tau_1 \rangle\) and \(\langle \tau_{12a} \rangle\), decreased with decreasing pH while the durations of level 2 blockages, \(\langle \tau_2 \rangle\) and \(\langle \tau_{12b} \rangle\), increased with decreasing pH. ......................... 105

5.12 The strength of the electro-osmotic flow as a function of the pH value of the trans-side solution. The cis-pH was kept at 7.5. ................. 106

5.13 Schematic free energy profile for a short NaPSS molecule translocating through αHL. With two absorbing boundaries, exiting on the left and right represent unsuccessful and successful translocations, respectively. ................................. 107

5.14 Free energy profiles of translocation with different polymer-pore interaction parameter \(\epsilon\). \(\delta = 1, \Delta F_1 = 10, \Delta F_2 = 5, M = 10, N = 50\) and \(qV = -1\). ......................... 108
5.15 Probabilities of successful translocation. The scatter points are the percentage of the sum of events 12 and 2 at 140 mV. The lines are evaluated using Eq. (5.9) with $\delta = 1, \Delta F_1 = 8, \Delta F_2 = 2, M = 17, N = 69, V = 140$ mV and $q = -0.3e$.

5.16 The average successful translocation time $\langle \tau_t \rangle$, the average unsuccessful translocation time $\langle \tau_u \rangle$ and the successful translocation probability $P_t$. $\delta = 1, \Delta F_1 = 8, \Delta F_2 = 2, M = 17, N = 69, V = 140$ mV and $q = -0.3e$ are used.

5.17 Voltage dependence of $\langle \tau_1 \rangle$ for NaPSS with different chain lengths. The cis-pH and trans-pH were 7.5 and 4.5, respectively.

6.1 Three common conformations of a chain confined in a spherical cavity. Left: Both of the chain ends are free to move in the cavity. The free energy of the confined chain is $F_S$. Center: One end is free to move and the other end is anchored near the surface with anchoring distance $c$. The free energy is $F_{S1}$. Right: Both ends are fixed near the surface exactly opposite to each other, with anchoring distances $c_1$ and $c_2$ respectively. The free energy is $F_{S2}$.

6.2 Anchoring free energies calculated from the Langevin dynamics simulations of a $N$-segment chain confined inside a spherical cavity with radius $R = 4.5d$, where $d$ is the diameter of the beads. Chains with and without excluded volume interactions were used. $\Delta F_a/k_BT$ calculated from Eq. (6.10) is 6.48, agreeing with the Gaussian chain simulations at large $N$.

6.3 1-2-1 translocation. (a) The chain first reside in a single cavity. (b) A chain end finds the hole that connects to the neighboring cavity. The free energy of the chain increases by $2\Delta F_a$. (c) The chain undergoes translocation. There are $p$ segments in the neighboring cavity and $N - p$ segments in the original one. (d) The translocation is complete when all the segments are transferred to the neighboring cavity.

6.4 Periodic free energy profile experienced by the polymer chain. To start the translocation process, the chain has to overcome an entropic free energy barrier $2\Delta F_a$ for find the hole to the next cavity. In the middle of the translocation process, there is a free energy local minimum due to the minimized excluded volume repulsion in the expense of decreased entropy. To finish one period, the chain has to overcome the excluded volume barrier to return to the initial state of occupying one cavity.
6.5 (a) The 2-3-2 translocation starts with a chain partitioning in two cavities with \( N/2 \) segments in each one. One of the chain ends then finds the hole to the adjacent cavity. In the middle of the process, the chain occupies three cavities with \( N/3 \) segments each. The second half of the process is identical to the reverse of the first half. (b) Under the cooperativity assumption, the 2-3-2 translocation can be mapped to a 1-2-1 translocation. The excluded volume of the segments in the donor and the receptor cavities are \( w/2 \) and \( w \) respectively. The new translocation process is also symmetric about \( p = N/3 \). 128

6.6 The probability of a \( N \)-segment chain occupying \( m \) cavities with radii \( R = 4 \). \( w = 0.2 \) and \( c = 1 \) are used. 129

6.7 The most probable time for a chain to diffuse a distance of one cavity. \( w = 0.2 \), \( c = 1 \) and \( \delta = 1 \) are used. 130

7.1 Pressure exerted by the confined chain on the cavity wall. The best fits of the pressures predicted by the mean field theory (Eq. (7.8a)), the scaling theory (Eq. (7.8b)) and the hard sphere chain theory (Eq. (7.8c)) are shown on top of the data. \( R = 3.5\sigma \) in the simulation. The pressure is in the simulation unit. 139

7.2 All data of the average escape times obtained from the simulations are plotted against the values predicted by Eq. 7.4. The solid line has a slope of 1 and passes through the origin. The error bars are smaller than the point size. 140

7.3 The escape trajectory of a chain with 100 beads from a cavity of radius 3.4\( \sigma \). The time is in the simulation unit and is measured from the start of the escape. The escape process is highly stochastic especially near the end of the process. The escape of the last 20 beads is responsible for about half of the escape time. 141

7.4 (a) The distribution of the waiting time for one end of the confined chain to find the hole and begin the escape. Solid line is the fit of the form \( \exp(-a\tau) \). (b) The distribution of the escape time for a chain with one bead already outside the cavity. Solid line is the fit of the form \( (\tau - a)^b \exp(-c(\tau - a)) \). \( N = 25 \) and \( R = 2.55\sigma \) in the simulations. 142

7.5 The average escape time of a Gaussian chain was well described by the prediction of equilibrium dynamics. 143
CHAPTER 1
GENERAL INTRODUCTION

Nature’s manifestation of the complexity of life through deterministic physical laws has long amazed mankind. It is hard to imagine the existence of life based on physical laws. Yet scientific progresses in molecular biology have confirmed the idea that physical laws are indeed solely responsible for how life works. Another interesting aspect of biological systems is the dominant roles played by various kinds of polymer. Some of the most notable biopolymers are DNA, RNA, protein and polysaccharide. Their dynamic properties govern the time scales of many biology processes.

This thesis is devoted to understanding how single polymers pass through geometrical constrictions. Examples in biology include DNA ejection from a viral capsid to a host cell, protein transport through membrane channels and gene swapping through bacterial pili. While one can be certain that physical laws are responsible for these seemingly miraculous biological functions, studying these systems in vivo have been proven difficult due to the limitations in instrumentation and the overwhelming complexity of the systems. Fortunately, advances in nanotechnology have made experimental studies of simplified systems possible. For example, biopolymer translocation through membrane channels was studied using α-Hemolysin protein nanopores [58] and synthetic nanopores [65] with the aids of applied electric fields, and the dynamics of DNA passing through periodic confined regions was visualized using fluorescence microscopy and nano-fabrication techniques [48]. We focus our attention on understanding simplified systems where single polymer dynamics are individually tractable and influenced by relatively few factors.
We applied concepts and methods in polymer physics to understand these simplified biological systems. We studied the process of polymer capture by and translocation through protein and synthetic nanopores. We also studied the dynamics of a chain inside various confining geometries. We hope that these fundamental results will serve as building blocks for understanding the more complex and highly cooperative biological systems.
CHAPTER 2
POLYMER CAPTURE BY ELECTROOSMOTIC FLOW
OF OPPOSITELY CHARGED NANOPORES

2.1 Introduction

Translocation of DNA through nanopores has a potential application of fast and economic DNA sequencing. Early efforts focused on driving single-stranded DNA or RNA molecules through α-Hemolysin protein nanopores with applied voltage [58, 2, 78, 79]. While the electric current signals of translocation events show DNA or RNA sequence dependence [78, 55], it is difficult to devise a reliable and systematic procedure to extract the sequence from the current signals. Moreover, the unique geometry and charged sites inside the protein nanopore makes the translocation process difficult to be controlled.

Much progress has been made in the fabrication of synthetic nanopores [64, 24, 65, 23, 119] as small as a few nanometers in diameter. Synthetic nanopores provide a promising way to fine tune the translocation process by changing the pore size, coating the pore with materials with different surface charges [24], and controlling the length of the nanopore. So far double-stranded (ds) DNA molecules have been studied in synthetic nanopore translocation experiments[65, 23].

In the literature, the main theoretical result of voltage driven polymer translocation through a nanopore is that the translocation time $\tau \sim N/V$, where $N$ is the chain length and $V$ is the applied voltage. The theories assume that the translocation process is slow enough such that the chain is able to equilibrate in the course of translocation [67, 85]. If the translocation is fast, a scaling theory [57] predicts
that $\tau \sim N^{1+\nu}/V$, where $\nu$ is the Flory size exponent. The scaling theories and simulations of polymer translocation do not take hydrodynamics into account.

With the recent advance of microfluidic systems [116], there is a renewed interest in studying electroosmotic flow (EOF), that is the flow of the fluid caused by the counterions of the charged surface under an applied electric field. In sub-micron systems, EOF plays an important role in fluid behavior under electric fields and is predicted to produce nontrivial phenomena such as local eddies induced by a varying channel diameter[102]. Binding of neutral molecules to sites within modified $\alpha$-Hemolysin protein nanopores suggested that EOF is prominent within charged protein nanopores [45]. The effect of EOF in voltage-driven polymer translocation through a charged nanopore, which is expected to be important due to nano-scale pore size and high electric field, is unexplored in the literature.

Recently, Chen et al. [23] observed an unexpectedly large (diameter$\approx 3\mu$m) DNA capturing region around the nanopore with 15nm in diameter. In the experiment, $\lambda$ DNA in 1M KCl solution with radius of gyration $R_g \approx 0.8\mu$m was forced to undergo translocation from the cis side to the trans side under the electric potential $V$. It was observed that whenever a DNA molecule diffused to the capturing region, it was absorbed to the nanopore immediately. The large capturing region was ascribed [23] to the extended electric field from the nanopore. However, the Debye screening length $\kappa^{-1}$ for the experimental system is only about 0.3 nm, and the electric field is unlikely to extend to the micron scale. A more careful analysis has to account for the fact that the system is not in equilibrium. Under the applied voltage, electric currents (positive and negative ions) pass through the nanopore continuously. The electric potential and salt concentration are given by the Poisson-Nernst-Planck (PNP) equations at the steady state.
\[ \nabla \cdot \mathbf{J}_i(r) = 0, \quad (2.1) \]
\[ \mathbf{J}_i(r) = -D_i \left[ \nabla c_i(r) + \frac{z_i c_i}{kT} \nabla \Phi(r) \right], \quad (2.2) \]
\[ \nabla \cdot [\epsilon(r) \nabla \Phi(r)] = -\Sigma_i z_i c_i(r), \quad (2.3) \]

where the subscript \( i \) denotes the ion species with valence \( z_i \), \( \mathbf{J}_i(r) \) is the ionic flux, \( D_i \) is the diffusion coefficient of the ion, \( \Phi(r) \) is the electric potential, \( c_i(r) \) is the concentration of the ion, and \( \epsilon(r) \) is the dielectric permittivity. For KCl solution, the two ion species are \( K^+ \) and \( Cl^- \). Eq. (2.1) states that each ionic flux has to be continuous. Eq. (2.2) states that the ionic flux is created by both the salt concentration gradient and the electric field. Eq. (2.3) is the Poisson equation. The three PNP equations have to be solved self-consistently to obtain \( \mathbf{J}_i(r) \), \( \Phi(r) \) and \( c(r) \). PNP equations have been studied extensively in the literature for ion channels. Numerical solutions have shown that the electric field vanishes after several Debye lengths away from the pore (for example, see Refs. [96, 21]). Nakane et al. [93] considered the extended electric field from a nanopore. Although realizing the 0.3nm Debye screening length, they failed to consider the salt gradient term of Eq. (2.2) (first term on the right hand side), leading to the conclusion that there is a long range electric field \( -\nabla \Phi \sim J_i \sim 1/R^2 \), with distance \( R \) away from the pore in the cis side. We conclude, based on PNP solutions reported in the literature, that the micron scale absorbing region is not caused by the extended electric field.

In this chapter, we provide a possible explanation of the large absorbing region. We propose that it is caused by the positive surface charge of the nanopore. If the surface charge is opposite to the charge of the polymer (negative for DNA), the EOF is in the same direction as the polymer translocation. The flow inside the nanopore creates a hydrodynamic flow field near the pore. For the high electric field (\( \sim 10^7 \) V/m) in the experiment [23], the flow field creates an absorbing region comparable to the size of the polymer. When a polymer enters this region, the fluid velocity gradient
is strong enough that coil-stretch transition [29, 62, 103] of the polymer occurs and the molecule undergoes a deformation. The polymer molecule is elongated, reducing the entropic barrier of translocation. Thus the polymer translocation is assisted by EOF, which is a hydrodynamic effect.

In the next section, we briefly review the theory of EOF inside a nanopore. The mechanism and size of the absorbing region are discussed in Section 2.3. The translocation rate is derived in Section 2.4. Relevance with experiments and implication on theories are discussed in the last section.

### 2.2 Electroosmotic flow inside a nanopore

In this section, we review the EOF created by a nanopore with surface charge. We consider a cylindrical nanopore of diameter $2a$ and length $L$, as shown in Fig. 2.1. An electrical potential $V$ is applied across the nanopore. There is a uniform surface charge density $\sigma$ on the surface inside the nanopore, creating a surface potential $\Phi_0$. The system is assumed to be in a symmetric $z : z$ electrolyte solution (eg. KCl), where the coions and counterions distribute near the wall and screen the surface potential. At the same time, they are driven by the electric field across the nanopore and move along it. The two processes, the radial distribution of the ions and the axial movement of the ions, are assumed to be independent of each other for sufficiently long cylindrical pores so that we can treat them independently [105].

The potential inside the pore is given by the Poisson-Boltzmann equation[107]

\[
\nabla^2 \Phi(r) = -\frac{\rho}{\epsilon}, \\
\nabla^2 \Phi(r) = -\frac{nz e}{\epsilon} \left[ \exp\left(\frac{-ze\Phi(r)}{kT}\right) - \exp\left(\frac{ze\Phi(r)}{kT}\right) \right] = \frac{2n ze}{\epsilon} \sinh\left[\frac{ze\Phi(r)}{kT}\right].
\]

\[ (2.4) \]
Φ(r) is the electrostatic potential, n is the bulk coion and counterion density, e is the absolute value of the electron charge, kT is the thermal energy and ε is the dielectric permittivity of the electrolyte solution. Under small surface potentials, Eq. (2.4) can be linearized using Debye-Hückel approximation

\[
\frac{1}{r} \frac{d}{dr} \left( r \frac{d}{dr} \Phi(r) \right) = \kappa^2 \Phi(r),
\]

where \( \kappa = \sqrt{2nz^2e^2/\epsilon kT} \) is the inverse Debye length. The potential \( \Phi(r) \) depends only on radial position \( r \) because of the cylindrical symmetry. The solution of Eq. (2.5) is

\[
\Phi(r) = \Phi_0 I_0(\kappa r)/I_0(\kappa a),
\]

where \( I_n(x) \) is the n-th order modified Bessel function of the first kind [1]. From Eqs. (2.4) and (2.6), the charge density is

\[
\rho(r) = -\epsilon \kappa^2 \Phi_0 \frac{I_0(\kappa r)}{I_0(\kappa a)}.
\]

An electric potential \( V \) is applied across the pore axis. The electric field is then \( E = V/L \). The motion of the electrolyte solution is coupled with the net charge density via the Navier-Stokes equation

\[
-\eta \nabla^2 v(r) = \rho(r) E,
\]

where \( \eta \) is the viscosity of the solution and \( v(r) \) is the local velocity of the solution. With the boundary condition \( v(a) = 0 \), the velocity is given by

\[
v(r) = -\frac{\Phi_0 E \epsilon}{\eta} \left[ 1 - \frac{I_0(\kappa r)}{I_0(\kappa a)} \right].
\]
Then the liquid flux created by the EOF is

\[ J_0 = \int_0^a v(r)2\pi r dr = -\frac{\Phi_0 E\varepsilon}{\eta}(\pi a^2)\beta, \quad (2.10) \]

where

\[ \beta = 1 - \frac{2}{\kappa a} \frac{I_1(\kappa a)}{I_0(\kappa a)} \approx 1 - \frac{2}{\kappa a} \quad \text{for} \quad \kappa a \gg 1 \quad (2.11) \]

is the correction factor for a finite salt concentration. At very high salt concentrations, the Debye layer near the wall is so thin that Eq. (2.8) can be solved with the slip boundary condition. In this case \( \beta = 1 \) and the velocity is uniform all over the cylinder [105]. The minus sign in Eq. (2.10) indicates that the EOF is opposite to the applied electric field if the surface charge is positive. The salt concentration dependent velocity profile and the strength of the flux are plotted in Fig. 2.2. At high salt concentrations, the velocity profile is flat except near the wall, reflecting the short length of the Debye double layer. At low salt concentrations, the Debye layers overlap and the velocity profile is parabolic. Most of the translocation experiments were conducted at 1M high salt concentration.

### 2.3 Polymer capture by EOF in the nanopore

#### 2.3.1 Absorption time estimation

Suppose the nanopore has a positive surface charge density \( \sigma \). The surface potential \( \Phi_0 \) is positive and there is a net negative charge density near the pore wall to compensate the positively charged surface. An electric potential is applied across the pore so that the negatively charged polymers are driven from the *cis* side to the *trans* side. In this way, the EOF is in the same direction as the polymer translocation, as depicted in Fig. 2.3.
Because of the continuity of the fluid flow, a liquid flow field is created near the pore in the *cis* side. The velocity of the fluid at distance \( R \) away from the pore is

\[
 u(R) = -\frac{|J_0|}{2\pi R^2}.
\]  

(2.12)

The negative sign is to remind the reader that the origin of \( R \) is at the entrance of the pore. Eq. (2.12) is a good approximation as long as \( R \gg a \). First we consider a point mass dragged completely by the flow without perturbing the velocity \( u(R) \). The time taken for that point mass to travel from a distance \( R \) away from the pore to the mouth of the pore is estimated by

\[
 u(R) dt = dR,
\]

(2.13)

\[
 \tau = \frac{2\pi}{3|J_0|} R^3.
\]

(2.14)

If this absorbing time is shorter than the longest relaxation time of the polymer, any segments dragged by the flow field are not able to escape before reaching the pore and immediate polymer translocation through the pore follows. The longest relaxation time of the polymer is given by the Zimm relaxation time \([29, 34]\)]

\[
 \tau_Z \approx 0.3 \frac{\eta R_g^3}{kT}.
\]

(2.15)

The unperturbed radius of gyration of the polymer \( R_g \approx \frac{N^{3/5}b}{\sqrt{6}} \) for a self-avoiding chain with Kuhn length \( b \) and \( N \) Kuhn segments. For dsDNA in high salt (>10mM) solution, \( b \approx 100 \text{nm}=300 \text{bp} \) (base pairs). At high applied potential, the entropic barrier of the polymer entering the pore is not important. Every polymer molecule that reaches the pore would undergo translocation. The translocation events are
diffusion limited. Equating the approach time Eq. (2.14) with the Zimm relaxation time Eq. (2.15), the radius of the absorbing region is

$$R_C = \left[0.45 \frac{\beta \Phi_0 E \epsilon a^2}{kT} \right]^{1/3} R_g. \quad (2.16)$$

### 2.3.2 Coil-stretch transition

We study the problem from another perspective in this section. In the presence of velocity gradient of the fluid, polymer chains undergo the coil-stretch transition if the velocity gradient $\dot{\gamma} = du(R)/dR$ is larger than the critical velocity gradient $\dot{\gamma}_C$ which is given by [29, 62, 103]

$$\dot{\gamma}_C \tau_Z \approx 0.5. \quad (2.17)$$

Then there is a critical radius $R_C$ outside which $\dot{\gamma} < \dot{\gamma}_C$ and the polymer remains in coil conformation, as depicted in Fig. 2.4. Within $R_C$, $\dot{\gamma} > \dot{\gamma}_C$ and coil-stretch transition occurs. The polymer molecule undergoes affine deformation and was absorbed into the pore. Using Eqs (2.12), (2.15) and (2.17), the absorbing radius is

$$R_C = \left[0.6 \frac{\beta \Phi_0 E \epsilon a^2}{kT} \right]^{1/3} R_g \equiv (0.6\alpha)^{1/3} R_g. \quad (2.18)$$

We get essentially the same answer as in the last section (Eq. (2.16)). For convenience, we define

$$\alpha \equiv \frac{\beta \Phi_0 E \epsilon a^2}{kT} = \left[1 - \frac{2 I_1(\kappa a)}{\kappa a I_0(\kappa a)} \right] \frac{\Phi_0 E \epsilon a^2}{kT} \quad (2.19)$$

as a measure of experimental parameters. The strength of the EOF increases with $\alpha$. Large $\alpha$ favors polymer capture by EOF.

### 2.3.3 Estimation of the absorption radius $R_C$

We estimate $R_C$ using experimental values from Chen and coworkers’ nanopore experiment [23]. From the experiment, $a = 7.5\text{nm}$, $E = 500\text{mV}/10^{-6}\text{cm} = 5 \times 10^7\text{V/m}$,
\[ \epsilon = 80 \epsilon_0, \epsilon_0 \text{ is the permittivity of the vacuum}, \ T = 300K, \ k^{-1} = 0.304 \text{nm}/\sqrt{\text{[KCl]}} = 0.304 \text{nm in 1M KCl}. \ \beta \text{ is given by Eq. (2.11) and is } 0.92 \text{ under the above experimental condition}. \] The surface potential was uncertain. We assume it to be positive and use the typical value of \( \Phi_0 = 0.1V \) [105] which corresponds to \( \sigma = 1.4 e/\text{nm}^2 \). Eq. (2.7) is used to estimate \( \sigma \) and is valid if there is no overlapping of the Debye layers. Using Eq. (2.18), \( \alpha = 44.3 \) and the absorbing radius \( R_C = 3.0 R_g \). For \( \Phi_0 = 10 \text{mV} \) (\( \sigma = 0.14 e/\text{nm}^2 \)) and \( 1 \text{mV} \) (\( \sigma = 0.014 e/\text{nm}^2 \)), \( R_C = 1.4 R_g \) and \( 0.6 R_g \), respectively.

\( R_g \approx 0.8 \mu m \) for \( \lambda \) DNA used in the experiment. Therefore the EOF of the nanopore, provide that there is a positive surface charge on the nanopore, is able to create an absorbing region with radius \( \approx 1.5 \mu m \) observed in the experiment.

On the other hand, if \( \Phi_0 \) is negative, EOF is opposite to the direction of the polymer translocation and no absorbing region exists.

The main result is that, under the nanopore experimental conditions, the absorbing region created by EOF is comparable to the radius of the polymer \( R_g \). This absorbing zone results in a depletion of polymer concentration and is observable using fluorescent polymers [23].

### 2.4 Diffusion limited translocation rate

As discussed in the last section, if the electric potential is high enough, every polymer molecule that reaches the nanopore would translocate through it immediately. The concentration of the polymer is zero within the absorbing radius \( R_C \) centered at the nanopore. Outside the region, the polymer molecules undergo biased random walk in the presence of the flow field \( u(R) = -|J_0|/2\pi R^2 \) (Eq. (2.12)), which is not strong enough to deform the polymer. The negative sign of \( u(R) \) indicates that the liquid is flowing into the pore. At steady state, the concentration of polymer \( c(R) \) with distance \( R \) away from the pore is governed by the steady state diffusion equation with a drift term [14]

\[ \frac{d^2 c}{dR^2} + \frac{d}{dR} \left( D \frac{d c}{dR} \right) = -u(R) c(R) \]
\[
D \frac{1}{R^2} \frac{d}{dR} \left[ R^2 \frac{dc}{dR} \right] - \frac{1}{R^2} \frac{d}{dR} \left[ R^2 uc \right] = 0, \quad (2.20)
\]

where
\[
D \approx 0.2 \frac{kT}{\sqrt{6R_g}\eta} \quad (2.21)
\]
is the diffusion coefficient of the polymer [34]. With the boundary conditions \(c(R_C) = 0\) and \(c(\infty) = c_0\), the bulk concentration of the polymer, the solution of Eq. (2.20) is
\[
c(R) = c_0 \frac{\exp \left[ -\frac{|J_0|}{2\pi D} (\frac{1}{R_C} - \frac{1}{R}) \right] - 1}{\exp(-\frac{|J_0|}{2\pi DR_C}) - 1}. \quad (2.22)
\]

Near \(R_C\), the polymer concentration drops exponentially to zero with decay length \(l_D \sim 2\pi DR_C^2/|J_0| = 0.12R_g/\alpha^{1/3}\), where the dimensionless factor \(\alpha\) is defined in Eq. (2.19). The stronger the EOF, the sharper the drop. Using the experimental parameters stated in Section 2.3.3, \(l_D/R_g = 0.03, 0.07\) and \(0.15\) for \(\Phi_0 = 0.1\)V, 10mV and 1mV respectively. The concentration profiles are plotted in Fig. 2.5.

The translocation rate is given by the number of polymer molecules reaching the hemisphere with radius \(R_C\) centered at the nanopore per unit time. That is the polymer flux at \(R = R_C\) times the surface area of the absorbing hemisphere,
\[
\Gamma = -2\pi R_C^2 \left[ -D \frac{\partial c}{\partial R} + uc \right]_{R=R_C}. \quad (2.23)
\]
The second term is zero because of the boundary conditions. Substituting Eq. (2.22) into Eq. (2.23), the diffusion limited translocation rate is
\[
\Gamma = \frac{c_0 |J_0|}{1 - \exp(-|J_0|/2\pi DR_C)} = \frac{\pi c_0 kT}{\eta} \frac{\alpha}{1 - \exp(-7.26\alpha^{2/3})}. \quad (2.24)
\]

If \(\alpha\) is of order \(10^{-1}\) or above, the exponential factor in the denominator of Eq. (2.24) is negligible and the translocation rate \(\Gamma \approx c_0 |J_0| \sim E = V/L\). It is
in contrast with the voltage driven polymer translocation through the entropic barrier imposed by a nanopore. In that case, the translocation rate $\Gamma \sim \exp(V)$ [52, 8]. We note that if $\alpha$ is small, the effect of entropic barrier is important and Eq. (2.24) is not applicable. The translocation rate is determined by the entropic barrier, which decreases with the applied voltage. If the applied voltage is low so that the entropic barrier $\gg kT$, there would be no translocation. Hence there is a cut-off value $\alpha_C$ below which there is no translocation.

2.5 Summary and discussions

In this chapter, we consider the situation where the surface charge of the nanopore is opposite to the charge of the polymer. The electroosmotic flow (EOF) is in the same direction as the polymer translocation. We show that the EOF creates an absorbing region with radius $R_C = (0.6\alpha)^{1/3}R_g$. Within this absorbing region, the velocity gradient of the fluid is larger than the critical velocity gradient necessary for chain stretching [29, 62, 103]. The coil-stretch transition occurs and the polymer molecule undergoes a deformation, and is unable to escape the absorption of the nanopore. Similar estimate (Eq. (2.16)) for $R_C$ is obtained by comparing the time taken for the polymer to approach the pore and the Zimm relaxation time for the polymer. In typical nanopore experimental conditions [23] (with positive nanopore surface charge), $(0.6\alpha)^{1/3}$ is of order unity. Because of the high electric field ($\sim 10^7$ V/m) across the nanopore, even a small positive surface charge density $\sigma \sim 0.1e/nm^2$ results in an absorbing region comparable to the radius of the polymer $R_g$. If the electric field is high enough that the entropic barrier [85] of translocation is not important, every polymer molecule that is elongated by the EOF near the pore would undergo immediate translocation through the pore. The translocation events are diffusion limited, and the polymers outside $R_C$ undergo biased random walk under a flow field $u(R) \sim -E/R^2$ (Eq. (2.12)). If $\alpha$ is of order $10^{-1}$ or higher, the translocation rate
\( \Gamma \sim E \). We note that Eq. (2.24) is only valid when the translocation events are one at a time. Multiple entries may disturb the EOF field significantly, making the estimate inaccurate.

This work provides a possible explanation of the unexpectedly large absorbing region \((R_C \approx 1.5 \mu m)\) of \( \lambda \) DNA \((R_g \approx 0.8 \mu m)\) observed in the nanopore (diameter \( \approx 15 \text{nm} \)) experiment of Chen et al.[23]. We propose that the capture was facilitated by the EOF of the nanopore with positive surface charge, opposite to the charge of the DNA. The EOF produces a capture radius comparable to the radius of gyration of the DNA as analyzed in Sec. 2.3.3.

The coil-stretch transition picture is consistent with the experimental observation that DNA molecules were elongated after capture and the percentage of linear translocation increased with applied voltage, which could not be explained by energetic consideration as high applied voltage should favor bending of the DNA molecules. Our analysis indicates that the absorption time should be comparable to the Zimm relaxation time of the captured polymer molecule, which is about 46 ms for \( \lambda \) DNA (from Eq. 2.15 and Ref. [103]). This agrees with the experimental observation that \( \lambda \) DNA molecules were absorbed within 50 ms.

Our diffusion limited translocation rate \( \Gamma \sim V \) is consistent with experimental observations, as opposed to the usual translocation rate across a free energy barrier \( \Gamma \sim \exp(V) \). With \( E = 5 \times 10^7 \text{V/m} \), \( \Phi_0 = 1 \text{mV} \) and \( 5 \mu g/mL \) \( \lambda \) DNA solution, the calculated translocation rate from Eq. (2.24) is \( \Gamma = 0.5 \) events/sec.

The pioneering nanopore translocation experiment of Li et al.[65] showed low translocation rates. As explained in Ref. [24], the negative surface charge of the silicon nitrate nanopore created an EOF opposite to the direction of the DNA translocation, reducing the translocation rate. They used atomic layer deposition to coat the nanopore with alumina, resulting in a more linear I-V curve [24]. Calculation from Goldman-Hodgkin-Katz(GHK) equation showed that the cationic selectivity of
the pore was about 50%, implying an uncharged nanopore. We note that at high salt concentration (1M KCl) and for large nanopores (∼20nm), the linearity of the I-V curve is not a good indicator of nanopore surface charge [113] because the surface charge cannot create a significant asymmetric electrostatic potential to most of the ions passing through the pore. Furthermore, GHK equation assumes a linear potential drop across the pore, and may not estimate the ionic selectivity accurately in the presence of surface charge. The isoelectric point of alumina is about pH 9.0 while the nanopore experiment was conducted in a pH 8.0 buffer solution. It is possible that there was positive surface charge on the nanopore, resulting in the micron scale absorbing region.

In the literature, a scaling argument predicts that translocation time \( \tau \sim N^{1+\nu}/V \) for voltage driven translocation through a pore [57] if the chain cannot equilibrate in the course of translocation. We note that this argument ignores the effect of hydrodynamics. In the presence of the nanopore surface charge, the EOF may elongate the polymer substantially before and after entering the pore. The chain is no longer in equilibrium conformations \( R_g \sim N^{\nu} \), before and after the translocation, making the scaling argument invalid. Even for a neutral pore, if the pore is substantially larger than the diameter of the polymer segment, the fast unidirectional movement of the chain during translocation would drag the fluid to move along, creating a hydrodynamic flow in the same direction. Some computer simulations have shown that this hydrodynamic effect speeded up the translocation [6]. Whether it would alter the scaling prediction requires further investigations.

The analysis of the EOF in nanopores presented in Section 2.2 is strictly correct only in infinitely long pores. End effects will lead to a correction to our results. The end result is expected to make the pore radius and membrane thickness effective values in our equations rather than the actual values.
In our analysis of the polymer capture, we used the longest relaxation time of the polymer to deduce the capture properties. The amount of hairpin translocations may be able to be estimated from the shorter relaxation times of the chain corresponding to the higher Rouse modes of the chain.

There is a possible practical importance of the EOF of the nanopore described in this chapter. Coating the nanopore deliberately with a material of opposite charge to the polymer increases the capture rate, and is especially beneficial for small nanopores. The coil-stretch transition increases the occurrence of linear translocation, which is essential for single-stranded DNA or RNA sequencing.
Figure 2.1. A cylindrical nanopore with radius $a$ and positive surface potential $\Phi_0$. 
Figure 2.2. EOF in a nanaopore. (a) Velocity profile at different monovalent ($z = 1$) salt concentrations. (b) Strength of the electroosmotic flux as a function of salt concentration. See Eqs. (2.10) and (2.11) for the definition of $\beta$. $\Phi_0 = 0.01V$, $a = 7.5\text{nm}$, $E = 5 \times 10^7\text{V/m}$, $T = 300\text{K}$ and $\eta = 1\text{cP}$ are used in both graphs.
Figure 2.3. If the surface charge of the nanopore is opposite to the charge of the polymer molecules, EOF is in the same direction as the polymer translocation, assisting the translocation process.
Figure 2.4. When the polymer passes the absorbing radius $R_C$ given by Eq. (2.18), the fluid velocity gradient is larger than the critical velocity gradient. The polymer is unable to relax and coil-stretch transition occurs.
Figure 2.5. Concentration profile at different $\alpha$ given by Eq. (2.22).
CHAPTER 3
POLYMER TRANSLOCATION THROUGH A CYLINDRICAL CHANNEL

Polymer translocation through a nanopore or nanochannel is essential for biological functions. Examples include protein and DNA transport through membrane channels and viral DNA injection into host cells [5]. Recent experimental realization of monitoring single molecule translocation through protein or synthetic nanopores leads to the possibilities of single DNA sequencing, single molecular characterization and single molecular detection [58, 2, 11, 79, 23, 65, 119, 129, 73, 41]. Many simulations [26, 59, 123, 3, 28, 66, 71, 69, 74, 109, 124, 125, 130] and theoretical works [121, 67, 85, 86, 88, 8, 13, 40, 57, 60, 80, 112, 114, 35, 36, 110] have attempted to elucidate the translocation process.

Many factors affect the translocation process. For example, the driving force can be thermal diffusion of the polymer molecule, an electric field, a chemical potential or a pulling force. The shape of the pore, such as the asymmetric interior of the α-hemolysin nanopore, affect translocation dynamics. For charged polymer molecules, the low dielectric constant of the membrane may impose an electrostatic energy barrier to the polymer entry.

We focus our attention on the translocation dynamics arising from the finite diameter of the channel. Previous theories [88, 40, 60], except Ref. [114] which estimated the effect of channel width using a scaling argument, assumed that the channel is very narrow so that polymer aligns itself in a linear fashion inside the channel. In other words, a channel of length $M$ could accommodate at most $M/b$ segments, where $b$ is
the segment length. In reality, a channel has a finite diameter $2a$. The number of segments allowed to reside in the channel depends on both $a$ and $M$. If the translocation process is in quasi-equilibrium, the chain would adjust the number of segments inside the channel at each step, according to the free energy of chain conformations.

In this chapter, we account for the finite diameter of the channel so that the number of segments inside the channel can be adjusted during translocation according to the free energy. Different conformations of the chain in the translocation process are shown in Fig. 3.1. The corresponding free energies are obtained explicitly for a Gaussian chain. All the states except state (f) are described by a single variable, $p$ or $q$, or without any variables, where $p$ and $q$ are the number of segments inside the channel and the receptor compartment, respectively. The translocation, from the donor compartment to the receptor compartment, is considered as a two-step process. The chain initially residing in the donor compartment starts the process by finding the channel entrance with a chain end. It takes time $\tau_1$ for that chain end to travel through the channel, reaching the interface between the channel and the receptor compartment. All the segments then transfer into the receptor compartment in time $\tau_2$. The total translocation time $\tau_1 + \tau_2$ is calculated using the confinement free energy of a Gaussian chain. As shown below, $\tau_1$ and $\tau_2$ exhibit rich dependencies on the radius of the donor ($R_1$), radius of the receptor ($R_2$), radius of the channel ($a$), length of the channel ($M$), and chain length ($N$). All of these parameters are expressed in units of the segment length $b$.

3.1 Theory

Consider a Gaussian chain of $N$ segments with Kuhn segment length $b$. The probability $P(r, r_0; N)$, relative to the free space, that the two ends are at $r$ and $r_0$, respectively, obeys the equation [34]
\[
\left( \frac{\partial}{\partial N} - \frac{b^2}{6} \nabla_r^2 \right) P(r, r_0; N) = 0
\] (3.1)

with the boundary condition \( P(r, r_0; 0) = \delta(r, r_0) \). If the allowed conformations of the chain are limited by some surfaces, Eq. (3.1) is solved with the additional boundary condition that the probability is zero for \( r \) or \( r_0 \) at the surfaces. We first consider two confining geometries, a sphere and a cylinder, before constructing the free energy landscape for the translocation process.

### 3.1.1 Polymer confined in a sphere

For a spherical compartment of radius \( R \), the solution of Eq. (3.1) is a function of \( r \) and \( r_0 \), which consist of 6 independent position variables. However, from the spherical symmetry, the solution should depend only on the radial distances \( r, r_0 \) and the angle \( \gamma \) between the two vectors \( r \) and \( r_0 \). This can be achieved by applying the Legendre addition theorem [9] to the solution. The probability of the chain inside the sphere is

\[
P_S(r, r_0; N) = \frac{1}{2\pi R^2} \sum_{n=0, \alpha_{n+1/2}}^{\infty} P_n(\cos \gamma)(2n + 1) \frac{J_{n+1/2}(\alpha_{n+1/2}r/R)J_{n+1/2}(\alpha_{n+1/2}r_0/R)}{(J_{n+1/2}(\alpha_{n+1/2}))^2 \sqrt{rr_0}} \exp \left( -\frac{b^2\alpha_{n+1/2}N}{6R^2} \right),
\] (3.2)

where \( P_n(x) \) is the Legendre polynomial of order \( n \), \( J_n(x) \) is the Bessel function of order \( n \), and \( \alpha_{n+1/2} \) are the zeros of the Bessel function of the first kind such that \( J_{n+1/2}(\alpha_{n+1/2}) = 0 \). Note that for each \( n \), there is an infinite number of \( \alpha_{n+1/2} \) to be summed over.

If the probability of the chain depends only on the radial distance \( r \) and \( r_0 \) but not the angle \( \gamma \), only the \( n = 0 \) terms in Eq. (3.2) are retained. It reduces to

\[
P_S(r, r_0; N) = \frac{1}{2\pi Rrr_0} \sum_{m=1}^{\infty} \sin \left( \frac{m\pi r_0}{R} \right) \sin \left( \frac{m\pi r}{R} \right) \exp \left( -\frac{m^2\pi^2 b^2 N}{6R^2} \right),
\] (3.3)
which was given in Ref. [88]. Eq. (3.2) should be used when the probability of the chain depends on the angle $\gamma$.

If one of the chain ends is free, and the other one is anchored at $r = R - c$ near the surface of the compartment, where $c$ is the anchoring distance and is comparable to $b$, Eq. (3.3) becomes [88]

$$P_{S1}(R, N) = \frac{2c}{R} \sum_{m=1}^{\infty} \exp \left( -\frac{m^2\pi^2b^2N}{6R^2} \right)$$

$$\equiv \frac{2c}{R} f_{S1}(R, N). \quad (3.4)$$

The higher $O(c^2)$ terms are ignored. If both ends are free inside the sphere, the probability is

$$P_S(R, N) = \int_0^R 4\pi r^2 dr \int_0^R 4\pi r_0^2 dr_0 P_s(r, r_0; N)$$

$$= \frac{8R^3}{\pi} \sum_{m=1}^{\infty} \frac{1}{m^2} \exp \left( -\frac{m^2\pi^2b^2N}{6R^2} \right). \quad (3.5)$$

### 3.1.2 Polymer confined in a cylinder

Similarly, the probability of a chain inside a cylindrical compartment of length $M$ and diameter $2a$ is, by solving Eq. (3.1),

$$P_C(r, r_0; N) = \frac{8}{\pi Ma^2} \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} \sum_{\mu_n} \left( \cos n\theta \cos n\theta_0 + \sin n\theta \sin n\theta_0 \right) \sin \left( \frac{m\pi}{M} \left( z + \frac{M}{2} \right) \right)$$

$$\times \sin \left( \frac{m\pi}{M} \left( z_0 + \frac{M}{2} \right) \right) \frac{J_n(\mu_n r/a)J_n(\mu_n r_0/a)}{J_{n+1}(\mu_n)} \exp \left( -\frac{b^2}{6} \left( \frac{n^2}{a^2} + \frac{m^2\pi^2}{M^2} \right)N \right), \quad (3.6)$$

where $\mu_n$ are the zeros of Bessel function of the first kind such that $J_n(\mu_n) = 0$. The origin of the coordinates is at the center of the cylinder. $z$ is along the axis of the cylinder. $r = (r, \theta, z)$ and $r_0 = (r_0, \theta_0, z_0)$. 

If the chain end $r_0$ can be anywhere inside the cylinder, from Eq. (3.6), the probability that the other chain end is at $r$ is

$$P_C(r; N) = \int_{-\pi}^{\pi} d\theta_0 \int_0^a r_0 dr_0 \int_{-M/2}^{M/2} dz_0 P_C(r, r_0; N)$$

$$= \frac{16}{\pi} \sum_{m \mu_0} \left( \frac{J_0(\mu_0 r/a)}{\mu_0 J_1(\mu_0)} \right) \left( 1 - (-1)^m \right) \sin \left( \frac{m \pi}{M} (z + \frac{M}{2}) \right)$$

$$\times \exp \left( -\frac{b^2}{6} \left( \frac{\mu_0^2}{a^2} + \frac{m^2 \pi^2}{M^2} \right) N \right).$$

(3.7)

Here $\mu_0$ are the zeros of $J_0(x)$. If both ends are free, the probability of the confined chain is

$$P_C(N) = \int_{-\pi}^{\pi} d\theta \int_0^a r dr \int_{-M/2}^{M/2} dz P_C(r; N)$$

$$= \frac{16Ma^2}{\pi^2} \sum_{m \mu_0} \left( 1 - (-1)^m \right)^2 \exp \left( -\frac{b^2}{6} \left( \frac{\mu_0^2}{a^2} + \frac{m^2 \pi^2}{M^2} \right) N \right).$$

(3.8)

If one end is free, while the other end is allowed to be anchored anywhere at the lateral surface at $z = -\frac{M}{2} + c$, from Eq. (3.7) the probability is

$$P_{C1}(N) = \int_{-\pi}^{\pi} d\theta \int_0^a r dr P_C(r; N)|_{z=-M/2+c}$$

$$= \frac{16ca^2}{M} \sum_{m \mu_0} \frac{1 - (-1)^m}{\mu_0^2} \exp \left( -\frac{b^2}{6} \left( \frac{\mu_0^2}{a^2} + \frac{m^2 \pi^2}{M^2} \right) N \right)$$

$$\equiv \frac{16ca^2}{M} f_{C1}(N),$$

(3.9)

where $O(c^2)$ terms are ignored. By symmetry, the probability that one chain end is anchored at $z = \frac{M}{2} - c$, while the other chain end is free to move, is also given by Eq. (3.9). Similarly, the probability that one end is anchored at $z = -\frac{M}{2} + c_1$ while the other end is anchored at $z = \frac{M}{2} - c_2$ is
\[ P_{C2}(N) = \frac{32\pi^3 c_1 c_2}{M^3} \sum_{m=0}^{\infty} \frac{m^2}{\mu_0^2} (-1)^{m+1} \exp \left( -\frac{b^2}{6} \left( \frac{\mu_0^2}{a^2} + \frac{m^2 \pi^2}{M^2} \right)N \right) \]

\[ \equiv \frac{32\pi^3 c_1 c_2}{M^3} f_{C2}(N). \quad (3.10) \]

### 3.1.3 Free energy of confinement

The Helmholtz free energy of a confined Gaussian polymer is given by

\[ F = -TS = -k_B T \ln P \quad (3.11) \]

for the probability of the chain conformation \( P \). We consider the spherical donor and receptor compartments of radius \( R_1 \) and \( R_2 \), respectively, being connected with a cylindrical channel of diameter \( 2a \) and length \( M \). The chain can partition among the three compartments, according to the free energy of the conformation. In the present calculations, we assume that there is only one segment at the interface between the cylinder and the sphere.

Fig. 3.1 shows all possible conformations of a chain during translocation. Their corresponding free energy can be obtained from the various probabilities calculated in the last sections. If the chain resides completely in the donor compartment as shown in Fig. 3.1(a), from Eqs. (3.5) and (3.11), the free energy is

\[ \frac{F_{(a)}}{k_B T} = -\ln P_S(R_1, N) \]

\[ = -\ln \frac{8R_1^3}{\pi} - \ln \left( \sum_{m=1}^{\infty} \frac{1}{m^2} \exp \left( -\frac{m^2 \pi^2 N}{6R_1^2} \right) \right). \quad (3.12) \]

If the polymer is confined in the donor compartment with one end anchored anywhere at the entrance of the channel at distance \( c \), as in Fig. 3.1(b), its free energy is

\[ \frac{F_{(b)}}{k_B T} \approx -\ln \left( P_{S1}(R_1, N) \pi a^2 \right) \]

\[ = -\ln \frac{2\pi ca^2}{R_1} - \ln \left( \sum_{m=1}^{\infty} \exp \left( -\frac{m^2 \pi^2 N}{6R_1^2} \right) \right). \quad (3.13) \]
Eqs. (3.4) and (3.11) are used, and the factor $\pi a^2$ accounts for the area of the channel entrance that the chain end can anchor at. Suppose $p$ segments reside in the cylindrical channel and $N - p$ segments reside in the donor compartment, as in Fig. 3.1(c). Now Eqs. (3.4) and (3.9) yield the probability as

$$P(c) = \int dr \int dr_0 \int' dr' P_C(r, r'; p) P_S(r', r_0; N - p) \simeq P_{C1}(p) P_{S1}(R_1, N - p).$$  \hspace{1cm} (3.14)$$

The integrations of $r$ and $r_0$ are over the whole corresponding compartments. $r'$ is integrated over the area connecting the two compartments. Integration of $r'$ is performed by assuming that $r'$ is at $z' = -\frac{M}{2} + c_1$ inside the cylindrical channel while it is at $r' = R_1 - c$ inside the spherical compartment, making $P_{S1}$ independent of the integration of $r'$. The free energy is therefore

$$\frac{F(c)(p)}{k_B T} \simeq -\ln (P_{C1}(p) P_{S1}(R_1, N - p)).$$ \hspace{1cm} (3.15)$$

For the conformation of Fig. 3.1(d), one end is anchored in the cylinder surface at $z = \frac{M}{2} - c_2$, and the other end can be anywhere inside the donor compartment. If there are $p$ segments inside the channel, the free energy becomes

$$\frac{F(d)(p)}{k_B T} \simeq -\ln (P_{C2}(p) P_{S1}(R_1, N - p)).$$ \hspace{1cm} (3.16)$$

Likewise, the free energies of conformations in Fig. 3.1 (e), (f), (g), (h), (i), and (j) are
\[
\frac{F_{(e)}}{k_B T} = - \ln P_{C1}(N), \quad (3.17)
\]
\[
\frac{F_{(f)}(p, q)}{k_B T} \simeq - \ln (P_{S1}(R_2, q)P_{C2}(p)P_{S1}(R_1, N - p - q)), \quad (3.18)
\]
\[
\frac{F_{(g)}(q)}{k_B T} \simeq - \ln (P_{C2}(N - q)P_{S1}(R_2, q)), \quad (3.19)
\]
\[
\frac{F_{(h)}(q)}{k_B T} \simeq - \ln (P_{C1}(N - q)P_{S1}(R_2, q)), \quad (3.20)
\]
\[
\frac{F_{(i)}}{k_B T} \simeq - \ln (P_{S1}(R_2, N)\pi a^2), \quad (3.21)
\]
\[
\frac{F_{(j)}}{k_B T} = - \ln P_{S}(R_2, N), \quad (3.22)
\]

where the various probabilities \(P\) are defined in Eqs. (3.4), (3.5) (3.9) and (3.10).

### 3.2 Translocation process

#### 3.2.1 Free energy profile of translocation

The chain diffuses from the donor spherical compartment of radius \(R_1\), through a cylindrical channel of diameter \(2a\) and length \(M\), to the receptor compartment of radius \(R_2\). We construct the free energy profile as a function of \(q\), the number of segments inside the receptor.

First, the anchoring distances \(c, c_1, c_2\) and \(c_3\) have to be determined. They are fixed by the continuity of the translocation free energy profile. The conditions are

\[
F_{(b)} = F_{(c)}(1), \quad (3.23)
\]
\[
F_{(d)}(M) = F_{(f)}(M, 1), \quad (3.24)
\]
\[
F_{(f)}(M, N - M - 1) = F_{(g)}(N - M), \quad (3.25)
\]
\[
F_{(h)}(N - 1) = F_{(i)}. \quad (3.26)
\]

Solving the equations gives
\[ c = \frac{R_1}{2} \frac{f_{S1}(R_2, N - M)}{f_{S1}(R_2, N - M - 1) f_{S1}(R_1, 1)} \] {\text{(3.27)}}

\[ c_1 = \frac{M}{16} \frac{f_{S1}(R_1, N)}{f_{S1}(R_1, N - 1) f_{C1}(1)} \] {\text{(3.28)}}

\[ c_2 = \frac{M}{16} \frac{f_{S1}(R_2, N)}{f_{C1}(1) f_{S1}(R_2, N - 1)} \] {\text{(3.29)}}

\[ c_3 = \frac{R_2}{2} \frac{f_{S1}(R_2, N - M)}{f_{S1}(R_2, N - M - 1) f_{S1}(R_1, 1)} \] {\text{(3.30)}}

\( f_{S1}, f_{C1} \) and \( f_{C2} \) are defined in Eqs. (3.4), (3.9), and (3.10), respectively. Note that this is not the only way to determine the anchoring distances. The number of segments inside the channel can be larger (but not smaller) than \( M \). However, a change from \( M \) to \( 2M \) led to only 0.03\% change in the values of \( c \) and \( c_3 \) for \( R_1 = 30, R_2 = 30, N = 300, M = 30 \) and \( a = 3 \) while \( c_1 \) and \( c_2 \) are unaffected. The results presented in this chapter are based on Eqs. (3.27)–(3.30). The ground state dominance approximation [30] here yields unreasonably large anchoring distances as much as \( R_1/2 \) or \( M/2 \).

Consider a chain that is about to enter the receptor compartment so that one of the chain ends is fixed at the end of the channel. Two situations are possible. (1) There are \( p \) segments in the channel and \( N - p \) segments in the donor (Fig. 3.1(d)), and (2) the whole chain resides within the channel (Fig. 3.1(e)). For situation (1), there are \( N - M \) states for partitioning into the channel and the donor compartment, because \( M \leq p \leq N - 1 \). The probability of each conformation is calculated in the previous sections. Since the free energy \( F/k_B T = -\ln P \) and \( P \) itself is the probability of the state, the average free energy of a chain that anchors at the entrance of the receptor compartment is

\[ \overline{F_{an}} = -\frac{\sum_{p=M}^{N-1} P(d)(p) \ln P(d)(p) - P(e) \ln P(e)}{\sum_{p=M}^{N-1} P(d)(p) + P(e)} \] {\text{(3.31)}}

Next, we calculate the free energy of translocation when there are \( q < N - M \) segments in the receptor compartment. Two conformations as depicted in Figs. 3.1(f)
and (h) are possible. To map the translocation into a single variable process which is completely specified by \(q\), we assume that the process is in quasi-equilibrium so that when the chain takes the conformation in Fig. 3.1(f), all possible values of \(p\) are accessible at each step \(q\). The average free energy is then

\[
\overline{F}_1(q) = -\frac{\sum_{p=M}^{N-q-1} P_{(f)}(p, q) \ln P_{(f)}(p, q) - P_{(h)}(q) \ln P_{(h)}(q)}{\sum_{p=M}^{N-q-1} P_{(f)}(p, q) + P_{(h)}(q)}.
\]  

(3.32)

For \(q \geq N - M\), the free energy of translocation is given by \(F_{(h)}(q)\) because the chain is no longer long enough to partition into the donor compartment.

Fig. 3.2 shows the free energy profiles of translocation. The free energies labeled in the figure are defined in Eqs. (3.12), (3.13), (3.31), (3.32), (3.20), (3.21) and (3.22). A typical translocation process goes as follows. In the beginning, the chain resides in the donor compartment with both chain ends free and the free energy is \(F_a\) (Eq. (3.12)). Then one of the chain ends finds the entrance of the channel and is fixed at the interface between the donor compartment and the channel. The free energy increases to \(F_b\) (Eq. (3.13)) due to the loss in the translational entropy of the chain end. The translocation into the receptor compartment starts when a chain end is fixed at the interface between the channel and the receptor compartment. At the present state, the chain can either partition between the channel and the donor (Fig. 3.1(d)), or reside completely in the channel (Fig. 3.1(e)), with average free energy \(\overline{F}_{an}\) (Eq. (3.31)). For \(0 \leq q < N - M\), where \(q\) is the number of segments inside the receptor compartment, the chain can either have \(p\) segments inside the channel and \(N - p - q\) segments inside the donor compartment (Fig. 3.1(f)), or \(N - q\) segments completely inside the channel (Fig. 3.1(h)). The average free energy of translocation \(\overline{F}_1(q)\) is given by Eq. (3.32). For \(N - M \leq q \leq N\), the chain end is not long enough to partition into the donor compartment, and the free energy of translocation is \(F_{h}(q)\), given by Eq. (3.20). The translocation process ends when all segments translocate into the receptor compartment, with one chain end fixed at the interface between the
channel and the receptor compartment, with free energy \( F_i \) (Eq. (3.21)). Finally, the chain resides completely in the receptor compartment with both ends free, with free energy \( F_j \) (Eq. (3.22)). In Figs. 3.2(a) and (b), the shape of the free energy profile shows the transition from the state (f) to the state (h) as depicted in Fig. 3.1. State (h) is more favorable whenever \( N - q \) is small or the channel \((M\) and \(a\)) is large. Note that the free energy curve is not symmetric even when \( R_1 = R_2 \) because of the definition of \( q \).

When the translocation starts, the chain takes the conformation in Fig. 3.1(b) so that it is at the entrance of the channel. It takes \( \tau_1 \) to translocate to the conformation in Fig. 3.1(d) or (e) so that the leading chain end is at the end of the channel. The free energy change is

\[
F^+ = \mathcal{F}_{an} - F_{(b)}.
\]

\( F^+ \) is the free energy barrier for the chain to pass through the channel and reach the receptor compartment. Its dependence on different variables is shown in Fig. 3.3. Crossover behaviors can be seen in Figs. 3.3 (a) and (b). For small values of the cylinder radius \(a\) and the channel length \(M\), state (d) is preferred. State (e) dominates for larger values of \(a\) and \(M\).

### 3.2.2 Average translocation time

The translocation time is defined as the time needed for the polymer to go from the state that only one segment enters the channel (Fig. 3.1(b)) to the state that only one segment is left in the channel (Fig. 3.1(i)). The translocation can be considered as a two-step process, (1) the chain overcomes the free energy barrier \( F^+ \) of channel-filling and (2) the chain fills up the receptor with the number of segments from \( q = 1 \) to \( q = N - 1 \). The average translocation time of step (1) is

\[
\langle \tau_1 \rangle \simeq \frac{1}{\kappa_0} \exp(F^+/k_BT).
\]

(3.34)
where $k_0$ is a phenomenological rate constant, representing the friction between one monomer and the channel.

For step (2), the average translocation time is calculated from the mean first passage time of the corresponding Fokker-Planck equation [121, 85, 42]

$$\frac{\partial W_m(t)}{\partial t} = \frac{\partial}{\partial m} \left( \frac{k_0}{k_B T} \frac{\partial F(m)}{\partial m} W_m(t) + k_0 \frac{\partial W_m(t)}{\partial m} \right),$$

(3.35)

where $W_m(t)$ is the probability that the polymer has $m$ segments in the receptor compartment at time $t$. It can be shown that the mean first passage time $\tau$ satisfies the equation [42]

$$k_0 \tau''(z) - \frac{F'(z)}{k_B T} k_0 \tau'(z) = -1,$$

(3.36)

where the prime denotes the derivative with respect to $z$, the initial position in free energy profile $F(q)$. The solution is

$$\tau(z) = \frac{1}{k_0} \int_z^s dx \int_r^x dy \exp \left( F(x) - F(y) \right) / k_B T,$$

(3.37)

with reflecting and absorbing boundary conditions at $r$ and $s$ respectively. Therefore the average translocation time of step (2) is [85]

$$\langle \tau_2 \rangle = \frac{1}{k_0} \int_1^{N-1} dx \int_1^x dy \exp \left( F(x) - F(y) \right) / k_B T,$$

(3.38)

where $F(q)$ is given by $F_1(q)$ (Eq. (3.32)) for $q < N - M$ and by $F(q)$ (Eq. (3.20)) for $q \geq N - M$. The translocation time is calculated numerically.

The average total translocation time $\langle \tau_1 \rangle + \langle \tau_2 \rangle$ varies non-monotonically with the channel length $M$. Fig. 3.4 shows some typical results for the average total translocation time. The presence of the channel increases the free energy barrier $F^+$ of entering the receptor compartment, by making the conformations in Figs. 3.1 (d)
and (e) less entropically favorable. The net effect is that $\langle \tau_1 \rangle$ increases with $M$, as shown in Fig. 3.4(a). The chain has to partition into the channel before entering the receptor, taking a more confined state. Nevertheless, for a Gaussian chain, the driving force of translocation increases with the increase in entropy per step of movement. The channel increases the entropy difference, and hence the driving force, to enter the more spacious receptor compartment, resulting in the decrease of $\langle \tau_2 \rangle$ with $M$. The interplay between these two effects results in the existence of a minimum in the average total translocation time. The location of this minimum depends on the parameters $R_1, R_2, a$, and $N$, as demonstrated in Figs. 3.4(a)–(c).

The finite size of the channel leads to two different driving forces. First, if the chain is long enough, the conformations of the chain changes from (f) to (h) as translocation proceeds. For the conformation (f), the driving force is given by the entropy difference between the donor and the receptor if the number of segments inside the channel remains constant, as it is likely to happen for a long chain. For the conformation (h), the tail of the chain resides completely in the channel. The driving force is given by the favorable entropy difference between a segment in the channel and that in the receptor. In general, this driving force is larger than the force involved in transferring segments between compartments. These two driving forces are evident in the free energy landscape, in terms of the change in the slope with respect to $q$. Note that the change occurs well before $m = N - M$ if the diameter of the channel is large.

### 3.2.3 Translocation time distribution

The probability distribution of the translocation time $\tau_2$ is

$$
P(\tau_2) = -\left(\frac{k}{k_BT} \frac{\partial F(m)}{\partial m} W_m(\tau_2) + \frac{\partial W_m(\tau_2)}{\partial m}\right)_{m=N},
$$

(3.39)
where the probability distribution of the polymer $W_m(t)$ is obtained by solving Eq. (3.35) numerically using the Crank-Nicolson method [104]. The result is shown in Fig. 3.5. The distribution of the translocation time is a non-Gaussian peaked distribution with a long tail, with the character of drift-diffusion stochastic processes. $\tau_2$ increases with the channel radius $a$ because the entropy gain of escaping the channel into the receptor decreases. Note that the total translocation time $\tau_1 + \tau_2$ decreases as $a$ increases, see Figs. 3.4(c) and (d). The reason is that $\tau_1$ decreases as $a$ increases, and this effect dominates over $\tau_2$.

### 3.3 Discussion and conclusion

A formalism of polymer translocation between two spherical compartments through a cylindrical channel is developed. The average translocation time and its distribution are obtained. As a test case, the translocation of a Gaussian chain, which is purely entropy driven, is studied. The merit of the system is that exact formulas of the confinement free energies can be obtained. The translocation is composed of two steps, (1) worming through the channel to reach the entrance of the receptor compartment and (2) translocating all segments into the receptor.

The driving force of translocation may change during step (2). If the number of segments inside the channel remains the same, from Eq. (3.4), the driving force of the step (2) is

$$f_2 \sim \pi^2 \frac{1}{6} \left( \frac{1}{R_1^2} - \frac{1}{R_2^2} \right)$$

under the ground state dominance approximation, which is qualitatively correct. If the space of the channel is large compared to the radius of gyration of the remaining $N - q$ segments, the segments would prefer to reside in the channel completely. Conformation (h) rather than (f), see Fig. 3.1, would dominate the translocation. Using Eqs. (3.4) and (3.9), the driving force of the step (2) becomes
\[ f'_2 \sim \frac{1}{6} \left( \frac{\mu_0^2}{a^2} + \frac{\pi^2}{M^2} - \frac{\pi^2}{R_2^2} \right), \quad (3.41) \]

where \( \mu_0 \approx 2.4048 \) is the first zero of Bessel function \( J_0(x) \). \( f'_2 \) is larger than \( f_2 \) in the dimension scale we are interested in. Therefore the limited space of the channel results in a larger driving force. Likewise, the longer the channel, the more dominant the conformation (h) in step (2), and the shorter \( \tau_2 \). Of course \( \tau_1 \), the time for filling up the channel, increases with the channel length. As a result there is a minimum in the translocation time as a function of \( M \), solely caused by entropy. Recently, this effect was observed in a computer simulation \([69]\), as was also predicted for an infinitely thin cylinder \([88, 60]\).

Our calculation of \( \tau_1 \) has to be used with caution. In our prescription, the free energy increase associated with the filling of the channel \( F^+ \) (Eq. 3.33) depends on \( M \) non-monotonically, because we do not consider the intermediate states of Fig. 3.1(c) in obtaining \( F_{an} \). As \( M \) increases, the conformation of the final state of step (1) changes from (d) to (e) in Fig. 3.1. \( F^+ \) increases linearly with \( M \) at first, comes to a maximum and then converges to a constant value at large \( M \). Our estimation of \( \tau_1 \), using Eq. (3.34), is valid only for the linear region of \( F^+ \). For large \( M \), \( F^+ \) becomes constant and Eq. (3.34) predicts that \( \tau_1 \) is independent of \( M \), which is clearly wrong. Instead of estimating \( \tau_1 \) by assuming chain anchoring with free energy \( F_{an} \), a translocation coordinate of the chain entering the channel (Fig. 3.1(c)) is necessary. For very long channels, the whole chain may translocate into the channel without reaching the end of the channel. The chain then diffuses along the channel until it reaches the end, or returns to the donor compartment. The translocation times shown in Fig. 3.4 are all calculated in the linear region of \( F^+ \). Our calculation of \( \tau_2 \) is valid for any \( M \).

The average total translocation time changes non-linearly with \( N \) for short chains and linearly for long chains, as shown in Fig. 3.4(d). The non-linearity in the limit
of short chains mainly arises from $\langle \tau_1 \rangle$. However, for long chains, $\langle \tau_1 \rangle$ is essentially independent of $N$, as the change in free energy of the conformations (d) is not significant. On the other hand, for short chains, the conformations change from Fig. 3.1(e) to Fig. 3.1(d) as $N$ increases, and the free energy increases rapidly with $N$. The transition can be seen in Fig. 3.3(d). It is to be emphasized that in our formalism, the friction coefficient of every monomer is assumed to be the same, independent of $N$, whether it is within the spheres or the cylinder. For longer chains, $\langle \tau_2 \rangle$ dominates over $\langle \tau_1 \rangle$ and now $\langle \tau_2 \rangle$ increases linearly with $N$, as expected for a process dominated by drift.

The formalism is applied to only a Gaussian chain, in the interest of getting exact results. These results need to be appended by a proper account of excluded volume effect, in order to compare with experimental systems. Even for an isolated spherical cavity, there is a continuing debate [20, 56, 111] for the form of the confinement free energy. However, when the excluded volume effect is removed, our exact results must be recovered.

Our estimation of the translocation time by calculating $\langle \tau_1 \rangle$ and $\langle \tau_2 \rangle$ separately and adding them is valid only in certain situations. Under this approximation, the chain is not allowed to return from step (2) to step (1). As shown in the Appendix, the approximation is good for a free energy profile of step (2) that is significantly downhill without any initial barrier, such as those in Figs. 3.2(c) and (d). In this case the upper bound of the correction factor to $\langle \tau_1 \rangle$ is related to the inverse of the gradient of the free energy profile. If the free energy profile is initially flat up to $x_m$ and then significantly downhill, $\langle \tau_1 \rangle$ would need to be corrected by a factor up to $\sim x_m$. Finally, if there is an initial free energy barrier of height $\Delta f$, then the upper bound of the correction factor to $\langle \tau_1 \rangle$ is $\sim \exp(\Delta f)$. Physically, the polymer is unlikely to retreat from step (2) to step (1) in the first case, so that the approximation
is valid. For the second and third cases, returning to step (1) cannot be ignored. The details of arriving at these conclusions are given in the Appendix.

The accuracy of the free energy expressions given in this chapter depends on the geometry of the system. Explicitly, we require that $c/R \ll 1$ and $c/(M/2) \ll 1$, where $c$ is the anchoring distance, $R$ is the radius of the corresponding spherical compartment, and $M$ is the length of the channel. These conditions are satisfied when $(a/R)^2 \ll 1$ and $a^2/RM \ll 1$, where $a$ is the radius of the channel. It should be noted that all of these dimensions are in units of the Kuhn length $b$.

It must be noted that the present formalism relies on the quasi-equilibrium assumption so that the free energy during each step of translocation is given by the equilibrium conformational entropy of the chain. There exists another regime where the translocation time is shorter than or comparable to the relaxation time of the chain. Simulations and scaling arguments addressing this regime show conflicting conclusions and the results are under active discussion [27, 57, 71, 69, 100, 131, 35, 36, 110]. In the present chapter, we have focused our attention on situations satisfying the quasi-equilibrium assumption.

Further, it must be noted that we ignored the possibility that the chain may enter or leave the channel by forming hairpins. The wider the channel, the more probable is the formation of hairpins. The occupancy of chain entrance into the channel as hairpins is influenced by the extent of excluded volume effect and the bending elasticity of the polymer. However, in the present chapter for Gaussian chains, we have assumed that the entry of the polymer into the channel and its exit occur only through chain ends.

### 3.4 Error estimation of two-step translocation time

In this chapter, we separate the translocation time into two parts $\tau = \tau_1 + \tau_2$. In fact, the average translocation time $\langle \tau \rangle = \langle \tau_1 + \tau_2 \rangle \neq \langle \tau_1 \rangle + \langle \tau_2 \rangle$ so it is necessary to
estimate the error introduced by calculating $\langle \tau_1 \rangle$ and $\langle \tau_2 \rangle$ separately. Physically, this approximation does not allow the translocating polymer to go back from step (2) to step (1). Intuitively, this approximation is valid when the free energy profile of step (2) is monotonically downhill with a large gradient, and becomes worse when the free energy profile is flat, uphill or has a free energy barrier. We estimate the correction quantitatively as follows.

We consider a hypothetical free energy profile shown in Fig. 3.6(a),

$$F(x) = \begin{cases} 0, & -\delta \leq x < 0 \\ \Delta F + G(x), & 0 \leq x \leq N \end{cases}, \quad (3.42)$$

where $\Delta F$ is the free energy barrier in step (1), $G(x)$ is an arbitrary free energy profile satisfying $G(0) = 0$, and $\delta$ is a small arbitrary positive constant representing the length of the trapped state. In our calculations, we treat the translocation as a two step process,

$$\langle \tau_1 \rangle = \frac{1}{k_0} \exp(\Delta F), \quad (3.43)$$

$$\langle \tau_2 \rangle = \frac{1}{k_0} \int_0^N dx \int_0^x dy \exp \left( G(x) - G(y) \right) / k_B T. \quad (3.44)$$

On the other hand, the average total translocation time $\langle \tau \rangle = \langle \tau_1 + \tau_2 \rangle$ is given by

$$\langle \tau \rangle = \frac{1}{k_0} \int_{-\delta}^N dx \int_{-\delta}^x dy \exp \left( F(x) - F(y) \right) / k_B T, \quad (3.45)$$

with reflecting and absorbing boundary conditions at $x = -\delta$ and $x = N$, respectively. Eq. (3.45) can be rewritten as

$$\langle \tau \rangle = \frac{\delta^2}{2k_0} + \left[ \delta \int_0^N \exp \left( \frac{G(x)}{k_B T} \right) dx \right] \langle \tau_1 \rangle + \langle \tau_2 \rangle \quad (3.46)$$
by using Eqs. (3.43) and (3.44). The first term with $O(\delta^2)$ is negligible since $\delta$ is small. Immediately, we identify that $\langle \tau_1 \rangle$ has to be corrected by a factor $C\delta$, where 

$$C = \int_0^N \exp(G(x)/k_B T) dx.$$

The correction factor accounts for the fact that the polymer can retreat from step (2) to the trapped state step (1), which is missing in our calculations. Note that $\langle \tau_2 \rangle$ is unchanged and still given by the reflecting boundary condition at $x = 0$. This is because after falling back to the trapped state step (1), the polymer has to start step (2) again from $x = 0$. The net effect is that the time spent in step (2) is exactly the same as if there is a reflecting boundary at $x = 0$, and only $\langle \tau_1 \rangle$ has to be modified because of multiple visits.

It is instructive to estimate the correction factor $C$ for some typical free energy profiles. If step (2) is purely diffusive, such that $G(x) = 0$ as shown in Fig. 3.6(b), $C = N$. Physically, this means that the probability of the polymer going from step (2) to step (1) is proportional to the length of step (2), $N$. Before touching $x = N$, the polymer on average retracts back to step (1) for $N$ times. If the free energy profile of step (2) is a straight line with slope $\mu$ as shown in Figs. 3.6(c) and (d), $C = 1/|\mu|$ for $\mu < 0$ and $|\mu|N \gg 1$ while $C = \exp(\mu N)/\mu$ for $\mu > 0$ and $\mu N \gg 1$. The results have clear physical meanings. For $\mu < 0$ and $|\mu|N \gg 1$, the free energy profile of step (2) is straightly downhill so the polymer is unlikely to go from step (2) back to step (1). For $\mu > 0$ and $\mu N \gg 1$, the free energy profile of step (2) is straightly uphill so that step (2) also acts like a free energy barrier with height $\mu N$. The total translocation time is then $\langle \tau \rangle \sim \exp(\Delta F + \mu N)$. Finally, we estimate $C$ when the free energy profile of step (2) has a maximum at $x_m$ (Fig. 3.6(e)), or it is diffusive until $x_m$ and goes down with a constant slope $\mu$ (Fig. 3.6(f)). For the profile shown in Fig. 3.6(e) with $\mu_1 x_m \gg 1$ and $|\mu_2| (N - x_m) \gg 1$, $C = \exp(\mu_1 x_m)(1/\mu_1 + 1/|\mu_2|)$. While for the profile shown in Fig. 3.6(f) with $|\mu|(N - x_m) \gg 1$, $C = x_m + 1/|\mu|$. Thus when there is a significant free energy barrier in step (2) with height $\mu_1 x_m$, $\langle \tau_1 \rangle$ is modified as if the total barrier height is $\Delta F + \mu_1 x_m$, so that $\langle \tau_1 \rangle \sim \exp(\Delta F + \mu_1 x_m)$ while $\langle \tau_2 \rangle$ is
unchanged. For a profile that is initially flat until $x_m$ and then significantly downhill, $\langle \tau_1 \rangle$ is corrected by a factor $\sim x_m$.

Finally, we note that our estimation of the correction factor $C$ to $\langle \tau_1 \rangle$ cannot be applied to our problem in a straightforward manner, because we have defined the starting point of step (1) as the conformation that a chain end has just entered the channel (Fig. 3.1(b)). There is no reason to assume that the chain cannot retreat back to the donor compartment (Fig. 3.1(a)). In other words, instead of reflecting boundary at $x = -\delta$, we should use the absorbing boundary condition. Then the average translocation time $\langle \tau \rangle$ is given by the conditional mean first passage time, starting from somewhere near $x = -\delta$ and ending at $x = N$. However, the time calculated this way is always shorter than that given by a reflecting boundary condition at $x = -\delta$ as assumed above. Therefore the above estimation serves as an upper bound of the correction to the total translocation time. Moreover, $\langle \tau_2 \rangle$ will also be modified due to the fact that some reflections at $x = 0$ belong to unsuccessful translocations which the chain goes from step (2) to step (1) and then back to the donor compartment. Nevertheless, the complications can be avoided if we define the beginning of step (1) as the trapped conformation shown in Fig. 3.1(a) instead Fig. 3.1(b). Then the correction is exact and can be calculated numerically for an arbitrary free energy profile.
Figure 3.1. All possible conformations of a polymer of $N$ segments during translocation. $p$ and $q$ are defined as the number of segments inside the channel and the receptor compartment, respectively.
Figure 3.2. The free energy profiles of translocation of a chain with length $N = 300$ and donor compartment radius $R_1 = 30$. The translocation coordinate $q$ denotes the number of segments inside the receptor compartment. The various free energy labels $F$ in (a) are defined in Eqs. (3.12), (3.13), (3.31), (3.32), (3.20), (3.21) and (3.22). (a) $M = 30$, $R_2 = 30$ and $a = 3$. (b) $M = 30$, $R_2 = 60$ and $a = 3$. (c) $M = 30$, $R_2 = 60$ and $a = 5$. (d) $M = 45$, $R_2 = 60$ and $a = 3$. 
**Figure 3.3.** The free energy barrier $F^+$ of channel-filling as a function of different parameters. (a) $M = 30$, $R_1 = R_2 = 30$ and $N = 300$. (b) $a = 5$, $R_1 = R_2 = 30$ and $N = 300$. (c) $a = 5$, $R_2 = 30$, $M = 30$ and $N = 300$. (d) $a = 3$, $R_1 = 10$, $R_2 = 15$ and $M = 6$. 
Figure 3.4. Total translocation time $\langle \tau_1 \rangle + \langle \tau_2 \rangle$. (a) $a = 3$, $R_1 = 30$, $N = 300$. Solid lines: $R_2=30, 40, 50$ and $60$. Dotted line: $\langle \tau_1 \rangle$ for $R_2 = 60$. Dashed line: $\langle \tau_2 \rangle$ for $R_2 = 60$. (b) $N = 300$, $a = 5$, $R_1 = 30$, $R_2=30, 40, 50$ and $60$. (c) $N = 300$, $R_1=30$, $R_2=50$, $a=3, 4, 5$ and $6$. (d) $R_1 = 10$, $R_2 = 15$, $M = 6$, $a = 3, 4$ and $5$. 
Figure 3.5. Translocation time distribution of $\tau_2$ for $a = 3, 4, 5$, respectively. $R_1=30$, $R_2=50$, $M=10$ and $N=300$. 
Figure 3.6. (a) Free energy profiles of translocation involve crossing a free energy barrier $\Delta F$ (step (1)) and an arbitrary profile $G(x)$ (step (2)). The corresponding correction factors to $\langle \tau_1 \rangle$ (in the limits explained in the text) are (b) $C = N$, (c) $C = 1/|\mu|$, (d) $C = \exp(\mu N)/\mu$, (e) $C = \exp(\mu_1 x_m)(1/\mu_1 + 1/|\mu_2|)$ and (f) $C = x_m + 1/|\mu|$, respectively.
CHAPTER 4
SCALING THEORY OF POLYMER TRANSLOCATION INTO CONFINED REGIONS

4.1 Introduction

Polymer translocation from a spacious region into a confined region is ubiquitous in biology and nanotechnology. For example, protein molecules translocate to other cellular compartments through narrow membrane channels [5], polymer molecules are driven from solutions into gel matrices in electrophoresis [128], and DNA/RNA molecules are forced into micro- or nano-fluidic channels and protein channels [49, 48, 106, 117, 58, 2, 83, 65, 119, 31, 76, 95].

The phenomenon of polymer translocation is controlled by many factors [89]. The chief among the contributing factors is the entropic barrier arising from the reduction of polymer conformations in the confined regions. Therefore, a driving force is required for successful translocation. While the entropic barrier imposed by the restricted space is common, the driving force can take different forms, or a combination of them. For instance, DNA molecules can be driven into a confined region by an electric field, a pressure driven flow [117], or an electro-osmotic flow due to the surface charge of the confined region [44, 45]. In biological cells, the specific binding of specific signal sequences initiates the entry of a protein molecule to a narrow transmembrane channel [5].

Microfluidic channels with periodic spacious and confined regions [49, 48] demonstrate such interplay between entropic barrier and driving force. In these devices, the migration time of DNA molecules showed dramatic molecular weight dependence,
with longer DNA molecules traveling faster under applied electric fields [49, 47, 48, 50]. A typical system consists of periodic strips of deep and shallow regions with heights 1.5 - 3.0 µm and 75 - 100 nm, respectively. The DNA molecules, in kilobase-pairs length scale, are driven through the periodic constrictions by an applied electric field, as depicted in Fig. 4.1. The molecules are unrestricted laterally so that they are confined between two horizontal parallel surfaces placed at distances much larger than the sizes of DNA molecules. For each period, which consists of a deep region followed by a shallow region, a DNA molecule first travels through the deep region with time \( \tau_1 \). It then stops momentarily at the interface between the deep and shallow regions, and takes \( \tau_2 \) to enter the shallow region. Finally, it takes \( \tau_3 \) to finish passing through the shallow region. A DNA molecule thus takes time \( \tau_1 + \tau_2 + \tau_3 \) to travel through one period [101]. Because of the height difference, the electric field in deep regions is much lower than that in shallow regions [122, 50], making \( \tau_3 \) negligible compared to \( \tau_1 + \tau_2 \). In general, confinement effect is unimportant in deep regions so that the electrophoretic behavior of the DNA molecules is the same as that in a free solution. Consequently, \( \tau_1 \) is independent of molecular weight of the DNA molecule, as is well-known in the capillary electrophoresis [118]. Therefore, the molecular weight dependence comes essentially from \( \tau_2 \), the time for the DNA molecule to enter the shallow region from the deep region. The above description is consistent with fluorescence microscopy observations [48].

In order to understand the molecular basis of the above described experimental results, several theoretical attempts [49, 50, 25, 112] and computer simulations [122, 25, 120, 101, 63] have been reported in the literature. Han et al. [49, 48] considered DNA translocation into the confined region as an activation process and assumed that translocation time was in the Arrhenius form \( \tau_2 \sim \tau_0 \exp(F^*) \), where \( F^* \sim 1/E \) is the free energy barrier of translocation into the shallow region and \( E \) is the electric field in the shallow region. They argued that the activation probability was proportional
to the contact area between the slit and DNA molecule $RD$, where $R \sim N^\nu$ is the size of the DNA with $N$ segments, and $D$ is the height of the shallow region. Using size exponent $\nu \simeq 0.59$ for a flexible self-avoiding chain, the prefactor $\tau_0 \sim 1/RD \sim 1/N^\nu D$. Thus $\tau_2$ always decreases with chain length $N$ in this model. Sebastian et al. [112] drew analogy between the present problem and the Kramers problem of a phantom polymer inside an asymmetric double potential well. In this model, $\tau_0 \sim 1/N\sqrt{E}$ if the polymer crosses the free energy barrier between the two wells in hairpin conformations, and $\tau_0$ is independent of $N$ if the polymer crosses the barrier in linear conformations [122]. Note that these theories are derived without accounting for the statistics of the DNA molecules inside the particular geometry of the system. By performing Monte Carlo simulations, Chen et al. [25] found that the free energy barrier for a Gaussian chain to enter the shallow region increased with chain length and approached a constant value in low electric fields, while decreased with chain length in high electric fields. Streek et al. [120] found that, by increasing the height of the shallow regions in their simulations, a fast and a slow state of DNA migration through the periodic regions resulted from non-equilibrium dynamics. Tessier et al. [122] performed detailed Monte Carlo simulations with system dimensions similar to that of the experiments [49, 48]. In agreement with experimental observations, the mobility of the DNA increased with molecular weight in their simulations. They attributed this effect mainly to the deformation of the DNA molecules taking place during $\tau_2$. Panwar et al. [101] simulated the system with a Gaussian chain and found that the mobility of the chain could change non-monotonically with molecular weight, depending on the relative magnitudes of $\tau_1$, $\tau_2$ and $\tau_3$. Recently, Lee et al. [63] studied the system with both linear and star-branched polymers using Brownian dynamics simulations of bead-spring chain models, and found that the size of the polymer appeared to be the determining factor of the total migration time.
In this chapter, we present an analytically tractable scaling theory and calculate $\tau_2$ using the statistics of self-avoiding chains under confinement. Two modes of polymer entry to the confined region are considered: linear translocation which is led by a chain end, and hairpin translocation which is led by a hairpin of chain segments. By constructing their free energies and the corresponding translocation times, we find that the $N$ dependence of $\tau_2$ is dictated by the statistics of chain tail(s) in the spacious region. For self-avoiding polymers, $\tau_2 \sim N^{0.31}$ for linear translocations and $\tau_2 \sim N^{-0.38}$ for hairpin translocations. Our calculation suggests that the decrease of translocation time with increasing chain length observed in experiments can be ascribed to the dominance of hairpin translocations. Furthermore, we predict that translocation time increases with chain length if linear translocations dominate, which is favored by short chains.

4.2 Theory

We consider a polymer chain entering a confined region composed of two parallel planes (shallow region) from a spacious region (deep region). When the polymer molecule is trying to enter the shallow region, part of the chain partitions into the shallow region and the rest of the chain still resides in the deep region, as shown in Fig. 4.2. The free energy of the whole chain is the sum of free energies of polymer segments inside the shallow region and those inside the deep region.

The free energy of segments in the shallow region has two competing contributions: (1) energy gain due to the electric field and (2) conformational entropy lost due to the geometric confinement. We assume that the electric field is negligible and does not perturb the chain conformation in the deep region. The deep region is assumed to be wide and deep enough that the confinement effects on the chain is unimportant in the spacious region.
A typical free energy profile [87, 89] for a chain entering the shallow region is illustrated in Fig. 4.3. The translocation process is described by the number of polymer segments in the shallow region $m$ with free energy $F(m)$. The process is at first unfavorable because of the entropic confinement imposed by the shallow region. As the chain proceeds further, after a critical number of inserted segments $m^*$, the critical insertion size, the electrostatic free energy gain dominates and the translocation becomes favorable. Therefore, the free energy barrier of translocation is $F^* = F(m^*) - F(0)$. In this section, we construct the free energy expression $F(m)$ for linear and hairpin conformations. In the rest of the chapter, the units of length, energy and force are the Kuhn length $l$, thermal energy $k_B T$, and $k_B T/l$, respectively. Since we use only scaling arguments in various key steps, the equality sign in the following equations omits all numerical prefactors arising from the nonuniversal aspects of the problem.

4.2.1 Confinement free energy in shallow region

Consider $m$ polymer segments in the shallow region with height $D$ and constant electric field $E$. In view of the simplicity and success of the blob model of polymer chains in confined regions [30], we adopt the blob model. We assume that the segments inside the shallow region form a series of blobs with diameter $D$ each, and the electric field aligns the blobs in linear but does not influence the segment statistics inside them. There are $g = D^{1/\nu}$ segments in each blob, and within the blobs, the chain statistics is unaltered by confinement[30]. Therefore, the number of blobs is $n = m/g = m/D^{1/\nu}$. It is to be noted that the size exponent $\nu$ is 0.59 for the experimental conditions of Refs. [49, 47, 48, 50], corresponding to the statistics of the self-avoiding flexible chains. The free energy of confinement is [30]

$$F_c(m) = \frac{m}{D^{1/\nu}}. \quad (4.1)$$
The blob picture is satisfactory only if the persistence length $l_p$ is much smaller than $D$ and if it contains enough number of segments $m$ to obey the self-avoiding walk statistics [82].

4.2.2 Electrostatic free energy in shallow region

The conformations of a chain during linear and hairpin translocations are shown in Fig. 4.2. Each segment in the shallow region experiences an electrostatic force $f = qE$, where $q$ is the net charge of each segment, and $E$ is the magnitude of the uniform electric field in the shallow region. The product $qE$ is positive. Each blob has $g = D^{1/\nu}$ segments.

Let us first calculate the electrostatic energy when there is only one blob (with $g$ segments) inside the shallow region. The electrostatic energy gain is the sum from each segment and can be written as $E_1 = f \sum_{i=1}^{g} x_i$, where $x_i$ is the distance of the $i$-th segment away from the interface between the deep and shallow regions. A blob of length $D$ ensures that $0 \leq x_i \leq D$. Alternatively, the sum can be expressed as $\sum_{i=1}^{g} x_i = D \sum_{i=1}^{g} (1/2 + \Delta x_i)$ where $-1/2 \leq \Delta x_i \leq 1/2$ is the normalized position of the $i$-th segment relative to the center of the blob. If the distribution of segments is symmetric about the center of the blob, then $\sum_{i=1}^{g} \Delta x_i = 0$. By making this assumption, the electrostatic energy of the blob becomes $E_1 = gfD/2$. Now suppose that there are two blobs in series aligning in the direction of the electric field. Starting with one blob, the situation is equivalent to translating all segments in that blob by a distance $D$ in the direction of the electric field and creating another blob in the place of that blob. Thus, starting with one blob, translating it into the new site of the blob and then creating another blob in the place of the original one, the electrostatic energy gain is $E_1 + gfD$. It follows that the electrostatic energy for two blobs is $E_2 = E_1 + E_1 + gfD$. Similarly, when there are $n$ blobs inside the shallow region, the electrostatic energy change with respect to $n - 1$ blobs is equivalent to translating the
first blob by a distance \((n-1)D\), so that it moves to the end, and creating another blob in place of the first blob. The electrostatic energy of the \(n\) blobs is therefore
\[ E_n = E_{n-1} + E_1 + g f(n-1)D = n E_1 + g f D [ (n-1) + (n-2) + \ldots + 1 ] = n^2 g f D / 2. \]
Since \(n = m/g\) and \(g = D^{1/\nu}\), the electrostatic free energy of \(m\) segments assuming a linear conformation of blobs in the shallow region is
\[ F_{fl}(m) = -E_n = - \frac{f m^2}{2 D^{1/\nu-1}}. \]  

The same result can be derived by an equivalent argument. By assuming that \(E\) is uniform, the energy of a blob (with \(g\) segments) at a distance \(Dn'\) from the entrance to the shallow region is \(-f g D n'\). Therefore, for \(n\) blobs, the net gain due to the electric field is \(-f g D \int_0^n n' dn'\), which results in the same expression as Eq. (4.2). The negative sign indicates that the electric field drives the chain into the shallow region and favors translocation by competing against the positive confinement free energy given by Eq. (4.1). The rod-like limit \((\nu = 1)\) eliminates the \(D\) dependence in Eq. (4.2), recovering the result for an infinitely narrow channel \([88]\). Note that Eq. (4.2) does not apply for Gaussian chains \((\nu = 0.5)\) (also see Ref. \([30]\)) as statistical independence among perpendicular directions demands that the electrostatic energy does not depend on \(D\). Therefore we will confine our discussion to non-Gaussian chains.

For hairpin translocations, the entry is led by segments somewhere in the middle of the chain as illustrated in Fig. 4.2(b). We assume that a hairpin of \(n\) blobs forms a loop of two linear strings of \(n/2\) blobs side by side due to the pulling by the electric field. The electrostatic energy for a hairpin conformation is then
\[ F_{fh}(m) = -2 \left( \frac{1}{2} \right) \left( \frac{n}{2} \right)^2 g f D = - \frac{1}{4} n^2 g f D = - \frac{f m^2}{4 D^{1/\nu-1}}. \]  

54
Note that it is exactly half of its linear counterpart (Eq. (4.2)). It is to be noted that in the present flexible chain model, relevant for chain lengths much larger than \( l_p \), bending energy associated with chain stiffness is neglected.

### 4.2.3 Free energies of chain tails in deep region

We model a chain tail in the deep region as a chain anchored to a hard wall in a semi-infinite half space [85]. The partition sum (with respective to the free state) of such a chain with \( j \) segments is [39, 38]

\[
Z_j \sim j^{\gamma' - 1},
\]  

(4.4)

where \( \gamma' = 0.69 \) for a self-avoiding chain. The free energy of a \( j \)-segment chain tail in the deep region is then

\[
F_a(j) = -\ln Z_j = (1 - \gamma') \ln j.
\]  

(4.5)

### 4.2.4 Free energies of linear and hairpin conformations

With the above elements, we are in a position to construct the free energy profiles of linear and hairpin translocations. A chain in a linear conformation partitions between the deep and the shallow regions in the conformation shown in Fig. 4.2(a). Using Eqs. (4.1), (4.2) and (4.4), the probability of realizing a linear conformation with \( m \) segments in the shallow region and \( N - m \) segments in the deep region is

\[
P_l(m) = (N - m)^{\gamma' - 1} \exp \left( \frac{mf^2}{2D^{1/\nu - 1}} - \frac{m}{D^{1/\nu}} \right).
\]  

(4.6)

The free energy of a state \( F \) gives the relative probability \( P \) of realizing that state, according to \( P = \exp(-F) \). The translocation process is assumed to be in quasi-equilibrium (viz., the typical relaxation times for conformational change of the poly-
mer are shorter than the typical translocation time) so that it is solely dictated by the relative free energies among various states.

For hairpin conformations, the chain enters the shallow region with segments somewhere between the two ends, forming a hairpin in the shallow region as shown in Fig. 4.2(b). If there are \( m \) segments in the hairpin and \( j \) and \( N - m - j \) segments in the two tails in the deep region, respectively, using Eqs. (4.1), (4.3) and (4.4), the probability of the hairpin conformation is

\[
P_h(m, j) = j \gamma' - 1(N - m - j) \gamma' - 1 \exp \left( \frac{f m^2}{4 D^{1/\nu - 1}} - \frac{m}{D^{1/\nu}} \right)
\]

(4.7)

It follows that the corresponding free energies of linear and hairpin conformations are

\[
F_l(m) = -\frac{f}{2} \frac{m^2}{D^{1/\nu - 1}} + \frac{m}{D^{1/\nu}} + (1 - \gamma') \ln(N - m),
\]

(4.8)

\[
F_h(m, j) = -\frac{f}{4} \frac{m^2}{D^{1/\nu - 1}} + \frac{m}{D^{1/\nu}} + (1 - \gamma') \ln j + (1 - \gamma') \ln(N - m - j),
\]

(4.9)

respectively. In general, the above free energies have a maximum as a function of \( m \), as illustrated in Fig. 4.3. Note that in writing Eqs. (4.7) and (4.9), we implicitly assume that the two chain tails in the deep region do not interact. If a monomer at the interface between the shallow and the deep regions is allowed to move up and down in the range of height \( D \), the probabilities in Eqs. (4.6) and (4.7) would have an additional front factor \( D \). It follows from free energy \( F = -k_B T \ln P \) that there would be terms of order \( \ln D \), which are ignored in Eqs. (4.8) and (4.9) for simplicity. Their effects are secondary compared to the \( D \) dependence appearing in the electrostatic and confinement terms.
4.3 Results

4.3.1 Critical insertion size

Translocation into the shallow region is initially unfavorable because of the geometric confinement with free energy $\sim m/D^{1/\nu}$. After a critical number of inserted segments $m^*$, the translocation process becomes favorable and the translocation free energy profile is dominated by the electrostatic energy $\sim -f m^2/D^{1/\nu-1}$. The translocation time is largely determined by the height of the free energy barrier $F(m^*)$ relative to the free state in the deep region. Note that this height is different for linear and hairpin translocations. For hairpin translocations, it has an additional dependence on $j$, that is $F_h(m^*, j)$. $m^*$ computed in physical distance corresponds to the critical hernia nucleation length in the simulations of Tessier et al [122]. We will refer $m^*$ as the critical insertion size for both linear and hairpin translocations.

We consider first the linear translocation. Using Eq. (4.8), the linear critical insertion size $m^*_l$ is given by solving $\partial F_l/\partial m = 0$. The solution is

$$m^*_l = \frac{1}{2fD} \left[ fDN + 1 - \sqrt{(fDN - 1)^2 + 4(1 - \gamma')D^{1/\nu+1}f} \right]. \quad (4.10)$$

For large $N$, such that $N \gg D^{(1-\nu)/2\nu} f^{-1/2}$, the second term in the square root of Eq. (4.10) becomes negligible. $m^*_l$ now takes the simple form

$$\tilde{m}^*_l \approx \frac{1}{fD}. \quad (4.11)$$

In the rest of the chapter, a variable with tilde on top denotes an approximation that is valid for large $N$.

For hairpin translocations, the critical insertion size $m^*_h$ depends on $j$, the length of a tail in the deep region. Solving $\partial F_h/\partial m = 0$ for a fixed $j$, we get

$$m^*_h(j) = \frac{1}{2fD} \left[ 2 + Df(N - j) - \sqrt{(Df(N - j) - 2)^2 + 8(1 - \gamma')D^{1/\nu+1}f} \right]. \quad (4.12)$$
Similarly, Eq. (4.12) reduces to

\[ \tilde{m}_h^* \simeq \frac{2}{fD} \]  

(4.13)

for large \( N \) and is independent of the tail length \( j \). Note that, from Eqs. (4.11) and (4.13), \( \tilde{m}_h^* \) is exactly twice \( \tilde{m}_l^* \). Since a hairpin forms a loop of two linear blobs side by side in our model, the critical length of a hairpin, in physical distance, is the same as its linear counterpart for \( N \gg D^{(1-\nu)/2\nu} f^{-1/2} \). The exact solution Eq. (4.10) converges to the asymptotic limit Eq. (4.11) for large \( N \). The exact result still neglects coefficients of \( O(1) \), due to the nature of the blob theory. Physically, the approximations, Eqs. (4.11) and (4.13), ignore the change of entropy of chain tail(s) in the deep region when the chain translocates into the shallow region, which is negligible for large \( N \).

### 4.3.2 Relative probabilities of translocation modes

The translocation is linear if the chain attains the conformation of Fig. 4.2(a), with the critical number of segments \( m_l^* \) in the shallow region and \( N - m_l^* \) segments in the deep region. Using Eqs. (4.6) and (4.11), the probability of linear translocation is

\[ \tilde{P}_l^* \simeq N^{\gamma' - 1} \exp \left( -\frac{1}{2fD^{1/\nu + 1}} \right) \]  

(4.14)

for large \( N \).

Nevertheless, we argue from a physical point of view that one more situation leads to linear translocations. Fig. 4.4 depicts a chain conformation with a hairpin of \( m \) segments in the shallow region and only one tail with \( N - m \) segments in the deep region, which results from an insertion led by a hairpin that is too close to one of the chain ends. If \( m_l^* < m < m_h^* \), the hairpin is too short for hairpin translocations but more than enough to undergo a linear one. Therefore the conformation shown on the left of Fig. 4.4 leads to a linear translocation. Effectively, this consideration accounts for the fact that hairpins that are too close to chain ends lead to linear translocations.
The probability of the hairpin conformation shown on the left of Fig. 4.4 is given by \( P_h(m, 0) \) from Eq. (4.7), that is

\[
P_{hl}(m) = (N - m)^{\gamma' - 1} \exp \left( \frac{f}{4} \frac{m^2}{D^{1/\nu - 1}} - \frac{m}{D^{1/\nu}} \right),
\]

(4.15)

This conformation will lead to a linear translocation if \( m_l^* < m < m_h^* \). The total probability of hairpin-to-linear translocation is

\[
P_{hl}^* = \int_{m_l^*}^{m_h^*} P_{hl}(m) \, dm,
\]

(4.16)

\[
\tilde{P}_{hl}^* \simeq \int_{m_l^*}^{2m_l^*} N^{\gamma' - 1} \exp \left( \frac{f}{4} \frac{m^2}{D^{1/\nu - 1}} - \frac{m}{D^{1/\nu}} \right) \, dm,
\]

(4.17)

\[
= N^{\gamma' - 1} D^{(1-\nu)/2\nu} \sqrt{\frac{x}{f}} \operatorname{erfi} \left( \frac{1}{2} \frac{D^{(1+\nu)/2\nu}}{\sqrt{f}} \right) \exp \left( -\frac{1}{f D^{1/\nu + 1}} \right),
\]

(4.18)

where \( \operatorname{erfi}(x) \equiv \operatorname{erf}(ix)/i = 2\pi^{-1/2} \int_0^x \exp(t^2) \, dt \) is the imaginary error function. The total probability of the conformations that lead to linear translocations is then \( P_l^* + P_{hl}^* \).

For hairpin translocations, the probability that a chain conformation leads to a hairpin translocation is

\[
P_h^*(j) = (N - m_h^* - j)^{\gamma' - 1} j^{\gamma' - 1} \exp \left( \frac{f}{4} \frac{m^2}{D^{1/\nu - 1}} - \frac{m_h^*}{D^{1/\nu}} \right).
\]

(4.19)

For large \( N \), it becomes

\[
\tilde{P}_h^*(j) = (N - j)^{\gamma' - 1} j^{\gamma' - 1} \exp \left( -\frac{1}{f D^{1/\nu + 1}} \right).
\]

(4.20)

Eqs. (4.7) and (4.13) were used in getting the above result. \( j \) is between 0 and \( N - m_h^* \), because there are already \( m_h^* \) segments in the shallow region. The total probability of hairpin translocation is
\[ P_h^* = \int_0^{N-m_h^*} P_h^*(j) dj \]  
\[ \simeq \int_0^N \tilde{P}_h^*(j) dj \]  
\[ \tilde{P}_h^* = AN^{2\gamma'-1} \exp \left( -\frac{1}{fD^{1/\nu+1}} \right) \]  
(4.21)

(4.22)

(4.23)

where \( A = \Gamma(\gamma')^2/\Gamma(2\gamma')=1.94 \), and \( \Gamma(x) \) is the Gamma function. The approximation is valid for large \( N \).

The relative probabilities of the three different entry mechanisms are shown in Fig. 4.5(a). Linear translocations, with probabilities \( P_l^* \) and \( P_{hl}^* \), are favored by short chains. Hairpin translocations, with probability \( P_h^* \), dominate for long chains. In general, \( P_h^* \) changes more rapidly with \( N \) compared to \( P_l^* \) and \( P_{hl}^* \).

### 4.3.3 Probability of linear translocations

\( P_l^* \), \( P_{hl}^* \) and \( P_h^* \) are relative probabilities of conformations that lead to their corresponding modes of translocation. Hence the probability of linear translocations among all translocation events is

\[ \tilde{P}_l = \frac{P_l^* + P_{hl}^*}{P_l^* + P_{hl}^* + P_h^*}. \]  
(4.24)

Using the approximated formulas Eqs. (4.14), (4.18) and (4.23), \( \tilde{P}_l \) is plotted in Fig. 4.5(b). The exact numerical results are plotted on the same graph for comparison. The exact result still neglects coefficients of \( O(1) \), due to the nature of the blob theory. From the integration that leads to Eq. (4.23), it is clear that long chains favor hairpin translocations simply because they have more chance of forming hairpins.

### 4.3.4 Linear and hairpin translocation times

When a polymer molecule is trapped at the end of the deep region and tries to enter the shallow region, electrostatic force, with the help of thermal fluctuations,
attempts to pull segments into the shallow region against the entropic confinement until the number of segments in the shallow region exceeds the critical insertion size $m^*$. Polymer segments may be inserted into and retracted back from the shallow region many times before a thermal fluctuation makes the translocation successful. The process can be mapped to a one dimensional random walk under a free energy profile similar to the one shown in Fig. 4.3, with the initial condition $m(t = 0) = 0$, reflecting boundary condition at $m = 0$ and absorbing boundary condition at $m = N$ (for linear translocations). The reflecting boundary condition at $m = 0$ accounts for the fact that whenever the last segment is retracted from the shallow region to the deep region, the whole chain would still stick to the entrance of the shallow region and try the insertion again due to the weak electric field in the deep region. The absorbing boundary condition at $m = N$ represents the complete insertion into the shallow region. The dynamics is governed by the Fokker-Planck equation

$$\frac{\partial W_m(t)}{\partial t} = \frac{\partial}{\partial m} \left( k_0 \frac{\partial F(m)}{\partial m} W_m(t) + k_0 \frac{\partial W_m(t)}{\partial m} \right),$$

where $W_m(t)$ is the probability that there are $m$ segments in the shallow region at time $t$ and $k_0$ is a phenomenological rate constant depending on the friction between one monomer and the confined region. Applying the initial condition and boundary conditions to Eq. (4.25), the average translocation time into the shallow region can be written as [85, 42]

$$\langle \tau \rangle = \frac{1}{k_0} \int_0^N dx \int_x^N dy \exp (F(x) - F(y)),$$

which can be evaluated numerically.

If the chain spends a significant amount of time entering the shallow region, as in most experiments [49, 47, 48, 50], the free energy barrier imposed by the shallow
region is considerably larger than the thermal energy $k_B T$. This condition allows us to take the saddle point approximation to Eq. (4.26). The result is

$$\langle \tau \rangle \simeq \frac{1}{k_0} \exp(F(m^*) - F(0)). \quad (4.27)$$

This is the same as the average time needed to cross a free energy barrier in Kramers’ reaction rate theory [51] and of the same form of $\tau_2$ assumed in the previous studies [49, 48, 50, 122]. The approximation is good for the case that $F(m^*) - F(0)$ is considerably larger than $k_B T$, realized by low electric forces $f$ and narrow gap heights $D$. Eq. (4.27) will be used for both linear and hairpin translocation times. We take $F(0)$ to be zero by assuming that no segments are fixed at the entrance of the shallow region. Therefore, $F(m^*)$ is the free energy barrier height for a particular free energy profile.

It is instructive to write down the free energy barriers for the two translocation modes explicitly, denoted by $F^*$ in Fig. 4.3. Since linear translocations can be achieved by two mechanisms, linear and hairpin-to-linear translocations as depicted in Figs. 4.2(a) and 4.4, the effective free energy barrier of linear translocation is

$$F^*_l = -\ln(P^*_l + P^*_h), \quad (4.28)$$

$$\simeq (1 - \gamma') \ln N + \frac{1}{2fD^{1/\nu+1}}$$

$$- \ln \left[ 1 + D^{-\frac{1-k}{2\nu}} \sqrt{\frac{\pi}{f}} \exp \left( -\frac{1}{2fD^{1/\nu+1}} \right) \erfi \left( \frac{1}{2\sqrt{fD^{1/\nu+1}}} \right) \right], \quad (4.29)$$

where Eqs. (4.14) and (4.18) are used in the last equation, valid for large $N$. Similarly, the effective free energy barrier for hairpin translocation is obtained, by using Eq. (4.23), as
\[ F_h^* = -\ln(P_h^*), \quad (4.30) \]
\[ \simeq (1 - 2\gamma') \ln N - \ln A + \frac{1}{fD^{1/\nu+1}}. \quad (4.31) \]

From Eqs. (4.29) and (4.31), we observe that the free energy barriers are composed of two parts, one depending on \( N \) and the other depending on \( f \) and \( D \). For hairpin translocations, the part that depends on \( f \) and \( D \) is at least twice as large as that for linear translocations. However, given that \( 1 - \gamma' = 0.31 \) and \( 1 - 2\gamma' = -0.38 \), the free energy barrier increases with \( N \) for linear translocations and decreases with \( N \) for hairpin translocations. As \( N \) increases, the free energy barrier for hairpin translocations can be lower than its linear counterpart and eventually can dominate the mode of translocation.

Using Eqs. (4.27) and (4.29), the average linear translocation time is

\[
k_0 \langle \tau_l \rangle \simeq \exp(F_l^*) \\
\simeq \frac{N^{1-\gamma'} \exp\left(\frac{1}{2fD^{1/\nu+1}}\right)}{1 + D^{1/\nu} \sqrt{\pi/f} \exp\left(\frac{1}{2fD^{1/\nu+1}}\right) \text{erfi}\left(\frac{1}{2\sqrt{fD^{1/\nu+1}}}\right)}. \quad (4.32)
\]

Note that the linear translocation time \( \langle \tau_l \rangle \sim N^{1-\gamma'} \sim N^{0.31} \), increases with \( N \) for self-avoiding flexible chains. Similarly, the average hairpin translocation time is

\[
k_0 \langle \tau_h \rangle \simeq \exp(F_h^*) \\
\simeq \frac{N^{1-2\gamma'} A \exp\left(\frac{1}{fD^{1/\nu+1}}\right)}{fD^{1/\nu+1}}. \quad (4.33)
\]

Eqs. (4.27) and (4.31) were used. \( \langle \tau_h \rangle \sim N^{1-2\gamma'} \sim N^{-0.38} \), decreases with \( N \) for self-avoiding flexible chains.

If both linear and hairpin translocations are allowed, as in actuality, the free energy barrier is \( F^* = -\ln(P_l^* + P_h^* + P_{hl}^*) \) and the translocation time is given by

\[
k_0 \langle \tau \rangle \simeq \exp(F^*), \quad (4.34)
\]
which is plotted in Fig. 4.6(a) using exact formulas. The exact result still neglects coefficients of $O(1)$, due to the nature of the blob theory. Indeed, Eqs. (4.32) and (4.33) can be viewed as the limiting cases of Eq. (4.34) for $P_{l}^* + P_{hl}^* \gg P_h^*$ and $P_h^* \gg P_{l}^* + P_{hl}^*$, respectively. Eq. (4.34) suggests that the translocation time increases with $N$ when linear translocations dominate and decreases with $N$ when hairpin translocations dominate. The electric field response of the translocation time is shown in Fig. 4.6(b).

4.4 Discussion

We have investigated polymer translocation from a spacious region into a confined region by using equilibrium self-avoiding chain statistics. When the free energy barrier of translocation is sufficiently higher than $k_BT$, the linear translocation time $\langle \tau_l \rangle \sim N^{1-\gamma'} \exp(1/2fD^{1/\nu+1}) \sim N^{0.31}$ and the hairpin translocation time $\langle \tau_h \rangle \sim N^{1-2\gamma'} \exp(1/fD^{1/\nu+1}) \sim N^{-0.38}$. The results are valid for low electric forces $f$, narrow gap heights $D$ and long chain lengths $N$, and are different from those predicted by previous simpler models [112, 50]. In our theory, the $N$ dependence on translocation time is due to the fact that linear translocations are led by chain ends which can only happen near the two chain ends, while hairpin translocations are led by hairpins which can be formed anywhere between the two chain ends so that the number of possible hairpins increases with $N$. In addition, hairpin translocations are favored by wide gaps and high applied voltages.

Our result can be applied to explain the $N$ dependence of the voltage-driven DNA translocation time in periodic microfluidic channels [49, 48]. Since the electrophoretic mobility is independent of $N$ in the deep regions and the time through the shallow regions is negligibly small, the $N$ dependence of the translocation time arises predominantly from the insertion from the deep region into the shallow region. Our major prediction is that the translocation time either decreases with $N$
(\(\tau \sim N^{-0.38}\)) if hairpin-translocation dominates or increases with \(N\) (\(\tau \sim N^{0.31}\)) if linear-translocation dominates. These predictions are to be contrasted with Ref. [112] where \(\tau \sim N^0\) and \(N^{-1}\), respectively, for linear and hairpin translocations, and with Refs. [49, 48], where \(\tau \sim N^{-0.59}\). Our prediction \(\tau \sim N^{-0.38}\) for long chains where hairpin translocations dominate, agrees with \(\tau \sim N^{-0.42}\) obtained by fitting the experimental data [48]. Nevertheless, direct quantitative comparison with experimental results is difficult due to many reasons described below.

We have assumed that the deep region is much larger than the radius of gyration of the translocating polymer so that confinement effect is unimportant. In some experiments, the deep regions, although decades deeper than the shallow regions, are comparable to the size of the DNA, imposing a weak confinement to the chain [49]. Furthermore, we assume that the chain behaves flexible inside the shallow region. In reality, the height of the shallow regions may be comparable to the persistence length of DNA [49, 106]. In this case, confinement free energy of semiflexible chains [106, 99] and bending energy of hairpins have to be taken into account. In our calculations, all surfaces are assumed to be neutral. If otherwise, the applied electric field would induce an electro-osmotic flow. Nevertheless, electro-osmotic flow was suppressed by high buffer concentrations in experiments [50]. In our model, the weak electric field in the deep region and the electric field gradient at the interface between the regions are ignored, which might influence the chain statistics. We have assumed that the translocation time is much longer than the longest relaxation time of the polymer, so that the chain is in quasi-equilibrium during translocation. Given that the relaxation times of the DNA used in the experiments are of order 10 ms while translocation into a shallow region happens in seconds, this assumption is justified. However, non-equilibrium dynamics is possible under high electric fields, resulting in fast translocations observed in simulations [122]. Higher order translocation modes, such as double hairpins, are ignored in the present theory. They serve as higher order
corrections since they require a higher degree of chain deformation and are entropically unfavorable. Most importantly, the translocation time expressions are valid only when the free energy barrier is higher than $k_B T$. If otherwise, the translocation time cannot be approximated by Eq. (4.27) and the full expression Eq. (4.26) has to be evaluated numerically. Then the translocation coordinates of a hairpin chain, which depends on both $m$ and $j$, are two dimensional which complicates the problem significantly. Finally, we note that there is another definition of hairpin translocation coordinates. Instead of the definition shown in Fig. 4.2(b), a symmetric definition of variables can be adopted such that there are $m$ segments in the shallow region, leaving two tails of $N - j - m/2$ and $j - m/2$ segments, respectively, in the deep region. The expressions of translocation times and probabilities are more involved under this definition. Nevertheless, the results remains the same for large $N$, which makes the change of entropy of the tails unimportant during translocation.

The hairpin conformation shown in Fig. 4.2(b) assumes that the two series of blobs align side by side. This idealized situation may be discouraged by the excluded volume interaction between the two series of blobs. One possible model to account for this effect is to assume that, at the expense of introducing one more parameter, the hairpin makes an angle $\alpha$ at the turning point so that the two series of blobs keep a distance away from each other. Then we need to substitute $f \rightarrow f \cos(\alpha/2)$ in expressions involving hairpins, such as Eqs. (4.7) and (4.9). The critical insertion length Eq. (4.13) measured in physical distance however will stay the same for large $N$ due to the cancelation of the factor $\cos(\alpha/2)$. Finally, additional parameters can be introduced in devising different boundary conditions in solving the Fokker-Planck equation, Eq. (25).

The main concept that arises from our present theory is that the molecular weight dependence of the translocation time of DNA is dictated by the relative propensity of hairpin conformations versus linear conformations. If hairpin conformations dominate...
at the entrance to the shallow regions, the translocation time decreases with molecular weight. If linear conformations dominate, the translocation time is predicted to increase with molecular weight.
Figure 4.1. The DNA molecule starts at the beginning of the deep region as shown in (a). Driven by the electric field $E$ (the arrow indicates the direction of the movement of the chain), it takes time $\tau_1$ to migrate to the end of the deep region as shown in (b). The polymer molecule then takes $\tau_2$ to enter the shallow region, as sketched in (c). Once the polymer molecule enters the shallow region successfully, it takes $\tau_3$ to finish passing through the shallow region, as shown in (d).
Figure 4.2. Two modes of translocation (from the top view of geometry in Fig. 1). (a) For a linear translocation, the chain enters the shallow region with a chain end, forming $n$ blobs in a linear series, with $m$ segments in the shallow region and a tail of $N - m$ segments in the deep region. (b) For a hairpin translocation, the chain enters the shallow region with a hairpin, with $m$ segments in the shallow region and two tails of $j$ and $N - m - j$ segments in the deep region. In the deep region, there are two linear chains of blobs, each with $n/2$ blobs.
Figure 4.3. Typical free energy profile of translocation as a function of the number of segments in the shallow region $m$. $m^*$ is the critical number of segments in the shallow region beyond which the translocation process is favorable.
Figure 4.4. A hairpin entry with $m^*_h < m < m^*_l$ and only one tail in the deep region leads to a linear translocation. The chain enters the shallow region as a hairpin but $m$ does not exceed the hairpin critical insertion size $m^*_h$. The chain end forming the hairpin has no entropic penalty to prevent itself from proceeding into the shallow region and gaining electrostatic energy. If $m$ is larger than the linear critical insertion length $m^*_l$, linear translocation occurs.
Figure 4.5. (a) Relative probabilities of different translocation modes. The probability of hairpin translocation increases with $N$. $D = 1.5$ and $f = 0.08$ are used. (b) Probability of linear translocations (Eq. (4.24)) decreases with increasing $N$ and $f$. $D=1.5$, $N=30, 60, 90$ and $120$ respectively. Solid lines are calculated using approximations Eqs. (4.14), (4.18) and (4.23). Dotted lines are calculated using exact formulas. The approximation works well for large $N$. 
Figure 4.6. (a) Average translocation time $\langle \tau \rangle$ exhibits a crossover from linear to hairpin translocations as $N$ increases for $D = 1.2$, while this feature is missing for wider gaps. Exact formulas with $f=0.1$ are used. (b) $\langle \tau \rangle$ as a function inverse electric force $1/f$ with $N=100$. The exact formulas are used.
CHAPTER 5
POLYMER TRANSLOCATION THROUGH NANOPORES WITH TUNABLE POLYMER-PORE INTERACTIONS

5.1 Introduction

Polymer translocation through a nanometer-scale pore is ubiquitous in biological systems. The translocation process has been studied using nucleic acids [58, 52, 78, 2, 77, 15, 17, 18, 76, 94, 126, 33, 31] and synthetic polymers [84, 16] through α-Hemolysin (αHL) protein nanopores. Several theoretical modelings [121, 85, 67, 114, 8] and computer simulations [72, 91] were devoted to understanding this process.

The experimental setup is illustrated in Fig. 5.1. The αHL protein was embedded onto a lipid bilayer supported by a ∼ 50 μm wide Teflon aperture. An electric field was applied across the nanopore and drove the charged polymer molecules to pass through the nanopore. The narrowest part of the nanopore was about 1.5 nm, allowing the charged polymer molecules to pass through the nanopore linearly without forming hairpins. If a polymer molecule resided in the nanopore, a brief blockage in the ionic current would be observed.

Most of the previous studies treated the blockage event as an single-level event characterized by the event duration and the average blockage current. In reality, it is not uncommon to observe multiple-level events and this treatment averaged out the multiple-level information. Butler et al. [18, 17] analyzed the multiple-level events produced by single-stranded RNA and DNA through αHL. They concluded that shallow blockage levels corresponded to polymer trapping inside the vestibule of αHL and deep blockage levels corresponded to polymer passage through the pore.
The molecular structure of αHL is shown in Fig. 5.2 [115]. To pass through the pore from the cis to the trans side, a polymer first enters the relatively spacious vestibule with 4.6 nm internal diameter through a 2.6 nm opening. It then passes through a 1.5 nm constriction, which restricts the NaPSS molecule to pass through only linearly, to enter a 5 nm long and 2 nm wide β-barrel.

Both the pore geometry and the charge distribution may influence the dynamics of polymer translocation. Mostly notably, the vestibule may act as an entropic trap to the polymer because it is relatively spacious and the 1.5 nm constriction allows the polymer to enter the β-barrel only with chain ends but not hairpins. The 1.5 nm constriction is composed of 7 lysines (Lys147) and 7 glutamic acids (Glu111). In pH 7.5, the Lys147 residues are positively charged and the Glu111 residues are negatively charged, making the constriction overall neutral. At the end of the β-barrel, there is a highly charged ring composed of 14 aspartic acids (Asp127 and Asp128) and 7 lysines (Lys131). In pH 7.5, the net charge of the ring is $-7$ electron charges ($e$), possibly impeding the translocation of negatively charged polymers.

Based on the molecular structure of αHL, we arrive at 2 possible models for the dynamics of polymer translocation illustrated in Fig. 5.3. In the entropic barrier model, the vestibule of the αHL pore acts as an entropic trap to the polymer before entering the narrow β-barrel. In order to pass through the pore, the chain sacrifices conformational entropy by partitioning one of its chain ends into the β-barrel. In the enthalpic barrier model, both the vestibule and the β-barrel may contribute to the enthalpic barrier for a successful polymer passage. The attraction between the vestibule surface and the chain discourages the chain from entering the β-barrel. The chain has to overcome this attraction, and hence an enthalpic barrier, before entering the β-barrel and translocating through the nanopore. This model springs from our experimental observation that poly(acrylic acid), a negatively charged polyelectrolyte with a hydrophilic backbone, failed to produce blockages in the ionic current despite
the fact that its repeating unit was smaller than that of NaPSS. We conjecture that poly(acrylic acid) adsorbed onto the vestibule surface and was unable to overcome this attraction to enter the β-barrel. Likewise, the net surface charge inside the β-barrel may also impose an enthalpic barrier for entering the β-barrel via the electrostatic repulsion between the polymer and the β-barrel. In pH 7.5, there is a highly charged ring of $-7e$ at the end of the β-barrel. Since the width of a NaPSS monomer is about 0.9 nm and the Debye screening length in 1M KCl solutions is about 0.4 nm, the electrostatic repulsion between the charged ring and the polymer may be significant inside the 2 nm wide β-barrel. In addition, it is unclear whether the mean-field description of ionic screening in continuous media is appropriate for such microscopic length scales.

In this chapter, we present detailed analysis of the event types that consisted of either or both of the two blockage levels. Agreeing with Butler et al. [18, 17], we found that the shallow blockages corresponded to polymer trapping inside the vestibule and the deep blockages corresponded to part of the chain inside the β-barrel. Furthermore, we were able to distinguish between successful translocation events and unsuccessful attempts. By tuning the enthalpic interaction between the polymer and the β-barrel by controlling the solution pH on the trans side of αHL, we showed that increasing the attraction between the polymer and the β-barrel increased the propensity of successful translocation events. Finally, we constructed a stochastic theory based on vestibule trapping and polymer-pore interaction to explain our experimental observations.

5.2 Materials and methods

The NaPSS polymers were purchased from Scientific Polymer Products with weight-averaged molecular weights $N=16$, 57.5, 126.7 and 505 kg/mol respectively. The polydispersity indices are 1.13, 1.10, 1.17 and 1.24, respectively. The pH 7.5 buffer
solution is 1M KCl solution with 10mM HEPES. Double deionized water with resistivity of 18MΩcm (Millipore) was used to prepare the buffer solutions.

As shown in Fig. 5.1, the experimental setup consisted of two compartments connected with a polyethylene tube with a 50 µm Teflon aperture on one end. We called the compartment that the Teflon aperture resided in as *compartment A* and the other one as *compartment B*. Note that the flush system was only available in compartment A. A pair of Ag/AgCl electrodes (In Vivo Metric), connected to Axopatch 200B patch clamp amplifier (Axon Instruments, CA), were placed in the two compartments to apply an electric potential across the compartments and measure the ionic current passing through the pore. All ionic current were recorded in the voltage-clamp mode through a 10kHz low-pass Bessel filter with 3 µs sampling intervals.

The experimental procedure for preparing a single αHL pore is as follows. The Teflon aperture was first coated with several layers of 2g/L 1,2-diphytanoyl-sn-glycero-3-phosphocholine (diphytanoyl-PC) (Avanti Polar Lipids) in hexane solution. After the lipid hexane solution had dried up completely, 1M KCl 10mM HEPES pH 7.5 solution was introduced into both compartments and the polyethylene tube. After a diphytanoyl-PC lipid bilayer were formed across the Teflon aperture, 0.5 µL (or less) of 0.01 g/L of α-toxin from Staphylococcus aureus (Calbiochem) were added to the *cis* compartment. The spontaneous toxin insertion into the lipid bilayer produced an abrupt, quantized increase in the ionic current such that it was ∼ 120 pA at 120 mV and ∼ −90 pA at −120 mV for a single αHL pore.

We used solution with pH ranged from 3.5 to 7.5 throughout this study. We used 10 mM HEPES and 10 mM citric acid as the buffering agent of pH 7.5 and 3.5-6.0 solutions, respectively, in various salt concentrations.

Polymer translocation experiments were performed by either adding 1-10 µL of 1 g/L NaPSS solutions to the *cis* compartment, or forming the αHL pore in NaPSS solution followed by flushing the *trans* compartment with clean buffer solution. The
ionic current traces were recorded by the Clampex 9.2 software and analyzed by our custom MATLAB software.

5.3 Experimental results

5.3.1 αHL in asymmetric pH solution

Misakian and Kasianowicz [81] studied the changes in the current-voltage (I-V) curves of αHL upon various asymmetric pH conditions. In particular, they showed by titration that a pore with +120mV to −120mV current ratio $\gamma_{120} \sim 1.06$ in 1M KCl 10mM citric acid pH 4.5 solution corresponds to the $\gamma_{120} \sim 1.33$ αHL pore in 1M KCl 10mM HEPES pH 7.5 solution, which was a well-formed heptamer pore. They ascribed the pH response of the I-V curves to the change in the protonation states of amino acid residues on the interior surface of the pore.

To prepare an αHL pore in a 1M KCl pH 7.5 solution on the cis side and pH 4.5 on the trans side, we started with a symmetric 1 M KCl 10 mM HEPES solution. Because the flush system was only available in compartment A, it was necessary to form an inverted αHL pore with ionic currents $\sim -120$ pA at $-120$ mV and $\sim 90$ pA at 120 mV. This was achieved by setting the voltage to $-120$ mV instead of 120 mV while waiting for the spontaneous pore formation. In case of inactivity or multiple-pore formation, the bilayer was destroyed using the ZAP function of the patch-clamp while keeping the voltage at $-120$ mV. After forming a new lipid bilayer membrane, a single, inverted pore might form within a few minutes. Since an inverted pore had its trans side in compartment A, we replaced the solution of compartment A with a 1 M KCl 10 mM citric acid pH 4.5 solution to achieve a 7.5 cis-pH and a 4.5 trans-pH. Similarly, an αHL pore with cis-pH 4.5 and trans-pH 7.5 was achieved by first forming a pore with the normal orientation (i.e. cis side in compartment A) and displacing compartment A with a 1M KCl 10 mM citric acid pH 4.5 solution. Finally, an αHL pore in a symmetric pH 4.5 solution was achieved
by the spontaneous insertion of the toxins. Misakian and Kasianowicz [81] showed by titration that a pore with $+120\text{mV}$ to $-120\text{mV}$ current ratio $\gamma_{120} \sim 1.06$ in 1 M KCl 10mM citric acid pH 4.5 solution corresponds to the $\gamma_{120} \sim 1.33$ αHL pore in 1 M KCl 10 mM HEPES pH 7.5 solution, which was a well-formed heptamer pore we achieved earlier. Therefore obtaining a pore with $\gamma_{120} \sim 1.06$ in pH 4.5 served our purpose. We note that spontaneous insertions of the toxins were much faster (and hence required lower toxin concentrations) in pH 4.5 than in pH 7.5. When a single toxin insertion was observed, compartment A was flushed with clean citric acid solution to prevent additional toxin insertions. The nearly symmetric current-voltage curve in the pH 4.5 citric acid solution was thought to be caused by the neutralization of the acidic amino acid groups inside the channel [81]. Compartment A was then flushed with 1M KCl 10 mM HEPES pH 7.5 solution. We then had a single αHL channel with pH 7.5 solution on the cis side and pH 4.5 solution on the trans side.

We studied the asymmetric pH responses of αHL achieved by simple solution displacement instead of titration. More precisely, for example, pH 7.5 on the cis side and pH 4.5 on the trans side of αHL was achieved by first forming an αHL in symmetric pH 7.5 solution followed by simply flushing the trans compartment with pH 4.5 solution. In this way, changing the solution pH on one side of the pore is convenient and reversible. The I-V curves under different asymmetric pH conditions are shown in Fig. 5.4. In accordance with Misakian and Kasianowicz [81], we observed that the I-V curve was nearly symmetric ($\gamma_{120} \sim 1.06$) in cis-pH 7.5 and trans-pH 4.5 solution. The I-V curves changed more dramatically upon a change in the trans-pH than a change in the cis-pH, suggesting that the ionic passages were largely controlled by the surface charges of the narrow β-barrel rather than that of the spacious vestibule. The result agrees with the molecular dynamics simulations of αHL pore [4] that most of the electric potential change occurred inside the β-barrel.
The result also suggested that the pH values on one side of the pore did not apply to the whole nanopore.

We now estimate the effective charge of the constriction and the ring at the end of the β-barrel. From the definition of the acid dissociation constant, pKa,

$$\frac{[A^-]}{[AH]} = 10^{\text{pH} - \text{pKa}}, \quad (5.1)$$

where \([A^-]\) is the concentration of the unprotonated residue and \([AH]\) is the concentration of the protonated one. Note that \(A^-\) is negatively charged if it represents aspartic acid or glutamic acid, and is neutral if it represents lysine. Assuming the equilibrium charges of the amino acid residues of αHL follow that in the bulk, the equilibrium net charge of the ring at the end of the β-barrel is

$$7 \frac{[\text{Lys-H}]}{[\text{Lys-H}]+[\text{Lys}]} - 14 \frac{[\text{Asp}]}{[\text{Asp}]+[\text{Asp-H}]} . \quad (5.2)$$

Similarly, the equilibrium net charge of the constriction is

$$7 \frac{[\text{Lys-H}]}{[\text{Lys-H}]+[\text{Lys}]} - 7 \frac{[\text{Glu}]}{[\text{Glu}]+[\text{Glu-H}]} . \quad (5.3)$$

Using Eqs. (5.1), (5.2) and (5.3), the equilibrium net charges of the ring at the end of the β-barrel and the constriction in different pH are plotted in Fig. 5.5. For pH 7.5 hosting solution, in which most of the previous experiments were conducted, the constriction is neutral while the ring at the end of the β-barrel is \(-7e\). Decreasing the solution pH leads to neutralization of negative charges and eventually turns both to be positive.
5.3.2 NaPSS translocation through nanopores in asymmetric pH condition

Misakian and Kasianowicz [81] suggested that the free energy landscape for small ions across the αHL pore was partially altered when the trans-pH was changed from 7.5 to 4.5. We extend this idea to study the effect of asymmetric pH on polymer translocation through nanopores. We applied an electric field to drive NaPSS from the cis to the trans side of the αHL pore. In pH 7.5 solution on both sides, the three most common event types that accounted for more than 95% of all events are shown in Fig. 5.6. They were composed of two distinct blockage levels, the shallow (level 1) and the deep (level 2) levels. We define a level 1 blockage as a steady blockage with a blockage current $I_b$ between threshold 1 and threshold 2. A level 2 blockage is defined as a steady blockage with $I_b$ below threshold 2. From these two blockage levels, we define an event 1 as an event that is composed of a single level 1 blockage. An event 12 is defined as an event that is composed of a level 1 blockage followed by a level 2 blockage. Finally, an event 2 is an event that is composed of a single level 2 blockage. We denote events 12a and 12b as the first and second sub-levels of event 12 with durations $\tau_{12a}$ and $\tau_{12b}$ respectively. Butler et al. [18, 17] concluded that level 1 blockages corresponded to polymer trapping inside the vestibule and level 2 blockages corresponded to polymer passing through the β-barrel. Here we provide further evidence for supporting this interpretation.

5.3.2.1 Relative frequency of blockage events

Figure 5.7 compares the typical ionic current blockages produced by 57.5 kg/mol NaPSS when the trans-pH values were 7.5 and 4.5 respectively. The cis-pH were 7.5 in both cases. Both ionic current traces consisted mainly of events 1, 12 and 2. However, we note that there are several differences between the two. The blockages were more frequent in trans-pH 4.5 than that in trans-pH 7.5. Both of the two blockage levels
in \textit{trans}-pH 4.5 were deeper than that in \textit{trans}-pH 7.5. Threshold 1 and threshold 2 were adjusted accordingly in different pH conditions to capture the two levels.

As shown in Fig. 5.8, the frequency of occurrence of event 1 decreased with the applied voltage while the opposite was true for events 12 and 2. Remarkably, the frequency of occurrence of the three event types changed dramatically in different \textit{trans}-pH. For 16\text{kg/mol} \textit{NaPSS} in \textit{trans}-pH 7.5, events 1 accounted for most the observed events. The second dominant type was event 12, which accounted for less than 20\% of the total events for applied voltages under 160 mV. Events 2 constituted less than 1\% of the total events. However, simply replacing the pH 7.5 solution with a pH 4.5 one on the \textit{trans} side increased the occurrences of events 12 and 2, which accounted for most of the observed events above 110 mV. Furthermore, events 12b and 2 appeared to overlap with each other in the event diagrams shown in Fig. 5.8 (a) and (b).

Owing to the more frequent occurrence of events 12 and 2, we could analyze the statistics of the three event types under the \textit{cis}-pH 7.5 and the \textit{trans}-pH 4.5 condition. Fig. 5.9 shows the average durations of events 12a, 12b and 2 for \textit{NaPSS} with $N=16$, 57.5, 126.7 and 505 kg/mol. The rationale behind plotting the average durations against $N/V$ is that the mean translocation time of a polymer passing a nanopore is known to follow this relation [58, 85, 67]. It is clear from Fig. 5.9 that $\langle \tau_2 \rangle$ and $\langle \tau_{12b} \rangle$, which were both level 2 blockages, followed $N/V$ and were fitted using a single straight line, while $\langle \tau_{12a} \rangle$ stayed almost constant over the same range.

Intuitively, increasing the applied voltage increases the driving force of polymer translocation through the pore. Base on our experimental observations that (1) the percentages of events 12 and 2 increased with applied voltages and (2) $\langle \tau_2 \rangle$ and $\langle \tau_{12b} \rangle$ followed $N/V$, we conclude that only the level 2 blockages were successful translocations. In other words, only events 12 and 2, which consisted of level 2 blockages, were successful translocations while events 1 were unsuccessful attempts.
Therefore, level 1 blockages, which appeared in events 1 and 12, corresponded to polymer trapping inside the nanopore. A level 1 blockage might lead to a successful translocation, such as in events 12, or an unsuccessful one such as in events 1.

5.3.2.2 Threshold voltage of successful translocation

As shown in Figs. 5.8(c) and (d), the lowest applied voltages that events 12 and 2 appeared were different under different trans-pH. For example, the threshold voltage was about 130 mV in trans-pH 7.5 and 90 mV in trans-pH 4.5. Since we now know that only events 12 and 2 are successful translocations, we define the threshold voltage of successful translocation as the lowest voltage that the sum of events 12 and 2 accounts for more than 1% in all recorded events.

A detailed study of threshold voltages of translocation for 16 kg/mol NaPSS under different trans-pH is shown in Fig. 5.10. The threshold voltage stayed almost constant at about 125 ± 0.5 mV when lowering the trans-pH from 7.5 to 6. Further lowering the trans-pH decreased the threshold voltage dramatically, down to 35 ± 0.5 mV in trans-pH 3.5. In other words, lowering the trans-pH below 5.5 decreased the free energy barrier of activating the translocation process.

5.3.2.3 trans-pH dependence of average blockage durations

Figure 5.11 shows how the average blockage durations (⟨τ₁⟩, ⟨τ₂⟩, ⟨τ₁₂ₐ⟩, and ⟨τ₁₂ₖ⟩) defined in Fig. 5.6) depended on the trans-pH while keeping the cis-pH constant at 7.5. The average durations of level 2 blockages (⟨τ₂⟩ and ⟨τ₁₂ₖ⟩), which represented successful translocations starting from the chain entering the β-barrel, increased when lowering the trans-pH. This is in contrast with the average durations of level 1 blockages (⟨τ₁⟩ and ⟨τ₁₂ₐ⟩) representing the trapping of the chain inside the vestibule, which decreased when lowering the trans-pH.

The above results can be explained by the change of the polymer-pore interaction inside the β-barrel in response to the trans-pH. Lowering the trans-pH increases the
attraction between the polymer and the \( \beta \)-barrel. This makes the polymer easier to enter the \( \beta \)-barrel and thus decreases the trapping time in the vestibule. However, the attraction also makes the chain harder to leave the \( \beta \)-barrel by imposing a free energy barrier for leaving the pore. The translocation time increases as a result. This effect was recently studied by a computer simulation \([70]\). We will construct a stochastic theory to describe the polymer-pore interaction and compare with our experimental results.

### 5.3.2.4 Effect of electro-osmotic flow on polymer translocation

Electro-osmotic flow, the coupled movement of water molecules with charged species under electric fields, are known to be able to drive small molecules into \( \alpha \text{HL} \) \([46, 44, 45]\). Since increasing the surface charge of the \( \beta \)-barrel increases the negatively charged ions inside the pore and NaPSS is also negatively charged, decreasing the trans\(-\)pH increases the EOF in the direction of the NaPSS translocation. Therefore, it is necessary to estimate whether the increase in EOF was responsible for the enhanced occurrence of successful translocations.

We follow the method of Gu et al. \([44, 45]\) to estimate the electro-osmotic flow (EOF) through the nanopore in asymmetric pH conditions. First we estimate the permeability ratio \( P_+/P_- \) of the two most dominant contributors of the ionic current, \( K^+ \) and \( Cl^- \) ions, across the nanopore. It is given by the Goldman-Hodgkin-Katz (GHK) voltage equation \([53]\)

\[
P_+ \over P_- = \frac{[Cl^-]_c - \exp(-eV_R/k_B T)[Cl^-]_t}{\exp(-eV_R/k_B T)[K^+]_c - [K^+]_t},
\]

(5.4)

where \([K^+]\) and \([Cl^-]\) denote the concentrations of \( K^+ \) and \( Cl^- \) ions, respectively, on one side of the nanopore. The subscripts \( c \) and \( t \) denote, respectively, the cis and the trans side. \( V_R \) is the reversal voltage which sets the ionic current passing through the
nanopore zero. $e$, $k_B$ and $T$ are the electron charge, the Boltzmann constant and the temperature, respectively.

To estimate the strength of the EOF, following Gu et al. [44, 45], we further assume that each ion passing through the nanopore is coupled with $N_w$ water molecules. $N_w \simeq 10$ according to an all-atom molecular dynamics simulation [4]. Therefore, the strength of the EOF (in number of water molecules per unit time) is given by

$$J_w = N_w I e \left( \frac{1 - P_+/P_-}{1 + P_+/P_-} \right),$$

(5.5)

where $I$ is the ionic current under a given voltage.

Equation (5.4) shows that an asymmetric salt condition, i.e. different salt concentrations on the *cis* and the *trans* sides, is necessary for measuring the reversal voltages $V_R$. To ensure the stability of the electrodes in the asymmetric salt conditions, we placed 1.5% Agarose salt bridges containing 3M KCl between the buffer solution and the electrodes. The $V_R$ measured were then corrected for the junction potentials using the Junction Potential Calculator, which used the Henderson equation [10], supplied by the Clampex 9.2 current recording program.

We used 1 M KCl solution on the *cis* side and 200 mM KCl solutions on the *trans* side when measuring the reversal voltage $V_R$. The *cis*-pH was kept at 7.5 while *trans*-pH ranged from 4.5 to 7.5. The measured $V_R$ and the corresponding permeability ratios $P_+/P_-$ are shown in Table 5.1. All measured values of $P_+/P_-$ were less than unity, meaning that negative ions passed through the pore more easily than positive ones. The strength of the EOF was then calculated using Eqs. (5.4) and (5.5), and is plotted in Fig. 5.12. We used 0.6 and 0.72 for the activities of 1 M and 200 mM KCl solution (For example [KCl] in Eq. (5.4) is corrected to 1M×0.6=0.6M for a 1 M KCl solution.) [53, 54]. As expected, lowering the *trans*-pH increased the strength of the EOF through $\alpha_{HL}$. We observed a linear relation between the strength of the EOF and the *trans*-pH.
If EOF is responsible for the increased propensity of successful translocations, we would expect a direct correlation between the strength of the EOF and the threshold voltage of successful translocation. Comparing Fig. 5.12 with Fig. 5.10, we find that the two did not correlate well. In particular, changing the trans-pH from 7.5 to 6.0, the strength of the electro-osmotic flow almost tripled but the threshold voltage remained unchanged. We conclude that electro-osmotic flow was not the main reason for the dramatic increase in successful translocation events when lowering the trans-pH. This conclusion is further corroborated by our experimental observation that a similar concentration of polyethylene glycol, a neutral water-soluble polymer, on the cis side was unable to produce any blockage signals. Therefore, although EOF could drive small molecules into $\alpha$HL [46, 44, 45], it did not assist macromolecular translocations through $\alpha$HL in a notable way.

5.4 Stochastic model of polymer translocation through $\alpha$HL

5.4.1 Formalism

To model the dynamics of polymer translocation through $\alpha$HL that capture both successful translocations and unsuccessful attempts, we use the Fokker-Planck equation

$$\frac{\partial W_m(t)}{\partial t} = \frac{\partial}{\partial m} \left( k_0 \frac{\partial F(m)}{\partial m} W_m(t) + k_0 \frac{\partial W_m(t)}{\partial m} \right),$$

with absorbing boundary conditions at both ends of the translocation coordinate $m$. Here $W_m(t)$ is the probability that the chain is in state $m$ at time $t$, $F(m)$ denotes the free energy profile of translocation and $k_0$ is a phenomenological rate constant depending on the friction between one monomer and $\alpha$HL. Starting from Eq. (5.6), it can be shown that the probability of successful translocation $P_t(z)$ and the corresponding mean successful translocation time $\langle \tau_t(z) \rangle$ follow [42]
\[ k_0 [P_t(z)\langle \tau_t(z) \rangle]'' - \frac{F'(z)}{k_B T} k_0 [P_t(z)\langle \tau_t(z) \rangle]' = -P_t(z), \quad (5.7) \]
\[ P_t''(z) - \frac{F'(z)}{k_B T} P_t'(z) = 0, \quad (5.8) \]

where \( z \) (\( a \leq z \leq b \)) is the initial position in the translocation coordinate and the prime denotes the differentiation with respect to \( z \). Starting at \( z \), we define absorption by \( a \) and \( b \) as unsuccessful and successful translocations, respectively. Eqs. (5.7) and (5.8) are solved with the boundary conditions \( P_t(a)\langle \tau_t(a) \rangle = P_t(b)\langle \tau_t(b) \rangle = 0 \), \( P_t(a) = 0 \) and \( P_t(b) = 1 \). The mean unsuccessful translocation time and probability follow similar equations. Explicitly, the successful translocation probability, the mean successful translocation time and the mean unsuccessful translocation time are

\[
P_t(z) = \frac{\Psi(a, z)}{\Psi(a, b)}, \quad (5.9)
\]
\[
\langle \tau_t(z) \rangle = \frac{1}{k_0} \frac{\Psi(a, z)\Phi_t(z, b) - \Psi(z, b)\Phi_t(a, z)}{\Psi(a, z)\Psi(a, b)}, \quad (5.10)
\]
\[
\langle \tau_u(z) \rangle = \frac{1}{k_0} \frac{\Psi(a, z)\Phi_u(z, b) - \Psi(z, b)\Phi_u(a, z)}{\Psi(z, b)\Psi(a, b)}, \quad (5.11)
\]

where

\[
\Psi(p, q) \equiv \int_p^q e^{F(y)} dy, \quad (5.12)
\]
\[
\Phi_t(p, q) \equiv \int_p^q dz \int_a^z dy \int_a^y dw \left[ e^{-F(y) + F(z) + F(w)} \right], \quad (5.13)
\]
\[
\Phi_u(p, q) \equiv \int_p^q dz \int_a^z dy \int_y^b dw \left[ e^{-F(y) + F(z) + F(w)} \right]. \quad (5.14)
\]

The probability of unsuccessful translocations \( P_u(z) \) satisfies \( P_u(z) + P_t(z) = 1 \), of course.

We describe the translocation of 16 kg/mol NaPSS using a one-dimensional pairwise free energy profile \( F(m) \) shown schematically in Fig. 5.13. The initial local minimum in \( F(m) \) at \( m = \delta \) represents the trapped state of the polymer inside the
vestibule. The free energy barrier $\Delta F_1$ from $m = \delta$ to $m = 0$ represents the free energy barrier for the polymer to escape from the vestibule back to the cis side. The free energy barrier $\Delta F_2 + z_0 + \epsilon$ from $m = \delta$ to $m = 2\delta$ represents the free energy change for partitioning one chain segment into the $\beta$-barrel. Here $z_0 \simeq 1.54$ is the mean coordination number of a self-avoiding chain [30] accounting for the total loss of entropy of segments inside the $\beta$-barrel. $\epsilon$ is the interaction energy between one monomer and the $\beta$-barrel, with negative values correspond to attractions between the polymer and the $\beta$-barrel. $\Delta F_2$ is the entropy loss when one of the chain end is fixed at the entrance of the $\beta$-barrel. For $2\delta < m < 2\delta + M - 1$, where $M$ is the length of the $\beta$-barrel, the chain end worms through the $\beta$-barrel in single file and reaches the $\textit{trans}$ side. For $2\delta + M - 1 < m < 2\delta + N - 1$, part of the chain resides in the $\textit{trans}$ compartment while keeping the number of segments inside the $\beta$-barrel the same. Finally, for $2\delta + N - 1 < m < 2\delta + M + N - 1$, the chain leaves the $\beta$-barrel while the other chain end is inside the $\beta$-barrel. Explicitly, the free energy of translocation is

$$F(m) = \begin{cases} -\Delta F_1 m/\delta & , 0 \leq m \leq \delta \\ F(\delta) + (\Delta F_2 + z_0 + \epsilon)(m - \delta)/\delta & , \delta < m \leq 2\delta \\ F(2\delta) + (qV + z_0 + \epsilon)(m - 2\delta) & , 2\delta < m \leq 2\delta + M - 1 \\ F(2\delta + M - 1) + qV(m - (2\delta + M - 1)) & , 2\delta + M - 1 < m \leq 2\delta + N - 1 \\ F(2\delta + N - 1) - (z_0 + \epsilon)(m - 2\delta - Nt + 1) & , 2\delta + N - 1 < m \leq 2\delta + M + N - 1. \end{cases}$$

(5.15)

We have expressed the energy in $k_B T$ and the length in the monomer length of NaPSS. We have assumed that the conformational entropy of the chain tails residing in the $\textit{trans}$ compartment is negligible and all the electric potential drop $V$ takes place at the constriction between the vestibule and the $\beta$-barrel. $q$ is the net charge of a monomer unit. Note that the free energy profile for $2\delta < m \leq 2\delta + M - 1$ should be non-linear because the confinement energy of a polyelectrolyte inside the vestibule $\sim N^x$ with $x \geq 2$ [43, 60, 19, 61]. In other words, free energy of the chain changes non-linearly when transferring segments from the vestibule to the $\beta$-barrel. However, we have
neglected this confinement energy for simplicity and retained only the entropy loss $(\Delta F_2)$ when entering the $\beta$-barrel. Owing to the numerical difficulty in handling Eqs. (5.10) and (5.11) when $F(m)$ is non-linear, we confine our calculations using the pair-wise free energy profile Eq. (5.15) which allows us to keep the expressions analytical until final evaluations.

5.4.2 Results

We interpret our experimental results in different trans-pH as the change in the polymer-pore interaction, which is described by $\epsilon$. As shown in Fig. 5.5, the equilibrium net charges of the constriction and the charge ring at the end of the $\beta$-barrel change dramatically with the pH of the hosting solution. In particular, decreasing the trans-pH increases the net charge of the $\beta$-barrel, which increases its attraction to the negative polymer. In our formalism, the decrease in trans-pH corresponds to the decrease in $\epsilon$. The free energy profiles of translocation, Eq. (5.15), with different polymer-pore interaction parameters $\epsilon$ are shown in Fig. 5.14. We note that the total free energy drop across the translocation profile does not depend on the value of $\epsilon$ because it is always given by the free energy barriers $\Delta F_1$ and $\Delta F_2$, and the potential difference across the $\alpha$HL pore. $\epsilon < 0$ represents attraction between the polymer and the $\beta$-barrel, which competes with the confinement free energy and gives the polymer more incentive to enter the narrow $\beta$-barrel. Mathematically, the free energy barrier $\Delta F_2 + z_0 + \epsilon$ for entering the $\beta$-barrel is lowered as $\epsilon$ decreases. However, the same attraction discourages the chain from leaving the $\beta$-barrel and may create a free energy barrier for leaving the pore.

The simple theoretical model is able to reproduce the features of our experimental results. Fig. 5.15 compares the experimental results with the probability of successful translocations calculated using Eq. (5.9). We used $N = 78$ corresponding to the weighted-averaged molecular weight of 16 kg/mol NaPSS with monomer length 0.3
nm. The length of the β-barrel is 10 nm which translates to \( M = 17 \). We have assumed the degree of ionization of the NaPSS molecule to be 0.3. Equivalently, the average net charge of each unit is \( q = -0.3e \). The probability of successful translocation in the experiment is the percentage of the sum of events 12 and 2 in all blockage events. Note that the threshold voltage of successful translocation emerges naturally in our theory, a unique feature in using two absorbing boundaries. Although the successful translocation probability describes the experimental results well, we note that the average translocation times only agree with the experiments qualitatively.

Fig. 5.16 shows the average successful and unsuccessful translocation times, \( \langle \tau_t \rangle \) and \( \langle \tau_u \rangle \), and the corresponding successful translocation probabilities \( P_t \). \( \langle \tau_t \rangle \) corresponds to \( \langle \tau_{12} \rangle \) and \( \langle \tau_2 \rangle \), while \( \langle \tau_u \rangle \) corresponds to \( \langle \tau_1 \rangle \). \( \langle \tau_1 \rangle \) decreases with decreasing \( \epsilon \) for \( \epsilon \lesssim 1 \) under the parameters used in Fig. 5.15, and the opposite is true for \( \langle \tau_u \rangle \). Both trends agree with that observed in our experiments as shown in Fig. 5.11. However, it is difficult to find a set of parameters that quantitatively describes the average successful translocation time, the average unsuccessful translocation time and the successful translocation probabilities at the same time. We attribute the quantitative discrepancy to our highly simplified pair-wise free energy profile of translocation.

### 5.5 Conclusions and discussions

In this chapter, we identified the molecular mechanisms that were responsible for the two-level ionic current blockages produced by single polyelectrolyte passage through αHL. Agreeing with Butler et al. [18, 17], we showed that only translocation events or sub-events with deep current blockages (level 2) are successful translocations which followed the relation \( \tau \sim N/V \). By changing the pH on the trans side of the pore, we tuned the effective charge of the β-barrel and showed that successful translocation events increased dramatically when increasing the polymer-pore attraction. Furthermore, we showed that electro-osmotic flow did not drive macromolecules
through the pore in a notable way, although it was shown to be able to drive small molecules into the pore [46, 44, 45]. Finally, we constructed a stochastic theory with pair-wise free energy profile with polymer-pore interaction. The mean successful and unsuccessful translocation times, and the successful translocation probability obtained in the theory agrees qualitatively with our experiments.

Our analysis showed that both the entropic and the enthalpic barrier models contributed to the dynamics of NaPSS translocation through αHL. The free energy barrier of translocation had both entropic and enthalpic contributions, and can be written as $\Delta F = \Delta H - T \Delta S$. Decreasing the pH in the trans compartment increased the attraction between the polymer and the β-barrel, which in turn decreased $\Delta H$. This caused a dramatic decrease in the threshold voltage of successful translocation shown in Fig. 5.10. Note that even though the level 1 blockage corresponds to the polymer trapping inside the vestibule, the enthalpic interaction between the polymer and the vestibule may also contribute to the energetics of the trapping, making the trapping not purely entropic.

Another interesting question is how a polyelectrolyte enters the vestibule of the αHL pore on the cis side. The opening of the vestibule of αHL is 2.6 nm wide and the persistence length of NaPSS in 1M KCl is 1.2 nm. The flexibility of NaPSS should allow the chain to enter the vestibule with an hairpin. Indeed we believe that it was true in our experiment. Figure 5.17 shows the voltage dependence of the average durations of event 1 for different chain lengths. $\langle \tau_1 \rangle$ stayed constant in the whole voltage range for the shortest 16 kg/mol NaPSS but increased with the applied voltage for higher molecular weights. Note that the sudden drops in $\langle \tau_1 \rangle$ at high voltages for 57.5 and 505 kg/mol were deceptive because many level 1 blockages were too long (> 5 s) that we needed to reverse the applied voltage to let the chain get back to the cis side, and hence the recordings were discarded. We observed this effect frequently in high molecular weights and high applied voltages but not in 16
kg/mol NaPSS in any voltages we measured. The volume fraction of a 16 kg/mol NaPSS molecule inside the vestibule is \( \sim 0.1 \). It is likely that a 16 kg/mol NaPSS molecule resides completely inside the vestibule. If this is true, there would be no difference between entering the vestibule near a chain end or with an hairpin. Both of the chain ends would be inside the vestibule and one of them might find its way to the \( \beta \)-barrel. However, for long chains, successful translocations would be impossible if the chain enters the vestibule with hairpins somewhere along the chain because the 1.5 nm constriction between the vestibule and the \( \beta \)-barrel does not allow hairpin passages. At low voltages, a long chain can simply retreat to the \( cis \) side and result in an event 1. At high voltages, however, the electrostatic energy gain for entering the vestibule discourages the chain from retreating to the \( cis \) side, resulting in a prolonged and unterminated trapping of the chain inside the vestibule. We believe that this is the molecular origin of the prolonged and unterminated level 1 blockages at high molecular weights and high voltages.

Recently, there have been a debate about the validity of the equilibrium condition during polymer translocation through nanopores [27, 57, 35, 100]. The Fokker-Planck formalism we used is only valid when the chain is able to equilibrate at every step of translocation. The longest relaxation times (the Zimm times) of NaPSS are \( \sim 10^{-5} \) to \( 10^{-2} \) ms for 16 kg/mol to 505 kg/mol NaPSS, which are several orders of magnitude shorter than the successful translocation times. Therefore, it is safe to assume that the chain is in equilibrium during translocation through \( \alpha \)HL. In fact, as shown in Fig. 5.9, the successful translocation times followed the relation \( \sim N/V \), a theoretical result that was derived under the equilibrium assumption [85, 67].

Our experimental results suggest a practical method of controlling single polymer passage through nanopores. The successful passage rate can be controlled by simply displacing the \( trans \)-side solution with solution of different pH. It is low-cost, convenient and reversible. The same method should apply to a synthetic nanopore if the
pore surface is coated with charged species with pKa within the range of the solution pH used on the *trans* side. However, since the diameters of synthetic nanopores are generally much larger than that of αHL, EOF might contribute to the polymer passage more significantly.

In summary, we extracted the translocation dynamics from the blockage events and found that they were consistent with both the entropic and the enthalpic barrier models. Lowering the *trans*-pH altered the protonation states of the αHL nanopore, increasing its attraction to the translocating NaPSS polyelectrolyte. This resulted in a dramatic increase in successful translocation events. By constructing an equilibrium stochastic theory, we reproduced many features of our experimental results.
Table 5.1. The reversal voltages ($V_R$) measured in pH 7.5 1 M KCl (cis) and 200 mM KCl with variable pH (trans). The permeability ratio $P_+/P_-$ were calculated using Eq. (5.4).

<table>
<thead>
<tr>
<th>trans-pH</th>
<th>$V_R$ (mV)</th>
<th>$P_+/P_-$</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.5</td>
<td>-2.2</td>
<td>0.87</td>
</tr>
<tr>
<td>6.0</td>
<td>-8.1</td>
<td>0.60</td>
</tr>
<tr>
<td>5.5</td>
<td>-13.9</td>
<td>0.40</td>
</tr>
<tr>
<td>5.0</td>
<td>-16.9</td>
<td>0.32</td>
</tr>
<tr>
<td>4.5</td>
<td>-18.0</td>
<td>0.30</td>
</tr>
<tr>
<td>4.0</td>
<td>-18.5</td>
<td>0.28</td>
</tr>
<tr>
<td>3.5</td>
<td>-19.6</td>
<td>0.26</td>
</tr>
</tbody>
</table>
Figure 5.1. The αHL protein nanopore were embedded onto a lipid bilayer in a 1M KCl solution. Charged polymers were driven by the applied voltage to pass through the nanopore.
Figure 5.2. The space-filling modeling of αHL protein nanopore in pH 7.5. Blue: positively charged residues. Red: negatively charged residues. Green: hydrophobic residues. Gray: hydrophilic residues.
**Figure 5.3.** Two possible models for translocation dynamics through αHL, the entropic barrier model and the enthalpic barrier model. Both entropy and enthalpy may be responsible for the trapping of polymer inside the vestibule.
Figure 5.4. The ionic current (I) is plotted against the applied voltage (V) for αHL pores in 1M KCl solution with different pH values on the cis and trans sides. Open and solid symbols denote symmetric and asymmetric solution, respectively.
Figure 5.5. The equilibrium net charges of the 1.5 nm constriction and the charge ring at the end of the β-barrel of αHL.
Figure 5.6. The three most prominent event types (event 1, event 12 and event 2) in NaPSS translocation through αHL. The blockages were produced by 57.5 kg/mol NaPSS in 1M KCl pH 7.5 solution under 150 mV.
Figure 5.7. Typical ionic current traces in trans-pH 7.5 and 4.5, respectively, for 57.5 kg/mol NaPSS under 160 mV. The cis-pH were 7.5 in both cases.
Figure 5.8. The event diagrams and the percentages of event types for pH 7.5 and 4.5 on the trans side of the pore. The pH on the cis side was always 7.5. (a) and (b): Event diagrams, the normalized current blockage against the event duration, of 16 kg/mol NaPSS under 140 mV in trans-pH 7.5 and 4.5 solution respectively. 1000 events are plotted in both diagrams. The levels 1 and 2 blockages of events 12 are treated separately and denoted as events 12a and 12b, respectively. (c) and (d): Percentages of event types under the same conditions as in (a) and (b), respectively, at different voltages.
Figure 5.9. The average event duration is plotted against the molecular weight ($N$) divided by the applied voltage ($V$). Both $\langle \tau_{12b} \rangle$ and $\langle \tau_2 \rangle$ obeyed the same linear relation against $N/V$, while $\langle \tau_{12a} \rangle$ was insensitive to the changes in $N$ and $V$. 
Figure 5.10. The threshold voltage of successful translocation for 16 kg/mol NaPSS in cis-pH 7.5 decreased as trans-pH deceased. The error bars are smaller than the symbols.
Figure 5.11. The average blockage times of 16 kg/mol NaPSS at 140 mV changed with trans-pH. The durations of level 1 blockages, $\langle \tau_1 \rangle$ and $\langle \tau_{12a} \rangle$, decreased with decreasing pH while the durations of level 2 blockages, $\langle \tau_2 \rangle$ and $\langle \tau_{12b} \rangle$, increased with decreasing pH.
Figure 5.12. The strength of the electro-osmotic flow as a function of the pH value of the trans-side solution. The cis-pH was kept at 7.5.
Figure 5.13. Schematic free energy profile for a short NaPSS molecule translocating through αHL. With two absorbing boundaries, exiting on the left and right represent unsuccessful and successful translocations, respectively.
Figure 5.14. Free energy profiles of translocation with different polymer-pore interaction parameter $\epsilon$. $\delta = 1, \Delta F_1 = 10, \Delta F_2 = 5, M = 10, N = 50$ and $qV = -1$. 
Figure 5.15. Probabilities of successful translocation. The scatter points are the percentage of the sum of events 12 and 2 at 140 mV. The lines are evaluated using Eq. (5.9) with $\delta = 1$, $\Delta F_1 = 8$, $\Delta F_2 = 2$, $M = 17$, $N = 69$, $V = 140$ mV and $q = -0.3e$. 
Figure 5.16. The average successful translocation time $\langle \tau_t \rangle$, the average unsuccessful translocation time $\langle \tau_u \rangle$ and the successful translocation probability $P_t$. $\delta = 1$, $\Delta F_1 = 8$, $\Delta F_2 = 2$, $M = 17$, $N = 69$, $V = 140$ mV and $q = -0.3e$ are used.
Figure 5.17. Voltage dependence of $\langle \tau_1 \rangle$ for NaPSS with different chain lengths. The cis-pH and trans-pH were 7.5 and 4.5, respectively.
CHAPTER 6
POLYMER DIFFUSION THROUGH PERIODIC
CAVITIES

6.1 Introduction

Single polymer diffusion in crowded environments is ubiquitous in both biology and nanotechnology. Macromolecules diffuse a few times slower in the cytoplasm than in water, presumably due to squeezing through similar-sized obstacles [127]. In biotechnology, gene therapy requires the diffusion of DNA molecules in the crowded cytoplasm to the target [68]. Recently, experimental studies of single polymer diffusion in well-defined periodic structures have been accessible [97, 75, 92, 49, 47, 48, 50].

The diffusion of a macromolecule is limited by its surroundings. The chain adjusts its conformation according to the excluded volume interactions with the confining surfaces and its conformational entropy. The chain dynamics will be even more complex if the obstacles are also diffusing and have other interactions with the chain. For highly crowded environments, a chain deforms significantly by the excluded volume interaction between the obstacles and the chain. In particular, chain segments would localize in relatively spacious regions separated by some very narrow, bottleneck regions.

To model polymer diffusion in highly crowded environments, we consider a simplified system of a single polymer confining inside an one-dimensional periodic array of connected spherical cavities with internal radii $R$. The linearly aligned cavities are connected with small holes that only allow the chain to pass through with one of its chain ends. If the equilibrium chain size in free solution $R_g$ is much smaller than $R$,
the chain would behave like a small particle diffusing inside the cavity [37, 90]. Diffusing to an adjacent compartment through the connector is similar to the diffusion of a polymer across a membrane through a small hole [121, 85, 8] with polymer concentration $\sim 1/R^3$. The translocation requires temporary decreases in both translational and conformational entropies of the chain, which imposes a free energy barrier to the process. If $R \approx R_g$, the chain has no translational entropy inside a single cavity. The entropic barrier for the chain to translocate to an adjacent cavity is due to the entropy loss of fixing one chain end near the hole and the deformation of the chain when passing through the small hole. For longer chains, the chain prefers occupying several cavities as its equilibrium state. If the cavities are so small that the chain is almost stretched completely, the chain dynamics would be similar to worming through a cylindrical tube [32].

In this chapter, we focus our attention on the regime $R \leq R_g$, where a chain does not possess translational entropy inside a single cavity and it may occupy more than one cavity in equilibrium. Using simplified mean-field confinement free energies, we calculate the diffusion time and the diffusion coefficient of a chain along a one-dimensional periodic array of connected cavities by mapping the diffusion process into an one-dimensional periodic free energy profile.

### 6.2 Confinement free energy

The conformational entropy of a Gaussian chain with $N$ segments and segment length $b$ is governed by the equation [34]

$$
\left( \frac{\partial}{\partial N} - \frac{b^2}{6} \nabla_r^2 \right) P(r, r_0; N) = 0,
$$

(6.1)

where $P(r, r_0; N)$ is the probability, relative to the free space, that the two ends are at $r$ and $r_0$ respectively. Eq. (6.1) is solved with the condition $P(r, r_0; N) = \delta(r - r_0)$. If
the chain is confined in a spherical cavity, using the above boundary conditions gives

\[
PS(r, r_0; N) = \frac{1}{2\pi R^2} \sum_{n=0, B_{n+\frac{1}{2}}} P_n(\cos(\gamma)) (2n+1) \frac{J_{n+\frac{1}{2}}(B_{n+\frac{1}{2}} r) J_{n+\frac{1}{2}}(B_{n+\frac{1}{2}} r_0)}{(J_{n+\frac{1}{2}}(B_{n+\frac{1}{2}}))^2 \sqrt{rr_0}} \exp\left(-\frac{-b^2 B_{n+\frac{1}{2}}^2 N}{6R^2}\right),
\]

(6.2)

where \( \gamma \) is the angle between the vectors \( r \) and \( r_0 \), \( P_n(x) \) is the Legendre polynomial of order \( n \), \( J_n(x) \) is the Bessel function of order \( n \), and \( B_n \) relates to the zeros of the Bessel function by \( J_n(B_n) = 0 \). Note that for each \( n \), there is a set of \( B_{n+1/2} \) to be summed over. If the chain assumes a conformation in spherical symmetry, the probability is independent of the angle \( \gamma \). Eq. (6.2) reduces to

\[
PS(r, r_0; N) = \frac{1}{2\pi r r_0} \sum_{m=1}^{\infty} \sin\left(\frac{m\pi r_0}{R}\right) \sin\left(\frac{m\pi r}{R}\right) \exp\left(-\frac{-m^2 \pi^2 b^2 N}{6R^2}\right).
\]

(6.3)

If both ends are free to move inside the cavity, the probability is

\[
PS(R, N) = \int_0^R 4\pi r^2 dr \int_0^R 4\pi r_0^2 dr_0 PS(r, r_0; N)
= \frac{8R^3}{\pi} \sum_{m=1}^{\infty} \frac{1}{m^2} \exp\left(-\frac{-m^2 \pi^2 b^2 N}{6R^2}\right).
\]

(6.4)

Similarly, if one end of the chain is free while the other end is anchored at \( r = R - c \) near the surface of the compartment, where \( c \) is the anchoring distance and is comparable to \( b \), Eq. (6.3) becomes [88]

\[
PS_1(R, N) = \frac{2c}{R} \sum_{m=1}^{\infty} \exp\left(-\frac{-m^2 \pi^2 b^2 N}{6R^2}\right)
= \frac{2c}{R} f_{S1}(R, N).
\]

(6.5)

The \( O(c^2) \) terms are ignored. If both ends are anchored near the surface exactly opposite to each other, that is \( \gamma = \pi \), with anchoring distance \( c_1 \) and \( c_2 \), the probability is
\[ P_{S2}(R, N) = \frac{c_1 c_2}{2\pi R^5} \sum_{n=0, B_{n+1/2}} (-1)^n (2n + 1) \frac{B_{n+1}^2 J_{n-\frac{1}{2}}(B_{n+\frac{1}{2}})^2}{J_{n+\frac{1}{2}}(B_{n+\frac{1}{2}})^2} \exp \left( -\frac{b^2 B_{n+1}^2 N}{6R^2} \right). \] (6.6)

Using a self-consistent mean-field theory, Kong et al. showed that the free energy of an excluded volume chain inside a spherical cavity \( \sim wN^2/R^3 \), where \( w \) is the excluded volume parameter [60], which is valid for low segment densities. Motivated by the fact that the free energy barrier of translation is largely contributed by the entropy loss of fixing one chain end near the hole, we postulate that the free energy of confinement inside a closed cavity is the sum of the Gaussian entropy term and the excluded volume term described above. The entropy terms can be obtained from Eqs. (6.4), (6.5) and (6.6) using the ground state dominance approximation and the relation \( F = -k_B T \ln P \), where \( F \) is the free energy of the confined Gaussian chain, \( k_B T \) is the thermal energy, and \( P \) is the solution of Eq. (6.1). This approximated confinement free energy was recently shown to describe the partitioning of a single polymer between two spherical cavities [98].

It is instructive to write down our confinement free energies for several common conformations explicitly. Fig. 6.1 shows three different conformations that are the elements of our theory. Under the approximations mentioned above, using Eq. (6.4), the free energy of a chain that both of the chain ends are free to move inside the cavity is

\[ \frac{F_S}{k_B T} = \ln \left( \frac{\pi}{8R^3} \right) + \frac{\pi^2}{6} \frac{N}{R^2} + w \frac{N^2}{R^3}. \] (6.7)

If one end is free to move and the other end is fixed near the surface with a distance \( c \), the free energy is

\[ \frac{F_{S1}}{k_B T} = \ln \left( \frac{R}{2c} \right) + \frac{\pi^2}{6} \frac{N}{R^2} + w \frac{N^2}{R^3}. \] (6.8)

If both ends are fixed near the surface and are exactly opposite to each other (\( \gamma = \pi \)), the free energy is

\[ \frac{F_{S2}}{k_B T} = \ln \left( \frac{R^5}{2\pi c_1 c_2} \right) + \frac{\pi^2}{6} \frac{N}{R^2} + w \frac{N^2}{R^3}. \] (6.9)
Thus, the free energy increase for fixing one chain end near a hole on the surface is

\[
\Delta F_a / k_B T = F_{S1} / k_B T - F_S / k_B T
\]
\[
= F_{S2} / k_B T - F_{S1} / k_B T
\]
\[
= \ln \left( \frac{4R^4}{\pi c} \right). \tag{6.10}
\]

We have set \(c_1 = c_2 = c\). Note that \(\Delta F_a\) of an excluded volume chain is the same as that of a Gaussian chain and is independent of \(w\) in our theory, which would not be the case in actuality. To estimate the error, we performed Langevin dynamics simulations of a single polymer confined inside a spherical cavity. The free energies \(\Delta F_a\) were measured using the histogram method. As we can see from Fig. 6.2, \(\Delta F_a\) of an excluded volume chain is generally lower than that of a Gaussian chain, presumably due to lower number of configurations of an excluded volume chain inside the cavity. The free energy increase for fixing one end of an Gaussian chain near the cavity surface did not depend on whether the other end was fixed for large \(N\) (See Eq. (6.10)). We found that it was also true for an excluded volume chain. However, unlike Gaussian chains, \(\Delta F_a\) changes with \(N\) for excluded volume chains.

### 6.3 Translocation time

#### 6.3.1 1-2-1 translocation

We first consider the simplest case of polymer diffusion in periodic cavities as shown in Fig. 6.3. Initially, the chain is confined in one cavity with free energy \(F_S\). To start the translocation to the adjacent cavity, one of the chain ends has to find the hole that connects the two cavities. The free energy increase \(\Delta F_a\) is given by Eq. (6.10). The chain then undergoes a \(N\)-step translocation process and partitions between the two cavities, with \(p\) segments in the receptor and \(N - p\) segments in the original one. The translocation finishes when all segments transfer to the adjacent
cavity, that is \( p = N \). We call this process the 1-2-1 translocation because the chain initially occupies one cavity, then partitions into two cavities and finally occupies one cavity again.

The free energy profile of the 1-2-1 translocation is

\[
F(p) = \begin{cases} 
F_S, & -\frac{\delta}{2} \leq p < 0 \\
F_{S1}(p) + F_{S1}(N - p), & 0 \leq p \leq N \\
F_S, & N < p \leq N + \frac{\delta}{2} 
\end{cases} 
\]  

(6.11)

Explicitly, it is

\[
F(p)/k_B T = \begin{cases} 
\frac{\pi^2 N}{6R^2} + w\frac{N^2}{R^2}, & -\frac{\delta}{2} \leq p < 0 \\
\frac{2\Delta F_a}{k_B T} + \frac{\pi^2 N}{6R^2} + w\frac{p^2 + (N-p)^2}{R^3}, & 0 \leq p \leq N \\
\frac{\pi^2 N}{6R^2} + w\frac{N^2}{R^2}, & N < p \leq N + \frac{\delta}{2} 
\end{cases} 
\]  

(6.12)

where \( \delta \) is a small arbitrary constant denoting the length (in the reaction coordinate \( p \)) of the state that the chain is trapped in one sphere with both ends free. There is a local free energy minimum at \( p = N/2 \) due to the local stable state of partitioning \( N/2 \) segments in each cavity. Therefore, the chain has to overcome the entropic barrier \( 2\Delta F_a \) before proceeding to the local minimum state at \( p = N/2 \), and then it has to overcome the excluded volume barrier to finish the translocation.

Suppose the chain diffuses across \( g \) identical cavities, the free energy profile is periodic with \( g \) consecutive copies of the \( N \)-step translocation. The average translocation time is \([42]\)
\begin{align*}
\langle \tau_g \rangle &= \frac{1}{k_0} \int_{-\delta/2}^{g N + (g-1/2)\delta} dx \int_{-\delta/2}^{x} dy \exp \left( \frac{F(x) - F(y)}{k_B T} \right) \\
&= \frac{g (g - 1)}{2k_0} \int_{-\delta/2}^{N+\delta/2} dx \int_{-\delta/2}^{N+\delta/2} dy \exp \left( \frac{F(x)}{k_B T} \right) \\
&\quad + \frac{g}{k_0} \int_{-\delta/2}^{N+\delta/2} dx \int_{-\delta/2}^{x} dy \exp \left( \frac{F(x) - F(y)}{k_B T} \right) \\
&= \frac{g^2}{2k_0} \int_{-\delta/2}^{N+\delta/2} dx \int_{-\delta/2}^{N+\delta/2} dy \exp \left( \frac{F(x)}{k_B T} \right) \\
&= g^2 \langle \tau_1 \rangle. \tag{6.13}
\end{align*}

We have used the properties that the free energy is periodic with period $N + \delta$ and each is symmetric about $N/2$. $k_0$ is a phenomenological rate constant specified by the friction between one segment and the hole. Note that the double integral becomes two single integrals which allow easier evaluation. Hence it is sufficient to calculate the translocation time across one cavity $\langle \tau_1 \rangle$.

The chain travels a total distance of $2gR$ after crossing $g$ cavities. The diffusion constant is

\begin{equation}
D = \frac{(2gR)^2}{2 \langle \tau_g \rangle} = \frac{2R^2}{g \langle \tau_1 \rangle}. \tag{6.14}
\end{equation}

$\langle \tau_1 \rangle$ can be evaluated explicitly.

\begin{align*}
k_0 \langle \tau_1 \rangle &= \frac{1}{2} \int_{-\delta/2}^{N+\delta/2} dx \int_{-\delta/2}^{N+\delta/2} dy \exp \left( \frac{F(x)}{k_B T} \right) \\
&= \frac{\delta^2}{2} + \frac{1}{2} \exp \left( -\frac{wN^2}{2R^3} + 2\Delta F_a \right) \sqrt{\frac{\pi R^3}{2w}} \text{Erf} \left( \frac{N}{R^{3/2}} \sqrt{\frac{w}{2}} \right) \\
&\quad + \frac{1}{2} \exp \left( \frac{wN^2}{2R^3} - 2\Delta F_a \right) \sqrt{\frac{\pi R^3}{2w}} \text{Erf} \left( \frac{N}{R^{3/2}} \sqrt{\frac{w}{2}} \right) \\
&\quad + \frac{\pi}{4w} R^3 \text{Erf} \left( \frac{N}{R^{3/2}} \sqrt{\frac{w}{2}} \right) \text{Erfi} \left( \frac{N}{R^{3/2}} \sqrt{\frac{w}{2}} \right). \tag{6.15}
\end{align*}

Erf$(x)$ and Erfi$(x)$ are error function and imaginary error function respectively.
6.3.2 2-3-2 translocation

A 2-3-2 translocation is defined as the chain occupying two cavities in the initial state, going through intermediate states of occupying three cavities and advancing one cavity by occupying two cavities again.

According to Eq. (6.10), the free energy increase of fixing a chain end near the hole is still $\Delta F_a$, which is independent of the number of segments inside the cavity and does not depend on whether there is a segment fixed at the opposite hole. When the polymer occupies three cavities at the same time during a translocation process, the state of the polymer is described by the two variables $p$ and $q$, the number of segments inside the first and the second cavities respectively and the last cavity has $N - p - q$ segments. In order to keep the theory analytically tractable, we assume that the chain moves cooperatively. Under this assumption, the polymer always takes the conformation that the free energy is the lowest for any given $p$. Then we have $q = (N - p)/2$ segments in the second cavity and also $N - p - q = (N - p)/2$ in the last cavity. This is equivalent to assuming that the translocation process follows the saddle point of the free energy profile. Thus, we have mapped the two-dimensional stochastic process into a one-dimensional one with $2N/3 + \delta$ steps per period. The first half of the free energy profile of translocation is

$$F^*(p)/k_BT = \begin{cases} 
\frac{2\Delta F_a}{k_BT} + \frac{\pi^2}{6} \frac{N}{R^2} + \frac{w}{2} \frac{N^2}{R^3}, & -\frac{\delta}{2} \leq p < 0 \\
\frac{4\Delta F_a}{k_BT} + \frac{\pi^2}{6} \frac{N}{R^2} + w \frac{p^2}{R^3} + \frac{w}{2} \frac{(N-p)^2}{R^3}, & 0 \leq p \leq N/3 
\end{cases} \quad (6.17)$$

When $p$ proceeds to $N/3$, the chain is in the most stable state for occupying three cavities. Each cavities now have $N/3$ segments. The second half of the translocation process is identical to the reverse of the first half. In other words, the free energy profile of translocation is symmetric about $p = N/3$. Therefore the translocation time for advancing one cavity is
\[
\frac{k_0}{2} \langle \tau_1 \rangle = \frac{4}{2} \int_{-\delta/2}^{N/3} \exp \left( \frac{F^*(x)}{k_B T} \right) dx \int_{-\delta/2}^{N/3} \exp \left( -\frac{F^*(y)}{k_B T} \right) dy.
\] (6.18)

Note that the rate constant is now \( k_0/2 \) because we have assumed that the whole chain moves cooperatively and hence the total friction equals to the number of holes that the chain passes through at the same time.

By inspecting the free energy expression, we observe that the 2-3-2 translocation can be mapped to a 1-2-1 translocation. The idea is depicted in Fig. 6.5. The 2-3-2 translocation is equivalent to a \( N \)-segment chain confined in the initial cavity with excluded volume \( w/2 \). The chain overcomes the entropic barrier and finds the hole connecting to the next cavity, in which the excluded volume of the chain is \( w \). The chain then proceeds until the mid-point of the translocation process with \( p = N/3 \) segments in the next cavity. Likewise, the second half is the reverse of the above process. The process has \( 2N/3 + \delta \) steps.

### 6.3.3 \( m-(m+1)-m \) Translocation

Using the same idea, we generalize the calculation of the translocation time that the chain occupies \( m \) cavities initially. It is equivalent to a chain initially residing in a cavity of radius \( R \) and excluded volume \( w/m \) and translocating to another cavity of the same radius and excluded volume \( w \). It takes \( \delta/2 + N/(m+1) \) steps to finish the first half of the translocation with a free energy profile

\[
F^*(p)/k_B T = \begin{cases} 
\frac{2(m-1)\Delta F_a}{k_B T} + \frac{\pi^2}{6} \frac{N}{R^2} + \frac{w}{m} \frac{N^2}{R^3}, & -\frac{\delta}{2} \leq p < 0 \\
\frac{2m\Delta F_a}{k_B T} + \frac{\pi^2}{6} \frac{N}{R^2} + w \frac{p^2}{R^3} + \frac{w}{m} \frac{(N-p)^2}{R^3}, & 0 \leq p \leq \frac{N}{m+1}.
\end{cases}
\] (6.19)

Similarly, the second half is the reverse of the first half. Using Eq. (6.19), the translocation time across one cavity is

\[
\langle \tau_1 \rangle = \frac{2m}{k_0} \int_{-\delta/2}^{N/(m+1)} \exp \left( \frac{F^*(x)}{k_B T} \right) dx \int_{-\delta/2}^{N/(m+1)} \exp \left( -\frac{F^*(y)}{k_B T} \right) dy
\] (6.20)
6.4 Most probable occupation number

So far we have not specified \( m \) in the previous section and have treated it as an independent parameter. In fact, the probability that the chain occupies \( m \) cavities depends on its overall free energy. The free energy of a chain partitioning into \( m \) cavities is

\[
F_m(p_1, p_2, \ldots, p_{m-1}) = \frac{\pi^2 N}{6 R^2} + w \sum_{i=1}^{m-1} p_i^2 + (N - \sum_{i=1}^{m-1} p_i)^2 + 2(m-1)\Delta F_a, \quad (6.21)
\]

where \( p_i \) is the number of segments inside the \( i \)-th cavity. We denote \( F_m^* \) as the minimum free energy given by \( p_1 = p_2 = \ldots = p_{m-1} = N/m \), which we denote as \( F_m^* \).

Thus the most probable occupation number \( m^* \) is given by solving \( \partial F_m^*/\partial m^* = 0 \), which is

\[
m^* = \sqrt{\frac{w^2}{2\Delta F_a R^3}} N. \quad (6.22)
\]

The probability that a \( N \)-segment chain occupying \( m \) cavities is

\[
P(m) = \frac{\exp(-F_m^*/k_BT)}{\sum_{i=1}^{\infty} \exp(-F_i^*/k_BT)} \quad (6.23)
\]

and is plotted in Fig. 6.6. From Eq. (6.22), the chain extends inside the periodic cavity linearly proportional to its chain length \( N \). This relation is expected to hold for sufficiently large \( N \) when the chain occupies more than one cavity.

6.5 Most probable translocation time

A chain with \( N \) segments prefers occupying \( m^* \) cavities given by Eq. (6.22). Furthermore, Fig. 6.6 shows that the chain is unlikely to occupy a number of cavities that is far away from \( m^* \). We calculate the most probable time for translocating across one cavities \( \langle \tau_1^* \rangle \) by considering a \( m^*-(m^*+1)-m^* \) translocation. Using Eqs. (6.20) and
(6.22), we plot the most probable translocation time for a chain traveling a distance of one cavity in Fig. 6.7. For $N \to \infty$, $\langle \tau^*_1 \rangle$ becomes

$$
\langle \tau^*_1 \rangle = \frac{\pi R^{3/2}N}{k_0 \sqrt{8\Delta F_a w}} \text{Erf}(\sqrt{2\Delta F_a}) \text{Erfi}(\sqrt{2\Delta F_a}).
$$

(6.24)

If $\Delta F_a \gg 1$, Eq. (6.24) simplifies to

$$
\langle \tau^*_1 \rangle = \frac{4R^{19/2}N}{k_0 c^2 \pi^{3/2} \sqrt{w \ln(4R^4/\pi c)}}.
$$

(6.25)

The diffusion constant is given by $D = 2R^2/\langle \tau^*_1 \rangle \sim 1/N$, resembling the Rouse dynamics [34]. The most probable time for diffusing over a length of the chain size is $\langle \tau^*_d \rangle = (2m^* R)^2 / D \sim N^3$, resembling the reptation dynamics [34].

### 6.6 Discussions and conclusion

In this work, we estimated the translocation time and the diffusion constant of a chain inside a one-dimensional array of cavities. Due to the complexity of the problem, we have made the following assumptions in order to keep the theory analytically tractable. (1) We used a mean-field type free energy which was the sum of the entropic free energy of a Gaussian chain ($\sim N/R^2$) and the free energy of an excluded volume chain ($\sim wN^2/R^3$). (2) We assumed that the chain only moved in a way that it occupied $m$ and $m+1$ cavities alternatively. (3) We assumed that the whole chain moved cooperatively so that it was always in its lowest free energy state. (4) The friction constant of translocation was proportional to the number of holes the chain was passing through at the same time since the chain moved cooperatively.

With the above assumptions, we mapped the multi-dimensional diffusion process into a one-dimensional periodic stochastic process. By considering the most probable conformation of the chain inside the cavities, we find that the diffusion constant
\[ D \sim \frac{1}{N} \] and the time for diffusing one chain size \( \langle \tau^*_d \rangle \sim N^3 \), resembling the Rouse and the reptation dynamics respectively.

We expect the translocation dynamics will be vastly different for short chains that its equilibrium size is smaller than the size of the cavity. In this case, the entropic barrier for translocating to the adjacent cavity is not only due to the entropy loss for fixing one chain end near the hole, but also due to the loss of translational entropy of the center of mass of the chain. We expect such short chains would be trapped inside the cavity significantly longer than longer ones.
Figure 6.1. Three common conformations of a chain confined in a spherical cavity. Left: Both of the chain ends are free to move in the cavity. The free energy of the confined chain is $F_S$. Center: One end is free to move and the other end is anchored near the surface with anchoring distance $c$. The free energy is $F_{S1}$. Right: Both ends are fixed near the surface exactly opposite to each other, with anchoring distances $c_1$ and $c_2$ respectively. The free energy is $F_{S2}$. 
Figure 6.2. Anchoring free energies calculated from the Langevin dynamics simulations of a $N$-segment chain confined inside a spherical cavity with radius $R = 4.5d$, where $d$ is the diameter of the beads. Chains with and without excluded volume interactions were used. $\Delta F_a/k_B T$ calculated from Eq. (6.10) is 6.48, agreeing with the Gaussian chain simulations at large $N$. 

\[ \Delta F_a/ k_B T \]
Figure 6.3. 1-2-1 translocation. (a) The chain first reside in a single cavity. (b) A chain end finds the hole that connects to the neighboring cavity. The free energy of the chain increases by $2\Delta F_a$. (c) The chain undergoes translocation. There are $p$ segments in the neighboring cavity and $N - p$ segments in the original one. (d) The translocation is complete when all the segments are transferred to the neighboring cavity.
Figure 6.4. Periodic free energy profile experienced by the polymer chain. To start the translocation process, the chain has to overcome an entropic free energy barrier $2\Delta F_a$ for find the hole to the next cavity. In the middle of the translocation process, there is a free energy local minimum due to the minimized excluded volume repulsion in the expense of decreased entropy. To finish one period, the chain has to overcome the excluded volume barrier to return to the initial state of occupying one cavity.
Figure 6.5. (a) The 2-3-2 translocation starts with a chain partitioning in two cavities with $N/2$ segments in each one. One of the chain ends then finds the hole to the adjacent cavity. In the middle of the process, the chain occupies three cavities with $N/3$ segments each. The second half of the process is identical to the reverse of the first half. (b) Under the cooperativity assumption, the 2-3-2 translocation can be mapped to a 1-2-1 translocation. The excluded volume of the segments in the donor and the receptor cavities are $w/2$ and $w$ respectively. The new translocation process is also symmetric about $p = N/3$. 

\[ \text{Excluded volume} \]

\[ \left(0 < p < N/3\right) \]
Figure 6.6. The probability of a $N$-segment chain occupying $m$ cavities with radii $R = 4$, $w = 0.2$ and $c = 1$ are used.
Figure 6.7. The most probable time for a chain to diffuse a distance of one cavity. 
\( w = 0.2, c = 1 \) and \( \delta = 1 \) are used.
CHAPTER 7
POLYMER CONFINEMENT AND ESCAPE FROM A SPHERICAL CAVITY

7.1 Introduction

The escape of a single polymer from a cavity is commonplace in biological systems [5]. For example, DNA molecules are packaged in and released from viral capsids. Inside a cavity, the number of allowed conformations are significantly reduced and the limited space promotes many-body interaction among segments.

The confinement free energy of a polymer inside a spherical cavity have been studied using various theoretical formalisms. Kong et al. [60] applied the self-consistent mean field theory with segment-segment interaction to the problem and suggested that the free energy of a chain with \( N \) segments confined in a spherical cavity with internal radius \( R \) is

\[
F_{\text{MF}}/k_B T \sim N/R^2 + wN^2/R^3, \tag{7.1}
\]

where \( k_B T \) is the thermal energy and \( w \) is the excluded volume. The first term represents conformational entropy of an ideal chain and the second term accounts for the segment-segment interaction. Grosberg [43] argued that the pressure exerted by the chain on the cavity \( P = -\partial F_{\text{Sc}}/\partial V \big|_{R,N} \) should be a function of \( N/V \) only, where \( F_{\text{Sc}} \) denotes the confinement free energy of the scaling theory and \( V = 4\pi R^3/3 \) is the volume of the cavity. Using the scaling ansatz that \( F_{\text{Sc}} \) is a function of \( lN^\nu/R \) only, where \( l \) is the segment size and \( \nu \approx 0.6 \) is the size exponent of the chain in good solvents, the scaling free energy is

\[
F_{\text{Sc}}/k_B T \sim (lN^\nu/R)^{3/(3\nu-1)} \sim N^{2.25}/R^{3.75}. \tag{7.2}
\]
Intuitively, there should be a maximum number of chain segments that can be packaged into a closed space such as a spherical cavity. An important measure of such effect would be the packing fraction $\eta = N\sigma^3/8R^3$, where $\sigma$ is the diameter of the spherical segment. For a hard-core chain, the correct confinement free energy should diverge at $\eta = 1$ because it is impossible to pack extra segments into a completely filled space. However, both free energy expressions, Eqs. (7.1) and (7.2), assert that any finite number of segments can be put into the cavity with a finite amount of energy. In fact, the argument that leads to the scaling free energy $\mathcal{F}_{Sc}$ is strictly valid only for chains composed of point particles.

On the other hand, the time for a polymer chain to escape from a cavity is still not well understood. Two theoretical formalisms have been suggested to calculate the average escape time. In the *equilibrium* translocation theory, the chain is assumed to be in equilibrium at every step of the escape process. If the reduction in confinement free energy for transferring one segment outside the cavity is a constant $\Delta\mu$ throughout the process, it can be shown that the mean escape time follows $\langle \tau \rangle \sim N/\Delta\mu$ [86]. The *non-equilibrium* translocation theory states that a sufficiently long chain cannot stay equilibrium throughout the escape process. The *lower bound* of the mean escape time is estimated to be

$$\langle \tau \rangle \sim N^{1+\nu}/\Delta\mu,$$  \hspace{1cm} (7.3)

for $\Delta\mu$ staying constant throughout the process, where $\nu \simeq 0.6$ is the size exponent of the chain [27, 57, 19]. By employing the scaling theory and setting $\Delta\mu = -\partial \mathcal{F}_{Sc}/\partial N$ using Eq. (7.2), Cacciuto et al. [19] deduced that the lower bound of the mean escape time followed

$$\langle \tau \rangle \sim N^{1+\nu-\frac{1}{3\nu-1}} R^{\frac{3}{3\nu-1}} \sim N^{0.35} R^{3.75}$$  \hspace{1cm} (7.4)

if the escape process is out of equilibrium. Interestingly, both the equilibrium and the non-equilibrium predictions were supported by computer simulations [86, 19].
As an attempt to resolve these theoretical discrepancies, we studied the confinement and escape of a polymer inside a spherical cavity using Langevin dynamics computer simulations. We found that the confinement free energy was more adequately described by the hard sphere chain model. We also evaluated the validity of the equilibrium assumption during the escape process.

### 7.2 Confinement of a hard sphere chain inside a spherical cavity

The Green’s function of a hard sphere chain, given by the path integral formalism, is

\[ G(r - r') = \mathcal{N} \int_{R_0=r}^{R_N=r'} \exp \left[ -\frac{3}{2\sigma^2} \sum_{n=1}^{N} (R_n - R_{n-1})^2 - \frac{1}{k_B T} \sum_{i<j} U(R_i - R_j) \right] \delta \mathbf{R}. \]  

(7.5)

\( \mathbf{R}_i \) denotes the position of the center of the \( i \)-th bead. \( r \) and \( r' \) are positions of the two chain ends. The integral runs over all beads and covers only the volume inside the cavity. \( \mathcal{N} \) is a normalization constant. \( G(r - r') \) is the probability of the chain relative to the free state. The hard sphere potential \( U(r) \) is

\[ U(r) = \begin{cases} \infty, & r < \sigma \\ 0, & r > \sigma \end{cases}. \]  

(7.6)

The confinement free energy is given by \( \mathcal{F} = -k_B T \ln \int d\mathbf{r} \int d\mathbf{r}' G(r - r') \). Evaluating Eq. (7.5) is, however, analytically formidable. We assume that the hard sphere chain confinement free energy can be written as \( \mathcal{F}_{\text{HSC}} = \mathcal{F}_{\text{id}} + \mathcal{F}_{\text{hs}} \). Equivalently, we approximate Eq. (7.5) by integrating the two terms in the integral separately. Then the entropic contribution to the free energy is \( \mathcal{F}_{\text{id}} \sim N/R^2 \) [34] and the hard
sphere multiple body interaction $F_{hs}$ is given by the excess free energy (in addition to the idea gas) of a hard sphere fluid, which is well studied and various approximation schemes are available [12]. We will employ the Carnahan-Starling approximation [22], which matches the exact hard sphere free energy up the sixth virial coefficient. The confinement free energy of a hard sphere chain is then in the form

$$F_{HSC} \sim a \frac{N}{R^2} + bN \left( \frac{1}{(1 - \eta)^2} + \frac{2}{1 - \eta} - 3 \right),$$

(7.7)

where $a$ and $b$ are unknown parameters. Note that Eq. (7.7) diverges at $\eta = 1$ as expected.

### 7.3 Pressure exerted by the chain on the cavity

To test the validity of Eq. (7.7), we performed Langevin dynamics simulations of a flexible polymer chain, composed of connected spherical beads with diameter $\sigma$, confined in a spherical cavity. The beads interacted with each other via the Lennard-Jones pairwise potential [7]. The bead-bead interaction strength was 0.2 kcal/mol. The wall-bead interaction strength was twice as large and purely repulsive. The bead diameter $\sigma$ was 5 Å. The viscosity was 1 cP. Temperature was 25°C. The mass of each bead was 120 atomic mass units. The parameters were set to simulate a protein molecule in water without secondary structures. The unit time in the simulation was 3 ps. The pressure exerted on the internal wall by the confined chain was measured directly in the simulation. On the other hand, the free energy of confinement $F$ is related to the pressure $P$ by $P = -\partial F/\partial V|_{R,N}$. Using Eqs. (7.1), (7.2) and (7.7), the pressure predicted by the mean field, the scaling and the hard sphere chain theories are
\[
\frac{P_{\text{MF}}}{k_B T} \sim \frac{N}{R^5} + \frac{N^2}{R^6},
\]
\[
\frac{P_{\text{Sc}}}{k_B T} \sim \frac{N^{3\nu/(3\nu-1)}}{R^{3/(3\nu-1) + 3}} \sim \frac{N^{2.25}}{R^{6.75}},
\]
\[
\frac{P_{\text{HSC}}}{k_B T} \sim \frac{a}{2\pi R^5} + b \frac{N}{V} (2\eta) \left( \frac{2 - \eta}{1 - \eta} \right)^3,
\]
respectively. The pressure obtained from the simulations and the corresponding fits of the expressions Eqs. (7.8a), (7.8a) and (7.8c) are shown in Fig. 7.1. The hard sphere chain theory fitted best to the rapidly increasing pressure with \( N \) due to many-body interactions. The pressure expressions derived from the mean field and the scaling theories did not increase rapidly enough with \( N \). The result is not surprising given that their \( N \) and \( R \) dependence are quite similar (\( N^2 R^{-6} \) and \( N^{2.25} R^{-6.75} \)). The plots of the \( P \) against \( R \) give similar results.

### 7.4 Polymer escape from a cavity through a small hole

We studied the escape of the polymer from the spherical cavity. The chain was first allowed to equilibrate 4\( N^2 \) simulation time units inside the cavity. It was roughly twice the longest relaxation time of the chain in the free space. A hole with round corner and 1.5\( \sigma \) diameter was then opened on the cavity wall to allow the spontaneous escape of the chain. The hole was small enough that the chain could only escape linearly starting from one of the chain ends. The wall of the cavity was 1\( \sigma \) thick. The escape time was defined as the duration from the last time when only one bead had escaped to the time when the whole chain had just escaped. \( N \) ranged from 25 to 200 and \( R \) ranged from 1.8\( \sigma \) to 4.1\( \sigma \). The initial packing fractions ranged from 0.27 to 0.53. The translocation times were averaged over at least 100 runs. The details of the escape dynamics were described elsewhere [86].

We observed that the escape processes were stochastic. A typical escape dynamics is shown in Fig. 7.3. In the beginning, the escape process behaved like a biased
random walk, suggesting a significant change in the confinement free energy as the escape proceeded. At the end of the escape, the change in confinement free energy was small and the process behaved like an unbiased random walk. This escape picture contradicts with the earlier theoretical treatments that the change in confinement free energy per escaped segment $\Delta \mu$ is constant throughout the process.

After the hole was opened, one of the chain ends had to find the hole to start the escape. We call this duration the waiting time. Fig. 7.4 shows the typical distributions of the waiting time and the escape time. The escape time had a peak at a non-zero value with a characteristic long tail of a stochastic process. In contrast, the waiting time distribution exhibited a simple decay.

A previous work [19] studied the escape process using Monte Carlo simulations of hard sphere chains. They used the scaling confinement free energy, Eq. (7.2), together with the lower bound of the translocation time [57] to predict Eq. (7.4) ($\langle \tau \rangle \sim N^{1+\nu-\frac{1}{2\nu+1}} R^{-3\nu} \sim N^{0.35} R^{3.75}$). Interestingly, this formula described our simulation results well, only with systemic deviations at small $\tau$. However, the scaling confinement free energy was shown to be inadequate in describing the pressure of the confined chain (see Fig. 7.1). In addition, the expression of $\Delta \mu$ used in deriving Eq. (7.4) is not a constant throughout the process. Its applicability to Eq. (7.3) is unclear.

### 7.5 Escape of a gaussian chain

In the study of the dynamics of polymer escape, both the confinement free energy and the equilibrium assumption of the escape are not clearly understood. To eliminate one unknown, we studied the escape of a Gaussian chain from the a cavity so that the confinement free energy is given by $F_{Gauss} \sim N/R^2$ for sufficiently long chains. Therefore $\Delta \mu \sim 1/R^2$ and the average equilibrium and non-equilibrium escape times are $N R^2$ and $N^{1+\nu} R^2$ respectively. We repeated the escape simulations using chains
without intra-chain interaction. The result is shown in Fig. 7.5. We found that the equilibrium dynamics described the escapes of Gaussian chains more adequately.

7.6 Discussions and conclusions

In this chapter, we studied the confinement free energy of a chain inside a spherical cavity. We found that, due to the limited space available inside the cavity, the hard sphere chain model was a better candidate in describing the confinement effect. Interestingly, the scaling confinement free energy together with the non-equilibrium escape dynamics described the averaged escape time of the polymer well, although the scaling confinement free energy failed to describe the pressure exerted by the confined chain on the cavity wall. We further studied the escape of a Gaussian chain out of a cavity. With the known confinement free energy, we found that the equilibrium dynamics described our simulation well.

Although the average escape times of real chains were well described by Eq. (7.4), the theoretical argument that led to the equation was unclear. Even we assume that the scaling confinement free energy is applicable to our situation, \( \Delta \mu = -\partial F_{sc}/\partial N \) calculated from the scaling free energy Eq. (7.2) cannot be used directly in the lower bound of the translocation time, Eq. (7.3), because it is not a constant throughout the process. Moreover, the scaling free energy was shown inadequate in describing the pressure of the chain inside the cavity. Therefore, the reasons for the success of Eq. (7.4) remain unclear.

We observed that the escape of a Gaussian chain followed the equilibrium translocation dynamics. One possibility is that the escapes of both real and Gaussian chains might be in equilibrium at every steps. In this case, we need to use a more realistic \( \Delta \mu \) and without assuming it as a constant in deriving the mean escape time. The other possibility is that the lower driving force of the escape of a Gaussian chain allowed the chain to equilibrate more easily during the translocation. Then we expect
the escape would become non-equilibrium for longer chains because of the longer relaxation times.

We note that we only used the confinement free energy of the polymer when calculating $\Delta \mu$. During the escape, there are segments both inside and outside the cavity and the confinement free energy only accounts for the segments inside the cavity. Moreover, the confinement free energy is likely to be invalid near the end of the escape, where only relatively few segments left in the cavity. As evident from Eq. 7.3, the escape time of the last part of the chain accounts for a significant part of the total escape time because of the lower driving force. A more realistic theoretical analysis should include the effect of the segments outside the cavity and the random-walk dynamics near the end of the escape.

In summary, we found that the scaling confinement free energy together with the non-equilibrium translocation theory described our simulation results successfully, although the scaling free energy failed to describe the pressure exerted by the chain on the wall of the cavity. We further found that the escape of a Gaussian chain followed the equilibrium dynamics. More in-depth theoretical studies are needed for these findings to be reconciled.
Figure 7.1. Pressure exerted by the confined chain on the cavity wall. The best fits of the pressures predicted by the mean field theory (Eq. (7.8a)), the scaling theory (Eq. (7.8b)) and the hard sphere chain theory (Eq. (7.8c)) are shown on top of the data. $R = 3.5\sigma$ in the simulation. The pressure is in the simulation unit.
Figure 7.2. All data of the average escape times obtained from the simulations are plotted against the values predicted by Eq. 7.4. The solid line has a slope of 1 and passes through the origin. The error bars are smaller than the point size.
Figure 7.3. The escape trajectory of a chain with 100 beads from a cavity of radius $3.4\sigma$. The time is in the simulation unit and is measured from the start of the escape. The escape process is highly stochastic especially near the end of the process. The escape of the last 20 beads is responsible for about half of the escape time.
Figure 7.4. (a) The distribution of the waiting time for one end of the confined chain to find the hole and begin the escape. Solid line is the fit of the form $\exp(-a\tau)$. (b) The distribution of the escape time for a chain with one bead already outside the cavity. Solid line is the fit of the form $(\tau - a)^b \exp(-c(\tau - a))$. $N = 25$ and $R=2.55\sigma$ in the simulations.
Figure 7.5. The average escape time of a Gaussian chain was well described by the prediction of equilibrium dynamics.
CHAPTER 8
GENERAL CONCLUSION

The studies on single polymer dynamics inside confined geometries presented in this thesis are far from complete. Below are the directions of research that are, in our opinion, worth pursuing.

In Chapter 2, we studied the polymer capture by the electro-osmotic flow induced by the surface charge of a nanopore. The next step would be studying the polymer translocation process after the capture. The electro-osmotic flow would likely to be altered because of the presence of the polymer inside the pore. Both the electric field and the electro-osmotic flow drive the polymer across the pore. It would be interesting to investigate the effect of the surface charges of the nanopore on the experimentally observable quantities such as the translocation times and the percentages of folded translocation events. The percentages of folded translocation events are likely to be related to the higher modes of the Zimm relaxation times of the polymer. Another unsolved problem is the translocation rate and the translocation time when the surface charge of the nanopore is of the same sign as the charge of the polymer. In this case, the electro-osmotic flow would oppose instead of facilitate the polymer capture.

In Chapter 3, we studied polymer translocation through a cylindrical channel. We employed an approximation to average out the degree of freedom of the number of segments inside the channel, so that we could map the process into a one-dimensional one. A more explicit but involved treatment would be studying the process using a two-dimensional Fokker-Planck equation. Furthermore, we studied the unrealistic process of the translocation of a Gaussian chain. Significant advancement can be made
by adding excluded volume interactions, hydrodynamic interactions and electrostatic interactions among a polyelectrolyte, its counterions and salt ions. In particular, the entropy of counterions and salt ions are expected to play an important role in the equilibrium partitioning and the translocation of a polyelectrolyte chain between two compartments.

In Chapter 4, we studied the electrophoretic behavior of a polyelectrolyte passing through periodic confined regions. We used the confinement free energy of a flexible chain inside the confined regions. In some experiments, the heights of the confined regions were comparable to the persistence length of double-stranded DNA. However, the analytical form of the confinement free energy in the crossover regime between confining a flexible and a semi-flexible chains is still unknown. On the other hand, most nano-fluidic channels have surface charges. The electro-osmotic flow inside the nano-fluidic channels can be significant in some cases. Depending on the sign and the magnitude of the surface charge, the flow would either assist or oppose the electrophoretic motion of the DNA due to the applied electric field. The effective electrophoretic behavior of a polyelectrolyte in the presence of both the electric field and the electro-osmotic flow is not well understood.

In Chapter 5, we studied the effect of pH conditions on polymer translocation through α-Hemolysin protein nanopores. We proposed that the free energy barrier of translocation could be written in the form $\Delta F = \Delta H - T\Delta S$. To obtain the quantitative entropic and enthalpic contributions, systematic studies at different controlled temperatures can be conducted. Also, we obtained the expressions for evaluating translocation times and probabilities by solving the Fokker-Planck equation using two absorbing boundary conditions. A systematic study of polymer translocation through an infinitely narrow channel with polymer-pore interaction using these expressions is still missing. Such study will provide insightful results for understanding the free energy barrier of polymer translocation through a narrow channel.
In Chapter 6, we employed the cooperativity assumption to evaluate the diffusion constant of a chain trapped inside a one-dimensional array of period cavities. In reality, it is unlikely that the chain would move cooperatively so that it always assumes the conformation with the lowest free energy. Instead, the movement is more likely to be driven by the density differences among the cavities that the chain occupies. The density inhomogeneity may propagate along the chain as a wave. The situation is similar to the reptation dynamics originally studied by De Gennes [32].

In Chapter 7, we studied the confinement and escape of a chain inside a spherical cavity. As in other similar studies, there were no explicit solvents in our simulations. In such confined environment, the solvent molecules are likely to have significant effects on the pressure measured in our simulations. In addition, the escape of a polyelectrolyte chain with explicit counterions is another interesting fundamental problem.
APPENDIX

DERIVATIONS OF THE MEAN FIRST PASSAGE TIMES

Our derivations follow closely with that given by Gardiner [42]. Starting with the one-dimensional Fokker-Planck forward equation,

\[ \frac{\partial p(x,t|x',t')}{\partial t} = -\frac{\partial}{\partial x} [A(x,t)p(x,t|x',t')] + \frac{\partial^2}{\partial x^2} [B(x,t)p(x,t|x',t')], \] (A.1)

where \( A(x,t) \) is the drift term and \( B(x,t) \) is the diffusive term. This equation describes the random motion of a Brownian particle under drifts at different locations. \( p(x,t|x',t') \) is the probability that the particle starts at position \( x' \) and time \( t' \) and ends at position \( x \) and time \( t \). Note that the derivatives are with respect to the present position \( x \) and time \( t \) (and hence the equation is called a forward equation).

It can be shown that Eq. (A.1) is equivalent to its backward equation

\[ \frac{\partial p(x,t|x',t')}{\partial t'} = -A(x',t') \frac{\partial p(x,t|x',t')}{\partial x'} - B(x',t') \frac{\partial^2 p(x,t|x',t')}{\partial x'^2}. \] (A.2)

Note that the derivatives are now with respect to the initial position \( x' \) and time \( t' \).

Consider a probability distribution \( p(x',t|x,0) \). If the time is homogeneous, both \( A \) and \( B \) do not depend on time and \( p(x',t|x,0) = p(x',0|x,-t) \). It follows that the backward equation Eq. (A.2) can be rewritten as

\[ \frac{\partial p(x',t|x,0)}{\partial t} = A(x) \frac{\partial p(x',t|x,0)}{\partial x} + B(x) \frac{\partial^2 p(x',t|x,0)}{\partial x^2}. \] (A.3)
Suppose the particle is restricted inside a region \( a \leq x \leq b \). A reflecting wall or an absorbing point to the particle may be located at \( a \) and \( b \). The survival probability, the probability that the particle is still in the region \( a \leq x \leq b \) at time \( t \), is

\[
G(x, t) = \int_{a}^{b} p(x', t|x, 0)dx',
\]

(A.4)

where \( x \) denotes the initial position at time \( t = 0 \). It is also the probability that the first passage time \( T(x) \), the time that the particle first exits the region \( a \leq x \leq b \), is larger than \( t \). That is

\[
\text{Prob}(T(x) > t) = G(x, t).
\]

(A.5)

Note that \( G(x, 0) = 1 \) and \( G(x, \infty) = 0 \) if there is at least one absorbing boundary. By integrating Eq. (A.3) with respect to \( x' \) and using Eq. (A.4), we get a differential equation for \( G(x, t) \):

\[
\frac{\partial G(x, t)}{\partial t} = A(x)\frac{\partial G(x, t)}{\partial x} + B(x)\frac{\partial^2 G(x, t)}{\partial x^2}.
\]

(A.6)

Since \( 1 - G(x, t) \) is the cumulative probability that the particle has exited at time \( t \), the probability distribution of the first passage time \( T(x) \) is \(-\partial G(x, t)/\partial t \). It follows that the mean first passage time (MFPT) is

\[
\langle T(x) \rangle = -\int_{0}^{\infty} t \frac{\partial G(x, t)}{\partial t}dt = \int_{0}^{\infty} G(x, t)dt,
\]

(A.7)

where integration by part and boundary conditions for \( G(x, t) \) are used. Hence we arrive at an ordinary differential equation for the MFPT by integrating Eq. (A.6)

\[
A(x) \frac{d\langle T(x) \rangle}{dx} + B(x) \frac{d^2\langle T(x) \rangle}{dx^2} = -1.
\]

(A.8)

Note that we have used \( \int_{0}^{\infty} \partial G(x, t)/\partial tdt = G(x, \infty) - G(x, 0) = -1 \). We reiterate that \( x \) in Eq. (A.8) denotes the initial position of the particle that time \( t = 0 \).
A.1 General solution of the MFPT equation

Eq. (A.8) is an inhomogeneous second order ordinary differential equation which can be solved by standard methods. The general solution is the sum of the complementary solution $\langle T_c(x) \rangle$ and the particular solution $\langle T_p(x) \rangle$.

To find $\langle T_c(x) \rangle$, we need to solve the homogeneous solution

$$A(x)\frac{d\langle T_c(x) \rangle}{dx} + B(x)\frac{d^2\langle T_c(x) \rangle}{dx^2} = 0. \quad (A.9)$$

This equation can be solved by making the coefficient of $d^2\langle T_c(x) \rangle/dx^2$ unity and multiplying an integrating factor

$$\psi(x) = \exp \left( \int_a^x \frac{A(x')}{B(x')} \, dx' \right) \quad (A.10)$$
on both sides. The solution is

$$\langle T_c(x) \rangle = C_1 \int \frac{dx}{\psi(x)} + C_2, \quad (A.11)$$

where $C_1$ and $C_2$ are constants to be fixed by the boundary conditions. $\langle T_p(x) \rangle$ can be found by the method of variation of parameters [108]. From Eq. (A.11), the two independent solutions of Eq. (A.9) are $T_1(x) = \int dx/\psi(x)$ and $T_2(x) = 1$. The particular solution is

$$\langle T_p(x) \rangle = -T_1(x) \int \frac{T_2(x)g(x)}{W(x)} \, dx + T_2(x) \int \frac{T_1(x)g(x)}{W(x)} \, dx, \quad (A.12)$$

where $g(x) = -1/B(x)$ and the Wronskian $W(x) = -1/\psi(x)$. Hence the particular solution is

$$\langle T_p(x) \rangle = -\left( \int \frac{dx}{\psi(x)} \right) \left( \int \frac{\psi(x)}{B(x)} \, dx \right) + \int \left[ \left( \int \frac{dx}{\psi(x)} \right) \left( \frac{\psi(x)}{B(x)} \right) \right] \, dx, \quad (A.13)$$
The general solution is given by

\[ \langle T(x) \rangle = \langle T_c(x) \rangle + \langle T_p(x) \rangle. \]  \hspace{1cm} (A.14)

### A.2 One absorbing and one reflecting boundaries

Suppose there is an reflecting wall at \( a \) and an absorbing point at \( b \). The particle starts at position \( x \) at time \( t = 0 \) and exhibits Brownian motions in the interval \( (a,b) \). The particle would reflect back to \( (a,b) \) if it touches \( a \) and exit the interval if it touches \( b \). The boundary conditions are

\[ \left. \frac{\partial G(x,t)}{\partial x} \right|_{x=a} = 0, \]  \hspace{1cm} (A.15)

\[ G(b,t) = 0. \]  \hspace{1cm} (A.16)

Using Eq. (A.7), the boundary conditions become

\[ T'(a) = 0, \]  \hspace{1cm} (A.17)

\[ T(b) = 0. \]  \hspace{1cm} (A.18)

Applying these boundary conditions to the general solution Eq. (A.14) to determine \( C_1 \) and \( C_2 \), we obtain

\[ \langle T(x) \rangle = \int_x^b dy \int_a^y dz \frac{\psi(z)}{\psi(y)B(z)}, \]  \hspace{1cm} (A.19)

where \( \psi \) is given in Eq. (A.10). In the Fokker-Planck equation used in this thesis, \( A(x) = -k_0 \partial F(x)/\partial x \) and \( B(x) = k_0 \) give

\[ \psi(x) = \exp[-F(x) + F(a)]. \]  \hspace{1cm} (A.20)
Using Eqs. (A.19) and (A.20), the MFPT is

\[ \langle T(x) \rangle = \frac{1}{k_0} \int_x^b dy \int_y^a dz \exp [F(y) - F(z)], \]  

(A.21)

where \( F(m) \) is the free energy profile and \( x \) is the initial position of the particle in the free energy profile.

**A.3 Two absorbing boundaries**

Suppose there absorbing points at both \( a \) and \( b \), starting at the initial position \( a \leq x \leq b \), the particle is allow to exhibit random motions between interval \((a, b)\). It will exit the interval once it touches either \( a \) or \( b \). The corresponding boundary conditions are

\[ T(a) = 0, \]  

(A.22)

\[ T(b) = 0. \]  

(A.23)

Applying these boundary conditions to Eq. (A.14) gives the MFPT

\[ \langle T(x) \rangle = \left( \int_a^x \frac{dy}{\psi(y)} \right) \left( \int_x^b \frac{dy \int_y^b dz \frac{\psi(z)}{\psi(y)B(z)}}{\psi(y)} \right) - \left( \int_a^b \frac{dy}{\psi(y)} \right) \left( \int_a^x \frac{dy \int_y^a dz \frac{\psi(z)}{\psi(y)B(z)}}{\psi(y)} \right) \int_a^b \frac{dy}{\psi(y)}. \]  

(A.24)

Note that this formula does not distinguish between the two exits \( a \) and \( b \).

**A.4 Conditional exit probabilities and mean exit times in two absorbing boundaries**

Depending on the initial position \( x \) and the drifts, the probability of exiting at one end is likely to be different from that at the other end. In this section, we will obtain the conditional exit probabilities and the corresponding mean exit times.
Consider two absorbing boundaries \(a\) and \(b\). The probability that the particle starting at \(x\) \((a \leq x \leq b)\) exits the interval at \(a\) is at a time greater than \(t\) is

\[
g_a(x,t) = -\int_t^\infty dt' J(a,t'|x,0),
\]

\[
= \int_t^\infty dt' \{-A(a)p(a,t'|x,0) + \frac{\partial}{\partial x}[B(a)p(a,t'|x,0)]\},
\]

where \(J(x,t|x',t')\) is the probability current satisfying the forward equation

\[
\frac{\partial p(x,t|x',t')}{\partial t} = -\frac{\partial J(x,t|x',t')}{\partial x}.
\]

The negative sign in Eq. (A.25) accounts for the probability current pointing to the left. Using Eq. (A.26) and the fact that \(p(a,t'|x,0)\) satisfies the backward Fokker-Planck equation Eq. (A.3), we obtain an equation for \(g_a(x,t)\),

\[
A(x)\frac{\partial g_a(x,t)}{\partial x} + B(x)\frac{\partial^2 g_a(x,t)}{\partial x^2} = \int_t^\infty dt' \left\{-A(a)\frac{\partial p(a,t'|x,0)}{\partial t'} + \frac{\partial}{\partial x}[B(a)\frac{\partial p(a,t'|x,0)}{\partial t'}]\right\}
\]

\[
= J(a,t|x,0)
\]

\[
= \frac{\partial g_a(x,t)}{\partial t}.
\]

Given the particle exits at \(a\), the probability that the particle exits after \(t\) is \(P(T_a > t) = g_a(x,t)/g_a(x,0)\). The probability distribution of the first passage time (given the particle exits at \(a\)) \(T_a(x)\) is then \(-\partial P(T_a > t)/\partial t\). It follows that the MFPT for exiting at \(a\) is

\[
\langle T_a(x) \rangle = \int_0^\infty dt \left(-t\frac{\partial P(T_a > t)}{\partial t}\right) = \int_0^\infty g_a(x,t)dt/g_a(x,0).
\]

Therefore, we get the equation for \(\langle T_a(x) \rangle\) by integrating Eq. (A.28) with respect to \(t\):

\[
A(x)\frac{\partial [P_a(x)\langle T_a(x) \rangle]}{\partial x} + B(x)\frac{\partial^2 [P_a(x)\langle T_a(x) \rangle]}{\partial x^2} = -P_a(x),
\]

\[
152
\]
where \( P_a(x) = g_a(x,0) \) is the probability of the particle exiting at \( a \). The boundary conditions for Eq. (A.30) are

\[
P_a(a)\langle T_a(a) \rangle = P_a(b)\langle T_a(b) \rangle = 0 \tag{A.31}
\]

since \( \langle T_a(a) \rangle = 0 \) and \( P_a(b) = 0 \). The equation for \( P_a(x) \) can be obtained by setting \( t = 0 \) in Eq. (A.28),

\[
A(x) \frac{\partial P_a(x)}{\partial x} + B(x) \frac{\partial^2 P_a(x)}{\partial x^2} = 0, \tag{A.32}
\]

with boundary conditions \( P_a(a) = 1 \) and \( P_a(b) = 0 \). Eqs. (A.30) and (A.32) can be solved by the same method as that of Eq. (A.8). In fact, Eq. (A.32) is in the same form of the homogenous equation of Eq. (A.30). The MFPT for the particle exiting at \( a \) is

\[
\langle T_a(x) \rangle = \frac{\left( \int_a^x \frac{dz}{\psi(z)} \right) \left( \int_a^b \frac{dz}{\psi(z)} \right) \left( \int_a^y \frac{dy}{\psi(y)} \right) \left( \int_a^b \frac{dy}{\psi(y)} \right) - \left( \int_a^b \frac{dz}{\psi(z)} \right) \left( \int_a^b \frac{dz}{\psi(z)} \right) \left( \int_a^y \frac{dy}{\psi(y)} \right) \left( \int_a^b \frac{dy}{\psi(y)} \right)}{\left( \int_a^x \frac{dz}{\psi(z)} \right) \left( \int_a^b \frac{dz}{\psi(z)} \right)} \tag{A.33}
\]

with probability of exiting at \( a \) as

\[
P_a(x) = \int_a^b \frac{dz}{\psi(z)} / \int_a^b \frac{dz}{\psi(z)}. \tag{A.34}
\]

Following a similar derivation, the MFPT for the particle exiting at \( b \) is

\[
\langle T_b(x) \rangle = \frac{\left( \int_a^x \frac{dz}{\psi(z)} \right) \left( \int_b^y \frac{dy}{\psi(y)} \right) \left( \int_a^b \frac{dy}{\psi(y)} \right) \left( \int_a^b \frac{dy}{\psi(y)} \right) - \left( \int_a^b \frac{dz}{\psi(z)} \right) \left( \int_a^x \frac{dz}{\psi(z)} \right) \left( \int_a^y \frac{dy}{\psi(y)} \right) \left( \int_a^b \frac{dy}{\psi(y)} \right)}{\left( \int_a^x \frac{dz}{\psi(z)} \right) \left( \int_a^b \frac{dz}{\psi(z)} \right)} \tag{A.35}
\]

with probability of exiting at \( b \) as

\[
P_b(x) = \int_a^x \frac{dz}{\psi(z)} / \int_a^b \frac{dz}{\psi(z)}. \tag{A.36}
\]
The two probabilities satisfy $P_a(x) + P_b(x) = 1$, of course. $\psi(x)$ is defined in Eqs. (A.10) or (A.20).
BIBLIOGRAPHY


