Knot Formation on Nanochannel-Confined DNA and Flow-Assisted DNA Disentanglement and Linearization

Susan Amin

Doctor of Philosophy

Department of Physics

McGill University

Montreal,Quebec

April, 2018

A Dissertation Submitted to the Faculty of Graduate Studies in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

ACKNOWLEDGEMENTS

I would like to thank my family: my parents and my sister for supporting me spiritually throughout my PhD and my life in general. I am extremely thankful to my husband for the unceasing encouragement, support and attention.

I would like to express my sincere gratitude to my supervisor Prof. Walter Reisner for sharing expertise, and invaluable guidance and effort extended to me. I would also like to thank the rest of my PhD committee: Prof. Robert Sladek and Prof. Paul Wiseman for their insightful comments and encouragement.

I wish to express my deep thanks to my collaborators: Lili Zeng, Ahmed Khorshid, Sara Mahshid, Philip Zimny, Peter Yao, Ilja Czolkos and Yi Cao for their great help with the projects. I am also grateful to Yuning Zhang, Mohammed Jalal Ahamed, Pierre Berube and Haig Djambazian who provided me with their generous help and insightful discussions. I take this opportunity to express gratitude to my all fellow lab mates with whom I shared my time and thoughts. I would also like to express my gratitude to Dr. Yoichi Miyahara for helping me with AFM categorization of my samples. Finally, I appreciate the help and patience from all my friends who supported me unconditionally.

ORIGINALITY AND CONTRIBUTION

This dissertation represents the work I have done as a PhD student at McGill University. The thesis is written in the article-based format and includes three manuscripts.

❄ Chapter 1 – a general introduction and motivation.

❄ Chapter 2 – background.

❄ Chapter 3 – a published manuscript that introduces DNA knot factory on a nanofluidic device:

A Nanofluidic Knot Factory based on Compression of Single DNA in Nanochannels

Susan Amin, Ahmed Khorshid, Lili Zeng, Philip Zimny and Walter Reisner; Nature Communications 9, no. 1506 (2018): 1506.

❄ Chapter 4 – the supplementary information for the DNA knot factory manuscript:

A Nanofluidic Knot Factory based on Compression of Single DNA in Nanochannels: Supplementary Material

❄ Chapter 5 – a manuscript on DNA knotting dynamics – ready for submission:

Complex Knot Dynamics on Hydrodynamically Compressed Nanochannel-Extended DNA

Susan Amin, Ahmed Khorshid and Walter Reisner.

❄ Chapter 6 – a manuscript on flow-assisted mega-base-pair DNA disentanglement and linearization – ready for submission:

Flow-Assisted Single-Cell Derived Genomic DNA Purification, Disentanglement and Linearization

Susan Amin, Peter Yao, Ilja Czolkos, Sara Mahshid, Robert Sladek and Walter Reisner.

❄ Chapter 7 – Conclusion

This dissertation presents results of two projects:

1. The DNA knot factory project, where a highly-efficient technique is introduced for knot formation on DNA molecules as well as a knotting free energy formalism that explains the experimental findings. This project led to two manuscripts (one published in Nature Communications and another one ready for submission). I am the first author in both manuscripts and the project is supervised and guided by Prof. Walter Reisner. In the first manuscript, Ahmed Khorshid, Philip Zimny and I contributed to the design, fabrication and characterization of the devices; I conducted the experiments and performed data analysis; Walter Reisner, Lili Zeng and I contributed in the coding and Walter Reisner and I wrote the paper. In the second manuscript, Ahmed Khorshid and I fabricated and characterized the devices; I performed the experiments and analyzed the data and Walter Reisner and I wrote the paper.
2. An applied project where fluid flow in microchannels is utilized for single-cell-derived mega-base-pair DNA purification, disentanglement and linearization. The manuscript for this research project is ready for submission. I am the first author and Prof. Walter Reisner and Prof. Robert Sladek have supervised and guided the project. I designed and fabricated the devices and performed the experiments; Sara Mahshid and Ilja Czolkos helped me with the device fabrication; Peter Yao contributed in performing the experiments and I wrote the paper.

ABSTRACT

Inspired by the knot abundance on biopolymers such as extracted DNA from phage capsids, DNA in nanofluidic systems and proteins, we introduce a knot factory on-chip based on hydrodynamic compression of single T4 DNA molecules against barriers in nanochannels. The proposed method provides an efficient way to form knots on DNA under compression and detects the independently evolving knots during the relaxation process following the compression. We observe that knotting probability increases with pressure as well as with the waiting time in the compressed state. Moreover, we develop a free energy formalism derived from scaling arguments, which describes our experimental findings. Our proposed model suggests that while Poissonian statistics is able to explain knot formation at low pressures, it breaks down at high pressures, which can be explained by the important role of knot-knot interactions on highly compressed molecules. Finally, we observe step-wise complex-knot unraveling on the relaxing molecules and find that in multiple-knot events, knots interact with each other via hard-core repulsion. Moreover, we develop a platform for flow-assisted DNA disentanglement. In our lab-on-chip device, long mega base-pair genomic DNA is extracted from single cell, purified and disentangled by hydrodynamic flow. The chip design allows for single-cell trapping in a microchannel, cell lysis and genomic DNA extraction and purification. The long entangled/knotted DNA molecules are subsequently trapped at the post-array entrance. Using hydrodynamic flow, the DNA molecules are disentangled and subsequently stretched in the post arrays. Our proposed chip design minimizes DNA fragmentation and is a high-throughput method, which can be further utilized for DNA sequencing.

ABRÉGÉ

Inspirés par l'abondance de nœuds présents dans les biopolymères tels que l'ADN extraite des capsides des phages, l'ADN de systèmes nano-fluidiques et l'ADN de protéines, nous introduisons un dispositif de fabrication de nœuds basée sur la compression hydrodynamique de molécules individuelles d'ADN T4 dans des nano-fentes. La méthode proposée fournit une manière efficace de former des nœuds dans l'ADN compressée ainsi que de détecter l'évolution indépendante des nœuds durant le processus de relaxation suivant la compression. De plus, sur la base d'un argument de mise à l'échelle, nous développons un formalisme d'ènergie libre qui décrit nos observations expérimentales. Le modèle proposé suggère que, même si les statistiques Poissoniennes peuvent expliquer la formation de nœuds sous faible pression, elles ne peuvent expliquer leur formation sous haute pression; leur formation peut être expliquée par le rôle important joué par les interactions entre les nœuds sur les molécules hautement compressées. D'autre part, nous observons un dénouement séquentiel complexe des nœuds sur les molécules au repos. De plus, nous notons que lors d'événement concernant plusieurs nœuds, les nœuds interagissent entre eux via une interaction de type interatomique. Finalement, nous introduisons une plateforme de démêlage des nœuds de l'ADN par courant hydrodynamique. Dans notre dispositif laboratoire sur puce, des séquences d'ADN longues de plusieurs mégabases sont extraites d'une cellule, purifiées et démêlèes par le courant hydrodynamique. La géomtrie de notre dispositif laboratoire sur puce permet l'emprisonnement de cellules individuelles dans un micro-tube, puis l'extraction et la purification de leur lyse ainsi que de leur ADN génomique. Les longues molécules d'ADN emmêlées dans le micro-tube sont ensuite emprisonnées à l'entrée de

nano-fentes. Les molécules d'ADN sont démêlées et étirées dans les nano-fentes à l'aide du courant hydrodynamique. La conception du dispositif laboratoire sur puce suggérée minimise la fragmentation de l'ADN et maintient une haute capacité de production. Notre dispositif laboratoire sur puce est donc un bon outil pour le séquenage d'ADN.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1
Introduction

1.1 Introduction

Employing knots as a motif in stamps and seals dates back to 3500 BC [1,2]. Knots have also been used as decorative objects in arts as well as in early Christian monuments and manuscripts, such as the Book of Kells and the 8th-century St. Teilo Gospels [3]. Beside their ornamental usage, knots are employed in different applications ranging from climbing and sailing to tying shoelaces and surgical sutures. Despite their usefulness, however, knots can be bothersome if they cause undesirable conditions, such as knots formed in long hair, electrical cords and catheters [4]. Knots also naturally exist in microscopic scale on biopolymers, in particular DNA and proteins [4–6]. Microscopic knots vary from simplest knot types detected in proteins [7] to very complex knots on the DNA extracted from tailless P4 phage capsids [8]. How these microscopic knots influence the physics of polymers in biological systems is a very important question in polymer physics. Many theoretical and experimental studies have been devoted to understanding knotting mechanisms, knot complexities and how knots affect the global polymer physics [4, 9–13]. Studies have shown that knots decrease the relaxation time of the stretched molecules in elongational fields [14] and reduce the tensile strength of chains [15, 16]. Finally, different knotting methods have been proposed [9, 15, 17, 18] in order to provide a platform for studying knots.

This dissertation presents the results of the two projects that I performed during my PhD. Chapter 1 provides a general introduction to the subject and introduces motivation for the performed projects. Chapter 2 explains the background and the literature of the projects. In chapter 3, a nanofluidic knot factory on chip is introduced. The theory and calculations are given in chapter 4. Chapter 5 presents some observations of knot dynamics on DNA molecules as well as knot complexities. In chapter 6, a lab-on-chip platform for single cell trapping, lysis, DNA extraction and purification and flow-assisted disentanglement of DNA molecules in post corridors are introduced. Finally, chapter 7 summarizes the obtained results followed by conclusions and perspectives. In the current chapter, knots as topological entities are defined; knot characterization methods and different knot types are explained and the instances and significance of knots formed in macroscopic and microscopic scales are discussed. Moreover, different knotting mechanisms on confined polymers are introduced and compared with each other. Finally, the concept of microfluidics and nanofluidics is explained and examples of the improvements in the field are provided.

1.2 Knot Definition

Knots are mathematically well defined on closed nonintersecting curves. A knot is a self-entanglement on a closed string that can *only* be untied if the string is cut [19]. On the contrary, a knot-free closed loop is called an unknot or a trivial knot. Figure 1–1 illustrates schematics of an unknot (Figure 1–1 (a,b)) and a knot (Figure 1–1 (c)) on closed strings. We can identify knot types by projecting a three-dimensional knot onto a two-dimensional plane and analyzing the crossings (nodes). In this respect, if an entangled state of a closed loop can be rearranged such that the resulting 2D projection has zero crossings, it is called an unknot or a trivial knot [4].

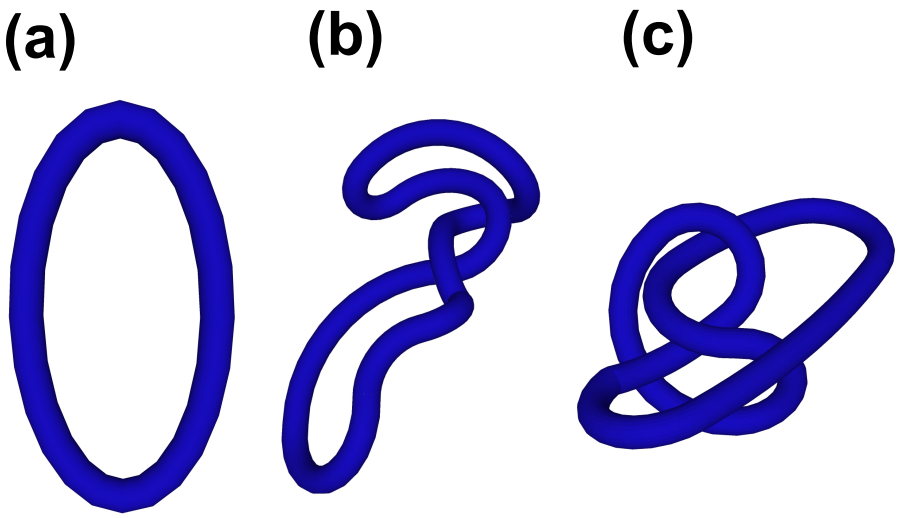

Figure 1–1: Diagrams of closed curves representing examples of an unknot (a,b) and a knot (c). According to the definition of knots on closed curves, the self-entanglement in (b) can be rearranged to (a), thus is considered an unknot. Panel (c) gives an example of a trefoil knot.

Topological entanglements on proteins [4,5] and linear DNA molecules [13,20] and their significant effect on the physics of linear polymers [9, 16] show the importance of systematic knot identification on open chains. Knots on open chains, however, are not rigorously defined. Taylor [21] introduces an intuitive way of knot detection on an open curve: an entanglement on an open chain is called a knot if it cannot be opened via holding and pulling the two ends of the chain. Figure 1–2 shows two simplest knot types on closed and open curves. Mathematically, in order to define knots on open chains, we imagine that the two ends are closed such that they do not pass through loops on the same chain prior to the closure (Figure 1–2).

1.3 Knot Formation on Macroscopic Chains

Knots on ropes, strings and cords have been used throughout history for human benefit and prosperity including human safety, building houses and boats, fishing and textile production [2]. Moreover, spontaneous knots are observed in long strings such

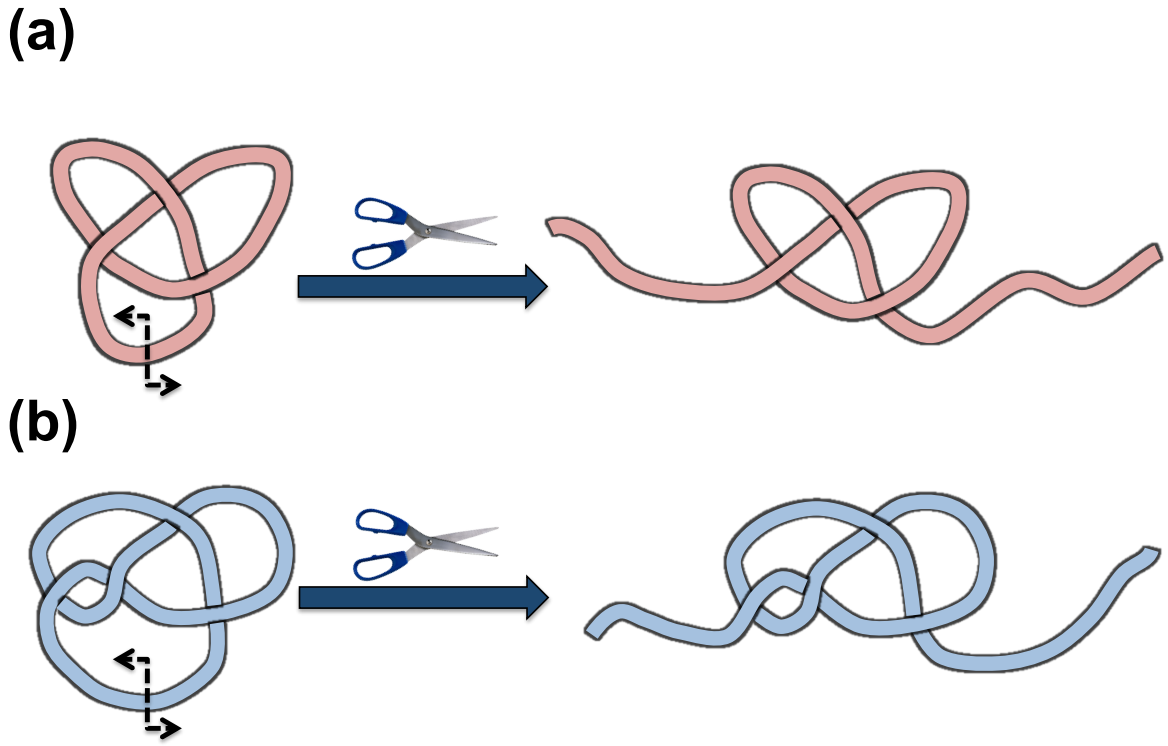

Figure 1–2: Schematic of trefoil (a) and figure eight (b) on closed (*left*) and open (*right*) chains. The arrows depict the process of cutting a closed knotted curve to yield equivalent knot type on the corresponding open chain. The closure of a knotted open curve would be in the reverse direction.

as in hair and electrical cords [4]. In biomedical context, knots appear spontaneously in surgically implanted catheters [22] as well as in umbilical cords [23]. These observations have led to studies on the origination of knots and the factors influencing knottedness as well as how knots affect the physical and mechanical properties of chains [24–27]. Some studies [18, 24, 26] show that there are similarities and differences between knotting in macroscopic chains and knot formation in microscopic scales. Some argue that the relative simplicity in the implementation and analysis of knot-related experiments in macroscopic scales can lead to a better understanding of the details in microscopic scales [28]. Some of these studies are introduced below.

Belmonte *et al.* [29] suspend a stainless steel chain of N-coupled pendula from an oscillating support, which imposes sinusoidal movement of the chain at different ranges of frequency. The results show that while the chain exhibits rigid rod-like and planar pendular motion respectively at very low and intermediate frequencies, it displays a vigorous motion at high frequency, resulting in a chaotic movement that leads to the self-knotting of the chain. Hickford *et al.* [24] perform a similar experiment on ball chains of different lengths sitting on a horizontal plate subjected to a sinusoidal movement . They find that knotting probability increases monotonically with chain length N for $N_{\mathrm{min}} < N < N_{\mathrm{threshold}}$. The knotting probability is zero for $N < N_{\mathrm{min}}$ and plateaus for $N > N_{\mathrm{threshold}}$. The authors argue that while one cannot eliminate the possibility of higher knotting probability obtained for much longer chains, their results are in contrast to the predicted equilibrium behavior of polymers in self-avoiding random walks, where the knotting probability approaches unity for sufficiently long polymers [30]. Raymer *et al.* [26] obtain similar qualitative results for a chain confined in a rotating cubic box. They observe that the knotting probability in their system increases with the length of the chains and levels off for long chains. Reference [4] presents the simulation results on knotting probability for agitated strings in a box as a function of chain length . In contrast to the experimental results obtained in [26], the simulation results show that the knotting probability is close to one for sufficiently long chains, which is consistent with the theoretical predictions [30]. Raymer *et al.* argue that the empirical probability could approach unity if the chains are made much longer and flexible.

Ben-Naim *et al.* explore the opening time of knotted granular chains subjected to vibration [28]. In their experiments, a simple (trefoil) tight knot is made at the center of a ball chain and the chain is subsequently positioned on a vibrating plate.

The ball chain is composed of N hollow spheres connected together via thin rods. The chain can be viewed as a bead-spring system, where the spheres interact via hard-core repulsion. They show that the unknotting time of a chain of size N scales with N^2, which implies the diffusive behavior of the knot on the chain. Ben-Naim *et al.* argue that their system can be used in understanding the dynamics of topological constraints in polymers, as the vibrating plate provides the motion induced by thermal energy in microscopic systems.

Finally, knots affect the strength of strings significantly [4]. It is well known in rock climbing and fishing that knots weaken ropes and fishing lines [31, 32]. Experiments on fishing lines and cooked spaghetti noodles confirm that knots decrease the tensile strength of the strings down to 50% [27] and suggest that the breakage always occur at the entrance of the knot region where the curvature is at its highest. These results are in good agreement with the observed decreased strength in the knotted actin filaments and their breakage within the knotted region [15].

1.4 Knot Formation on Microscopic Chains

Frisch and Wasserman [33] and Delbruck [34] were the first ones who conjectured that the knotting probability for a polymer at equilibrium with n monomers approaches unity as n goes to infinity. A theoretical study on random self-avoiding walks on a three-dimensional cubic lattice later confirmed that the probability of knot existence on a sufficiently long chain is almost one [30]. Knots on biomolecules were first characterized by electron microscopy on single-stranded DNA in 1976 [35]. Knots have been observed on partially replicated plasmids [36, 37] as well as in the cyclized DNA extracted from P2 and P4 phage capsids [8, 38] with the knotting probability $> 95\%$. Knots have been

also identified in proteins [6], although knot identification in proteins is possible only after closing and smoothing the backbone using numerical methods [21].

Observation of knots in biophysical systems has encouraged scientists to develop knotting techniques in order to provide platforms for studying knots and knotting effects on biophysical systems. Knots have been produced via chemical ligation of DNA with stems and loop regions [39] as well as self assembly of nucleic acids in single-stranded DNA and RNA [17]. Knots have also been constructed mechanically via manipulation of polystyrene beads attached to the ends of actin filaments [15] and DNA molecules [9] using optical tweezers. The relatively slow knotting process (on the order of tens of seconds) in the early application of optical trapping technique, where the microscope stage and the optical tweezers were moved manually, required the use of non-Newtonian viscous fluid in order to overcome the rapid relaxation of the polymer [15], which in turn led to the suppression of polymer dynamics [9]. Later, the electronically programmed movement of stage and optical tweezers reduced the knotting time and made the use of linearly viscous solution possible [9].

Some knotting techniques have been inspired by the effect of confinement on knot formation. Numerical studies revealed that knotting probability in ring polymers increases as the spherical confinement increases [40]. Observation of highly knotted DNA molecules extracted from tailless phage capsids [8] with the knotting probability of over 0.95 in contrast with 0.03 for randomly cyclized P4 phage DNA in free solutions [41], suggested the possible role of confinement in knot formation. Later, Arsuaga *et al.* [10] showed that more than half of the knots form when the phage DNA is still inside capsids and concluded that DNA compaction in restricted volume indeed promotes knotting probability. Consequently, many theoretical [11,12,42] studies and some experimental work [13,18] focused on the effect of restricted volumes on

knot formation in polymers. Tang *et al.* [18] compressed single T4 DNA molecules in microchannels to sphere globules via exposure to AC electric field. Upon stopping the electric field and DNA expansion, knots were detected on the relaxing molecules. Compared to the collapsed molecules in AC electric field, DNA molecules in capsids are in a much tighter environment and experience a much higher level of compaction. Thus, Tang *et al.* concluded that knot compactification alone cannot induce knots on moderately-confined molecules, but another effect is required; electric field-induced rotational motion (*tumbling*) of polymer segments [18]. The *Tumbling* effect was first introduced in [26], where macroscopic strings were agitated in a rotating box for different rotating times. The results revealed that at low agitation times, knotting via tumbling of the strings is kinetically limited, while at long enough agitation times, knotting probability plateaus at the equilibrium values for a semiflexible self-avoiding chain in a box. Tang *et al.* [18] argued that electrohydrodynamic instability in their system induces the tumbling required for knot formation.

1.5 Knot Formation in Nanochannels

Knots form on DNA molecules in nanochannels and inhibit the scanning of base sequences in DNA mapping techniques by concealing the knotted contour [20]. Metzler *et al.* [13] observed that DNA collision with defects on nanochannel walls induces knots on the polymer contour. The knots were visualized as strongly localized fluorescent spots that diffuse on the polymer contour and unravel at one of the molecule ends with the diffusion time scaling quadratically with the DNA contour length. Knot formation, however, is not systematically controlled and the knotting probability is not quantified. In chapter 3, we introduce a simple hydrodynamic knot factory on chip, where all knotting parameters are well controlled. The proposed knotting approach uses an ultra low Reynolds number flow to compress T4 DNA molecules in nanochannels against

slit barriers followed by pressure release and relaxation of DNA molecules. Knots are present on the relaxing DNA molecules, visualized as localized bright fluorescent features that diffuse on the polymer contour. We argue that in the compressed state the Kuhn segments of DNA molecule have zero segmental current, thus we can assume equilibrium condition for the DNA. We introduce a free energy formalism for knotted states, which predicts knotting probability as a function of compression and waiting time at the compressed state. Moreover, we investigate the knot formation kinetics and measure the knot spatial distribution on the chain. Finally, formation of multiple knots on DNA molecules enables us to study knot-knot interactions. We show that knotting probability for composite knot states at low pressures can be described by Poisson statistics, while at high compression Poisson model breaks down due to knot-knot interactions. We argue that our knotting set-up has the capability of forming knots with high probability in a highly parallel system, where several molecules can be knotted in an array of nanochannels. Compared to the approach of Tang *et al.*, where knots form in a highly out-of-equilibrium environment as a result of electrohydrodynamic instabilities in DNA globules leading to complications in analyzing the dynamics of the system, in our proposed method, the zero segmental current of the compressed molecule in an ultra low Reynolds number flow leads to an inhomogeneous equilibrium state, where the knot formation can be modeled via a simple free energy formalism.

Knots are divided into two types: *prime knots* and *composite knots*. Prime knots are the building blocks of composite knots and cannot be decomposed to simpler knots, while composite knots form as the result of concatenation of two or more prime knots [43]. Prime knots are further categorized to simple and complex knots depending on the essential number of crossings and their topology. Knot complexity plays a significant role in the static and dynamic properties of knots [44]. The radius of gyration

decreases weakly with increasing knot complexity [44] and knots with the same complexity have the same mobility [45]. Moreover, increasing knot complexity results in a larger friction factor, thus smaller diffusion constant in the process of self-reptation [9]. If we visualize the path of contour length inside the knot region, we can confidently determine knot complexity. Among different proposed methods of microscopic knot construction, mechanical tying of knots via optical tweezers [9, 15] has the advantage that knot types are exactly known, as the contour path can be tracked in the knotting process. In general, however, due to the relatively low resolving power of imaging techniques, determination of knot complexity has remained a challenge in the other microfluidic and nanofluidic knotting approaches since the topological information about microscopic knots is not accessible. In chapter 5, we introduce the method of step-wise knot unraveling via which we can categorize knots by eliminating the possibility of a knot belonging in certain knot groups.

There has been occasional reports of multiple (composite) knots on DNA [46], however knot-knot interactions in such structures have not been studied yet. In our knot factory set-up, we observed several cases of multiple-knot formation on DNA. Our free energy formalism proposed in chapters 3 and 4, suggests the existence of a linear potential leading to single-file ordering of knots on nanochannel-confined DNA molecules, which restricts knot movement on the polymer such that knots cannot pass each other in the contour. In chapter 5, we present our experimental observations of multiple knotting events. We show that hard-core repulsion and single-file ordering exist between knots. Finally, we illustrate instances of entangled knots visualized as a single knot. The knots split up after a few seconds and diffuse on the chain due to the thermal fluctuations of the molecule.

1.6 Microfluidics and Nanofluidics

Microfluidics and nanofluidics refers to a relatively new scientific field that studies fluid flow in micro and nano scales. Microfluidic devices have been designed to manipulate and transport fluid as well as chemical and/or biological entities through micron-sized channels. Microfluidic-based devices provide a convenient platform for performing a sequence of operations on-chip, thus called *lab-on-chip* systems. They consist of microchannels for transportation of reagents and other components such as valves [47, 48], mixers [49–51] and pumps [52]. Miniaturization of fluidic devices has several advantages such as portability and compactness of the systems, parallelization of several operations on one device, provision of single-molecule resolution for studying proteins and nucleic acids, elimination of the need for large volumes of reagents, lower costs and faster processing times [53].

Microfluidic-based technologies have led to significant developments in several domains such as in analytical chemistry, biology, optics and information technology. The earliest applications of microfluidic systems were in molecular analysis due to their high efficiency in terms of the amount of reagents used, analysis time, cost and sensitivity of the devices and techniques. Other major pioneer applications were in the domains of biodefence, molecular biology and microelectronics [54]. Molecular biology in particular, gained popularity in the 1980s due to an escalation in the recognition of genomics. The demand for high-throughput DNA sequencing with higher resolution drove the field of microfluidics toward addressing the limitations of the then-existing sequencing methods.

1.6.1 Single-Cell Studies

Genetic heterogeneity is common in tumor cells [55–57] and is believed to be an important cause of resistance to cancer treatments. Thus, many studies have emphasized

the importance of analyzing cells individually [58, 59] and several works have focused on single-cell derived genomic DNA sequencing [60–64]. Microfluidic-based devices are very promising for single-cell studies [65, 66].

Single-cell manipulation requires single-cell trapping as the first step, especially for long run times. There are different cell-trapping methods in the literature based on which various designs have been developed. Mechanical traps, such as microdams or microwells are one of the extensively used features for single-cell isolation and trapping [67–69]. Kobel *et al.* [70] proposed a trapping method based on differential fluid resistance for single-cell culture. Later, a similar study utilized the same concept in a highly parallel system for homotypic and heterotypic single-cell co-culture [71]; the method was later optimized to avoid the need for very long serpentine channels [72]. Dielectrophoretic traps [73, 74], optical traps [75–77] and microdroplet traps [78, 79] are some other examples that are extensively-used in trapping techniques [80].

Cell lysis is the next step in single-cell analysis. There exist three distinct types of cell lysis in the literature: (1) mechanical, (2) electrical and (3) chemical cell lysis. In mechanical cell lysis methods, a cell is subjected to a mechanical shear force via sharp objects such as nanoblades [81] or nanowires [82]. In electrical lysis, small pores are created in cell membranes via application of electric fields [83–85]. Chemical cell lysis is performed by injecting cell lysis solution following single-cell trapping [86, 87].

Single-cell analysis on microfluidic-based devices is done with different objectives, such as improvement in the accuracy of protein analysis, metabolites analysis or gene analysis [80]. As mentioned earlier in this section, one of the most significant applications of on-chip single-cell analysis is in genome sequencing. Several studies have been devoted to single-cell whole-genome amplification (WGA), where single-cell derived

and purified DNA molecules are amplified and sequenced via different sequencing techniques [64, 88, 89]. Single-molecule mapping techniques have also emerged in the past few decades [90–93]. The main feature of these approaches is that they do not require DNA amplification, and thus do not introduce bias to the DNA readings [94]. Integration of on-chip single-cell genomic DNA extraction techniques and single-molecule mapping is still challenging.

1.6.2 DNA Sequencing and DNA Linearization

The earliest on-chip DNA sequencing technique dates back to 1993 when Fodor *et al.* [95] invented a semiconductor-based technology with DNA microarrays for analysing millions of proteins in a cell or tissue. Later Woolley and Mathies [96] introduced a microfabricated capillary electrophoresis device for performing DNA separation and sequencing. Several studies enhanced the accuracy of the technique and reduced the sequencing time [97–100]. Fully-integrated microfluidic devices, where DNA molecules are extracted, amplified and sequenced were later developed [101–103]. The main drawback of those techniques is the need for DNA amplification so that enough DNA molecules are provided to perform the sequencing process. DNA amplification introduces bias to the DNA readings, which reduces the genome sequencing coverage [94]. Thus, some later studies moved toward single-molecule sequencing, where no amplification is required [93, 104–107].

Among the state-of-the-art techniques, there exist some that perform mapping on nanochannel-confined DNA. In the approach proposed by Lam *et al.* [92], the authors fluorescently label specific sequence motifs in single DNA molecules and perform fluorescence imaging of nanochannel-confined DNA molecules. The physical distances between the sequence motifs provide maps for hundreds of DNA-molecule fragments ~ 50 kb in length. In a coarse-grained mapping technique proposed by Reisner *et al.* [91],

DNA molecules are fluorescently labeled and in a solution containing formamide they are driven into nanochannels. The authors then subject the molecules to heat treatment via which the molecules are denatured and provide sequence-dependent barcodes. Both proposed methods require linearization of DNA molecules in nanochannels. On the other hand, some theoretical and experimental studies show that increasing the concentration of long DNA molecules leads to DNA entanglement [108,109], which hinders DNA linearization and can result in knotted structures [30]. Moreover, DNA elongation and confinement require a large free energy in order to reduce entropy and overcome the energy barrier between the less confining area and the highly-confining region [110]. Lam *et al.* [92] used electric field and pre-stretched DNA-molecule fragments, ~ 50 kb in size, between post arrays before the nanochannels entrance. However, one of the necessary requirements for the arrangement of the pillars is provision of enough level of confinement for DNA molecules, yet enough sparsity between the pillars so that the molecules can uncoil [92]. The latter feature limits the applicability of pillars for very long DNA molecules (on the order of Mbp) as long molecules tend to entangle between the posts [111, 112].

In chapter 6, we introduce a microfluidic platform for extraction, purification, disentanglement and linearization of long Mbp DNA molecules extracted from single cells. In particular, our system provides a high-throughput flow-assisted genomic DNA disentanglement and linearization in post corridors. To ensure an efficient purification of DNA, the protein molecules in our system are visualized and subsequently digested and removed. Finally, using a dense set of post arrays we anchor long DNA molecules to prevent their escape into the deep microchannels.

CHAPTER 2
Background

2.1 Knots

Knots were first approached mathematically in the 19th century by Gauss [113, 114]. Knots are defined as topological entanglements on closed loops that cannot be untied without opening the loop [19]. Knots on open chains, on the other hand, are not mathematically well defined. An intuitive approach to discriminating knots and unknots on open strings is to hold the two ends of the chain and pull them; knots maintain their entangled state while unknots are disentangled [21]. A more rigorous method of characterizing knots on open chains is imaginary transformation of the open chain to a corresponding closed loop. The closed chain can subsequently be studied for characterization of potential knots and their complexities.

2.1.1 Knot Types

The ability of distinguishing different knot types is crucial in studying effects of knot complexity on the physics of different systems including biophysical processes. Knot theory provides the essential tools for characterizing knots [19]. Knots are categorized according to their complexity, determined by the essential number of crossings in a knot type as a knot-invariant. In order to analyze a knot, the 3D knot is projected onto a 2D plane. The number of the chain intersections (points or crossings) determines the complexity of the knot [4]. By convention, a knot type is shown as C_s, where C denotes the essential number of crossings and s represents the s^{th} knot topology in the sequence of knots with C number of crossings. The size of the set of

knots with the same number of crossings increases rapidly with C. There are two knot types with five number of crossings, three knot types with six number of crossings, seven different knot types with seven number of crossings, but 21 different knot types in the category $C = 8$. Figure 2–1 (a-d) illustrates schematic of simplest knot types from $C = 3$ to $C = 5$. Trefoil (3_1) knot and figure eight (4_1) knot are the simplest *torus* and *twist* knots, respectively. Torus knots can be thought of as the knots formed by wrapping the chain around a torus, while twist knots are constructed when the two ends of a twisted closed loop are linked together. The knots 5_1 (Figure 2–1(c)) and 5_2 (Figure 2–1(d)) are other examples of torus and twist knots, respectively. Every

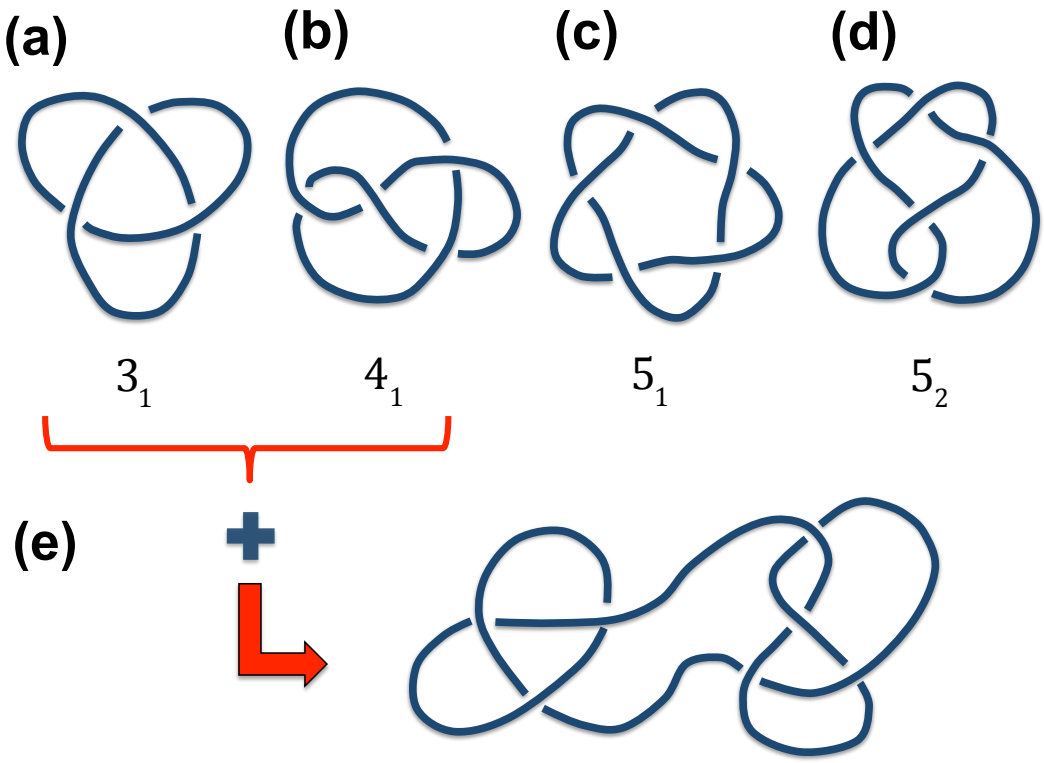

Figure 2–1: 2D projection of the simplest knot types. There is only one knot type in the category $C = 3$, called trefoil knot (a), and in the category $C = 4$, called figure eight knot (b). The set of knots with $C = 5$ essential crossings consists of two knot types (c,d). (e) A composite knot presented by the knot sum of a trefoil and a figure eight knot.

knot can be decomposed into the knot sum of *prime* knots [115, 116]. Prime knots are the knot types that upon decomposition into knot sums yield a factor of *unknot* or trivial knot [117]. Rolfsen table [118] presents the set of *prime* knot diagrams up to $C = 10$. A knot that is not a prime knot is called a *composite* knot. Composite knots are composed of two or more non-trivial prime knots (Figure 2–1(e)).

2.1.2 Unknotting Number

Unknotting number is the minimum number of times the end of a knotted string must be passed through the knot to unravel it [19]. All knots with unknotting number 1 are prime knots [119] and the unknotting number of composite knots is at least 2. In general, the unknotting number of a knot with number of crossings C is less than $C/2$ [120]. The unknotting numbers for the first four knot types in figure 2–1 are 1 (Fig. 2–1(a)), 1 (Fig. 2–1(b)), 2 (Fig. 2–1(c)) and 1 (Fig. 2–1(d)), respectively. Finally, all twist knots have unknotting number 1, since they can be unknotted by unraveling the twisted link.

2.1.3 Knot Formation Mechanisms

Knots form either by breaking and reconnecting of chain segments, or via passing loose ends of a string through loops on the same string [4]. In biological systems, knotting via the first method is triggered by enzymes such as topoisomerases and recombinases. There exist two types of topoisomerases: type I and type II. In double-stranded DNA, type I topoisomerases break a single strand temporarily and allow it to pass through the complementary strand [121], while type II topoisomerases break both strands at certain points on the molecule and let the segments pass through loops on the string [122, 123]. The broken ends are subsequently chemically rejoined. Similarly, recombinases induce knots on DNA molecules via breaking and rejoining segments.

However, unlike topoisomerases, recombinases insert, excise or invert a segment instead of simply cutting a site on DNA molecules.

The second mechanism relies on passing the loose ends of a chain through loops on the same chain. Random cyclization of linear molecules such as P2 and P4 bacteriophage DNA is an example of knot formation via the second mechanism [8, 38]. In this dissertation, we focus on the conditions that favor knot formation via second mechanism. In particular, we investigate the effect of confinement on knot formation in double-stranded T4 DNA.

2.2 Physics of Polymers

Polymers are molecules made of covalently bonded sub-units, called monomers. The most well-known biopolymers are nucleic acids (DNA and RNA) and proteins. Polymer solutions are constantly subject to thermal fluctuations, which subsequently change the configuration of polymers in space. Temperature, different sources of extrinsic forces, chemistry of solution and concentration of polymer segments are some other important factors that strongly affect the physics of polymers. Polymers are long molecules, thus in order to model and explain their physical behavior, polymer physics employs statistical physics to overcome the complications of studying polymers deterministically. This section summarizes the main concepts and models in polymer physics as well as the physics of knots on DNA molecules. In addition, discussion of the DNA knotting free energy formalism is presented.

2.2.1 Polymer Models

Polymer models are categorized into two main groups: *ideal chains* and *real chains*. The ideal chain is a simplified polymer model that ignores interactions between the chain segments [108]. The chain structural units (*bond vectors* in figure 2–2(a)) are

consequently formalized as a simple random walk: the orientations in any two segments are independent of each other (*i.e.* the monomers can take any orientation in space). As a result, segments can intersect each other and occupy the same point in space.

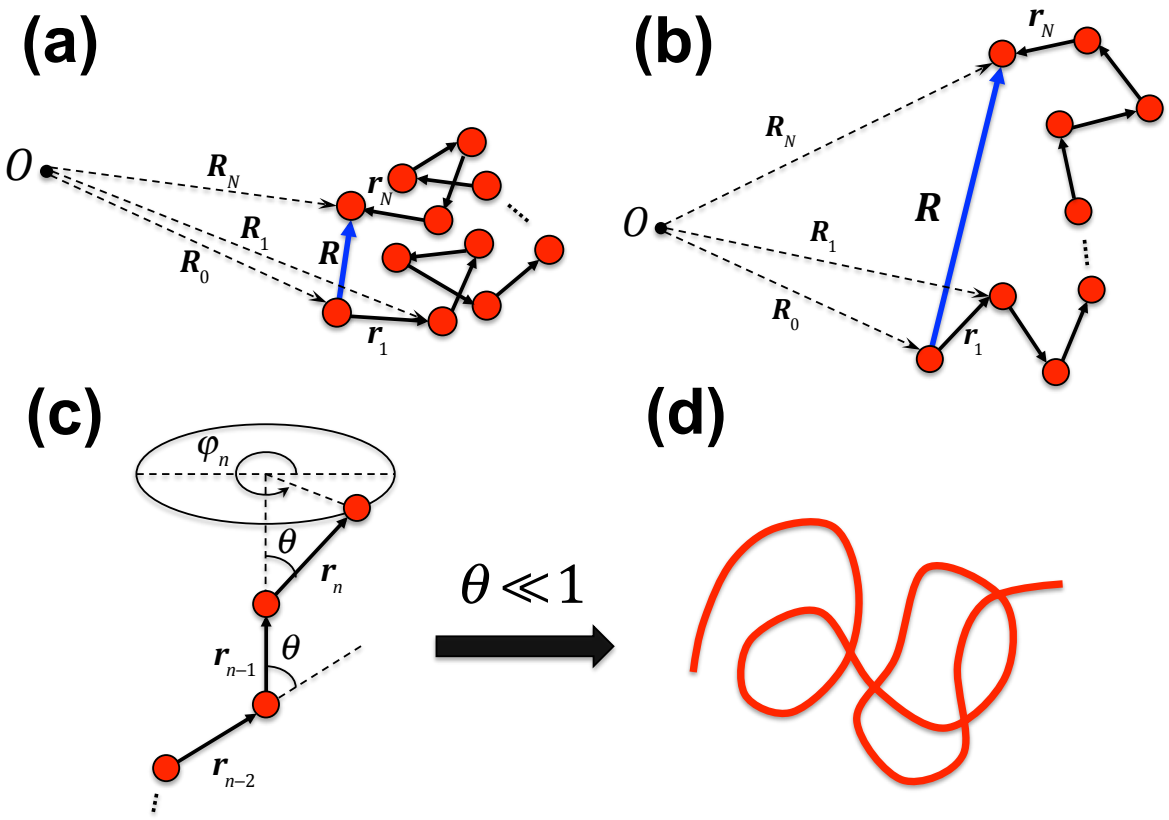

Figure 2–2: Discussion of principle polymer models. (a) An ideal chain: the monomers can intersect each other. (b) A real chain: there exists no pair of monomers that occupy the same volume in space. The chains in (a) and (b) are made of N consecutively connected monomers, depicted by bond vectors $\boldsymbol{r}_i$, where $i = 1, 2, \ldots N$. The position vectors are denoted by $\boldsymbol{R}_j$, where $j = 0, 1, \ldots N$. The end-to-end vector is shown by $\boldsymbol{R}$. (c) Representation of a freely-rotating chain. In the limit of small θ, a freely-rotating chain can be approximated as a continuous ideal chain, also called semi-flexible or worm-like chain (d).

Chains are represented as a sequence of N bond vectors $\{\boldsymbol{r}_i\}_{i=1:N}$ with fixed size $||\boldsymbol{r}_i||_2 = b_o$ and $N+1$ position vectors $\{\boldsymbol{R}_i\}_{i=0:N}$. The end-to-end vector of the chain is thus $\boldsymbol{R} = \sum_{i=1}^{N} \boldsymbol{r}_i$ (See figure 2–2(a,b)). The extension of the maximally extended chain

gives the chain *contour length*, $L = Nb_o$. One representation of ideal chains is *freely-jointed chains* (FJC's) (figure 2–2(a)), where the bond vectors $\{\boldsymbol{r}_i\}_{i=1:N}$ are oriented randomly in space. A FJC with fixed contour length L can take many configurations in space. At thermal equilibrium, since the interactions between the monomers are ignored, the total energy of the system is independent of the shape of the polymer. Therefore, the end-to-end vector being $\boldsymbol{R} = \tilde{\boldsymbol{R}}$ and $\boldsymbol{R} = -\tilde{\boldsymbol{R}}$ are equally probable. Consequently, the average end-to-end vector $\langle \boldsymbol{R} \rangle$ in a FJC equals zero ($\langle \boldsymbol{R} \rangle = 0$) [124]. Thus, we describe the size of a polymer via the root-mean-square value of the end-to-end vector $R \equiv \sqrt{\langle \boldsymbol{R}^2 \rangle}$, or alternatively the root-mean-square distance from the center of mass R_g,

$$R_g^2 \equiv \frac{1}{N} \left\langle \sum_{i=0}^{N} (\boldsymbol{R}_i - \boldsymbol{R}_{cm})^2 \right\rangle \tag{2.1}$$

where $\boldsymbol{R}_{cm} \equiv \frac{1}{N} \sum_{j=0}^{N} \boldsymbol{R}_j$ is the mean position of the monomers. The angular brackets denote the ensemble average. Radius of gyration R_g has two major advantages over root-mean-square value of the end-to-end vector R: it is experimentally more accessible and can be meaningfully defined for various types of chains including ring polymers and branched polymers. For a pure random walk with number of segments $N \gg 1$, the root-mean-square value of the end-to-end distance and the radius of gyration are respectively computed as follows [125],

$$R^{\mathrm{FJC}} = \sqrt{L b_o} = b_o \sqrt{N} \tag{2.2}$$

$$R_g^{\mathrm{FJC}} = \frac{b_o \sqrt{N}}{\sqrt{6}} = \frac{R^{\mathrm{FJC}}}{\sqrt{6}}. \tag{2.3}$$

Equation 2.3 shows that radius of gyration R_g^{FJC} changes linearly with R^{FJC} and scales with $N^{1/2}$.

Another type of ideal chains is *freely-rotating chains* (FRC's) (figure 2–2(c)). FRC model introduces correlation between bond vectors via constraining the angle between two consecutive bond vectors $\theta = \bar{\theta}$, while picking the azimuthal angle φ uniformly at random. We can show that the correlation between the i^{th} and j^{th} bond vectors is a decaying exponential computed as [125],

$$\langle \boldsymbol{r}_i \cdot \boldsymbol{r}_j \rangle = b_o^2 \left(\cos\theta\right)^{|j-i|} = b_o^2 e^{-|j-i|/s_p} \tag{2.4}$$

where θ is the angle between the i^{th} and j^{th} bond vectors and $s_p = -1/\left(\ln(\cos\theta)\right)$ is the correlation length defined as the number of consecutive bond vectors after which the chain forgets its orientation [126]. Finally, the root-mean-square end-to-end distance of a FRC is found as [125],

$$R^{\text{FRC}} = b_o \sqrt{N \left(\frac{1+\cos\bar{\theta}}{1-\cos\bar{\theta}}\right)}. \tag{2.5}$$

Equation 2.5 shows that $R \sim \sqrt{N}$ (where the symbol "$\sim$" reads as "scales with") obeys random walk statistics (equation 2.2). The step-size in a FRC, however, is $\sqrt{\left(1+\cos\theta\right)/\left(1-\cos\theta\right)}$ times larger than that in a FJC. In the limit of small bond angle $\theta \ll 1$, a FRC is approximated as a continuous chain, called *worm-like chain* (WLC) (Figure 2–2(d)), where the bond vectors are replaced with unit tangent vectors $\hat{t}(0,\tau)$ and $\hat{t}(s,\tau)$ at respective positions 0 and s on the polymer contour at time τ,

$$\left\langle \hat{t}(0,\tau) \cdot \hat{t}(s,\tau) \right\rangle = e^{-s/P} \tag{2.6}$$

where P is called the *persistence length* of the polymer and represents the correlation length of the chain tangent vectors. The mean-square end-to-end distance of a worm-like chain R^{WLC} is computed by the "Kratky-Porod" formula [109],

$$R^{\mathrm{WLC}} = \sqrt{2PL\left(1 + \frac{P}{L}\left[\exp\left(-\frac{L}{P}\right) - 1\right]\right)} \tag{2.7}$$

where L is the total polymer contour length. When $L \gg P$, $R^{\mathrm{WLC}} \simeq \sqrt{2PL} = 2P\sqrt{N}$ and we retrieve equation 2.2 with a larger step size $b \equiv 2P$, called the *Kuhn length*. On the other hand, when $L \ll P$, $R^{\mathrm{WLC}} \simeq L$, indicative of a rigid rod-like polymer. Worm-like chains are also called *semi-flexible chains* in contrast with *flexible chains* (*e.g.* FJC). A semi-flexible chain behaves like a rigid rod at short length scales and is governed by random-walk statistics at long length scales.

In the real chain model, interaction between monomers is modeled by introducing the notion of *excluded volume*. Excluded volume refers to the physical property of a real chain, that is no two segments can occupy the same volume in space. In other words, chains cannot intersect themselves, leading to a decrease in the conformational possibilities of the chain. Thus, real chains are *swollen* compared to ideal chains and obey *self-avoiding* random walks statistics. *Flory* treated the problem of conformation of real chains at equilibrium in dimension d by a balance between two competing forces: the excluded volume interactions, which lead to a repulsive energy and a restoring force prompting an elastic energy [108].

The repulsive energy is derived from the number of excluded volume interactions in the chain. Excluded volume scales as $\vartheta \sim b^2 w$, where b is the Kuhn length and w is the width of polymer segments. With a chain of unknown size R with mean density of monomers $\rho \sim \frac{N}{R^d}$ assumed, the number of excluded volume interactions for each

monomer is found as $\vartheta N/R^d$. The corresponding energy for each monomer is subsequently $k_B T\vartheta N/R^d$, where k_B is the Boltzmann constant and T is the temperature. Thus, the total repulsive energy for N monomers is calculated as [108],

$$F_{\text{repulsive}} = k_B T\vartheta \frac{N^2}{R^d}. \tag{2.8}$$

The elastic energy tends to shrink the coil. Flory approximated the elastic energy for a self-avoiding chain to be the same as the entropic energy for ideal chains,

$$F_{\text{elastic}} \sim k_B T \frac{R^2}{Nb^2}. \tag{2.9}$$

The total energy of a self-avoiding chain is then calculated by summing the two equations 2.8 and 2.9,

$$\begin{aligned} F_{\text{real}} &= F_{\text{repulsive}} + F_{\text{elastic}} \\ &\sim k_B T \left(\vartheta \frac{N^2}{R^d} + \frac{R^2}{Nb^2} \right) \end{aligned} \tag{2.10}$$

The chain size at equilibrium, R_F, is derived via minimizing the total energy of the real chain F_{real} with respect to R [108],

$$R_F^{d+2} \sim \vartheta b^2 N^3. \tag{2.11}$$

The root-mean-square end-to-end distance for a self-avoiding chain R_F is strongly dependent on the dimensionality of the space d and is found as,

$$R_F \sim N^{\nu} \tag{2.12}$$

where $\nu = 3/(2+d)$ [127] and is called the *Flory* exponent. By setting $\nu = 1/2$, we retrieve equation 2.2 for ideal chains.

2.2.2 Polymers under Confinement

Thus far, the conformation of ideal and real chains in the absence of external forces has been discussed. Experimentally, there are instances that polymers are confined in nanochannels. Some of these studies aim at accessing the genetic information in DNA molecules such as the denaturation mapping of nanochannel-confined single DNA molecules [91] and some others investigate the effect of confinement on the physical properties of polymers [128, 129] . While nanochannel confinement of ideal chains does not alter their equilibrium extension parallel to the nanochannel axis ($R_{\text{bulk}}^{\text{ideal}} = R_{||}^{\text{ideal}}$) [108], different confinement regimes substantially influence the conformation and extension of real chains. In the absence of confinement (in bulk) ($D \gg R_F$), a self-avoiding polymer has an equilibrium radius of gyration $R_g \sim R_F$, where R_F is calculated from equation 2.12 (See figure 2–3(a)). When the polymer is confined in a nanochannel, excluded volume effect becomes much more significant.

Blob Theory. De Gennes introduced the *blob* theory [108, 130] via which he explained the conformation and physics of confined self-avoiding polymers in the classic confinement regime. In the blob model, the chain is encapsulated by a series of imaginary blobs of size D (Figure 2–3(b)). The blobs interact via hard core repulsion and pack linearly along the nanochannel. Inside the blobs, the chain segments do not feel the confinement and obey the self-avoiding random walks statistics in bulk (Equations 2.8, 2.9 and 2.12). Thus, the polymer extension in each blob is found as,

$$R_b \sim D \simeq bN_b^{\nu} \tag{2.13}$$

where N_b is the number of Kuhn segments in a blob. The extension of polymer parallel to the channel axis is found as the size of each blob D multiplied by the number of blobs L/L_b, where $L_b = 2PN_b$ is the contour length contained in each blob. By solving

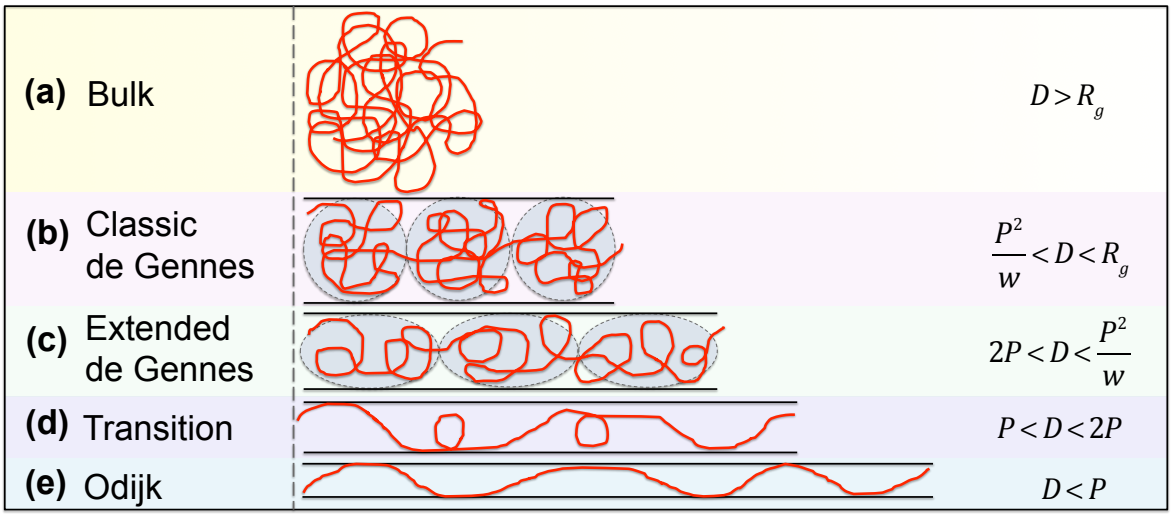

Figure 2–3: Illustration of polymer conformation in different nanochannel confinement regimes. Level of confinement increases from top to bottom, leading to a gradual increase in the polymer extension as the channel diameter decreases. (a) The polymer is unconfined and obeys bilk equations when $D \gg R_g$. (b) The polymer is confined in a channel with the size on the order $P^2/w \ll D \ll R_g$. In the classic de Gennes regime, the polymer is divided into imaginary symmetric blobs of size D. (c) By decreasing the channel size to the values smaller than P^2/w, the polymer enters the extended de Gennes regime with elongated blobs. (d) Transition from extended de Gennes to Odijk regime starts at $D = 2P$. Isolated backbends are apparent on the polymer. (e) At channel sizes $D < P$, hairpins disappear and the polymer deflects back and forth between the walls.

equation 2.13 for L_b,

$$L_b \sim \frac{D^{5/3}}{(Pw)^{1/3}} \tag{2.14}$$

and the total extension of the chain along the nanochannel axis R_{dG} is found as,

$$R_{\mathrm{dG}} \simeq L\frac{(wP)^{1/3}}{D^{2/3}}. \tag{2.15}$$

Equation 2.15 shows the strong dependence of polymer extension on the channel size ($R_{\mathrm{dG}} \sim D^{-2/3}$). The confinement free energy in classic de Gennes regime is found by multiplying k_BT by the number of blobs L/L_b, yielding $F_{\mathrm{dG}} \sim D^{-5/3}$.

As the channel diameter is decreased, the isometric blob model breaks at $D \simeq P^2/w$ [109]. In this new regime, called *extended de Gennes* regime, the anisometric blobs still interact via hard-core repulsion while the contour within each blob obeys ideal chain statistics (Figure 2–3(c)) [131]. The extent of each blob parallel to the channel axis is found as,

$$D_{||} = \sqrt{bL_b}. \tag{2.16}$$

In addition, in the extended de Gennes regime we have $F_{\text{repulsive}} \sim k_B T$ in equation 2.8 for Flory coils, indicating that the free energy due to excluded volume interactions is neither very small nor very large, so that the polymer inside the blob is on the cross-over between ideal and Flory coil [126]. This gives:

$$\vartheta \frac{N^2}{R^d} = \frac{(2PN_b)^2 w}{D^2 D_{||}} = \frac{L_b^2 w}{D^2 D_{||}} \sim 1. \tag{2.17}$$

Equations 2.16 and 2.17 yield,

$$L_b \sim \frac{b^{1/3} D^{4/3}}{w^{2/3}} \tag{2.18}$$

and the total extension of the polymer in extended de Gennes regime is,

$$R_{\text{ext-dG}} \sim D_{||} \frac{L}{L_b} = L \frac{(wP)^{1/3}}{D^{2/3}} \tag{2.19}$$

which gives the same result as in equation 2.15. The free energy in the extended de Gennes regime is still obtained by multiplying $k_B T$ by the number of elongated blobs L/L_b, which yields $F_{\text{ext-dG}} \sim D^{-4/3}$. Note that the free energy in the extended de Gennes regime is in fact dominated by the free energy arising from the transverse confinement of the chain, which looks like the confinement free energy of an ideal chain.

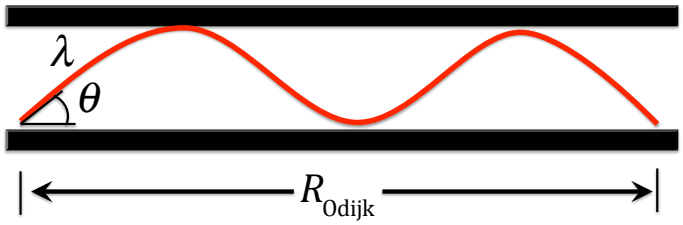

Figure 2–4: Illustration of contour deflection in Odijk regime.

When the channel size is decreased to $D \sim P$, the de Gennes blob theory fails to explain the chain conformation. Odijk found that at $P < D < 2P$, the sudden rise in the bending energy as a result of reduction in the channel size leads to rare isolated hairpins in the contour (Figure 2–3(d)) [131]. This new regime is called *transition regime* and will be discussed later in this section.

When the channel size is further decreased to $D < P$, hairpins disappear due to the large bending energy in the polymer segments (Figure 2–3(e)). This new regime is called the *Odijk* regime, where segments of length $\lambda = (PD^2)^{1/3}$ (Odijk deflection length) [132, 133] deflect successively between the channel walls (Figure 2–4). The polymer extension R_{Odijk} is calculated as $\lambda \cos\theta$ times the number of deflected segments L/λ, where θ is the angle that the deflected segments make with the wall and is assumed to be small. Thus, we have

$$\begin{aligned} R_{\text{Odijk}} &= \lambda(\cos\theta)\frac{L}{\lambda} = L\cos\theta \\ &\simeq L\left(1 - \frac{1}{2}\theta\right). \end{aligned} \tag{2.20}$$

For small θ, we can approximate it as $\theta \simeq \sin\theta = D/\lambda$. We can rewrite equation 2.20 as [129],

$$R_{\text{Odijk}} = L\left(1 - A\left(\frac{D}{P}\right)^{2/3}\right) \tag{2.21}$$

where A depends on the geometry of the nanochannels. The free energy of confinement in the Odijk regime $F_{\text{Odijk}} \sim k_B T \frac{L}{\lambda}$, that is

$$F_{\text{Odijk}} \sim k_B T \frac{L}{(PD^2)^{1/3}}. \tag{2.22}$$

Equation 2.22 shows the dependency on the channel diameter $F_{\text{Odijk}} \sim D^{-2/3}$.

Flory Theory. Flory free energy formalism for real chains introduced in equation 2.10 can be used with a small modification for nanochannel-confined polymers. The volume occupied by a confined polymer $V \sim RD^2$ replaces the denominator in the repulsive energy term (equation 2.8). The Flory free energy is then,

$$F_{\text{Flory}} \sim k_B T \left(\vartheta \frac{N^2}{RD^2} + \frac{R^2}{Nb^2} \right). \tag{2.23}$$

In the extended de Gennes regime, the polymer can be considered as a chain with L/L_b number of segments of size $b = D_{||}$ and $w = D$. Then, the excluded volume $\vartheta \sim D_{||} D^2$ and equation 2.23 can be rewritten as,

$$\begin{aligned} F_{\text{Flory}} &\sim k_B T \left(\frac{D^2 D_{||} N^2}{RD^2} + \frac{R^2}{N D_{||}^2} \right) \\ &= k_B T \left(D_{||} \frac{(L/L_b)^2}{R} + \frac{R^2}{(L/L_b)\, D_{||}^2} \right). \end{aligned} \tag{2.24}$$

If we replace L_b in equation 2.24 with 2.18, we retrieve the relation $F_{\text{ext-dG}} \sim D^{-4/3}$ found from the blob argument for the extended de Gennes regime. By setting $D_{||} = D$ in equation 2.24 and replacing L_b with equation 2.14, we recover the relation $F_{\text{dG}} \sim D^{-5/3}$ for the classic de Gennes regime. Minimization of the derived Flory free energies for the classic de Gennes and extended de Gennes regimes with respect to the polymer extension R yields the corresponding extension relations derived from blob arguments (Equations 2.15 and 2.19). In the transition regime between the extended de Gennes

and Odijk regimes, there exists a *global persistence length* G, which arises from the average distance between chain back-folds [134] and is found as,

$$G \sim r \exp\left(\frac{F_{\text{coil}}(r)}{k_B T}\right) \tag{2.25}$$

where r is the radius of the coiled segments and $F_{\text{coil}}(r)$ is the coil bending energy. The Flory free energy from equation 2.23 for confined polymer in the transition regime is thus,

$$F_{\text{Flory-Trans}} \sim k_B T \left(\vartheta \frac{N^2}{RD^2} + \frac{R^2}{GL}\right) \tag{2.26}$$

where GN is replaced with L and $\vartheta = \lambda^2 w \left(D/P\right)^{2/3}$.

In summary, under proper assumptions blob theory and Flory theory yield same predictions for the free energy and the polymer extension in the two classic and extended de Gennes regimes. Blob theory breaks down as the channel size $D < 2P$ due to the large increase in the bending energy, however the Flory theory can still be used to derive the confinement free energy as long as $D > P$. Nevertheless, the polymer extension predicted by equation 2.26 is different from $R_{\text{Trans.}} \sim 1/D$ found from simulations [135,136]. In the Odijk regime ($D < P$), Flory free energy is not able to explain the polymer behavior and the confinement free energy is found via equation 2.22. Note that neither blob nor Flory theory provides exact predictions; they are both scaling arguments and ignore numerical prefactors, yet are powerful tools for explaining a wide range of experimental observations.

Physics of Compressed Nanochannel-Confined Chains

The renormalized free energy obtained in equation 2.24 for nanochannel-confined polymers can be further generalized to describe the free energy of compressed chains in nanochannels for both classic and extended de Gennes regimes [137]. Note that

replacing L_b with equation 2.18 in equation 2.24 gives the free energy of a confined polymer in the extended de Gennes regime and by setting $D_{||} = D$ in equation 2.24 and substituting L_b with equation 2.14 the free energy equation 2.23 for the classic de Gennes regime is recovered. According to the equations 2.15 and 2.19, the expected equilibrium chain size for both classic and extended de Gennes regimes scales as,

$$R_o \sim LD^{-2/3}. \tag{2.27}$$

If the chain is compressed to the extension R, equation 2.24 can be rewritten in terms of R_o and R as,

$$F_{\text{comp}} \sim k_B T \frac{L}{L_b} \left(\frac{R^2}{2R_o^2} + \frac{R_o}{R} \right). \tag{2.28}$$

Equation 2.28 works perfectly well for a uniformly extended chain in a nanochannel in both classic and extended de Gennes regimes.

2.2.3 Physical Properties of DNA

DNA is a biopolymer and is composed of sugar-phosphate backbone and a sequence of bases *Adenine* (A), *Thymine* (T), *Cytosine* (C) and *Guanine* (G). In a double-stranded DNA (dsDNA) molecule, two backbones run in opposite directions and form a double helix. The nitrogen bases on each strand form pairs with the complementary bases on the other strand; A pairs with T with two hydrogen bonds and C pairs with G with three hydrogen bonds. Each base pair has an average contour length $b_o \simeq 0.34\,\text{nm}$ [138] and the total contour length of a dsDNA is $L = b_o N$, where N is the number of base pairs. Dimeric-cyanine nucleic acid stains such as YOYO-1 and BOBO-3 are common fluorescent labels used for DNA visualization [139]. These dyes modify DNA properties such as persistence length and increase the contour length [140, 141].

For instance, unstained T4 bacteriophage DNA is about 166 kbp or $L \simeq 56\,\mu$m long. When stained with YOYO-1, the contour length increases to $L \simeq 63.7\,\mu$m [142].

The persistence length P of DNA depends strongly on the ionic condition of the DNA solution, the bending rigidity of the polymer and the temperature T [143]. Persistence length of dsDNA at high buffer salt concentration $I_s > 10\,\text{mM}$ is $P_o \simeq 51\,\text{nm}$. The Odijk-Skolnick-Fixman (OSF) theory estimates P for different salt concentration as [144, 145],

$$P = P_o + \frac{32\,\text{mM}}{I_s}\,\text{nm}. \tag{2.29}$$

The DNA effective width (w) represents the range of interaction for DNA segments and is larger than the intrinsic width $w_o \simeq 2\,\text{nm}$. The Stigter theory considers DNA segments as charged cylinders and estimates w based on the excluded volume interactions between the segments [145, 146]. The effective width for a dsDNA in an 8 mM buffer is $w \simeq 17\,\text{nm}$.

2.2.4 Polymer Dynamics

Thus far, we have considered the static configuration and the corresponding free energy of polymers both in bulk and under confinement. Polymer solutions are also subject to different time-dependent phenomena such as relaxation and constant random deformation and movement due to the thermal agitation of the molecules or density gradient [109]. Polymer Dynamics models the time-dependent processes in polymer solutions, which helps elucidate the nature of some experimentally-observed phenomena such as chain relaxation [147] and interaction with defects [148]. In this section, we first introduce the concept of diffusion in polymers and subsequently explain the dynamics of a nanochannel-confined chain subjected to compression by a sliding gasket.

Diffusion in Polymers

Diffusion in polymers can happen due to the thermal motion of a molecule with respect to the other ones, described by the *self-diffusion* coefficient D_s, or due to a density gradient in the polymer solution, characterized by *cooperative diffusion* coefficient D_c [149]. In the following part, we derive the diffusion coefficients using two different approaches.

The Langevin equation. We derive the equation of Brownian motion for a point particle with mass m in free space [109, 150]. For simplicity, we assume that the particle moves only in one dimension. Let $x(t)$ be the position of the particle at time t and $v = \dot{x}$ be its velocity. Thus, the particle feels a drag force F_f from the fluid. Moreover, the solution molecules exert a fluctuating force on the particle, presented as the stochastic variable $F_r(t)$. If we use the classic equation of motion $\sum F = m\ddot{x} = m\dot{v}$ to describe the particle dynamics, we get

$$m\frac{dv}{dt} = F_f + F_r(t) \tag{2.30}$$

where $F_f = -\zeta v$ and ζ is the friction factor. The ensemble average of the random force $\langle F_r(t)\rangle = 0$ and the time correlation function for macroscopic time-scales is,

$$\langle F_r(t_1)F_r(t_2)\rangle = A\delta(t_1 - t_2), \tag{2.31}$$

which denotes the ultra short correlation time on the order of the molecular collision time (psec). A is a constant and is found as [150],

$$A = 2\zeta k_B T. \tag{2.32}$$

By solving equation 2.30 for $v(t)$, we have

$$v(t) = \frac{1}{m}\int_{-\infty}^{t} dt_1 e^{-(t-t_1)/\tau_v} F_r(t_1) \tag{2.33}$$

where $\tau_v = m/\zeta$ is the velocity correlation time. Thus, $\langle v(t)v(t')\rangle$ can be calculated using equations 2.31, 2.32 and 2.33 as,

$$\langle v(t)v(t')\rangle = \frac{k_B T}{m} e^{-|t-t'|/\tau_v}. \tag{2.34}$$

In addition, we can show that the time correlation of velocity can be written as a function of the time correlation function of the position [150],

$$\begin{aligned}\langle \dot{x}_i(t)\dot{x}_j(0)\rangle &= -\frac{\partial^2}{\partial t^2}\langle \Delta x_i(t)\Delta x_j(0)\rangle \\ &= \frac{1}{2}\frac{\partial^2}{\partial t^2}\langle (x_i(t)-x_i(0))\,(x_j(t)-x_j(0))\rangle.\end{aligned} \tag{2.35}$$

Thus, we can write,

$$\langle (x_i(t)-x_i(0))\,(x_j(t)-x_j(0))\rangle = 2\int_0^t dt_1 \int_0^{t_1} dt_2 \langle \dot{x}_i(t_2)\dot{x}_j(0)\rangle, \tag{2.36}$$

which gives,

$$\langle (x(t)-x(0))^2\rangle = 2\int_0^t dt_1 \int_0^{t_1} dt_2 \langle v(t_2)v(0)\rangle. \tag{2.37}$$

Assuming that τ_v is negligible compared to the time scale under consideration, t_1 in equation 2.37 can be approximated as $t_1 \approx \infty$. Therefore, we can define the constant D_s as,

$$D_s = \int_0^{\infty} dt \langle v(t)v(0)\rangle. \tag{2.38}$$

Equations 2.37 and 2.38 give,

$$\langle (x(t) - x(0))^2 \rangle = 2D|t|. \tag{2.39}$$

Substitution of equation 2.34 in 2.38 yields,

$$D_s = \frac{k_B T}{m} \tau_v = \frac{k_B T}{\zeta}. \tag{2.40}$$

D_s is called the *self-diffusion* coefficient and equation 2.40 is referred to as the *Einstein relation*. D_s is inversely proportional to the friction constant ζ, thus shows that the fluctuation in the particle position is restricted by the friction between the particle and the surrounding fluid.

The Smoluchowski equation. We use the generalization of diffusion equation to describe the diffusion of particles in the presence of density gradient as well as an external potential $U(x)$ [109]. The process of diffusion starts when concentration is non-uniformly distributed. When $(U(x) = 0)$, the process of diffusion is described by the Fick's law as,

$$j(x,t) = -D_d \frac{\partial c(x,t)}{\partial x}, \tag{2.41}$$

where $j(x,t)$ is the diffusion flux and $c(x,t)$ is the concentration at position x and time t. D_d is called the diffusion constant, also known as the *collective diffusion* constant. The conservation of concentration leads to the continuity equation,

$$\frac{\partial c}{\partial t} + \frac{\partial j}{\partial x} = 0. \tag{2.42}$$

Substitution of equation 2.41 in 2.42 gives the diffusion equation [109],

$$\frac{\partial c}{\partial t} = D_d \frac{\partial^2 c}{\partial x^2}. \tag{2.43}$$

If the external potential $U(x) \neq 0$, a force $F = -\frac{\partial U}{\partial x}$ is exerted on the system, leading to an average velocity $v = -\frac{1}{\zeta}\frac{\partial U}{\partial x}$ and an additional flux cv. Fick's law is then modified as,

$$j = -D_d \frac{\partial c}{\partial x} - \frac{c}{\zeta}\frac{\partial U}{\partial x}. \tag{2.44}$$

When the system is at equilibrium, the flux vanishes and the concentration c_{eq} is obtained by the Boltzmann distribution,

$$c_{\text{eq}} \sim e^{-U(x)/k_B T}. \tag{2.45}$$

From equations 2.44 and 2.45 for $j \neq 0$, we retrieve the Einstein relation,

$$D_d = \frac{k_B T}{\zeta}. \tag{2.46}$$

Note that D_d is generally different from the self-diffusion constant D_s, but the two constants are essentially the same in dilute solutions [150]. By substituting equation 2.46 in 2.44 and 2.42, we obtain the Smoluchowski equation,

$$\frac{\partial c}{\partial t} = \frac{\partial}{\partial x}\frac{1}{\zeta}\left(k_B T \frac{\partial c}{\partial x} + c\frac{\partial U}{\partial x}\right). \tag{2.47}$$

Cooperative diffusion coefficient D_c is closely related to the collective diffusion coefficient D_d and is defined for semidilute polymer solutions [108], where the osmotic pressure gradients induces the cooperative diffusion of polymer segments [108]:

$$D_c(c) = \frac{c\xi^2}{\zeta}\frac{\partial \Pi}{\partial c}. \tag{2.48}$$

ξ is the symmetric blob size. The cooperative current is found as,

$$J_D = -D_c \frac{\partial c}{\partial x}. \tag{2.49}$$

D_c is dependent on concentration c and describes the dissipation of concentration fluctuations. The main difference between the self-diffusion coefficient D_s and the cooperative diffusion coefficient D_c is that D_c increases with concentration, while D_s decreases remarkably.

The Sliding Gasket Model

In the sliding gasket experiments, a nanochannel-confined DNA is dynamically compressed by an optically trapped bead with a fixed velocity V relative to the polymer [151, 152]. Prior to the compression, the polymer has the equilibrium extension and concentration r_o and c_o, respectively. Once the bead hits the DNA, the molecule experiences transient compression followed by a final compressed steady-state. Upon bead removal, the polymer relaxes back to the equilibrium state. During the transient compression and relaxation, the chain concentration is a function of both the position on the chain and the time ($c(x,t)$), while the compressed steady-state is described by a time-independent concentration profile ($c_s(x)$).

Compressed steady state. Khorshid *et al.* [151] show that there exists a critical sliding speed $V^* = k_B T/(D\eta r_o)$ below which the polymer slides with the bead and is not compressed. The chain concentration and extension are the same as those in the equilibrium state (Figure 2–5 (a,b)). The solution viscosity is denoted by η and D is the channel size. At $V > V^\star$, the molecule reaches the compressed steady-state, with time-independent average extension r and concentration profile $c_s(x)$. At the bead speed $V^* < V < V^{**} = k_B T/D^2\eta$, the chain is slightly compressed and the concentration profile is composed of a flat portion and a linear ramp. The concentration at the free

edge of the molecule equals the equilibrium concentration c_o (Figure 2–5 (c,d)). At higher velocities ($V > V^{**}$), the DNA is highly compressed (Figure 2–5 (e)) and the concentration profile is only a linear ramp. The concentration of the free edge of the molecule at high compression is larger than the equilibrium concentration c_o (Figure 2–5 (f)).

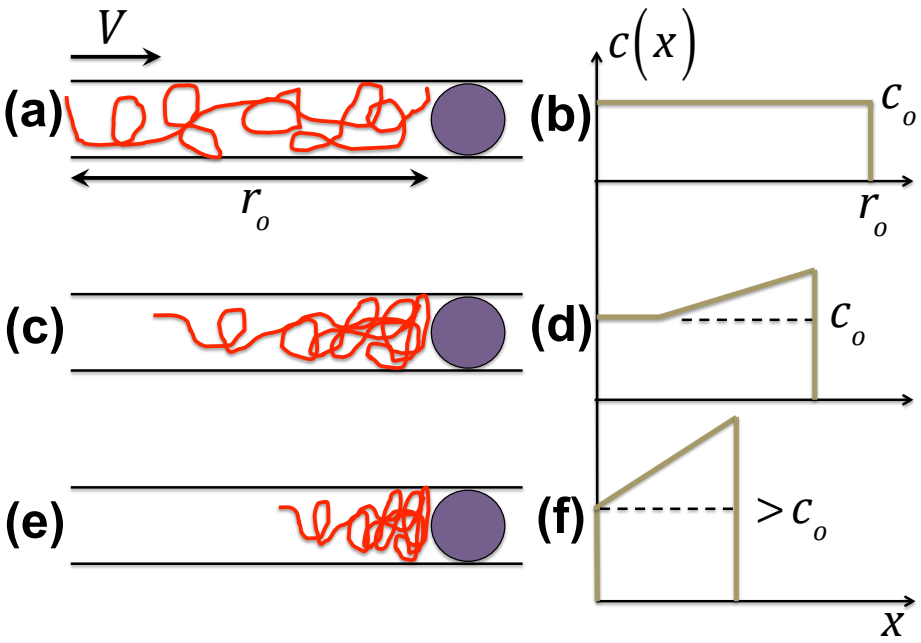

Figure 2–5: A schematic of the sliding gasket experiment. (a) When the sliding speed V is very small ($V < V^*$), the chain is not compressed and slides with the bead. The extension of the polymer remains r_o. (b) The concentration profile of the polymer is the same as that in the equilibrium state (c_o). (c) At $V^* < V < V^{**}$, the chain is slightly compressed; (d) the concentration profile includes a flat portion and a linear ramp. The concentration at the free edge of the molecule is still the same as the equilibrium concentration c_o. (e) At higher velocities ($V > V^{**}$), the chain is highly compressed; (f) the concentration profile is only a linear ramp. The concentration of the molecule free edge at highly compressed state is larger than its equilibrium concentration c_o.

Transient dynamics. During the transient compression and relaxation phases, the continuity equation 2.42 holds, as the concentration is conserved. The current is

found as,

$$j = -D_c(c(x,t))\frac{\partial c}{\partial x} + cV, \tag{2.50}$$

where D_c is the cooperative diffusion coefficient and is a function of concentration and the osmotic pressure gradient (equation 2.48). Here, the cooperative diffusion coefficient is utilized to describe the internal dynamics of a *single* chain. This is possible due to the fact that semidilute solutions and single chain are both composed of packings of blobs that have a size small compared to the radius of gyration of a single chain. Note that at steady state, the cooperative current j_D (equation 2.49) balances the convective current $j_c = cV$, thus j in equation 2.50 vanishes,

$$D_c\frac{\partial c}{\partial x} = cV. \tag{2.51}$$

In general, equations 2.42 and 2.50 give,

$$\frac{\partial c}{\partial t} - \frac{\partial}{\partial x}\left(D_c\frac{\partial c}{\partial x} - cV\right) = 0. \tag{2.52}$$

Solving the partial differential equations 2.51 and 2.52 gives the concentration profile $c_s(x)$ in the steady-state [151] and the evolution of the concentration profile $c(x,t)$ in the transient cases [152], respectively.

2.2.5 Physics of Knot Formation on Polymers

Entanglements are caused by physical constraints such as excluded volume effects and the connectivity between chain segments. Knots form when entanglements are subjected to topological (non-crossing) constraints.

Chains in bulk – Grosberg and Rabin [153] have developed a formalism for knot formation on worm-like chains in bulk. Their theory suggests that non-crossing

constraints on a semiflexible chain with zero thickness is analogous to self-confinement of the chain in a virtual tube of finite thickness D_k (Figure 2–6).

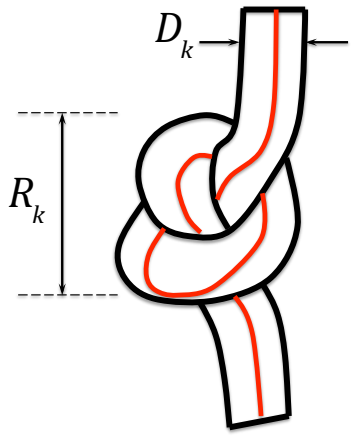

Figure 2–6: A knot on a worm-like chain with size R_k is self-confined in a virtual tube with diameter D_k.

Two terms contribute the free energy cost of knot formation: (1) bending energy $F_{\text{bend}}^{\text{WLC}}$ as a result of an increase in the curvature of the chain in the knotted region, and (2) confinement free energy $F_{\text{conf}}^{\text{WLC}}$ due to the non-crossing restriction within the knotted region [153, 154],

$$F_{\text{bend}}^{\text{WLC}} \sim \frac{L_k P}{R_k^2} \tag{2.53}$$

$$F_{\text{conf}}^{\text{WLC}} \sim \frac{L_k}{D_k^{2/3} P^{1/3}} \tag{2.54}$$

where $R_k \sim D_k$ is the knot size and $L_k \sim D_k$ is the contour length within the knot. By replacing D_k and R_k in the equations 2.53 and 2.54 with L_k, the total free energy cost of knot formation on a worm-like chain in bulk is obtained as,

$$\begin{aligned} F_{\text{b}}^{\text{WLC}} &= F_{\text{bend}}^{\text{WLC}} + F_{\text{conf}}^{\text{WLC}} \\ &\sim \left(\frac{L_k}{P}\right)^{-1} + \left(\frac{L_k}{P}\right)^{1/3}. \end{aligned} \tag{2.55}$$

Equation 2.55 can be extended to describe knot formation free energy cost for polymers with finite thickness w [154]. Due to the repulsion between the chain and the virtual tube walls, D_k in equation 2.54 is replaced with the effective diameter of the tube $D_k^{\text{eff}} = D_k - w$. Consequently, equation 2.54 can be written for self-avoiding chains in the following form,

$$F_{\text{conf}}^{\text{real}} \sim \frac{L_k}{(D_k - w)^{2/3} P^{1/3}}. \tag{2.56}$$

Replacing D_k with L_k yields,

$$\begin{aligned} F_{\text{conf}}^{\text{real}} &\sim \frac{L_k}{(D_k - w)^{2/3} P^{1/3}} \\ &\sim \frac{L_k}{(L_k - pw)^{2/3} P^{1/3}} \end{aligned} \tag{2.57}$$

where $p \sim L_k/D$ is a numerical prefactor. By replacing $F_{\text{conf}}^{\text{WLC}}$ in 2.55 with equation 2.57, we get the free energy cost of forming a knot in a real chain,

$$F_{\text{b}}^{\text{real}} = k_1 \left(\frac{L_k}{P}\right)^{-1} + k_2 L_k \left(L_k - pw\right)^{-2/3} P^{-1/3} \tag{2.58}$$

where $k_1 = 17.06$, $k_2 = 1.86$ and $p = 16$ are found from simulations for trefoil knots [154]. The first term (bending energy) and the second term (confinement free energy) in equation 2.58 tend to swell and shrink the knot, respectively. The competition between these two terms minimizes the free energy at $L_k \simeq 12P$ referred to as *metastable* knot size [154]. The notion of metastable knot size has been controversial among different studies. While many experimental [9, 13, 18, 46, 155] and theoretical [11, 153, 154, 156–159] studies agree that knots on chains are localized as tight knots, other studies [160–162] believe that knots can form which are not tightly localized and can spontaneously expand along chains.

Chains under confinement – Dai *et al.* [11] have shown that when a semiflexible chain with finite thickness w is confined in a square channel of size D, the total free energy cost of knot formation is derived as,

$$F_{\rm tot}(L_k) = F_{\rm b}(L_k) + F_{\rm wk}(L_k) - F_{\rm wuk}(L_k) \tag{2.59}$$

where $F_{\rm b} \equiv F_{\rm b}^{\rm real}$ (equation 2.58), $F_{\rm wk}(L_k)$ is the energy cost for knot confinement (L_k

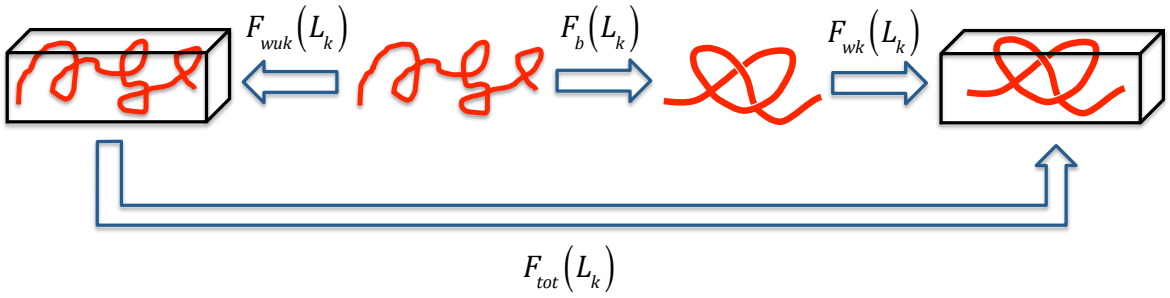

Figure 2–7: A diagram representing the free energy cost for knot formation on nanochannel-confined DNA. The change from the state of a confined unknotted chain to that of a confined knotted chain costs $F_{\rm tot}(L_k)$, which is equal to $F_{\rm b}(L_k) + F_{\rm wk}(L_k)$ (pointing in the same direction as that of $F_{\rm tot}(L_k)$) $-F_{\rm wuk}(L_k)$ (opposite direction).

is the knot contour length) and $F_{\rm wuk}(L_k)$ is the confinement free energy of an unknotted chain with contour length L_k. Figure 2–7 illustrates the knot formation free energy diagram. For simplicity, changes from one state to another are shown by arrows, which determine the signs in equation 2.59.

In the de Gennes regime where $P \ll D \ll R_F$, the confinement free energy of an unknotted chain with contour length L_k is derived as [11],

$$F_{\rm wuk}(L_k) \sim \frac{(wP)^{1/3}\, L_k}{D^{5/3}}. \tag{2.60}$$

Assuming that the confined knot is much smaller than the channel size ($g_k \ll D$, where g_k is the knot gyration radius), we can approximate the knot as a ball on a string of

identical balls with radius g_k. The confinement free energy for one ball is thus,

$$F_{\mathrm{wk}}(L_k) \sim \frac{g_k^{5/3}}{(D-2g_k)^{5/3}} \tag{2.61}$$

where $2g_k$ is the diameter of the knot. The denominator in equation 2.61 has replaced $D^{5/3}$ due to the fact that a ball of size $2g_k$ has occupied the channel and thus decreased the effective channel diameter. Dai *et al.* [11] find that the knot diameter is approximately proportional to L_k ($2g_k = \alpha L_k$, α is a numerical prefactor). Replacing $2g_k$ in equation 2.61 with αL_k gives,

$$F_{\mathrm{wk}}(L_k) \sim \frac{L_k^{5/3}}{(D-\alpha L_k)^{5/3}}. \tag{2.62}$$

After substituting equations 2.58, 2.60 and 2.62 in 2.59 we get,

$$F_{\mathrm{tot}}(L_k) = k_1 \left(\frac{L_k}{P}\right)^{-1} + k_2 L_k \left(L_k - pw\right)^{-2/3} P^{-1/3} + \beta \frac{L_k^{5/3}}{(D-\alpha L_k)^{5/3}} - \gamma \frac{(wP)^{1/3} L_k}{D^{5/3}} \tag{2.63}$$

where β and γ are numerical prefactors that are found from simulations. Simulations and theoretical results have shown that minimization of $F_{\mathrm{tot}}(L_k)$ with respect to L_k results in a metastable knot size $L_k^\star$ [11], which holds for knots much smaller than the channel size ($g_k \ll D$). In general, the metastable knot size $L_k^\star$ changes nonmonotonically with the channel size D [11]. Decreasing D leads to larger $L_k^\star$ until $D = 12P$; at larger D, the metastable knot size $L_k^\star$ decreases drastically.

CHAPTER 3
A Nanofluidic Knot Factory based on Compression of Single DNA in Nanochannels

In the previous chapters, knots were defined as topological entities that can form in macroscopic and microscopic scales. The mathematical definition of knots in closed chains was given and the definition of knots on open curves was provided. Formation of knots on biopolymers was discussed and examples of knot observations in biological systems such as in phage capsids were provided. Moreover, instances of knot formation *in vitro* were presented, such as knots formed on DNA molecules in nanochannels [20] and induced via DNA collision with defects on nanochannel walls [13]. In this chapter, a knot factory on chip is introduced for efficient formation and detection of knots. The proposed knotting method provides an experimental tool to assess the impact of multiple variables influencing knotting probability. Knots are produced during hydrodynamic compression of single DNA molecules against barriers in a nanochannel; subsequent extension of the chain enables direct assessment of the number of independently evolving knots. Knotting probability increases as the chain is compressed. In addition, we observe that knotting probability increases with waiting time in the compressed state, enabling direct access to knot formation kinetics. At high compression, our results suggest that knots do not form in our system via an independent Poisson process, hinting that knot-interactions may play a role in modifying the knot-formation statistics, an observation supported by the observed structure of knot spatial distributions. Using a free energy derived from scaling arguments, we develop a knot formation

model that can quantify the effect of interactions and the breakdown of Poisson statistics at high compression. Our model suggests that knotted states at high compression are stabilized by a decreased free energy as contour stored in the knots contributes a lower self-exclusion derived free energy than if the same contour was allowed to interact freely in the nanochannel.

This chapter is the integral text from:

A Nanofluidic Knot Factory based on Compression of Single DNA in Nanochannels

Susan Amin, Ahmed Khorshid, Lili Zeng, Philip Zimny and Walter Reisner; Nature Communications 9, no. 1506 (2018): 1506 [163].

3.1 Introduction

Knots naturally exist in DNA, proteins, umbilical cords and catheters [4,5]. Knots can form when an initially linear chain passes its loose free ends through one or multiple loops on the same chain, giving rise to a knot if the polymer is subsequently cyclized. For example, random cyclization of linear DNA in bulk [41] forms knotted chains with low probability; or knots can be directly tied via optical tweezers [15]. Chain compaction, induced via either spatial confinement, compression, or molecular crowding [164], tends to enhance the tendency for chains to self-entangle, and thus enhances knotting probability. Knot formation on DNA is a particular challenge in biology, due to high degree of compaction experienced by packaged genomes, and is consequently tightly regulated by enzymes like topoisomerases and recombinases that remove knots by breaking and rejoining of either single or double strands [4]. An extreme example is the high level of compaction experienced by viral genomes [10], resulting in a correspondingly high knotting probability for DNA extracted following capsid rupture [10]. Knots on genomic DNA in nanofluidic systems interfere with mapping by preventing complete linearization of contour stored in the knot, giving rise to an artifact resembling a deletion [20]

Consequently, there has been intense theoretical focus on knot production mechanisms [4] and physics of confined knots [11, 42, 165, 166]. Yet, while single-molecule techniques for knot sensing are advancing rapidly [167], and single knot diffusion and

size dynamics have been explored [15, 46], systematic experimental studies probing conditions enhancing knot formation in microscopic systems are limited. Knotting in DNA extracted from the P4 phage system has been extensively studied, but an *in vivo* system has inherent disadvantages, including a fixed parameter space and difficulty of determining whether knotting occurs inside the capsid or following rupture. There have been reports of knot formation in nanochannels [13]; coil collapse in an AC field has been used to induce knotted and self-entangled states of a single chain [18], but these experiments did not systematically quantify knot formation as a function of compaction. Understanding of knot-formation in microscopic chains is framed [4, 18] by a classic experiment involving tumbling of a macroscopic string inside a rotating box [26]. In this experiment, knots were formed when successive tumbles drove parallel concentrically coiled strands near the chain ends to cross. At low agitation times, knot formation was observed to be kinetically limited; at longer agitation times, the knotting probability saturated at a value that approached unity for longer, highly flexible strings [4]. An intriguing question is whether experiments probing knot formation in microscopic chains might reveal a similar kinetically limited regime at low times and a saturating knotting probability at long-times.

In this chapter, we introduce a knot factory on chip using low Reynolds number flow to compress single DNA molecules against slit-barriers in nanochannels (Fig. 3–1(a-f)). The chain is initially extended (Fig. 3–1(d)). After compression (Fig. 3–1(e)), the flow is released and the DNA molecules relax (Fig. 3–1(f)); knots are present along the relaxed DNA, visualized as sharply localized regions of high intensity on the extended molecule (Fig. 3–1(g-j). The knot-factory enables efficient knot formation and

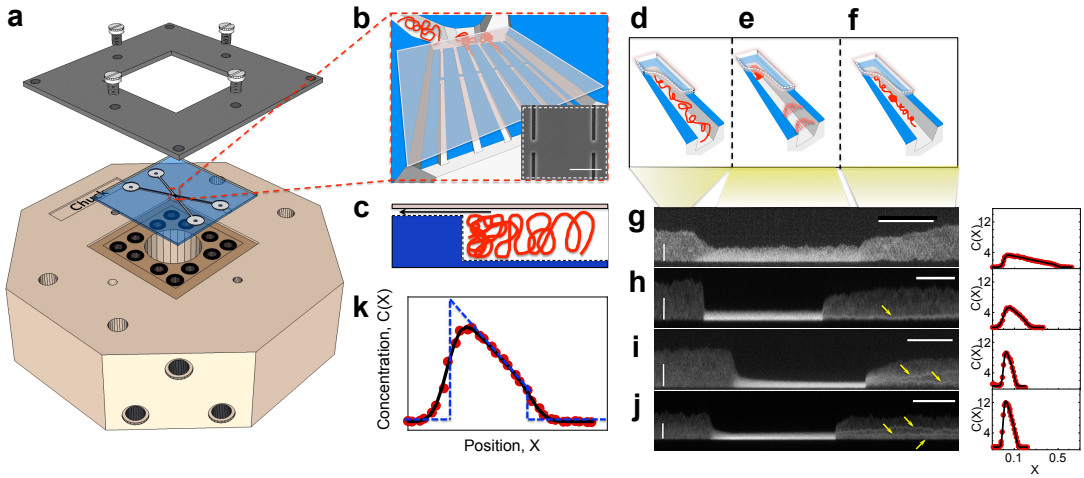

Figure 3–1: (**a**) The nanofluidic device is mounted on a chuck containing inputs for application of pneumatic pressure. Pneumatic pressure is used to transport DNA molecules in micro and nanochannels and enable hydrodynamic compression against slit barriers. (**b**) A magnified view of the center of the chip. The device is composed of two 1-μm deep loading channels spanned by a nanochannel array with blunt-ended barriers fabricated in the channel centers. A 30-nm deep slit, etched over the array, allows for solvent to escape while preventing the passage of large DNA molecules. *Inset*: an SEM image of the nanochannels with barrier (the scale bar is 3-μm). (**c**) A magnified cross-sectional view of a nanochannel at the device center showing the slit-barrier. The black arrow depicts the flow direction through the slit. (**d**)-(**f**) A three dimensional cartoon showing the process of knot formation detailing (**d**) DNA confinement, (**e**) compression against the barrier via hydrodynamic flow induced by applying a pressure drop across the nanochannels and (**f**) relaxation of a knotted chain. The red arrows in (**e**) depict the velocity profile of the flow during compression. (**g**)-(**j**) Examples of kymographs for knot-formation events with increasing degree of compression. Intensity along the nanochannel (vertical axis, scale bar 10 μm) is plotted versus time (horizontal axis, scale bar 10 s). Each molecule is compressed, held at a minimum extension for a waiting time t_w, and then relaxed. Normalized chain concentration profiles corresponding to the kymographs are illustrated on the right. (**g**) No knot is formed; (**h**) one knot is formed; (**i**) two knots are formed and (**j**) three knots are formed. The yellow arrows depict the knot locations. The second bright spot in (**h**) does not maintain its size and unravels in the chain mid-section shortly after pressure release so we do not count it as a knot [153]. (**k**) Normalized chain concentration profile $C(X)$, averaged over waiting time at compressed state; red circles are experiment; black line is a fit to a linear ramp concentration profile (Eq. 3.1) convolved with a Gaussian point-spread function (section 4.2.2). The blue-dashed line shows an estimate of the real, i.e. prior to convolution, concentration profile estimated from the fit.

detection in an *in-vitro* system where all parameters are well controlled, guiding development of models to quantify conditions favoring knot production. In particular, we can measure knot-formation probability as a function of compression and probe knot formation kinetics by relaxing the chain after a well-defined waiting period in the compressed state. By measuring knot position along the chain a short time after pressure release, we can gain insight into the knot spatial distributions. Finally, we can access conditions where more than one knot is formed, enabling investigation of the formation of composite knot states. Our results suggest that strong interaction between prime knots exist in composite knot states, leading to a breakdown of independent Poisson knot-formation statistics observed for extended chains at equilibrium in the absence of compression [12]. To rationalize our findings, we argue that the compressed chain is in a steady state with zero segmental current, equivalent to a state of inhomogeneous equilibrium, so that a generalized free energy can be developed to quantify the probability of knot-formation. This approach complements existing theories for knot-production by clarifying knotting free energy landscapes for compressed chains and explains our observations if topological barriers for forming knots are sufficiently small relative to thermal or flow-driven agitation for the chain to sample knotted states over measurement time-scales.

T_4 DNA are driven into nanochannel arrays from loading microchannels via a burst of pneumatic pressure (Fig. 3–1(b)). Low ionic strength conditions (10 mM Tris, pH 8.0) are used to ensure negligible knotting probability in bulk by ensuring a high DNA effective width (section 4.1, [41]). Once the molecules enter the nanochannel, the pressure is released in order to acquire movies of freely fluctuating extended DNA (Fig. 3–1(d)). This data is later used to determine the molecule equilibrium extension in the absence of flow r_o and concentration c_o via a standard fitting model based on

the convolution of a box with a Gaussian point-spread function [126] (see section 4.2.2, note $r_o = 14.3 \pm 0.3\,\mu$m for our channels). The molecules are then driven to the array center, compressed against the slit-barriers to an extension r and held at this extension for a waiting time t_w (Fig. 3–1(e)) until the pressure is released and the molecules relax (Fig. 3–1(f)).

The knotting state can be assessed by counting the number of knots present. Figure 3–1(g-j) gives example kymographs for individual compression-relaxation events at different degrees of compression. The sharply localized regions of high intensity in the kymographs represent knots [13, 18]. As we increase the pressure and compress the chain, the knot formation probability rises (Fig. 3–1(g-j)). In addition, for very high compression we observe that more than one knot can be formed (Fig. 3–1(i, j)). After each compression event, molecules are driven out of the nanochannels, new molecules are introduced in order to avoid possible effects of entanglement [18] that might lead to hysteresis.

Knots can be distinguished from other topological events such as "folds" [168] or "trivial knots" (including entangled segments and complex unknots, like slip-knots [169], that do not possess true knot topology) as knots formed on semiflexible chains quickly adopt a characteristic compact and time invariant structure [11, 153, 154]. A knot's topology creates an effective network of non-crossing constraints that is equivalent to confining the chain in an effective tube. The balance between bending and confinement free energy of polymer in the tube leads to a metastable knot size with a high free energy barrier for knot loosening, giving rise to knots possessing a soliton-like structure [153] with a stable shape that diffuse along the chain through self-reptation of contour. For example, knots tied by tweezers in fluorescently labeled chains will quickly adopt a bright (highly concentrated) and localized (sub-diffraction limit) structure that

can only be removed when the knot diffuses to the chain ends [9]. In contrast, "trivial knots" or "unknots," such as folds or entangled regions, are expected to gradually unravel under the influence of entropic forces driving contour to less confined regions [168]. Unknots can also unravel in mid-chain. Thus, in contrast to other topological events, knots are objects that once formed on the polymer: (1) are persistent, localized and bright features; (2) do not exhibit large-scale size fluctuations after reaching their final state; and (3) can unravel only at the molecule ends. Our approach, like those explored in references [13, 18], cannot form knots of known topology (in contrast to tweezers based approaches [9, 15]). Directly tying knots with tweezers, however, is extremely challenging and low-throughput [9] and non-trivial to apply in confined systems.

3.2 Methods

3.2.1 Device Fabrication and Experimental Set-up

The nanochannels are fabricated on fused silica substrates (HOYA) by electron beam lithography as described in [126]. The slit barriers are formed by patterning the nanochannels with blunt ends in the array center (Fig.3–1(b)). A 30-nm deep slit (measured using surface profilometry) is subsequently etched over the nanochannel array, transforming the blunt ends into barriers that will permit buffer flow but trap the DNA. In addition, adjoining the nanochannel array, the device contains two U-shaped microchannels (1μm deep, 50μm wide): these microchannels convey molecules from sand-blasted loading holes to the nanochannels. The 1×1 cm chips are then bonded directly to fused silica coverslips (Valley Design)(Figure 3–1(b)). The cover slip seals the channels while the slit introduces an opening at the barrier end of the nanochannels, which allows flow to pass, but traps DNA molecules (Figure 3–1(c)). Upon imaging

the chip cross section using SEM, the nanochannels have horizontal dimension $D_1 = 325\,\text{nm}$ and vertical dimension $D_2 = 415\,\text{nm}$ (Section 4.1 gives more detail on dimension acquisition). The loading buffer consists of 10 mM Tris titrated with HCl to pH 8.0. In addition, 2% β -mercaptoethanol (BME) is added to suppress photobleaching and photonicking. The DNA constructs used consist of T_4 bacteriophage DNA (Nippon Gene, 166 kbp), stained with YOYO-1 (Life Technologies) at an intercalation ratio of 10:1, resulting in a contour length of about 63.7μm [142]. The wet-chip is mounted on a chuck via o-ring seals with inlets for applying pneumatic pressure (Fig. 3–1(a)). The chuck-chip assembly is then mounted on an inverted microscope (Nikon Eclipse Ti-E) with a 100X N.A. 1.5 oil immersion objective. Imaging is performed via an EMCCD camera (ixon, Andor) with excitation illumination provided by a metal-halide lamp (Xcite).

3.3 Results

3.3.1 DNA Concentration Profile during Compression

The nanoscale dimensions of our channels give rise to ultra-low Reynolds number hydrodynamics ($\text{Re} \sim 10^{-8}$) that necessitate rigorously laminar and steady streamlines in the presence of a constant pressure drop [170]. Note that formation of nano-vortices at the slit-barrier requires a Reynolds number exceeding $\text{Re} = 0.055$ [171]. In contrast to [18], we do not apply an external electric field. We estimate that any streaming potential difference [170] across the nanochannel resulting from our flow is less than 1 mV, affecting the DNA velocity by less than 1% (see section 4.2.4).

We find the laminar flow leads to physics analogous to that of our optical piston experiments [151, 152], where an optically-trapped bead is used as a sliding gasket to compress single double-stranded DNA molecule with fixed velocity V (see section 4.2.5 for extended discussion). Like the sliding-gasket experiments, during the first

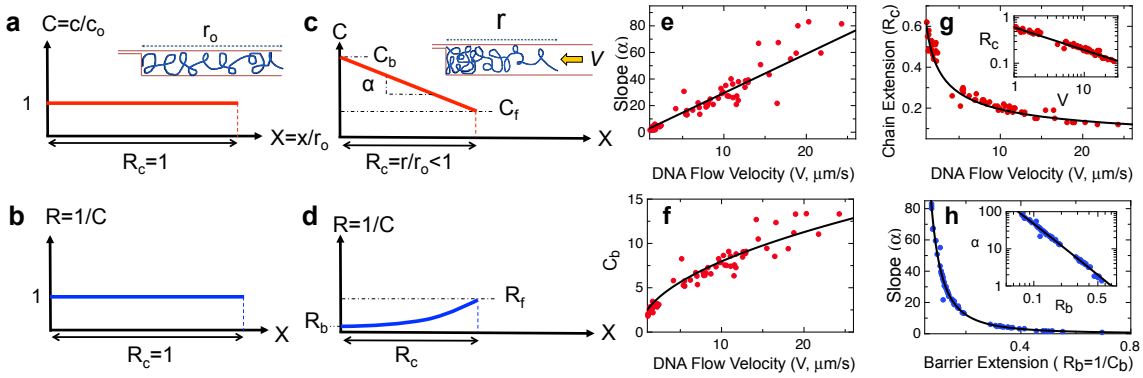

Figure 3–2: (**a**) The nanochannel confined chain in no-flow equilibrium has a uniform concentration profile $C(X) = 1$. Inset: schematic of no-flow equilibrium chain with extension r_o. (**b**) Local extension $R(X) \equiv 1/C(X)$ for chain in no-flow equilibrium. (**c**) When a flow V is applied, in the long-time limit the chain reaches a steady-state with a concentration profile that ramps linearly towards the barrier: $C(X) = C_b - \alpha X$. Inset: schematic of chain in flow-constrained equilibrium with extension r. (**d**) Local extension $R(X) \equiv 1/C(X)$ for chain in flow-constrained equilibrium. Note that for a uniform profile $R_b = R_c$. (**e**) Profile slope α, (**f**) barrier concentration C_b and (**g**) chain extension R_c versus V with fits to scaling relations predicted by piston theory ((**e**) $\alpha \sim V$, (**f**) $C_\mathrm{b} \sim \sqrt{V}$ and (**g**) $R_c \sim 1/\sqrt{V}$). (**h**) Combining data in (**e**) and (**f**) yields α versus R_b, described well by the scaling $\alpha \sim 1/R_\mathrm{b}^2$. The insets in (**g**) and (**h**) give the results on a log log-scale.

phase of compression, a "shock-wave" of segmental concentration builds up at the molecule edge abutting the barrier [152] (Fig. 3–1(g-j)). In this phase, the position of the molecule edge opposite the barrier (the 'free edge') is unconstrained and has constant speed V, a measure of the buffer flow speed in the channel (see Fig. 4–4 and section 4.2.6 for measurement of V). The second phase begins when the shock-wave reaches the free edge. In this second phase, the laminar flow forces the chain immobile against the slit barrier with forces due to the osmotic pressure gradient everywhere balancing hydrodynamic forces so that the net polymer current is zero (i.e. zero net movement of Kuhn segments). This zero current steady-state is equivalent to a state of inhomogeneous or force-constrained equilibrium [172]. In this state, the compressed

molecule, spanning the range from $x = 0$ to $x = r$, adopts a ramped concentration profile (Fig. 3–1(k) and Fig. 3–2(c)). Sliding gasket theory suggests the ramp is linear; in terms of the normalized variables $C \equiv c(x)/c_o$, $X \equiv x/r_o$ and $R_c = r/r_o$ the ramp has the form [152]

$$C(X) = C_\mathrm{b} - \alpha X, \tag{3.1}$$

with X ranging from zero to R_c. The quantity $C_\mathrm{b} \equiv C(0)$ is the (maximal) concentration at the slit barrier and α is the ramp slope (Fig. 3–2(c)). Equation 3.1, once convolved with a Gaussian point-spread function (see section4.2.2), describes experiment well (Fig. 3–1(k)). The parameters α, C_b and R_c are extracted from the experimental profiles via fitting to the convolved linear ramp (Fig. 3–1(k)) and plotted as a function of V (Fig. 3–2(e-g)). The plots show that the V-dependence is indeed consistent with gasket theory, which predicts the scalings $C_\mathrm{b} \sim \sqrt{V}$, $\alpha \sim V$ and $R_c \sim 1/\sqrt{V}$ [151, 152] (see section 4.2.5 for a review of the derivation of these scalings).

In addition, we choose to introduce a *local extension* $R(X) \equiv 1/C(X)$ (Fig. 3–2(b, d)). The local extension measures how locally compressed ($R < 1$) the chain is relative to the no-flow equilibrium where $C = R = 1$ everywhere along the chain. In particular, we use the *local extension at the slit barrier*, or 'barrier extension', defined by $R_\mathrm{b} \equiv 1/C_\mathrm{b}$ (Fig. 3–2(d)), to parameterize the compression profile in lieu of V or R_c. The barrier extension has useful properties; like R_c it becomes strictly smaller with increasing compression, is proportional to, but less than R_c ($R_b/R_c = 0.62 \pm 0.05$, see section 4.2.5) and directly characterizes chain properties at the slit barrier where knots are found with highest probability. Figure 3–2(h) combines the data in Fig. 3–2(e,f) and gives α as a function of R_b; this data is well described by the gasket scaling $\alpha \sim 1/R_b^2$.

3.3.2 Knotting Probability Measurement

The time-dependent knotting probability can be described by introducing constant transition rates k_{ij} from a state with i knots to a state with j knots (Fig. 3–3(a)), resulting in a set of coupled rate equations. Figure 3–3(b) gives knotting probability for single and multiple knot states as a function of waiting time. The rate equations are solved (see section 4.3) for the time dependent probabilities and fitted to the experimental results. The knotting probability rises with t_w and then asymptotes to a constant value at long-times ($t > 17\,\mathrm{s}$), suggesting a gradual equilibration of the knotting state. This equilibration time-scale compares on order of magnitude to the extensional relaxation time of confined T_4 DNA in channels of our size in no-flow equilibrium ($\sim 10\,\mathrm{s}$, obtained from scaling values for the λ-DNA relaxation time in [173] to T_4). Comparable transition rates for the no-knot to one-knot transition and the one-knot to two-knot transition suggest that the presence of an existing knot does not alter the energy barriers involved in forming the second knot ($k_{01} = 0.21 \pm 0.02\ \mathrm{s}^{-1}$ and $k_{12} = 0.19 \pm 0.13\ \mathrm{s}^{-1}$).

Figure 3–4(a) gives measurements of knotting probability as a function of R_b. The equilibrium knotting probability increases as R_b decreases. In particular, the one-knot states increase in frequency until reaching a peak at around $R_\mathrm{b} \approx 0.12$. The two-knot formation probability rises and becomes equal to the one-knot formation probability at $R_\mathrm{b} \approx 0.09$. We also observe a very small number of three-knot events (two in total).

Micheletti *et al.* suggest that formation of composite knots in nanochannel-confined DNA should arise from independent knotting events along the chain, leading to a description via Poisson statistics [12]. In particular, for a chain in no-flow

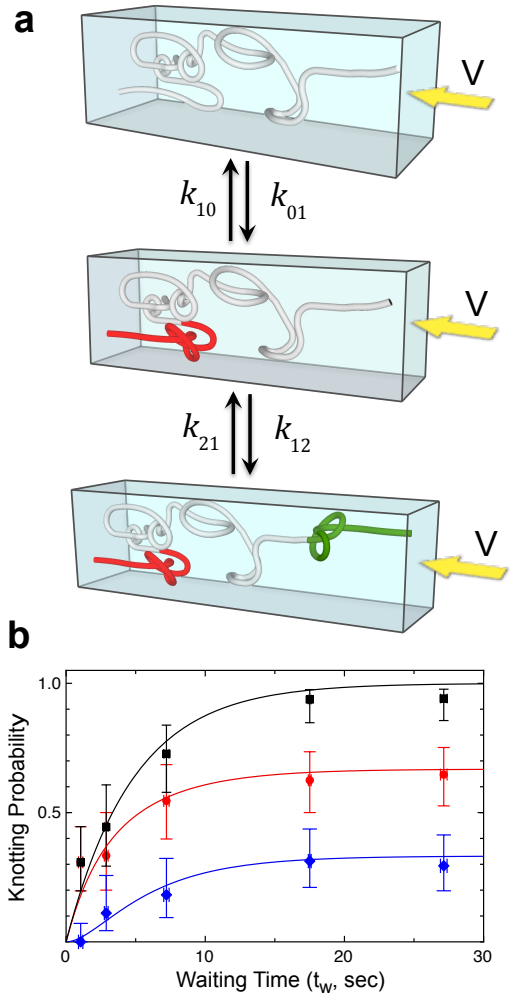

Figure 3–3: (**a**) The probability for finding no knots (*top*), forming 1 knot (*middle*), and 2 knots (*bottom*) are related through the transition rates k_{ij} from a state with i knots to a state with j knots. The transition rates define set of coupled rate equations (see section 4.3). (**b**) The probability of knot formation as a function of waiting time at average barrier extension $R_{\rm b} \approx 0.13$ with fits to the kinetic model. The black squares give experimental measurements for total probability of forming an event with any number of knots. The red circles and blue diamonds give respectively measurements of one-knot and two-knot event probabilities. The continuous lines represent the fits to the time-dependent knotting probabilities predicted by the rate equations. Each data point is determined from the average result of ∼10-15 events. The vertical error bars on probability have been calculated using a Wilson-score interval with a one-sigma confidence interval [174] (See section 4.12); the horizontal error bars show the error on the mean for t_w.

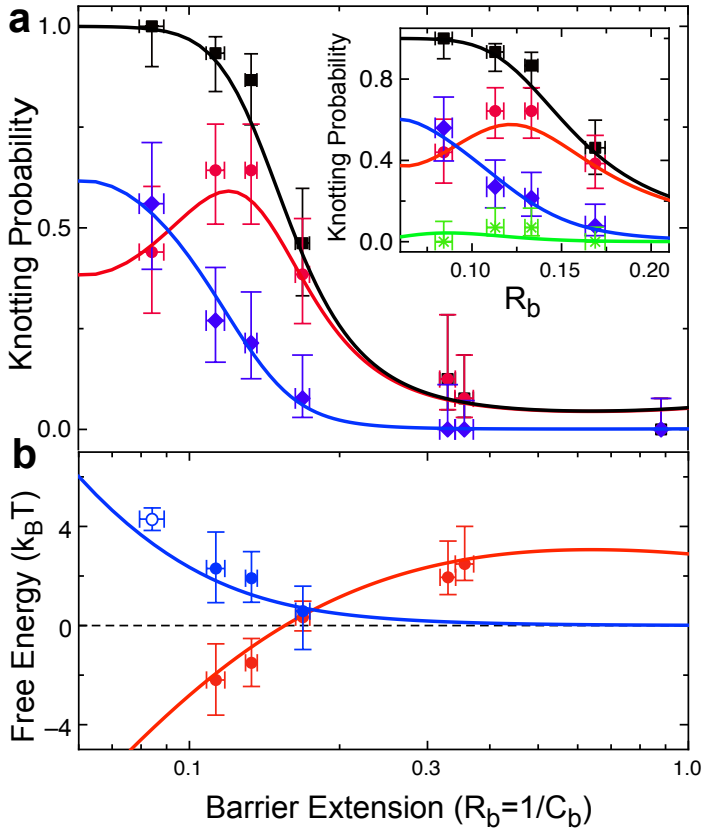

Figure 3–4: (**a**) The probability of knot formation as a function of barrier extension (y-scale linear, x-scale log). The inset in (**a**) illustrates the predicted probability of knotting for a sample space including three-knot formation probability (linear-linear scale). The black squares give experimental measurements for total probability of forming an event with any number of knots. The red circles and blue diamonds give respectively measurements of one-knot and two-knot event probabilities. The continuous lines indicate fits to the free energy model. (**b**) Free energy of single knot states $F_{\rm tot}(1, R_{\rm b})$ (red, circles) and two-knot interaction free energy $F^{\rm int}_{2,\rm tot}(R_{\rm b})$ (blue, circles) deduced assuming the profile is in a state of inhomogeneous equilibrium, with theoretical overlay using same fitting parameters for (**a**). Each data point is determined from the average result of ∼10-15 events. The vertical error bars on probability have been calculated using a Wilson-score interval with a one-sigma confidence interval [174] (See section 4.12); the horizontal error bars represent the error on the mean for $R_{\rm b}$ measurements for the corresponding events. For the blue open circle in (**b**) as no 0-knot states are observed, we estimate $F^{\rm int}_{2,\rm tot}$ by finding the difference $F_{\rm tot}(2, R_{\rm b}) - 2F_{\rm tot}(1, R_{\rm b})$ from experiment and extrapolating model predictions (red curve) to estimate the extra factor of $F_{\rm tot}(1, R_{\rm b})$.

equilibrium such as studied by Micheletti *et al.*, the Poisson model suggests that the probability of forming a composite knot based on m number of prime knots of the same topology is governed by

$$P_m = n^m \frac{e^{-n}}{m!} \tag{3.2}$$

with $n = \frac{L}{L_0}$ where L is the DNA contour length and L_0 a characteristic length scale depending on the channel width D. While the concentration profile is uniform for a chain in no-flow equilibrium, concentration uniformity is *not* required for Poisson statistics to hold; Poisson statistics requires only that the prime knots form *independently*. In an inhomogeneous Poisson process [175], the knot formation probability can vary along the chain, leading to a distribution identical to Eq. 4.65 but with n expressed as an integral of the varying knot formation probability along the chain. For both the uniform and non-uniform cases, we can eliminate n and express Eq. 4.65 purely in terms of the no-knotting ($m = 0$) probability P_0:

$$P_m = (-\log\ P_0)^m \frac{P_0}{m!}. \tag{3.3}$$

Figure 3–5 shows Eq. 3.3 plotted for $m = 1,\ 2$ and 3 against the experimental data. Values of P_0 on the horizontal axis are calculated from the observed total knotting probability $P_0 = 1 - P_{\text{total}}$ for different values of R_b. Higher values of P_0 in Fig. 3–5 correspond to lower compression. The Poisson model describes the data well when the molecules are only slightly compressed and P_0 is close to unity. However, for high compression ($P_0 < 0.2$) the Poisson model breaks down. The breakdown in Poisson statistics can be explained in two ways: (1) at high compression, constituent prime knots might interact so that their formation is no longer independent; (2) at high compression knots of complex topology are formed with higher probability, so that a single Poisson distribution does not reflect the overall knotting probability. We believe

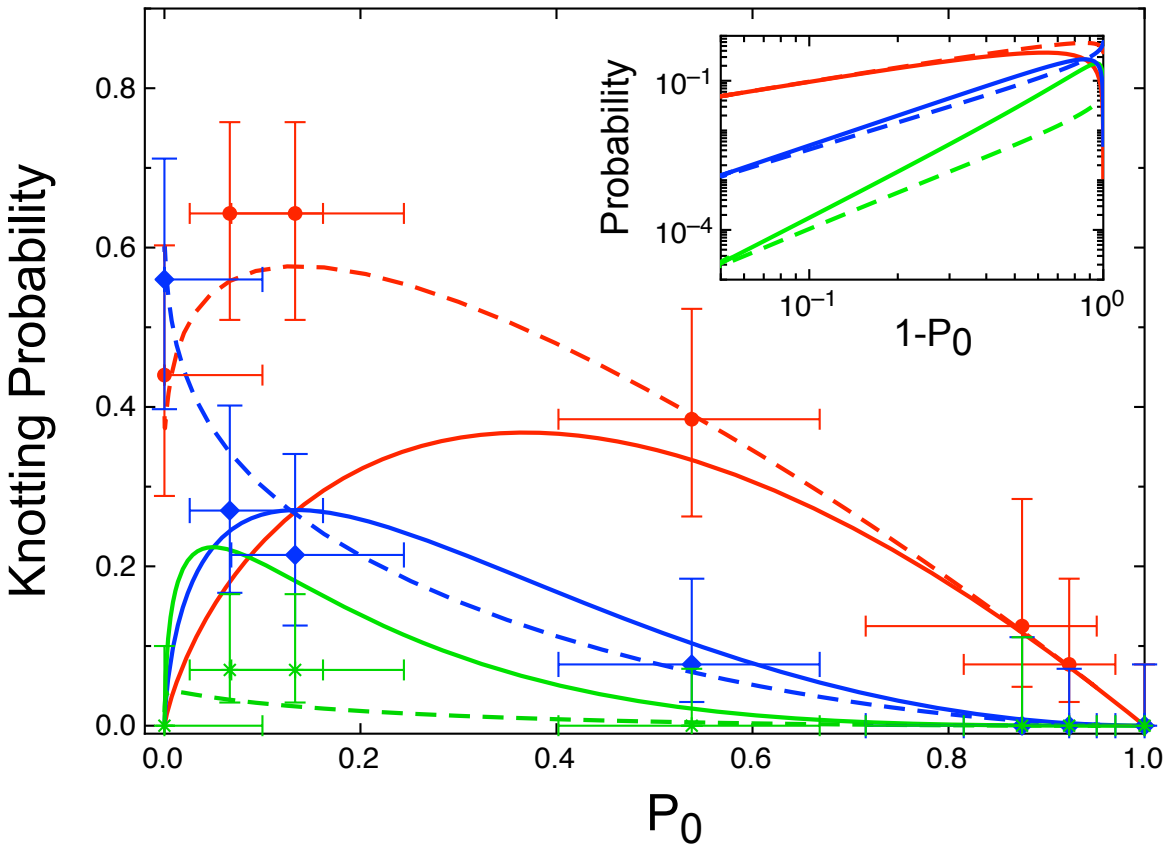

Figure 3–5: The red circles, blue diamonds and green stars respectively give measured probabilities for forming one, two and three knots. The solid lines indicate predictions of pure Poisson statistics (i.e. following from Eq. 3.3); the dashed lines indicate the predictions of the free energy model, with the red, blue and green curves respectively corresponding to $m = 1$, $m = 2$ and $m = 3$. The inset, which shows the same theory curves on a log-log scale versus $1 - P_0$, shows that the free energy model asymptotes to the Poisson description when the compression becomes very low and the profile approaches no-flow equilibrium.

that knot interactions at high compression are the likely explanation, due to the absence of composite knots states featuring many prime knots. This point can be made clear by quantifying the free energy of the knotting states.

Knot-formation is no longer kinetically limited at long-times where knot-formation probability asymptotes (Fig. 3–3(b)). In addition, the compressed chain is in a state

of inhomogeneous equilibrium. Fluctuations of the chain can be analyzed via a generalized free energy change that is equivalent to the minimum work required to drive the system out of the inhomogeneous equilibrium state [172,176,177]. This generalized free energy change includes the change in equilibrium free energy plus work performed by external forces [172], work which in our case arises from the viscous force exerted by hydrodynamic flow on the knots.

We can use our knotting probability measurements to make estimates of the free energy changes associated with knot formation. Let $F_{\rm tot}(m, R_{\rm b})$ be the total free energy change for forming a state with m knots on a profile with barrier extension $R_{\rm b}$ (in units of k_BT): $F_{\rm tot}(m, R_{\rm b}) = -\log Z(m, R_{\rm b})$ with $Z(m, R_{\rm b})$ the corresponding partition sum. The probability of forming m knots is then

$$P(m, R_{\rm b}) = Z(m, R_{\rm b}) / \sum_{i=0}^{n_k} Z(i, R_{\rm b}) \tag{3.4}$$

Note that $Z(0, R_{\rm b}) = 1$ as there is no free energy change for forming zero knots; n_k is the maximum number of knots observed to occur (we find that $n_k = 3$). Using Eq. 3.4, we can directly extract knot formation free energies from experiment using $F_{\rm tot}(m, R_{\rm b}) = -\log Z(m, R_{\rm b}) = -\log\left(P(m, R_{\rm b})/P(0, R_{\rm b})\right)$. In addition, we introduce an interaction free energy for two knots $F^{\rm int}_{2,\rm tot}$. The interaction free energy gives the increased free energy of the two knot state over the free energy of the two knot state satisfying pure Poisson statistics. For example, if the two knots obey Poisson statistics, their partition sum $Z_{\rm P}(2, R_{\rm b}) = Z(1, R_{\rm b})^2/2!$, leading to $F^{\rm int}_{2,\rm tot}(R_{\rm b}) \equiv F_{\rm tot}(2, R_{\rm b}) - 2F_{\rm tot}(1, R_{\rm b}) - \log(2)$.

Figure 3–4(b) gives the extracted free energies for forming a one knot state and knot interaction free energies as a function of $R_{\rm b}$. The single-knot free energy becomes increasingly negative for small $R_{\rm b}$. The interaction free energies are remarkably high (on order of several k_BT), suppressing multiple knot states. Knot-knot interactions,

for example, could arise through the excluded volume of one knot restricting the configuration space of the other knots ('knot-knot' excluded volume); this effect would scale as $k_B T g_k^3 / r D_1 D_2$ (with g_k knot gyration radius). Yet, we expect the volume of a single knot to be very small relative to the volume occupied by the chain: with $g_k \sim 100\,\mathrm{nm}$ [11] we find $g_k^3 / r D_1 D_2 \sim 10^{-2}$. The interactions must therefore have a more subtle origin.

3.3.3 Knot Spatial Distribution

Insight into the nature of the interactions can be gained by measuring the knot spatial distribution, which can be accessed a short-time following pressure release. Figure 3–6 shows the histogrammed position of knots for one-knot (Fig. 3–6(a)) and two-knot events (Fig. 3–6(b)), normalized to the chain extension r_{relax}, 2 sec after pressure release. For the two-knot events, the position of the knot closest to the slit-barrier ('lower-knot') and the knot farthest from the slit-barrier ('upper knot') are separately histogrammed. In addition, we show the cumulative histograms for the single (Fig. 3–6(c)) and two knot case (Fig. 3–6(d)). As the cumulative histograms are insensitive to binning, we perform all quantitative analysis on the cumulative histograms.

The single-knot distribution (Fig. 3–6(a,c)) is non-uniform and well described by an exponential probability density function (Fig. 3–6(c)), suggesting that knots are found preferentially in concentrated regions of the chain. While we do not expect the probability distributions after release to quantitatively mirror the distributions for a compressed chain (there could be considerable complexity in how the evolving chain profile during relaxation affects the knot distribution), we can say that the distribution observed after pressure release represents a lower-limit on the degree of spatial non-uniformity present in the knot distribution prior to release (the relaxation process

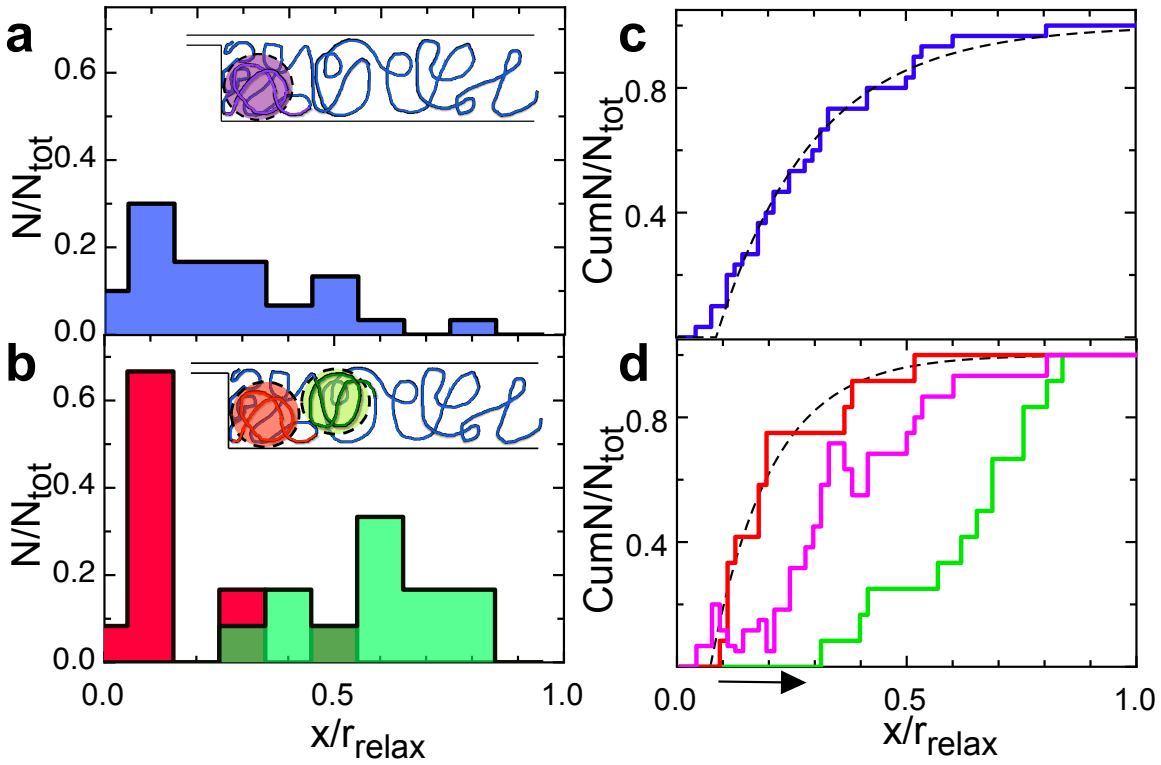

Figure 3–6: (**a**) Knot position histogram for one-knot states and (**b**) two-knot states 2 sec after pressure release. (**c**) Cumulative knot position histogram for one-knot states and (**d**) two-knot states normalized to the total number of counts. The x-axis is normalized to the extension $r_{\rm relax}$ of the relaxing molecule measured at the time for which the knot-position was obtained. The data used includes events with $R_{\rm b} \approx 0.11, 0.13, 0.17$. For the two-knot states, we separately histogram the position of the knot closest to the barrier ('lower knot', shown in red) and the knot farthest from the barrier ('upper knot', shown in green). The dashed curves in (**c**) and (**d**) are best-fits of respective data to the cumulative distribution corresponding to an exponential probability distribution. The arrow in (d) indicates that the upper knot distribution is shifted relative to the estimated non-interacting distribution (bold purple).

will smooth out an initially non-uniform distribution but it will not introduce non-uniformity).

The two-knot spatial distributions have structure indicative of knot-interactions. Note that the upper knot distribution is shifted to larger X relative to the lower knot distribution (as indicated by the arrow on the X-axis in Fig. 3–6(d)). If the two

knots do not interact (i.e. so that they are statistically independent) and both satisfy a distribution peaked near the barrier (as observed for single knots) we would expect that *both* of the knots would be found with high probability near the barrier. Instead, there is 'gap' in positions where only one knot is found. This gap could arise, for example, if the two knots interact like hard spheres over their gyration radius and satisfy a no-passing constraint, introducing a range in positions near the barrier where the second knot is physically excluded (see inset to Fig. 3–6(b)). Knot passing might be prohibited due to the large physical size of the knots, on order of the channel diameter [11, 165], prohibiting knot crossing mechanisms that require knot expansion [178].

To demonstrate rigorously that interactions exist, let us assume that the knots do not interact and are consequently statistically independent. Let the non-interacting lower knot distribution be $P_L(x)$ and the upper be $P_U(x)$. Moreover, let us assume that both knots follow an identical underlying probability distribution equal to the one knot distribution $P_1(X)$. These assumptions lead to (see section 4.4):

$$P_L(x) + P_U(x) = 2P_1(x) \tag{3.5}$$

Using the measured single knot and lower knot distributions in Fig. 3–6(c, d), from Eq. 4.29 we can infer the resulting upper knot distribution if the assumption of knot-independence (no-interactions) is valid. The result, shown in Fig. 3–6(d) (purple curve) differs from the shifted upper distribution (green curve), indicating that interactions are present.

3.3.4 Free Energy Model to Quantify Knot Formation

Our results suggest that knot-interactions are present at high compression, causing a pronounced deviation in observed knotting probability from Poisson statistics. The

interactions may arise from a hard-core repulsion mechanism preventing knot crossing in the channel. Yet, it is unclear how the hard-core repulsion translates into a higher free energy cost for formation of multiple knots and leads to a breakdown in Poisson statistics. Here we develop one possible model to quantify the effect of a no-crossing constraint on the knot free energies, elucidate the role of compression in increasing knot-formation probability and explain the breakdown in Poisson statistics at high compression. Qualitatively, our model suggests that knot free energy is lowered during chain compression by a novel excluded-volume mechanism: knots tightly localize the contour they contain, avoiding the free energy cost that would be introduced by releasing the stored contour to interact with the rest of the compressed, concentrated chain. Moreover, our model suggests that the free energy of a knot should vary with position along the compressed profile, with the free energy lowest at the barrier-edge. If a no-crossing constraint exists, multiple prime knots cannot all occupy the position of minimum free energy, but instead will stack single-file, leading to an increased free energy of a composite knot state relative to the free energy of the independently formed prime knots.

Dai *et al.* argue three types of free energy contribute to knot-formation [11]: the energy of forming a knot on the chain in bulk $f_{\mathrm{b}}(L_k)$, where L_k denotes the contour length of the knot; the energy of confining the knot between the channel walls $f_{\mathrm{wk}}(L_k, R)$ ('wall knot') and the energy *saved*, as contour stored in the knot no longer contributes the confinement free energy $f_{\mathrm{wuk}}(L_k, R)$ ('wall unknot') associated with an unknotted section of polymer of contour L_k. In addition, the flow exerts a constant drag force $\zeta_k V$, leading to a free energy contribution $f_h \equiv \zeta_k V X/k_B T$ (in units of $k_B T$ with ζ_k a knot friction factor). The total free energy change $f_1(L_k, R_{\mathrm{b}}, X)$ upon forming one knot of contour L_k in the nanochannel at position X on a profile with

barrier extension R_b is then

$$f_1 = A_b f_\mathrm{b} + A_{wk} f_\mathrm{wk} - f_\mathrm{wuk} + A_h f_h. \tag{3.6}$$

This equation makes explicit three dimensionless scaling constants A_b, A_{wk} and A_h to be determined via least squares fitting. The term f_b is the free energy of trefoil knots in bulk, an R-independent contribution that can be obtained from a balance of bending and confinement free energies [153, 154] (See section 4.5 for explicit functional form). We expect trefoil knots to dominate [166]. The contribution f_wuk is obtained from the Flory free energy of a confined chain [137]:

$$f_\mathrm{wuk} = A_\mathrm{wuk}(L_k/L_b)\left(R^2/2 + 1/R\right). \tag{3.7}$$

This contribution varies with position X and R_b through $R(X) = 1/C(X)$ (Eq. 3.1). The quantity L_b is the contour per blob in the equilibrium chain, which we obtain from extended de Gennes theory: $L_b = (2P/w^2)^{1/3} D_\mathrm{eff}^{4/3}$ with P the persistence length and w the effective width (we estimate $P = 55\,\mathrm{nm}$ and $w = 17\,\mathrm{nm}$ (Section 4.1). The quantity $D_\mathrm{eff} = \sqrt{(D_1 - w)(D_2 - w)}$ is an effective diameter taking into account the non-unity aspect ratio of the channel and wall-DNA interactions (estimating the electrostatic depletion near the channel walls by w [126]). The quantity A_wuk is a numerical constant ($A_\mathrm{wuk} = 2.81$ [137]). As the knot formation probability is only appreciable below $R = 0.5$, the R^2 term in Eq. 4.35 is negligible. Dai *et al.* show that the channel-confined knots can be approximated as spherical regions with a radius of gyration $g_k = 0.1 L_k$. Following Dai *et al.* [11], we argue that f_wk is given by the confinement free energy of one knot unit on a chain of identical such units,

$$f_\mathrm{wk} = L_k g_k^{2/3} D_\mathrm{eff}'^{-5/3} R^{-5/4}. \tag{3.8}$$

The quantity $D'_{\text{eff}} = \sqrt{(D_1 - \delta)(D_2 - \delta)}$, where δ takes into account the increase in effective confinement due to finite knot size. We use $\delta = \gamma g_k \sim L_k$ (Dai *et al.* find that $\gamma = 1$ describes their simulation results well). The R-dependent factor accounts for the increase in the free energy of a flexible chain with compression for $R < 0.5$ [137]. Lastly, we estimate ζ_k as the friction factor of a sphere of radius $\zeta_k = 6\pi\eta g_k$ with η viscosity ($\sim 1\,\text{mPs}$).

Our model suggests that knot stabilization is driven by a large *negative* single knot free energy at the slit barrier. Combining Eq. 3.1-3.8, we find f_1 varies linearly with position (Fig. 3–7(b)):

$$f_1(L_k, R_\text{b}, X) = f_1(L_k, R_\text{b}, 0) + \beta(L_k, R_\text{b})X \tag{3.9}$$

(see section 4.6 for explicit form of β and $f_1(L_k, R_\text{b}, 0)$; the $1/R^{5/4}$ term in Eq. 3.8 can be linearized by Taylor expansion for small X where knotting probability is appreciable). If f_{wuk} is large enough to ensure the free energy at the barrier $f_1(L_k, R_\text{b}, 0) < 0$, knots of size L_k will exist with a spatial distribution $P_1(X) \sim \exp(-f_1(L_k, R_\text{b}, X))$ leading to an exponential accumulation near the barrier, consistent with the observed single knot distribution (Section 4.7, Fig. 3–6(a, c) and Fig. 3–7(d)).

The barrier ($x = 0$), where the free energy is minimized, is the most probable location for a single knot to form, but knots can form at all x. Let the partition function $z_1(L_k, R_\text{b})$ count the number of ways a single knot of size L_k can form on a profile characterized by R_b. The number of statistically independent sites at which a knot can form is estimated by $n_{\max} = r/2g_k$, each site i weighted by a Boltzmann factor $\exp\left(-f_1(L_k, R_\text{b}, i2g_k/r_o)\right)$, leading to a partition function that can be summed geometrically (see section 4.8, we can approximate the summation limit as infinity as

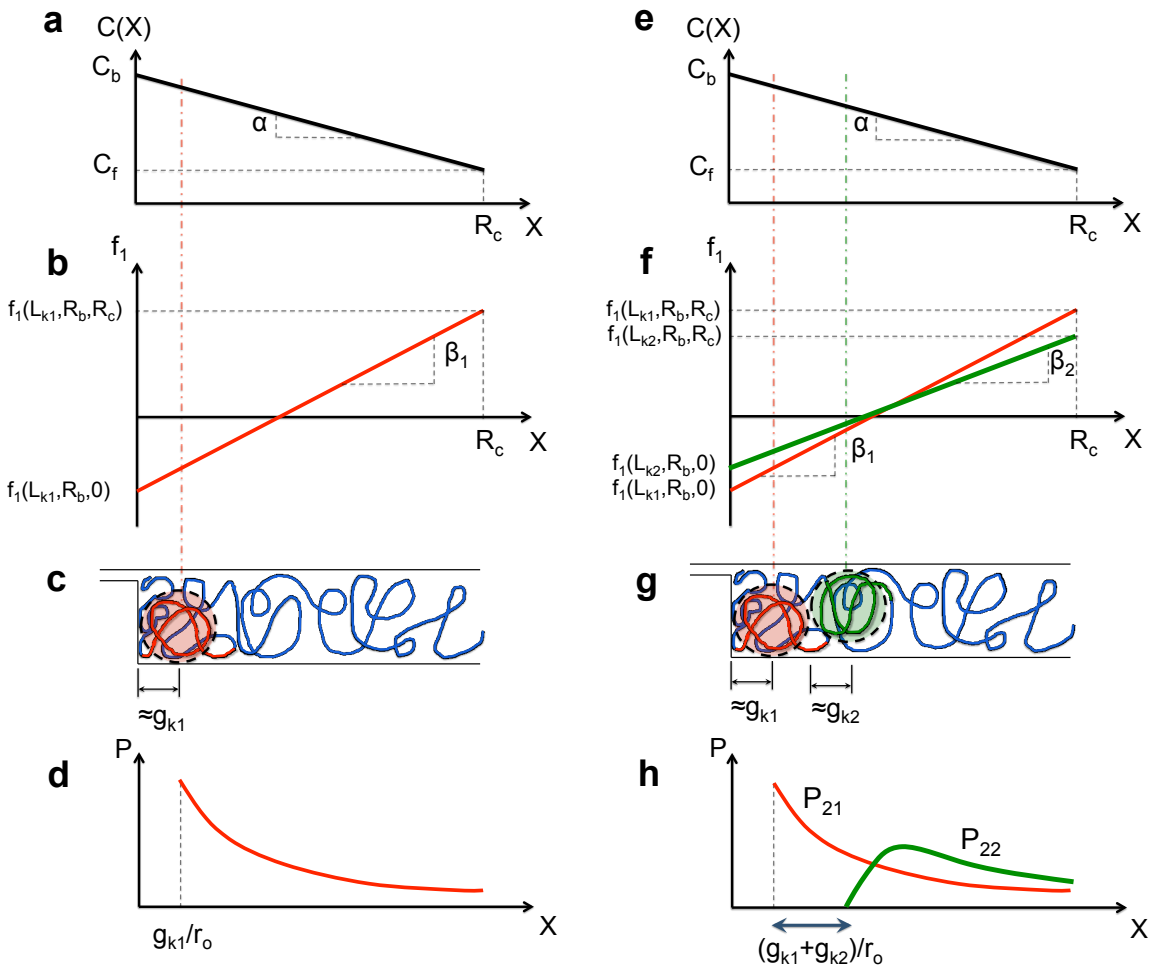

Figure 3–7: (**a**, **e**) Schematic of concentration profile $C(X)$, (**b**, **f**) resulting free energy profile $f_1(L_k, R_\mathrm{b}, X)$ and (**d**, **h**) resulting knot probability distributions for single (**a**-**d**) and two-knot (**e**-**h**) states. Note that in the two-knot state the two knots can experience different free energy profiles due to their varying size (e.g. L_{k1} does not necessarily equal L_{k2}): this is indicated in (**f**) by drawing two free energy profiles with slopes β_1 and β_2. In addition, in the two-knot state the knots will have different probability distributions: P_{21} for the 'lower' knot closest to the barrier and P_{22} for the 'upper' knot farthest from the barrier (see section 4.6 for detailed derivation of the probabilities) (**h**). (**c**) Cartoon of single-knot on chain, displaced by radius of gyration g_{k1} from barrier edge. The dashed line relates the knot position to the corresponding concentration and free energy. (**g**) Cartoon of two knots on chain. Knot 1 (red) is displaced by radius of gyration g_{k1} from barrier edge; knot 2 (green), by assumption of single-file ordering, is displaced by $2g_{k1} + g_{k2}$ from barrier edge.

the probability of finding knots far from the barrier is very small). To compute the partition function of a state containing multiple (m) prime knots $z_m(L_{k1} \cdots L_{km}, R_\mathrm{b})$ we must include knot interactions. Interactions equivalent to a no crossing constraint are incorporated by requiring that the summation respect knot ordering (i.e. so that the knots maintain single file). At high compression, these partition functions contain only one state, a 'ground state' configuration consisting of knots stacked single file, with no gaps, directly abutting the barrier (e.g. see Fig. 3–7(c) for one knot ground state; see Fig. 3–7(g) for two knot ground state). In this high compression limit Poisson statistics does not hold as only one state is accessible and strong interactions imply knots do not form independently. For low compression, knots can be excited away from the barrier. Interactions are weak as the knots are well separated and many states are accessible, leading to an emergence of Poisson statistics (see section 4.9).

Lastly, we must integrate over all knot sizes, forming a partition function:

$$Z(m, R_\mathrm{b}) = (2P)^{-m} \int z_m(L_{k1} \cdots L_{km}, R_\mathrm{b}) dL_{k1} \cdots dL_{ki} \tag{3.10}$$

In practice, we obtain $Z(i, R_\mathrm{b})$ from direct numerical integration. Equations 3.4, 4.33-3.8 and 3.10 then enable determination of knotting probabilities as functions of R_b. Simultaneous least-square fitting of model predictions to the experimental one and two-knot formation probabilities (Fig. 3–4(a)) yields $A_\mathrm{b} = 1.43 \pm 0.05$, $A_\mathrm{wk} = 0.98 \pm 0.12$ and $A_h = 1.12 \pm 0.07$, on order of unity suggesting that the approach is self-consistent. Equivalently, we can fix $A_\mathrm{wk} = 1$ and $A_h = 1$ and perform a one-parameter fit of the parameter A_b, which yields equivalent results (see section 4.10). Our theory captures the increasing knot formation probability with increasing compression, the non-monotonic behavior of the single knot formation probability (Fig 3–4(a)) and the energy scales of knot interactions and single knot stabilization (Fig 3–4(b)). Our

model also predicts the very small number of observed three-knot events (Fig. 3–4(a, inset)), a consequence of the large interactions. Finally, our model quantitatively captures the transition from Poisson statistics at weak compression to the non-Poisson regime at high compression (Fig. 3–5). The A_{b} value required to get agreement with experiment is slightly higher than unity, possibly arising from physical effects, such as knot compression, that lead to higher knot free energy and are not included in the model.

In detail, note that the non-monotonic variation of single knot probability with compression in our model arises due to the interplay of the increasingly negative single knot formation free energy and the normalization condition (Fig 3–4(b)). For moderate R_{b}, the single knot free energy falls and the single knot probability rises. As the single knot probability approaches unity, the probability of two-knots states also necessarily rises and the probability of one knot states falls to preserve normalization of total probability. Strong interactions suppress the probability of finding three-knot states.

3.4 Discussion

In conclusion, we show that hydrodynamic compression induces DNA knotting in nanochannels with high probability. This is remarkable as it demonstrates that moderate confinement, two orders of magnitude weaker than that found in capsids, can also induce knot formation, suggesting a knot formation mechanism qualitatively different from what has been proposed in [179,180] for capsids, where nematic ordering in strong spherical confinement can form toroidal knots with high probability. We show that the free energy scales for knotting under compression in the long-time limit

can be estimated by extending known free energy scales for confined knots in a no-flow equilibrium. In addition, we find that knot-interactions likely exist, arising from a hard-core repulsion between knots preventing knot crossing and lead to single-file ordering of knots. Our model suggests knot interactions suppress multi-knot states and lead to a pronounced deviation from Poisson statistics expected for knot formation in a no-flow equilibrium limit.

In a recent study, Tang *et al.* [18] introduced a technique for inducing knots on DNA molecules via application of an AC electric field. From a practical point-of-view, our approach has the advantage that it is inherently parallel; many molecules can be simultaneously compressed in an array of nanochannels and their relaxed, nanochannel-extended states analyzed. From a physical point of view, the approach of Tang *et al.* occurs in a much more complex, strongly non-equilibrium environment, with both solvent and DNA exhibiting complex dynamics resulting from the hydrodynamic instabilities induced by the DEP force. In particular, in the DEP approach the DNA tumbling dynamics leads to finite segmental current throughout the coil. In addition, the DNA is driven into a globule state, which is less well understood due to the complexity of the DEP-induced attractive interactions that drive the compression. In our system, based on geometric confinement and pressure-driven low Reynolds number flow, the DNA adopts a highly reproducible concentration profile that corresponds rigorously to an inhomogeneous equilibrium state with zero segmental current. The simplicity of the underlying DNA conformation in our system may facilitate modeling of knot generation processes. The transverse confinement in our system provides an additional parameter that can be used to tune knot-formation, with lower channel diameter in the extended de Gennes regime predicted to produce knots with greater probability

at an equivalent degree of compression (see section4.11). In addition, the channel diameter likely sets an upper limit on knot size for channels below about 500 nm. Our approach may thus lead to composite knot states formed from smaller stacked prime knots distributed towards one molecule edge. In contrast, we speculate the approach of Tang *et al.* may lead to easier production of larger, topologically complex knots in the molecule center [46].

A complete understanding of knot formation in our system requires understanding the physics behind the lowered topological barriers leading to favorable kinetics at experimentally accessible time-scales. We do find that knotting probability rises with waiting time in the compressed state, with a kinetically limited regime at low waiting times. This appears to confirm the picture suggested by [4] regarding the dependence of knotting probability on effective 'agitation time.' In our microscopic experiment, for example, thermal fluctuations could supply the necessary agitation, or thermal fluctuations could be assisted by additional hydrodynamic effects. In their DEP-based compression experiments Tang *et al.* hypothesize that a tumbling-like agitation is created by electric-field induced hydrodynamic instabilities.

In particular, lowering the topological barriers for knot-formation requires a mechanism for knot-ends to invade the main coil so that the chain ends can be threaded through internal loops in the chain [166]. In our experiment, one possibility is that the chain ends are forced in during the transient compression ("shock-wave") process, although this does not explain the long observed waiting time. A second possibility is that subtleties of the steady-state hydrodynamic flow, perhaps curving streamlines near the slit barrier, might play a role in helping drive the chain ends into the coil. We feel, however, that this mechanism would need to be more subtle than the flow-induced tumbling described in [18], as we expect the flow in our nanofluidic channels

to be steady and laminar, leading to a static packing of DNA against the barrier rather than continuous recirculation or agitation. We do not apply an electric field, and we expect effects of electrohydrodynamic coupling to be very weak, so there is no clear candidate for a physical effect that could create the recirculatory flow necessary to drive DNA tumbling. A third possibility is that thermal fluctuations alone are sufficient to drive the chain ends into the coil. At high compression linear ordering of blobs breaks down and the free energy barriers preventing long-range chain looping disappear. [137]. Brownian dynamics simulations of our system [181] would help clarify which mechanism is correct. Yet, whatever physics drives the favorable kinetics, once we deduce that the kinetics *are* favorable by observing time-dependent saturation of knotting, our free energy approach is valuable as it enables extraction of long-time knotting probabilities in a systematic way from knowledge of equilibrium behavior.

In the future the knot factory could be further exploited to study the effect of channel width, ionic strength, DNA contour length and to generate knots for further dynamic studies. In particular, we expect the physics of knot formation to be very different in the transition ($D < 100\,\mathrm{nm}$) and Odijk confinement regimes ($D < 50\,\mathrm{nm}$) due to the qualitatively distinct underlying chain conformation in these regimes [126]. While we predict knotting probability increases with decreased channel width throughout the extended de Gennes regime, the situation for smaller channel width is unclear and a fascinating question for further theoretical and experimental study.

CHAPTER 4
The Nanofluidic Knot Factory: Detailed Methods, Derivations and Arguments

In chapter 3, a knot factory for efficient formation of knots on nanochannel-confined DNA molecules was introduced. A free energy model for describing knot formation as well as quantifying the effect of knot interactions was provided and knot spatial distributions on DNA molecules were studied. This chapter contains the supplementary information to chapter 3. A review of past relevant work is provided; the device details and error analysis methods are discussed; A free-standing discussion of the spatial distribution of knots in both one and two knot states is introduced and detailed derivation of knot formation rate equations and knot interaction free energy and calculation of partition sums for assessment of knotting probabilities are provided. Finally, discussions with the referees during the review process are embedded in the text.

4.1 Experimental Methods

In order to obtain the shape and dimensions of the nanochannel cross section, a chip is cut carefully at the center where the nanochannels are located: the sample is then diced half-way through using a wafer saw and then broken by hand about the saw cut to avoid destroying the nanochannels with the saw blade. A platinum layer of 4-nm thickness is sputtered on the cleaved sample. Upon imaging the chip using SEM, the cross-section shape is observed to be trapezoidal with base dimensions $D_{\mathrm{b1}} = 250\,\mathrm{nm}$ and $D_{\mathrm{b2}} = 386\,\mathrm{nm}$ (Fig. 4–1). Adding the slit depth (measured with surface profilometer) to the measured depth of the unbonded channel gives $D_2 \sim 415\,\mathrm{nm}$ (Fig.

4–1(inset)). Note that the trapezoidal area $A = D_2(D_{b1} + D_{b2})/2 \equiv D_2 D_1$ with the dimension $D_1 \equiv (D_{b1} + D_{b2})/2 = 325\,\text{nm}$ (i.e. the average of the two base lengths). In other words, a rectangular cross-section of dimension D_1 and D_2, having the same area as the original channel, is the best rectangular approximation to our trapezoidal cross-section. The use of a non-unity aspect ratio requires that the channel width D must be replaced by the geometric average $D_{av} = \sqrt{D_1 D_2}$ =367 nm [126].

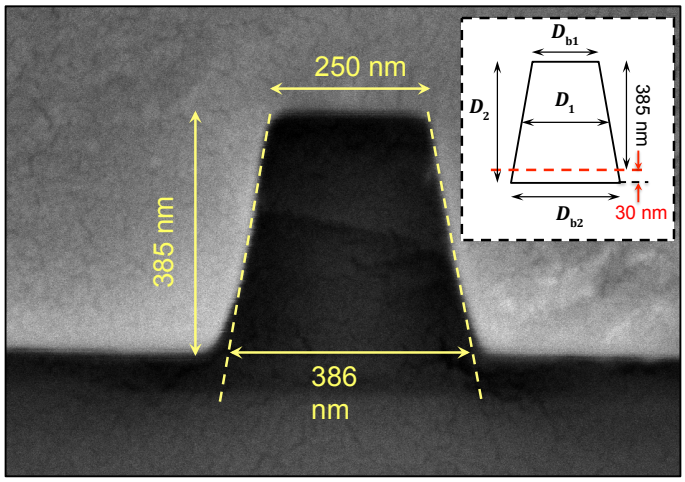

Figure 4–1: Cross section of a nanochannel imaged using SEM; the cross section is a trapezoid with base dimensions D_{b1}, D_{b2} and height dimension D_2. The yellow dashed lines show the trapezoid edges. The inset shows a schematic of how the dimensions of the bonded chip are calculated. Adding the slit depth (measured with surface profilometer and shown in red) to the measured depth of the unbonded channel gives $D_2 \sim 415\,\text{nm}$. The length $D_1 = 325\,\text{nm}$ is the average of the base-lengths $D_{b1} = 250\,\text{nm}$ and $D_{b2} = 386\,\text{nm}$. The channel cross section is approximated as a rectangle with dimensions D_1 and D_2.

We briefly discuss our approach for estimating the physical value of P and the effective width w for our buffer conditions. Our buffer has an estimated ionic strength of $8\,mM$, a value we obtain from solving the coupled chemical equilibria for Tris-HCl and BME ($\text{pK}_{\text{tris}} = 8.1$ and $\text{pK}_{\text{BME}} = 9.6$). The persistence length is estimated using

Odijk-Skolnick-Fixman (OSF) theory [144, 145]:

$$P = P_o + \frac{32\,\mathrm{mM}}{I_s}\,\mathrm{nm} \tag{4.1}$$

with P_o the high salt persistence length ($P_o = 51\,\mathrm{nm}$). We find $P = 55\,\mathrm{nm}$ at $I_s = 8\dot{\mathrm{m}}\mathrm{M}$. The Stigter theory, which estimates the effective width of DNA based on the excluded-volume of charged rods, is used to evaluate the effective w [145, 146]. We find $w = 17\,\mathrm{nm}$. The high value of the effective width suppresses bulk knotting [41]; the only knots observed are those formed through compression.

4.1.1 Knot Identification Criteria

Using an optical trap to tie knots on polymers can create knots with a topology that is precisely known. The method employed in the current study and some previous ones [13, 18], on the other hand, don't produce knots of known topology. While we cannot directly trace the chain path, there exist criteria that make "knots" on polymers distinguishable from "folds" or "trivial knots" (including tangled segments and complex unknots that do not possess true knot topology [9, 13, 18]). These criteria are motivated by A. Grosberg's tight knot theory [153], now well-established by simulation [11, 154] and the experimentally observed behavior of objects with well-defined knot structure created via tweezers [9]. In Grosberg's theory of tight-knot structure on semiflexible chains, a knot's topological structure creates an effective network of non-crossing constraints that is equivalent to confining the chain in an effective tube. The knot free energy arises from two sources: (1) bending of the tube and (2) confinement free energy arising from polymer constrained in the tube. If the knot is small the bending energy is high as the tube is squeezed tight. Large knots have high confinement free energy as more polymer contour, pulled from less constrained regions of the chain, is forced to lie confined within the effective tube. The balance between bending

and confinement free energy leads to a metastable knot size, giving rise to knots possessing a soliton-like structure [153] with a stable shape that diffuse along the chain through self-reptation of contour through the effective tube. In particular, in the Grosberg theory the metastable knot shape is stabilized by a high free energy barrier for introducing contour into the knot that prevents spontaneous knot loosening on experimentally accessible time scales (this barrier is estimated to be $\sim k_B T$ [153]). Knots tied by tweezers in fluorescently labeled chains, such as in [9], will adopt a bright (highly concentrated), localized (sub-diffraction) and stable structure that diffuses along the chain and *can only unravel at the chain ends*, consistent with the predictions of the Grosberg theory.

An entanglement in the contour is called a "knot" if it cannot be untied when the string is closed [4, 21]. In contrast, "trivial knots" or "unknots," such as the "bunched up" or "tangled" regions can unravel mid-chain. *Unknots* can be simple (in the form of $\mathcal{S}$ or φ also known as "folds") or complex, formed via concatenation of folds ("complex unknots"). Unknots, unlike knots, are not self-confined, thus they fluctuate in size significantly [13] and are eventually removed by thermal fluctuations. In particular, in a nanochannel, a concentrated region of DNA will have a higher free energy than the surrounding chain (due for example to self-exclusion interactions). Entropic forces will exist that will drive transport of contour from this concentrated region to non-concentrated regions, leading to large-scale unraveling dynamics (the simplest example is a confined polymer with a single-fold at the chain ends, which unravels via a simple kinetics described in [168]). If the structure possess true knot topology, this process will be halted when the object reaches metastable knot-size, but unknots will continue unraveling until all contour is removed. Note that, for unknot structures, there is no built-in topology preventing the opening of tight loops by thermal fluctuations that

rotate adjacent strands forming the loop. We thus expect small unknot structures to be highly unstable at the molecular level [4]; frictional effects do not play a role in stabilizing microscopic knots (unlike macroscopic knots), and these small unknot structures will possess high bending energy. Large, highly entangled structures may lead to arrested relaxation and take longer to unravel, as observed in [18], but our view is that only structures that posses a true knot topology will unravel to a final state possessing a diffraction limited "spot-like" structure.

We summarize our knot identification criteria as follows:

- Knots are persistent, localized and bright features on extended chains.
- Knots do not exhibit large-scale size fluctuations after reaching their final (metastable) state.
- Knots unravel only at the molecule ends.

The above knot identification criteria have been used in similar experimental studies [13, 18] and have been followed rigorously in the current study. Note that in the current study an event with multiple knots is reported *only* when the knots diffuse apart so they can be separately identified. Figure 3–1(i, j) gives an example of a two-knot and three-knot event where knot diffusion separates the knots on the chain so they can be distinguished.

4.2 Single-Chain Concentration Profiles: Data Analysis and Theoretical Models

4.2.1 Waiting time

Waiting time is defined as the duration for which a molecule is held in a compressed state. Each molecule is compressed to a certain minimum extension for a certain waiting time and then relaxed. To create a consistent measurement of the waiting

time, we first find the average extension of the molecule at equilibrium at compressed state r and the standard deviation σ_r in the extension. Starting from the first frame of the movie, once the extension of the molecule reaches a value within the interval $[r - \sigma_r, r + \sigma_r]$, we set that frame as $t = 0$. The point of pressure release determines the waiting time $t = t_w$.

4.2.2 Intensity Profile Fitting Functions

A nanochannel-extended molecule in its no-flow equilibrium conformation has a uniform concentration profile. Upon hydrodynamic compression against the slit barriers, the DNA molecules will evolve towards a steady-state concentration profile that has a ramp-like character (see Fig. 3–1(g-j)). We find that this profile is well-described by a linear ramp form (Eq. 3.1 and Fig. 3–1(k)). An experimental difficulty is that we cannot resolve the exact theoretical concentration profile $c(x)$ as our optics are diffraction limited. Instead, we image broadened profiles formed by convolving the theoretical profile with a point-spread function $f_{\mathrm{PSF}}(x)$ [151, 152]. A second difficulty is that we measure fluorescence intensity $I(x)$, not concentration. To deduce the broadened experimental profiles from a given predicted profile $c(x)$, we use:

$$I(x) = C \int_{-\infty}^{\infty} c(y) f_{\mathrm{PSF}}(x - y)\, dy. \tag{4.2}$$

For simplicity, we choose a Gaussian form for $f_{\mathrm{PSF}}(x)$:

$$f_{\mathrm{PSF}}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right). \tag{4.3}$$

The PSF width σ is estimated by fitting to the profiles. The constant C in Eq. 4.2 is a calibration factor representing the conversion between concentration and intensity level (depending on factors such as stain ratio, illumination power-levels, stain quantum

efficiency and the stain extinction coefficient). By using $I/I_o = c/c_o$, we remove the calibration factor from consideration.

We use two functional forms for the underlying concentration profiles $c(x)$. The uniform-profile shape, appropriate for an equilibrium no-flow profile, is defined by:

$$c_{\text{flat}}(x) = \begin{cases} c_o & \text{if } x > x_1 \text{ and } x < x_2 \\ 0 & \text{otherwise} \end{cases} \tag{4.4}$$

The extension is simply the difference between the profile-end points: $r = x_2 - x_1$. The ramp profile shape is defined by:

$$c_{\text{ramp}}(x) = \begin{cases} c_{\text{b}} - A_r x & \text{if } x > x_1 \text{ and } x < x_2 \\ 0 & \text{otherwise} \end{cases} \tag{4.5}$$

The quantity A_r is the profile ramp-rate. By applying Eq. 4.2 and Eq. 4.3 to Eq. 4.4 and Eq. 4.5, we can deduce analytic forms (expressed in terms of error-functions and Gaussians) for the broadened profiles. All profile parameters, such as extension, ramp-rate, maximum/minimum concentration and σ can be obtained by fitting experimental profiles to these broadened functional forms (see Fig. 3–1(k) for an example fit to experiment of a broadened ramp along with an estimate of the underlying ramp-profile).

4.2.3 DNA Equilibrium Conformation at Compressed State

Experimentally, we observe that once the transient compression phase is over, the compressed DNA undergoes thermal fluctuations about a well-defined profile shape. In particular, Fig. 4–2 shows the normalized extension (Fig. 4–2(a,b)) and ramp-slope (Fig. 4–2(c,d)) versus time of DNA molecules compressed via two different flow velocities (for the steady-state portion of the compression, i.e following completion of the transient). Evidently, the molecule extension in the compressed state and the

profile slope appear to fluctuate around a fixed average. We see no evidence of a slow-relaxation over the time-scales of the experiment (i.e. the data in both cases is flat).

The fact that we reach a steady-state does not mean that the steady-state is an equilibrium state. Our system is maintained in a *zero-current steady state*: this is equivalent to a state of *inhomogeneous equilibrium*, or *force-constrained equilibrium* where local forces everywhere are in balance. There exists a local mechanical equilibrium between the osmotic pressure gradient and applied hydrodynamic force, *leading to zero segmental current J everywhere along the profile*, or a zero segmental drift velocity (by the word "segment" we refer to Kuhn segments making up the DNA). The key points of agreement are that we observe: (1) a linear ramp and (2) quantitative scalings for the compressed extension and ramp slope as a function of V that are consistent with equilibrium predictions (see Fig. 3–2(g-j)). Note that when we use the word equilibrium, we are referring to the DNA molecule alone. The solvent of course is not in equilibrium as the solvent is undergoing steady ultra low Reynolds number flow, constantly penetrating the molecule, but the effect of the steady-solvent flow on the DNA is simply to supply a source of static forcing.

As has been rigorously demonstrated for a broad class of systems, including systems described by underlying hydrodynamic/diffusive equations (such as the nonlinear convective-diffusion formalism we use to describe the compressed DNA) [172,182], fluctuations from such steady-states can be analyzed via a generalized free energy change that is equivalent to the minimum work required to drive the system from the state of inhomogeneous equilibrium [176,177]. This generalized free energy includes the change in equilibrium free energy (work in absence of external forcing) plus work performed by external forces.

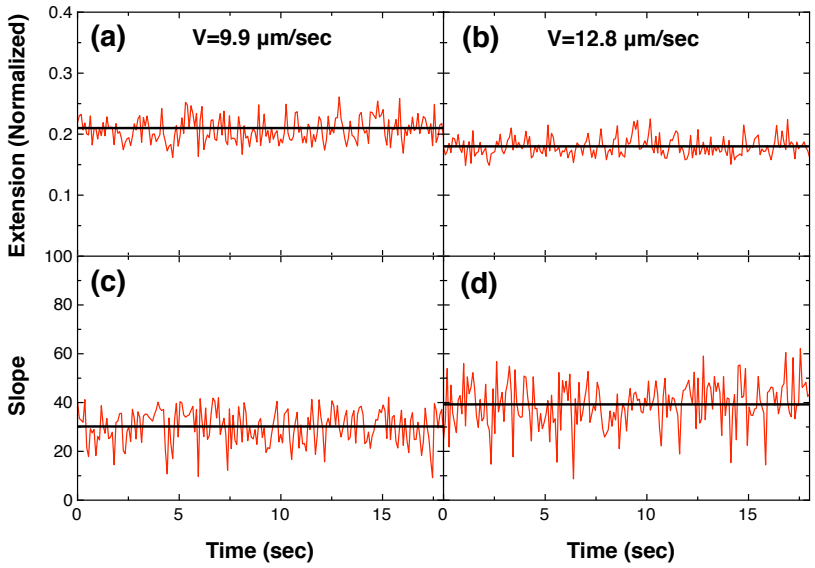

Figure 4–2: Measurements of normalized extension and slope of intensity profiles at compressed state for two different flow velocities $V = 9.9\mu\ m/sec$ (left) and $V = 12.8\mu\, m/sec$ (right). (a,b) Extension of DNA molecules at compressed state normalized to the initial extension of the molecules. The solid black lines depict the average extensions $r_c = 0.21 \pm 0.02$ (a) and $r_c = 0.18 \pm 0.01$ (b). (c,d) The slope of intensity profiles at compressed state. The solid black lines show the average slopes $\alpha = 30.2 \pm 3.7$ (c) and $\alpha = 39.25 \pm 4.6$ (d).

While flow is constantly penetrating the DNA coil and passing through the slit (in this sense the system is out-of-equilibrium), this flow is steady and necessarily laminar due to the ultra low Reynolds number of the nanochannels. Such a flow-field creates a source of *static external forcing* on the DNA, driving our chain against the barrier. Our problem is thus analogous to a chain pushed against a barrier by a static force, and bears resemblance to other problems in soft-matter where a system is driven by a force against a hard-wall, such as sedimentation, or centrifugation. Critically, the long-time-limit of such a problem corresponds to a special type of steady-state where the current J (that is current of DNA Kuhn segments in our case) is zero. In these systems the current vanishes even in the presence of a non-uniform concentration profile.

A classic example would be the barometric distribution of concentration in a constant gravitational field. In general, the vanishing of the current gives rise to an ODE that describes the steady-state profile. In the context of the nonlinear convective-diffusion formalism we use to describe the compressed DNA, the zero-current condition gives rise to a linear ramp. Moreover, our experimental evidence suggests that physical effects that would violate a simple zero current condition (e.g. recirculatory flows) do not exist, or at least do not occur at scales that we can resolve.

Nematic ordering of compressed DNA molecules

Some studies [179, 180] have shown that in order for knots to form, confined DNA needs to be stiffened by nematic ordering of DNA segments, similar to the knotting process in capsids. Such mechanism requires very high level of confinement. In our system, on the other hand, the concentration achieved is not high enough to induce nematic ordering. Our observations show that knots form at much weaker confinement than is observed in capsids. In particular, our view is that the mechanism described by Reith *et al* applies directly only to strong confinement in a spherical cavity and cannot necessarily be extrapolated to weaker confinements regimes with different geometries (e.g. the nanochannel or tube-like regime considered here). Reith *et al* argue that in order to produce knots with high probability, in situations of extreme confinement like viral capsids, loops must be formed that are sufficiently large for the free end to pass through. Formation of large loops is enhanced by making the polymer stiffer because, while flexible chains will fill the sphere, stiffer polymers tend to move towards the periphery of the sphere, maximizing the loop diameter. In particular, this configuration tends to produce torus knots. However, it is unclear whether this mechanism applies if the confinement is weaker. For example, Reith *et al* and Virnau et al. suggest that

the probability of knotting in bulk is more complicated, with stiffer polymers forming knots with lower probability above a critical stiffness value, and below the critical stiffness forming knots with higher probability (DNA is above the stiffness threshold, according to the authors). The current study is intermediate between the bulk an strong confinement, so it is unclear which type of argument is appropriate. Lastly, the geometry should definitely play an important role: in a long tube, such as used in our study, loops can form along the channel axis, so large loops could form even if the polymer does not follow the periphery of the tube. As noted above, even at the highest compression created, the extension is still an order of magnitude larger than the channel diameter.

The Tumbling Effect

In accordance with the assumption that knots form in the proposed knot factory at equilibrium and that knot formation can be explained via a free energy model, one might wonder that other effects such as hydrodynamic-flow induced rotational tumbling of DNA segments have led to the formation of knots. Tang *et al.* [18] argue that in their proposed knotting system, the presence of hydrodynamic instabilities, created by their applied AC-field, leads to *local* rotational motion (vortices). This local rotation leads to the DNA being continually "jostled" by the flow, resulting in knots. Such an instability-based mechanism cannot exist in our experiment. In nanofluidic devices, the Reynolds number is orders of magnitude below unity (in our device, we find that $\mathrm{Re} \sim 10^{-8}$). Such ultra low Reynolds number flows cannot create motion that in anyway can be described as "jostling" or "agitation". In the absence of any "special conditions," such as the high AC-fields applied by Tang *et al*, the stream lines must be perfectly laminar with no turbulence or vortices or any other time-dependent local

flows. To make this point precise, some recent studies have explored how to create nano vortices in micro and nanochannels for efficient mixing of different fluids. One proposed approach is consecutive contraction and expansion in the cross section of channels [171]. In this study, with a geometry very similar to what is used by us, the authors find that vortices only form above a critical $\mathrm{Re} = 0.055$. We are six orders of magnitude below this threshold!

Now, what about the hydrodynamic transition zone near the slit barrier? While for most of the extent of the compressed molecule the flow-lines will be perfectly straight and parallel to the channel axis, there will be a small region on order of one channel diameter from the slit barrier (Fig. 4–3 (top)), where the flow will deviate from perfect alignment with the channel axis and bend toward the slit. If the polymer was quite small, with a coil size on order of the channel diameter like the DEP condensed globules of Tang *et al*–so that it could fit completely within the transition region–we feel some rotation could conceivably occur. However, our molecules–even when fully compressed–have an extension along the channel that is ten times the size of this transition region (e.g. the extension of DNA molecules at the most compressed would be $\sim 2000\,nm$). Thus, it is not plausible that bending flow over such a small transition region could cause our molecule to tumble as a whole. To make this point precise, the polymer cannot rotate about the axis longitudinal to the channel: there are no net forces present acting tangent to this axis that could lead to rotation. The polymer cannot rotate about an axis parallel to vertical dimension, because again there are no net forces acting tangent to this axis and the molecule is constrained by the channel. Can the polymer rotate about an axis perpendicular to the vertical and longitudinal dimension? We feel this could happen if the polymer coil was smaller than the channel width (e.g. so it would be like a ball that tumble around in the transition region),

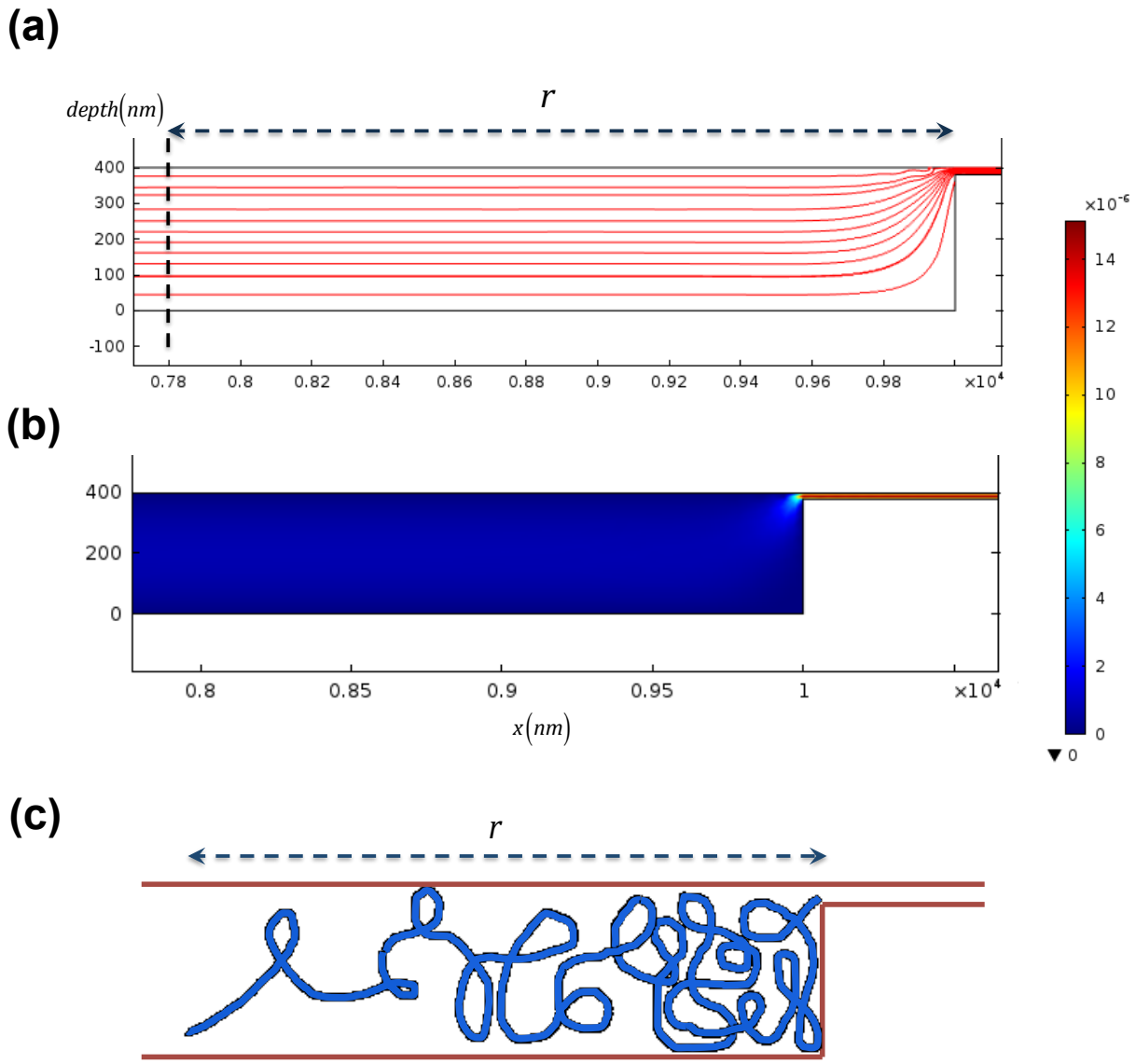

Figure 4–3: The simulation results illustrate the streamlines (a) and velocity field (b) in a nanochannel, obtained by solving Navier-Stokes equation in Comsol software. The streamlines are uniform at distances larger than 300-400 nm from the barrier. The horizontal distance from the dashed line in (a) to the slit represents the extension of highly compressed DNA. (c) A schematic representing a compressed DNA in a nanochannel with extension r.

but not for a large chain that extends beyond the transition region and would be constrained by the channel. What seems far more plausible to us, in the absence of any

strongly non-equilibrium mechanism to induce local agitation, is that the segments will simply 'pile up' and become highly concentrated near the barrier, with strong self-exclusion interactions preventing large-scale local motion of segments (e.g. the upwards motion of segments created by the bending flow-lines would be blocked by segments already present near the slit). Formally, this is to say that the chain is in a *steady-state* with the segmental current (current of Kuhn segments) equal to zero, so that the chain undergoes no *net* motion and thermal fluctuations are the only source of local agitation. Analysis of the chain concentration profile at long times suggests that the profile-shape is indeed quite well described by the assumption of zero segmental current. Formally, a steady-state with zero current can be characterized using equilibrium free energies [182], albeit with two very important provisos: (1) while the profile is static, the steady chain concentration profile is not uniform (e.g. chain concentration is highest near barrier) and (2) the presence of flow does couple into knot free energy, as work is required to drive knots away from the barrier against the flow. The flow is most certainly not creating violent (e.g. highly time-dependent) agitation in the transition region, the basis for knot production via the "tumbling-by-flow-jostling mechanism."

4.2.4 Streaming Potential Estimation

In our experiments no external electric field is applied. While an electric field will be introduced by electrokinetic cross-coupling due to streaming current, these effects are expected to be very small. To estimate the order of magnitude of the streaming potential, we scale classic measurements of the streaming conductance performed in fused silica nanoslit devices [183]. The streaming conductance S_{str} gives the resulting current per unit of pressure applied across the channel, resulting from motion of ions

in the Debye layers. Van der Heyden *et al.* [183] found that the streaming conductance S_{str} for 50 μm wide, 4.5 mm long fused silica nanoslits, with height between 279 nm and 563 nm, was in the range 20-30 pA/bar. To estimate the resulting streaming conductance for our nanochannel devices, we scale the nanoslit streaming conductance by the ratio of the nanochannel to nanoslit width (325 nm/50 μm) and multiply by 2 to take into account the Debye layer from the horizontal dimension of our channels (this is of course crude, but we are interested in an order of magnitude estimate). In our experiment, as there is no external load circuit, we expect that the streaming potential [184]

$$\Delta V = -S_{\mathrm{str}} \Delta p R_{\mathrm{ch}} \tag{4.6}$$

where Δp is the pressure drop across the nanofluidic channel and R_{ch} is the electrical resistance of the nanochannel ($R_{\mathrm{ch}} = L_c/D_1 D_2 \sigma_c$, with L_c channel length and σ_c the buffer conductivity). Using a buffer conductivity of $\sigma_c = 1\,\mathrm{mS/cm}$, $L_c = 200\mu\mathrm{m}$ and a pressure drop $\Delta p = 30\,\mathrm{mbar}$, corresponding to the highest flow achieved ($V = 30\mu\mathrm{m/s}$), we find that $\Delta V \sim 0.2\,\mathrm{mV}$. Alternatively, we can scale the direct streaming potential measurement in [184], taking into account our lower pressure and smaller channel length; this leads to $\Delta V \sim 0.3\,\mathrm{mV}$, which agrees on order of magnitude. The eletrokinetic mobility of DNA in our channels is around 2 μm cm/sV. The electrokinetic DNA velocity resulting from dropping the streaming potential across the channel length is then $\sim 0.02\mu$m/s, three orders of magnitude below the flow-speeds we observe.

4.2.5 Physics of Steady-State Compression

Here we justify in detail the analogy between hydrodynamic and piston-based compression and review relevant details from the past work [151, 152]. In our experiments, a nanochannel-extended chain is compressed against a barrier via steady hydrodynamic flow induced by creating a pressure drop across the channel. As the channel dimensions

are much smaller than the total molecule extension, even at full compression, we adopt a coarse-grained one-dimensional approach, with the concentration $c(x)$ referring to an average of the segmental concentration over the channel cross-section. This scenario is highly analogous to compression of nanochannel-confined DNA via a translating piston-gasket moving with fixed sliding speed [151,152] viewed in a reference frame co-moving with the piston. In the co-moving reference frame, the piston's sliding speed V is transformed into a uniform flow of speed V compressing the polymer against the piston-gasket (analogous to the slit-barrier). Now, if we could apply a perfectly uniform flow field in our channels, our hydrodynamic experiments would be identical to compression via the sliding piston. The difference is that, in the case of pressure-driven flow, the flow-profile is parabolic [170]. In this case, the relevant flow-speed is V_{av} arising from an average of the product of the parabolic Poiseuille flow profile across the channel and the molecule's transverse segmental distribution [185]. While the physical details that set the value of V_{av} are more complex, it behaves in our 1-D formalism analogous to the translation speed V. For convenience of notation, we choose to drop the subscript 'av' and write $V \equiv V_{\text{av}}$. As in the case of the sliding piston, this flow creates a 1-D convective segmental current along the channel $J_c = -Vc$ (the negative sign arises as in our coordinate system, with $x = 0$ at the barrier, flow drives segments to smaller x).

The convective flow is balanced by osmotic pressure gradients that drive segments from regions of high to low concentration. These osmotic pressure gradients lead to cooperative diffusion, characterized by the current:

$$J_D = -D_c(c)\frac{\partial c}{\partial x}. \tag{4.7}$$

The quantity $D_c(c)$ is a *cooperative diffusion constant* and itself is a function of polymer concentration:

$$D_{\mathrm{c}} = \frac{cD_{\mathrm{av}}^2}{\zeta}\frac{\partial \Pi}{\partial c} \tag{4.8}$$

where ζ is a friction factor per unit length. D_{av} is the geometric average of the channel cross section. The osmotic pressure can be obtained from the Flory free energy via an argument detailed in [152]:

$$\Pi = \Pi_o \left(-1/C + C^2\right) \tag{4.9}$$

The quantity $\Pi_o \equiv Ak_BT/(D_{\mathrm{av}}^2\xi_{\|})$ with A a numerical constant ($A = 2.81$) and $\xi_{\|}$ the blob extent. The quantity $C \equiv c/c_o$, where c_o is the concentration of the chain in the absence of flow. Experiment [152] and scaling considerations [151] suggest that $\zeta = \zeta_o C$, which inserted in Eq. 4.8 and along with Eq. 4.9 lead to:

$$D_c = \frac{2\Pi_o D_{\mathrm{av}}^2}{\zeta_o} C \left(1 + \frac{1}{2C^3}\right). \tag{4.10}$$

Note that the $1/2C^3$ term becomes negligible for $C > 2$ (true for the high degree of compression used in our knotting study) so that $D_c \sim C$.

The dynamics of the compression process can be described using the nonlinear diffusion PDE formalism developed in [152]. We insert our expression for the total current $J = J_D - cV$ (with J_D defined via Eq. 4.7) into a continuity equation expressing local conservation of polymer segments,

$$\frac{\partial c}{\partial t} + \frac{\partial J}{\partial x} = 0. \tag{4.11}$$

leading to a nonlinear partial differential equation for predicting the time-evolution of a chain's concentration profile $c(x,t)$:

$$\frac{\partial c}{\partial t} - \frac{\partial}{\partial x}\left(D_c \frac{\partial c}{\partial x} + cV\right) = 0. \tag{4.12}$$

When a molecule is driven by a steady flow against a barrier, one of the molecule edges will be pinned at the barrier ($x = 0$). The second edge, which we call the "free edge" x_f, will slide towards the barrier. Following [152], we argue that the position of the free edge $x_f(t)$ can be described via the equation:

$$\zeta_f \frac{dx_f}{dt} = \Pi(C(x_f))D_{\text{av}}^2 - \zeta_f V. \tag{4.13}$$

In this picture, the free edge is driven by a combination of the flow-induced velocity and osmotic pressure exerted at the chain free edge. When $c = c_o$ at the free edge so that $C(x_f) = 1$, Eq. 4.9 predicts that $\Pi = 0$ and the free edge is driven at the speed V. The quantity ζ_f is the free-edge friction; we find that treating ζ_f as constant gives good description of experimental data for transient compression [152]. Initially, during the compression process, the free edge remains at concentration c_o and is observed to slide towards the barrier at constant velocity [152]. This edge velocity can then be accessed by finding the slope of a position versus time plot of the polymer free edge x_f during transient compression, providing a way to measure V experimentally that self-consistently incorporates unknown details regarding the transverse averaging (see section II(E) for detail on the velocity measurement).

While the free edge slides towards the barrier, contour builds up locally at the barrier, creating a transient shockwave. At a time on order of r_o/V, the free edge reaches the barrier and the profile evolves towards a steady-state ramp in the limit of long times. The steady-state profile can be described by setting $J = 0$:

$$D_{\text{c}}(c)\frac{\partial c}{\partial x} = -Vc \tag{4.14}$$

Note that the ramp profile is determined by the functional form of $D_{\text{c}}(c)$. If we let $D_c \sim C$ the ramp is linear with a slope $\alpha \sim V$. In terms of the normalized position variable

$X \equiv x/r_o$, with the barrier edge at $X = 0$ corresponding to the most concentrated point on the chain with concentration $C_\mathrm{b} \equiv C(0) \geq C(x)$:

$$C(X) = C_\mathrm{b} - \alpha X. \tag{4.15}$$

The steady-state concentration at the free edge $c_\mathrm{f} \equiv c(r)$ is provided by the solution of Eq. 4.13 with $dx_f/dt = 0$:

$$\Pi(C_f)D^2 = \zeta_f V. \tag{4.16}$$

With $\Pi(C_f) \sim C_f^2$ (a good approximation for large C_f) and treating ζ_f as constant, we find that $C_f \sim \sqrt{V}$.

The profile extension r can be found from conservation of contour (i.e. the integral of $c(x)$ is constant [151]). We first define the normalized *chain extension* $R_c \equiv r/r_o$. Contour conservation, in terms of the dimensionless variables, can be expressed [152]:

$$\int_0^{R_c} C(X)\, dX = 1. \tag{4.17}$$

The chain extension R_c can be determined as a function of V by combining Eq. 4.15 and Eq. 4.17 and working the integral. For high V we find $R_c \sim 1/\sqrt{V}$ [151]. Using Eq. 4.15 we can relate the concentration at the barrier edge to the concentration at the free edge: $C_\mathrm{b} = C_f + A_1 V R_c$, so that $C_\mathrm{b} \sim \sqrt{V}$.

We choose to introduce a new variable $R(X) \equiv 1/C(X)$, which we call the *local extension*. The local extension measures how locally stretched ($R > 1$) or in our case how compressed ($R < 1$) the chain is relative to the no-flow equilibrium (where $R = 1$ at every point along the chain). We define the 'local barrier extension' $R_\mathrm{b} \equiv R(0)$. Note that $R_\mathrm{b} = 1/C_\mathrm{b} \sim 1/\sqrt{V}$. As $\alpha \sim V$, we can in turn write α as a function of R_b: $\alpha \sim 1/R_\mathrm{b}^2$. Holding all other physical variables constant (channel dimensions, effective width, persistence length) we need to select only one variable to completely

parameterize the profile for varying degrees of compression (i.e. completely determine the slope and the edge concentration and thus the concentration at all positions via Eq. 4.15). While a natural choice is V or R_c, we choose the local barrier extension as this is the variable most closely linked to the knotting free energy (see section VI) but like R_c it transparently characterizes the chain's total degree of compression (e.g. small R_b corresponds to high compression, with $R_\mathrm{b} \leq R(X)$ for all X). In fact, our theory predicts R_b is proportional to R_c, which we will show in the final paragraph of this section.

Figure 3–2 (e-h) shows experimental results for the steady-state profiles. We show results for all compression events without binning. Figure 3–2 (e) confirms that the data is consistent with a linear relation between α and V; Fig. 3–2 (f) confirms that our results are consistent with $C_\mathrm{b} \sim \sqrt{V}$ and Fig. 3–2 (h) confirms $\alpha \sim 1/R_\mathrm{b}^2$. For definiteness, let

$$\alpha = A_1 V \tag{4.18}$$

$$C_\mathrm{b} = A_2 \sqrt{V} \tag{4.19}$$

and

$$\alpha = A_3 / R_\mathrm{b}^2. \tag{4.20}$$

From least squares-fitting we find $A_1 = 3 \pm 0.1$, $A_2 = 2.52 \pm 0.04$ and $A_3 = 0.464 \pm 0.004$.

Lastly, we will prove that R_b is proportional to R_c and use the values of A_1 and A_2 to estimate the proportionality factor. Integrating Eq. 4.17 with the concentration profile determined by Eq. 4.15 and using Eq. 4.18 and Eq. 4.19 we find that:

$$1 = C_\mathrm{b} R_c - \frac{1}{2}\alpha R_c^2 \Rightarrow R_\mathrm{b} = 1/C_\mathrm{b} = \left[\frac{A_1}{A_2}\frac{1}{A_2 - \sqrt{A_2^2 - 2A_1}}\right] R_c \equiv A_4 R_c \tag{4.21}$$

Eq. 4.21 gives the proportionality factor $A_4 = 0.62 \pm 0.05$. We can also estimate the proportionality factor directly from experiment, from the slope of measurements of R_{b} plotted versus R_c. We find $A_4 = 0.64 \pm 0.01$, which agrees with the estimate from Eq. 4.21.

4.2.6 Molecule Free-Edge Speed Measurement

We use the motion of the free molecule edge, unconstrained during the transient compression phase, as a measure of the buffer flow-speed. Note that during the first phase of the compression process, the free edge remains at concentration c_o and is expected to move towards the barrier at constant speed V [152] (See Eq. 4.9 and Eq. 4.13). In agreement with this prediction, we observe that the measured free-edge position vs. time is linear (See Fig. 4–4). The free edge speed V is then extracted from the slope of a linear fit to the free edge position (via *lsqcurvefit* in Matlab).

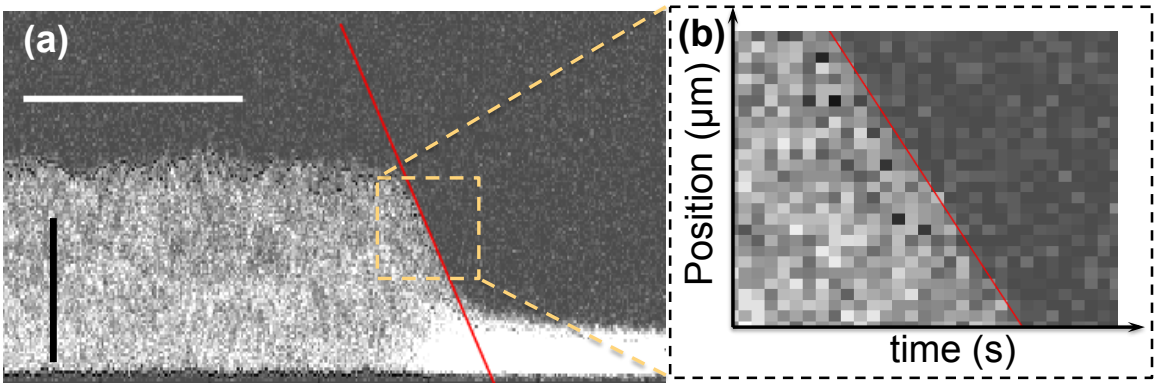

Figure 4–4: Our approach for measuring speed of molecule free edge. The kymograph plot is created by reslicing the raw image to show intensity along the nanochannel versus time (horizontal axis shows time and the vertical axis position). The edge speed is calculated from a linear least-squares fit to the displacement of the molecule free end with time during transient compression. The horizontal and vertical scale bars are 0.5 s and 10μ m, respectively. (b) Magnified image of the kymograph; the slope of the line gives the speed.

4.3 Knot Formation Rate Equations

The knot formation kinetics can be described by a set of coupled rate equations. Defining P_i as the probability of formation of i knots and k_{ij} as the transition rate from a state with i knots to a state with j knots, we have

$$P_0 \underset{k_{10}}{\overset{k_{01}}{\rightleftharpoons}} P_1 \underset{k_{21}}{\overset{k_{12}}{\rightleftharpoons}} P_2. \tag{4.22}$$

The explicit system of rate equations following from Eq. 4.22 is

$$\frac{dP_0}{dt_w} = -P_0 k_{01} + P_1 k_{10} \tag{4.23a}$$

$$\frac{dP_1}{dt_w} = P_0 k_{01} + P_2 k_{21} - P_1(k_{10} + k_{12}) \tag{4.23b}$$

$$\frac{dP_2}{dt_w} = P_1 k_{12} - P_2 k_{21} \tag{4.23c}$$

Note that as $t_w \to \infty$ the system reaches equilibrium and the time derivatives vanish giving rise to the three equations:

$$P_0(\infty)k_{01} = P_1(\infty)k_{10} \tag{4.24a}$$

$$P_0(\infty)k_{01} + P_2(\infty)k_{21} = P_1(\infty)(k_{10} + k_{12}) \tag{4.24b}$$

$$P_1(\infty)k_{12} = P_2(\infty)k_{21} \tag{4.24c}$$

Our experimental waiting time data was taken at high compression for which $P_0(\infty) \approx 0$. Thus the backwards rate $k_{10} \approx 0$. In addition, we have $P_1(\infty)/P_2(\infty) = k_{21}/k_{12}$.

Letting $\lambda = k_{21}/k_{12}$ the system can then be solved exactly,

$$P_1(t_w) = \frac{\left[\left(e^{-k_{01}t_w} - e^{-(1+\lambda)k_{12}t_w} + \lambda\left(e^{-k_{01}t_w} - 1\right)\right)k_{01} + \left(1 - e^{-k_{01}t_w}\right)\lambda(1+\lambda)k_{12}\right]}{(1+\lambda)(-k_{01} + k_{12}(1+\lambda))} \tag{4.25}$$

$$P_2(t) = \frac{[k_{01}(-1 + e^{-k_{12}(1+\lambda)t_w}) + k_{12}(1+\lambda)(1 - e^{-k_{01}t_w})]}{(1+\lambda)(-k_{01} + k_{12}(1+\lambda))} \tag{4.26}$$

The total probability of knot formation is $P_{\text{tot}} = 1 - P_0(t_w) = 1 - e^{-k_{01}t_w}$. The three parameters k_{01}, k_{12} and λ are found by least-square fitting of the model to the experimental data (Fig. 3–3 (b)), also using *lsqcurvefit*.

4.4 Knot-Spatial Distribution

The spatial distribution of knots can be accessed at a short-time following pressure release. Figure 4–5 (which has identical data to Fig. 3–6) shows the histogrammed position of knots for one-knot (Fig. 4–5(a,c)) and two-knot events (Fig. 4–5(b,d)) 2 sec after pressure release. For the two-knot events, the position of the knot closest to the slit-barrier ('lower-knot') and the knot farthest from the slit-barrier ('upper knot') are separately histogrammed. The figure shows knot histograms (Fig. 4–5(a,b)) yielding an estimate of the knot spatial probability distribution $P_{ij}(X_j)$ (i gives knot-state, i.e. $i = 1$ for one knot state, $i = 2$ for two knot; j indexes the knot number in the state). The cumulative knot histogram $C_{\text{cum},ij}(X_j)$ (Fig. 4–5(c,d)) is simply the total number of knots observed at positions X' with $X' < X_j$. The cumulative distribution has the advantage that it is insensitive to binning and is preferred for use in data interpretation.

Here we derive equation 4.29, the basis for how we estimate the expected upper knot distribution in the absence of knot interactions. To begin this argument, assume the knots do not interact so that their spatial probability distributions are statistically independent. Moreover, assume that the positions of knots 1 and 2 follow the same probability density function $P(x)$. Note that, even in the absence of interactions (i.e. the knots can freely pass and cannot be distinguished based on their ordering along the profile), the lower and upper knot distributions will have different functional forms. The upper non-interacting knot distribution $P_U(x)$ is given by

$$P_U(x) = P(x_1 = x)P(x_2 \leq x) + P(x_2 = x)P(x_1 \leq x). \tag{4.27}$$

The first term is the probability that position x is the highest spatial position of a knot on the profile and that $x_1 > x_2$. This term is the product of the probability $P(x_1 = x) = P(x)$ that knot 1 is at position x and $P(x_2 \leq x)$ the probability that knot 2 is at a position $x_2 \leq x$ (e.g. given by $\int_0^x P(x_2)\, dx_2$). The second term is the symmetric contribution assuming $x_2 > x_1$. By similar reasoning, the lower knot non-interacting knot distribution $P_L(x)$:

$$P_L(x) = P(x_1 = x)P(x_2 \geq x) + P(x_2 = x)P(x_1 \geq x) \tag{4.28}$$

with the probability that knot 2 is at position $x_2 \geq x$ given by $P(x_2 \geq x) = \int_x^\infty P(x_2)\, dx_2$. Note that, using $P(x_1 \leq x) + P(x_1 \geq x) = 1$, Eq. 4.27, 4.28 imply:

$$P_L(x) + P_U(x) = 2P(x) \tag{4.29}$$

This is equation 4.29.

To make this argument concrete, let $P(x)$ have an exponential form, as expected for the one-knot distributions in steady-state:

$$P(x) \sim \exp(-\beta x). \tag{4.30}$$

Evaluating Eq. 4.27 and Eq. 4.28 with $P(x)$ given by Eq. 4.30, we find that

$$P_L(x) \sim \exp(-2\beta x) \tag{4.31}$$

and

$$P_U(x) \sim (1 - \exp(-\beta x)) \exp(-\beta x) \tag{4.32}$$

Thus, shape-differences between the upper and lower distributions do not necessarily indicate that the knots interact, we expect P_U and P_L to have different forms even if no interactions are present. Note that the cumulative lower and upper knot distributions C_L and C_U can be formed from P_L and P_U via $C_{\text{cum}}(x) = \int_0^x P(x')\, dx'$.

How then can we tell if interactions are present between the knots? The key is to look for differences between the predicted *non-interacting* upper knot distribution P_U, and the measured upper knot distribution. If the knots do not interact by assumption, then the underlying distribution P must be identical for both knots and is given by the one knot distribution P_1; P_L can be obtained from the measured lower knot distribution and P_U is then necessarily determined by Eq. 4.29. If the obtained distribution P_U fails to describe the experimental upper knot distribution, then we can say interactions are present.

To test for knot interactions using our empirical knot distributions, we obtain P_U using the non-interacting assumption and compare it to the measured upper knot distribution (Fig. 4–5(d)). The comparison is best made using the cumulative distributions (C_U–purple bold curve, green curve–measured cumulative upper knot distribution).

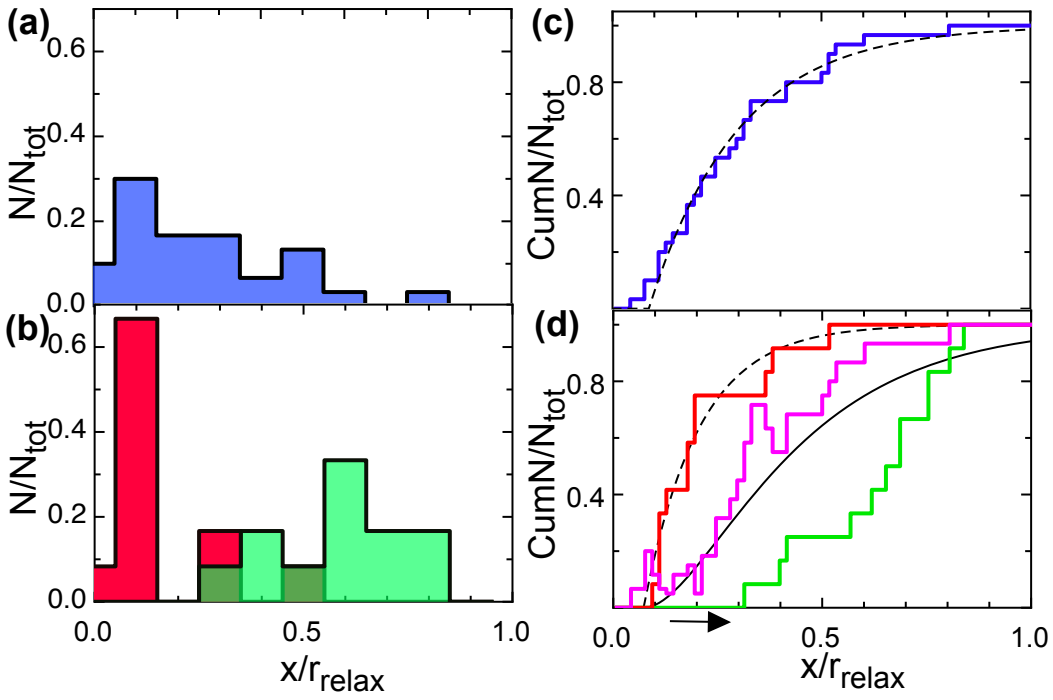

Figure 4–5: Knot position histogram for one-knot states (a) and two-knot states (b) 2 sec after pressure release. Cumulative knot position histogram for one-knot states (c) and two-knot states (d) normalized to the total number of counts. The x-axis is normalized to the extension $r_{\rm relax}$ of the relaxing molecule measured at the time for which the knot-position was obtained ($x = 0$ gives barrier edge, $x = 1$ gives molecule free edge). The data used includes events with $R_{\rm b} \approx 0.11, 0.13, 0.17$. The ratio of the event averaged $r_{\rm relax}$ to the steady-state extension r is 2.9; $R_{\rm relax} \equiv r_{\rm relax}/r_o = 0.5$. For the two-knot states, we separately histogram the position of the knot closest to the barrier ('lower knot', shown in red) and the knot farthest from the barrier ('upper knot', shown in green). The dashed curve in (c) is a fit of the cumulative single-knot distribution to the cumulative distribution corresponding to an exponential probability distribution; the dashed curve in (d) is a fit of the cumulative lower-knot distribution to the cumulative distribution for an exponential probability distribution. The purple bold curve in (d) is the estimated cumulative non-interacting upper knot distribution C_U formed directly using the single knot and lower knot measurements. The bold curve is the estimated cumulative non-interacting upper knot distribution C_U based on the exponential fit to the lower cumulative knot distribution. The arrow indicates that the upper knot distribution is shifted relative to the estimated C_U (bold black, bold purple), indicating that, for at least at this particular time-point following relaxation, the lower- and upper-knots interact and they cannot pass.

Alternatively, we can estimate P_U by noting that an exponential shape well describes the lower-knot distribution (dashed curve Fig. 4–5(d)), consistent with Eq. 4.31. The extracted parameter β (see Eq. 4.31) can then be used to estimate the distribution P_U via Eq. 4.32 (bold-line, Fig. 4–5(d)). We see that the measured distribution is shifted to higher X relative to P_U obtained with both approaches (again, the comparison is best made with the cumulative distributions). This proves interactions exist at 2 sec following pressure release and these interactions are strong enough to prevent knot crossing, consistent with our hypothesis regarding knot-interactions in steady-state and single-file ordering of knots.

Furthermore, note that our basic interaction test–which is independent of model assumptions–is not altered by averaging over knot sizes (we expect a distribution of knot sizes to be present during the experiment). Let the predicted non-interacting lower, upper and single-knot distributions averaged over all knot sizes be denoted by P_L^{av}, P_U^{av} and P^{av}. Then, averaging both sides of Eq. 4.29, we find that $P_L^{\text{av}}(x) + P_U^{\text{av}}(x) = 2P^{\text{av}}(x)$. As P^{av} and P_U^{av} will be obtained from the experimental knot data automatically (the experimental data is necessarily averaged over all knot sizes!), our argument in fact yields P_U^{av}, so the comparison is already being made with the averaged distributions.

4.5 Expanded Discussion of Knot Formation Free Energy

In this section, we expand the discussion of the knot free energy model developed in chapter 3. Our objective is to compute the probability of forming i knots on an extended chain in terms of the edge extension R_{b}. Our model is based very closely on the approach of Dai *et al.* [11]. In this section, we discuss in detail the types of free energy associated with a nanochannel-confined knot. The free energy of a single knot will be expressed as a function of knot contour length L_k, position along the profile

X and the edge extension R_b. Note that the edge extension serves to completely parameterize the concentration profile $C(X)$ obtained at a given flow speed V. In the following sections, we will discuss the procedure for spatial averaging over X, including interaction between knots for multi-knot states and forming partition functions that sum over L_k and can be compared directly to experimental knotting probabilities.

Three types of free energy are associated with a knot of size L_k on a nanochannel-confined chain [11]. The first contribution is the free energy *cost* of forming the knot on the chain in bulk $f_\mathrm{b}(L_k)$. The second contribution is the free energy *cost* of confining the knot between the channel walls $f_\mathrm{wk}(L_k, R)$ ('wall-knot'). The third contribution is the free energy *saved* as contour stored in the knot no longer contributes the confinement free energy $f_\mathrm{wuk}(L_k, R)$ ('wall-unknot') associated with an unknotted section of polymer of size L_k. Note that the bulk contribution depends only on the knot size L_k. The wall-knot and wall-unknot contributions depend on L_k *and* the local extension $R(X) = 1/C(X)$. The local extension is in turn a function of position X and the edge extension R_b through the profile (Eq. 4.15). Note that we see no evidence that the existence of knots alters the profile-shape. In addition, the knot will have an extra free energy $f_h(L_k, V, X)$ due to the flow V, which exerts a constant viscous drag force pulling knots to lower X and thus raising their free energy when they are displaced from the slit-barrier. Combining the terms, one knot of size L_k at position X on a profile with edge extension R_b has a total free energy:

$$\begin{aligned} f_1(L_k, R_\mathrm{b}, X) = A_\mathrm{b} f_\mathrm{b}(L_k) + A_\mathrm{wk} f_\mathrm{wk}(L_k, R(X, R_\mathrm{b})) \\ - f_\mathrm{wuk}(L_k, R(X, R_\mathrm{b})) + A_\mathrm{h} f_\mathrm{h}(L_k, R_\mathrm{b}, X). \end{aligned} \tag{4.33}$$

We choose to explicitly include the dimensionless scaling prefactors A_b, A_wk and A_h associated with the bulk, wall-knot and hydrodynamic terms: these will be determined

by fitting to experimental knotting probability data. Following Dai *et al.*, the existing theory will be used to explicitly calculate the scaling prefactor A_{wuk} as this term is the well-known confinement free energy of a nanochannel-confined polymer. Also, note that the free energy of the unknotted chain represents the base-line free energy for this theory: explicitly f_1 gives the excess (or decrease) in free energy arising from the presence of one knot on the chain. By definition we have $f_0 = 0$. In order to simplify the notation, all free energy terms will be expressed in units of k_BT. Note also that we express f_1 and f_h as a function purely of X, L_k and R_{b}: the flow speed V can be eliminated using Eq. 4.19 and Eq. 4.20.

Dai *et al.* have used a combination of scaling arguments and simulation to deduce the detailed forms for the free energy contributions in Eq. 4.33. The term f_b gives an R-independent contribution that can be obtained from a balance of bending and confinement free energy [153, 154]:

$$f_{\mathrm{b}}(L_k) = k_1(L_k/P)^{-1} + k_2 L_k \left(L_k - pw\right)^{-2/3} P^{-1/3}, \tag{4.34}$$

where P and w are the persistence length and effective width of the chain, respectively. The quantities k_1, k_2 and p are numerical constants determined from simulations of equilibrium trefoil knots ($k_1 = 17.06$, $k_2 = 1.86$ and $p = 16$ [154]).

The contribution f_{wuk} can be obtained from the Flory free energy for a confined chain [137]:

$$f_{\mathrm{wuk}}(L_k, R) = A_{\mathrm{wuk}} \frac{L_k}{L_b} \left(R^2/2 + 1/R\right) \tag{4.35}$$

This equation is derived for a uniformly extended chain (e.g. $R =$ constant), but we can interpret Eq. 4.35 in a local sense for a chain with a non-uniform concentration profile as giving the free energy of confinement associated with a small contour L_k extracted from position X along the profile with local extension $R(X)$. The quantity L_b is the contour

per blob in the equilibrium chain, which we obtain from extended de Gennes theory (the appropriate confinement regime for channel in our size-range): $L_b = (2P/w^2)^{1/3} D_{\text{eff}}^{4/3}$. The quantity $D_{\text{eff}} = \sqrt{(D_1 - w)(D_2 - w)}$ is an effective diameter taking into account the non-unity aspect ratio of the channel and wall-DNA interactions (estimating the electrostatic depletion near the channel walls by w [126]). Note that, in the classic de Gennes regime, $L_b = D_{\text{eff}}^{5/3}/(2Pw)^{1/3}$. The quantity A_{wuk} is a numerical constant ($A_{\text{wuk}} = 2.81$ [137]). Note that Eq. 4.35 does differ subtly from the unknot free energy used in [11]; Dai *et al.*, focusing mostly on channels of larger size than those used in our study, use a classic de Gennes assumption, while an extended de Gennes assumption is more appropriate for the channel size used here. Lastly, as the knot formation probability is only appreciable below $R = 0.5$, we argue that the R^2 term in Eq. 4.35 is negligible for our experiments.

Dai *et al.* observe in their simulations that the knots can be viewed as spherical regions with radius of gyration g_k. The relation between g_k and L_k can be determined from Monte Carlo simulation of knotted semiflexible and self-avoiding chains. Dai *et al.* find that the relation $g_k = 0.1L_k$ describes their simulation results well (see in particular the supplementary material to [11]). The free energy cost f_{wk} arises from two sources: the knot, acting like a hard sphere, interacts via excluded volume with the rest of the chain; the knot also experiences greater confinement due to its physical size, effectively reducing the channel width. In order to quantify these effects, we adopt the 'knot-chain' approach of Dai *et al.*. In the knot-chain approach, the confinement free energy of the knot is given by the confinement free energy of one bead on an effective chain made up of bead-units of radius g_k. For the case of an uncompressed chain,

where $R = 1$, Dai *et al.* argue that

$$f_{\text{wk}} = L_k \frac{g_k^{2/3}}{D_{\text{eff}}'^{5/3}} \tag{4.36}$$

The quantity $D'_{\text{eff}} = \sqrt{(D_1 - \delta)(D_2 - \delta)}$ takes into account the increase in effective confinement felt by the knot, arising from the fact that the cross-section of the channel occupied by the chain must be reduced by the finite size of the knot, introducing an offset $\delta = \gamma g_k$ with the value $\gamma = 1$ describing their simulations well. We have generalized Eq. 4.36 from the result given by Dai *et al.* to account for our channel's non-unity aspect ratio. Eq. 4.36 exactly reproduces Eq. 6 in [11] in the limit that $D_1 = D_2$.

Dai *et al.*'s original result (Eq. 4.36) requires modification when the chain is compressed ($R < 1$). The confinement free energy f_{wk} arises in part from excluded volume interactions between the knot and the rest of the chain. This free energy cost increases as the chain is compressed, due to the decreasing available chain volume, introducing a dependence on R (again, f_{wk} is interpreted locally as the free energy associated with a knot of size L_k at position X with local extension $R(X)$). For $R > 0.5$, reference [137] suggests

$$f_{\text{wk}}(L_k, R) = L_k \frac{g_k^{2/3}}{D_{\text{eff}}'^{5/3}} \left(R^2/2 + 1/R\right) \tag{4.37}$$

Note that the compression factor has the same form as Eq. 4.35. For higher compression, [137] suggests the physics enters a 'semidilute' regime where the dependence changes to $\sim 1/R^{5/4}$ [137]:

$$f_{\text{wk}}(L_k, R) = L_k \frac{g_k^{2/3}}{D_{\text{eff}}'^{5/3}} \frac{1}{R^{5/4}} \tag{4.38}$$

We will use Eq. 4.38 as in our data the knot formation events were mostly in the high compression regime. Note that this semidilute regime does not apply to f_{wuk},

the confinement free energy of the unknotted polymer, which should follow extended de Gennes statistics (e.g due to the large P/w ratio and consequent anisotropy of the polymer segments) and enter a mean-field regime at higher compression for which the $\sim 1/R$ scaling is correct.

Assuming that hydrodynamic interactions are likely present over the knot-size, which is on order of the channel diameter, we estimate the friction factor of the confined knot as that of a sphere of radius g_k. The knot friction factor is then $\zeta_k = 6\pi\eta g_k$ leading to a drag force $6\pi\eta g_k V$ so that the hydrodynamic term f_h (in units of k_BT) is:

$$f_{\mathrm{h}}(L_k, V, X) = 6\pi\eta g_k V r_o X/k_BT \tag{4.39}$$

where η is the solution viscosity. Note that we can choose to write the flow speed V as a function of the edge extension using Eq. 4.19 and Eq. 4.20, so that the single knot free energy f_1 can be completely expressed as a function of L_k, X and R_{b}.

4.6 Spatial Variation of Knotting Free Energy

The free energy of confined knots depends on their position along the profile (see Fig. 4–6). This dependence arises from $R(X) = 1/C(X)$ and the explicit X-dependence of the hydrodynamic term $f_h(L_k, V, X)$. While Eq. 4.15, 4.35 and 4.39 indicate that f_{wuk} and f_{h} depend linearly on X, Eq. 4.38 implies that $f_{\mathrm{wk}} \sim R^{-5/4}$, introducing a slight nonlinearity. We treat the nonlinearity by Taylor expanding f_{wk} for small X:

$$f_{\mathrm{wk}}(L_k, R(X, R_{\mathrm{b}})) = L_k g_k^{2/3} D_{\mathrm{eff}}^{-5/3} R^{-5/4} \approx L_k g_k^{2/3} D_{\mathrm{eff}}^{-5/3} \left[R_{\mathrm{b}}^{-5/4} - \frac{5}{4}\alpha R_{\mathrm{b}}^{-1/4} X \right] \tag{4.40}$$

This approximation is justified by the exponential dependence of one-knot formation probability on f_1 (Eq. 4.33) through Boltzmann statistics, ensuring that the probability is only appreciable for small X where the Taylor expansion is valid. Combining

equations 4.15, 4.33, 4.34, 4.35, 4.38, 4.39 and 4.40 leads to:

$$f_1(L_k, R_\mathrm{b}, X) = f_1(L_k, R_\mathrm{b}, 0) + \beta(L_k, R_\mathrm{b})X \tag{4.41}$$

with

$$f_1(L_k, R_\mathrm{b}, 0) = A_b f_\mathrm{b}(L_k) + A_\mathrm{wk} f_\mathrm{wk}(L_k, R_\mathrm{b}) - f_\mathrm{wuk}(L_k, R_\mathrm{b}) \tag{4.42}$$

and the slope of the free energy profile $\beta(L_k, R_\mathrm{b})$ given by:

$$\beta(L_k, R_\mathrm{b}) = \alpha \left(f_\mathrm{wuk}(L_k, 1) - \frac{5A_\mathrm{wk}}{4} f_\mathrm{wk}(L_k, 1) R_\mathrm{b}^{-1/4} \right) + A_\mathrm{h} f_\mathrm{h}(L_k, 1) \tag{4.43}$$

Both $\alpha \sim V \sim 1/R_\mathrm{b}^2$ and $f_\mathrm{h} \sim V \sim 1/R_\mathrm{b}^2$ (see Eq. 4.18, 4.19 and 4.20) so that $\beta \cong 1/R_\mathrm{b}^2$ (the scaling is only approximate due to the factor of $R_\mathrm{b}^{-1/4}$ from the Taylor expansion of f_wk).

4.7 Probability Distributions for Knot Position for Single and Two-Knot States

Equation 4.41 implies that single knots have an exponential probability distribution in position-space: $P_1(L_k, R_\mathrm{b}, X) \sim \exp\left(-\beta(L_k, R_\mathrm{b})X\right)$ (see Fig. 4–6(d)). While the knots are likely to be found near $X = 0$ and $f_1(L_k, R_\mathrm{b}, 0)$ provides a good estimate of the free energy of single-knot states, the free energy will be slightly higher due to thermal excitation of knots to higher X (described by $P_1(L_k, R_\mathrm{b}, X)$) and the finite size of the knots (the knots will 'bump' into the slit barrier at $x = g_k$ forcing them to have higher free energy, Fig. 4–6(b, c)).

In order to gain a quantitative description of states with multiple knots, we must include knot interactions. Including interactions prevents the formation of states with a very large number of knots even if the single knot formation probability is high. Knot-knot interactions, for example, could arise through the excluded volume of one knot restricting the configuration space of the other knots ('knot-knot' excluded volume);

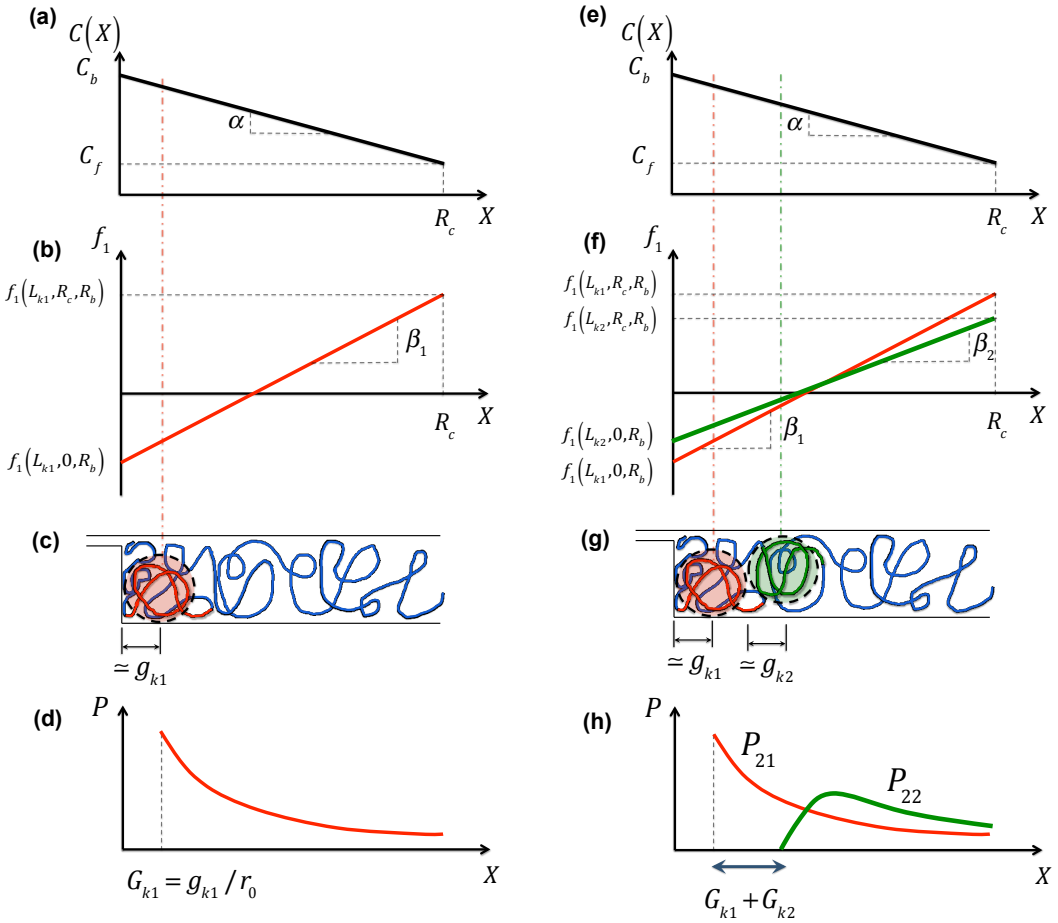

Figure 4–6: (a, e) Schematic of concentration profile $C(X)$, (b, f) resulting free energy profile $f_1(L_k, R_b, X)$ and (d, h) resulting knot probability distributions for single (a-d) and two-knot (e-h) states. Note that in the two-knot state the two knots can experience different free energy profiles due to their varying size (e.g. L_{k1} does not necessarily equal L_{k2}): this is indicated in (f) by drawing two free energy profiles with slopes β_1 and β_2. In addition, in the two-knot state the knots will have different probability distributions: P_{21} for the 'lower' knot closest to the barrier and P_{22} for the 'upper' knot farthest from the barrier (h). (c) Cartoon of single-knot on chain, displaced by radius of gyration g_{k1} from barrier edge. The red dashed line relates the knot position to the corresponding concentration and free energy: note that the steric interaction between the molecule and barrier edge leads to an increase βG_{k1} in the molecule free energy (recall that $G_{k1} \equiv g_{k1}/r_o$). (g) Cartoon of two knots on chain. Knot 1 (red) is displaced by radius of gyration g_{k1} from barrier edge; knot 2 (green), by assumption of single-file ordering, is displaced by $2g_{k1} + g_{k2}$ from barrier edge. The black dashed lines in (c, g) show the extent of each knot site on the compressed chain, used to calculate partition sums (See section VIII). i and j denote site indices used in equations 4.50 and 4.52 for the first and second knots, respectively.

this effect would scale as $k_BTg_k^3/rD_1D_2$. Yet, the volume of a single knot is very small relative to the volume occupied by the chain: with $g_k \sim 100\,\mathrm{nm}$ we find $g_k^3/rD_1D_2 \sim 10^{-2}$. The knot-knot interactions must be at least on order of a fraction of k_BT as they have an appreciable effect after only two knots are present (see Fig. 3–4 (b)). We believe the interactions must thus have a more subtle physical origin.

We argue that the knots do interact via hard-wall repulsion over their diameter $2g_k$, as in a typical excluded-volume scenario, but that the effect of this interaction is greatly magnified by the channel confinement. In particular, we argue that inside the channel the knots satisfy a *no-crossing* condition: knots cannot pass in either contour or position space, leading to single-file diffusion of knots in the linear potential $f_1(L_k, R_\mathrm{b}, X)$. While it is believed that molecular knots can pass through each other in bulk by diffusion along the chain contour [178], this mechanism requires an overall swelling of knot-size. In order to sufficiently lower the free energy barrier, so one knot can diffuse through the second knot, one knot must swell to almost its original size plus that of the second knot. Inside a channel, this mechanism cannot be operative if the knots have a diameter on order of the channel width. Thus, it is reasonable to argue that the knots behave like interacting hard spheres subject to a no-crossing condition via excluded-volume considerations alone.

With single-file ordering assumed, the knot interaction energy arises from a very simple mechanism: multiple knots stack in the linear potential, pushing knots to higher X and thus higher position in the free energy ramp (see Fig. 4–6(f, g)). We can make this idea rigorous using the approach of [186]. Firstly, we define the notion of normalized knot gyration radius $G_k \equiv g_k/r_o$. Say two knots of contour L_{k1} and L_{k2} with corresponding normalized gyration radii G_{k1} and G_{k2} are at normalized positions X_1 and X_2. In our linear approximation, we can say the knots have free energies $f_{11}(X_1) =$

$f_{11}^o + \beta_1 X_1$ and $f_{21}(X_2) = f_{21}^o + \beta_2 X_2$ with $f_{11}^o \equiv f_1(L_{k1}, R_\mathrm{b}, 0)$, $f_{21}^o \equiv f_1(L_{k2}, R_\mathrm{b}, 0)$, $\beta_1 \equiv \beta(L_{k1}, R_\mathrm{b})$ and $\beta_2 \equiv \beta(L_{k2}, R_\mathrm{b})$. In addition, the knots interact via a potential $f_\mathrm{int}(X_2 - X_1)$ which has the hard-core form:

$$f_\mathrm{int}(X_2 - X_1) = \begin{cases} 0 & \text{if } X_2 - X_1 > G_{k1} + G_{k2} \\ \infty & \text{if } X_2 - X_1 \leq G_{k1} + G_{k2} \end{cases} \tag{4.44}$$

The total free energy of the two interacting knots is then:

$$f_2(L_{k1}, L_{k2}, R_\mathrm{b}, X_1, X_2) = f_{11}(X_1) + f_{21}(X_2) + f_\mathrm{int}(X_2 - X_1). \tag{4.45}$$

Defining $\delta X \equiv X_2 - X_1$, we can rewrite Eq. 4.45:

$$f_2 = \underbrace{[f_{11}^o + f_{21}^o + (\beta_1 + \beta_2)X_1]}_{=f'(X_1)} + \underbrace{[\beta_2 \delta X + f_\mathrm{int}(\delta X)]}_{=f''(\delta X)} \tag{4.46}$$

In other words, when rewritten in terms of X_1 and δX, the free energy can be expressed as the sum of the contributions $f'(X_1)$ and $f''(\delta X)$ so that the probability distributions of the variables X_1 and δX are independent. Using the notation $P_{21}(X_1)$ to indicate the probability distribution of knot 1 in the two-knot state, and using Eq. 4.44, we can write:

$$P_{21}(X_1) = \begin{cases} 0 & \text{if } X_1 \leq G_{k1} \\ N_{21} \exp\left(-(\beta_1 + \beta_2)X_1\right) & \text{if } X_1 > G_{k1} \end{cases} \tag{4.47}$$

The probability distribution $P_\mathrm{diff}(\delta X)$ for δX is:

$$P_\mathrm{diff}(\delta X) = \begin{cases} 0 & \text{if } \delta X \leq (G_{k1} + G_{k2}) \\ N_\mathrm{diff} \exp\left(-\beta_2 \delta X\right) & \text{if } \delta X > (G_{k1} + G_{k2}) \end{cases} \tag{4.48}$$

with N_diff and N_{21} normalization constants.

Lastly, we can deduce the probability distribution of knot 2 in the two-knot state P_{22}. Recalling $X_2 = \delta X + X_1$, we observe that X_2 is expressed as the sum of the independent random variables δX and X_1 described respectively by the probability distributions $P_{\text{diff}}(\delta X)$ (Eq. 4.48) and $P_{21}(X_1)$ (Eq. 4.47). Thus, $P_{22}(X_2)$ is determined by the *convolution* of $P_{\text{diff}}(\delta X)$ and $P_{21}(X_1)$. Using Eq. 4.47 and Eq. 4.48,

$$\begin{aligned} P_{22}(X_2) &= \int P_{\text{diff}}(Y) P_{21}(X_2 - Y)\, dY \\ &= N_{21} N_{\text{diff}} \int_{G_{k1}+G_{k2}}^{X_2 - G_{k1}} e^{-\beta_2 Y} e^{-(X_2 - Y)(\beta_1 + \beta_2)}\, dY \\ &= N_{22} e^{-(\beta_1+\beta_2)X_2} \left(-e^{(G_{k1}+G_{k2})\beta_1} + e^{(-G_{k1}+X_2)\beta_1} \right) \end{aligned} \tag{4.49}$$

with N_{22} a normalization factor. The lower integration limit is determined by the hard-wall cut-off in Eq. 4.48. The upper limit is the maximum value of δX for fixed X_2 (δX is maximized at fixed X_2 when $X_1 = G_{k1}$). Note that P_{21} and P_{22} do not reduce to their respective one-knot distributions $P_1(x) \sim \exp(-\beta_1 X)$ when $G_{k1} = G_{k2} = 0$ as the integration limits still enforce $\delta X > 0$ so that $X_2 \geq X_1$ (i.e. the model still assumes that the knots are not allowed to pass). The lower distribution P_{21} and upper distribution P_{22} are shown in Fig. 4–6(h). Note that P_{22} is shifted from the origin.

4.8 Partition Functions for Knot Formation

In order to compare with our knot probability measurements we must compute partition functions that count the number of ways multiple knots can form on a compressed DNA profile: these partition functions involve summations over knot positions along the profile *and* knot size. We will start by considering summation over possible knot formation positions. Let the partition sum $z_1(L_k, R_b)$ count the total number of ways a single prime knot of size L_k, assumed to be of trefoil topology, can form on a profile characterized by R_b. The number of statistically independent sites at which a

knot can form along the profile is estimated by $n_{\max} = r/2g_k = R/2G_k$, each site i displaced by $\Delta X_i = 2G_k(i + 1/2)$ from $X = 0$. Note that i runs from 0 to $n_{\max} - 1$, with the first site at $i = 0$ displaced by a factor of G_k from the barrier due to the knot's finite spatial extent (see figure 4–6(c)). Each site is weighted by a Boltzmann factor $\exp(-f_1(L_k, R_{\rm b}, \Delta X_i))$. Equation 4.41 then leads to a single knot partition function:

$$z_1(L_k, R_b) = \sum_{i=0}^{n_{\max}-1} e^{-f_1(L_{k1}, R_b, \Delta X_i)} = e^{-f_1(L_{k1}, R_b, 0)} \sum_{i=0}^{n_{\max}-1} e^{-2\beta G_k\left(i+\frac{1}{2}\right)} \tag{4.50}$$

where $\beta \equiv \beta(L_k, R_b)$. As the probability of finding a knot away from the barrier is suppressed exponentially, it is a good approximation to set the sum limit to infinity, in which case we can sum Eq. 4.50 geometrically to find:

$$z_1(L_k, R_b) = \frac{e^{-f_1(L_{k1}, R_b, 0)+\beta G_k}}{e^{2\beta G_k} - 1}. \tag{4.51}$$

In the limit that the slope of the free energy profile is very steep, i.e. $\beta(L_k, R_b)$ is large, Eq. 4.51 leads to a single knot free energy $F_1(L_k, R_b) = -\log(z_1(L_k, R_b)) = f_1(L_{k1}, R_b, 0) + \beta(L_k, R_b)G_k$. This limit corresponds to a "ground state" configuration where knot formation at the minimum free energy position abutting the barrier dominates so that the knot free energy is simply the free energy at $X = 0$ ($f_1(L_{k1}, R_b, 0)$) plus the free energy increase arising from the knot's displacement by a factor of G_k "up the ramp" from $X = 0$ ($\beta(L_k, R_b)G_k$, see figure 4–6(c)).

The partition function $z_2(L_{k1}, L_{k2}, R_b)$ counts the total number of ways two trefoil knots can form on a profile characterized by R_b. When multiple knots form we must construct partition sums that preserve the linear ordering of the knots along the profile. Let $\Delta X_i = 2G_{k1}(i + 1/2)$ represent the displacement of the knot closest to the barrier from $X = 0$ (with i running from 0 to $n_{\max} - 1$ as before). The second knot must have a displacement $\Delta X_j \geq \Delta X_i + G_{k1} + G_{k2}$ leading to $\Delta X_j =$

$2G_{k1}(i+1) + 2G_{k2}(j+1/2)$. Each state (i, j) will then be weighted by a Boltzmann factor $\exp\left[-f_1(L_{k1}, R_\mathrm{b}, \Delta X_i) - f_1(L_{k2}, R_\mathrm{b}, \Delta X_j)\right]$. Making the approximation that the summation limits can be extended to infinity, Eq. 4.41 gives:

$$z_2(L_{k1}, L_{k2}, R_\mathrm{b}) = e^{-[f_1(L_{k1},R_b,0)+f_1(L_{k2},R_b,0)]} \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} e^{-[2\beta_1 G_{k1}(i+1/2)+2\beta_2(G_{k1}(i+1)+G_{k2}(j+1/2))]} \tag{4.52}$$

with $\beta_1 \equiv \beta(L_{k1}, R_b)$ and $\beta_2 \equiv \beta(L_{k2}, R_b)$. The same procedure gives the partition function of a three-knot state:

$$\begin{aligned} z_3(L_{\mathrm{k1}}, L_{\mathrm{k2}}, L_{\mathrm{k3}}, R_\mathrm{b}) = e^{-\left[f_1(L_{\mathrm{k1}},R_\mathrm{b},0)+f_1(L_{\mathrm{k2}},R_\mathrm{b},0)+f_1(L_{\mathrm{k3}},R_\mathrm{b},0)\right]} \\ \times \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{\infty} \Bigg(e^{-2\beta_1 G_{\mathrm{k1}}(i+1/2)} \\ \times e^{-2\beta_2\left(G_{\mathrm{k1}}(i+1)+G_{\mathrm{k2}}(j+1/2)\right)} \\ \times e^{-2\beta_3\left(G_{\mathrm{k1}}(i+1)+G_{\mathrm{k2}}(j+1)+G_{\mathrm{k3}}(l+1/2)\right)} \Bigg). \end{aligned} \tag{4.53}$$

The partition sums Eq. 4.52, 4.53 can be evaluated by geometric summation:

$$z_2(L_{k1}, L_{k2}, R_\mathrm{b}) = \frac{e^{-f_1(L_{k1},R_b,0)-f_1(L_{k2},R_b,0)+G_{k1}\beta_1+G_{k2}\beta_2}}{\left(e^{2G_{k2}\beta_2}-1\right)\left(e^{2G_{k1}(\beta_1+\beta_2)}-1\right)} \tag{4.54}$$

and

$$z_3(L_{k1}, L_{k2}, L_{k3}, R_\mathrm{b}) = \frac{e^{-f_1(L_{k1},R_b,0)-f_1(L_{k2},R_b,0)-f_1(L_{k3},R_b,0)+G_{k1}\beta_1+G_{k2}\beta_2+G_{k3}\beta_3}}{\left(e^{2G_{k3}\beta_3}-1\right)\left(e^{2G_{k2}(\beta_2+\beta_3)}-1\right)\left(e^{2G_{k1}(\beta_1+\beta_2+\beta_3)}-1\right)} \tag{4.55}$$

In the limit that the slopes are very high, we have the following free energy for the two-knot and three-knot states:

$$F_2(L_{k1}, L_{k2}, R_b) = -\log(z_2) = f_1(L_{k1}, R_b, 0) + f_1(L_{k2}, R_b, 0) + G_{k1}\beta_1 + (G_{k2} + 2G_{k1})\,\beta_2 \tag{4.56}$$

and

$$F_3(L_{k1}, L_{k2}, L_{k3}, R_b) = -\log(z_3) = f_1(L_{k1}, R_b, 0) + f_1(L_{k2}, R_b, 0) + f_3(L_{k3}, R_b, 0) + G_{k1}\beta_1 + (2G_{k1} + G_{k2})\beta_2 + (2G_{k1} + 2G_{k2} + G_{k3})\beta_3 \quad (4.57)$$

Equations 4.56 and 4.57 correspond to the free energy of "ground states" where the knots are stacked in single file, with no gaps, directly abutting the barrier (figure 4–6(g) gives the ground state for a configuration with two-knots).

As our knot-formation measurements include all knot sizes, the final step is to compute partition sums formed from integrating over the space of all knot sizes. We define $Z(m, R_\mathrm{b})$, the partition function for a system with m knots:

$$Z(m, R_\mathrm{b}) = (2P)^{-m} \int z_m(L_{k1}, \cdots, L_{km}, R_\mathrm{b}) dL_{k1} \cdots dL_{km} \quad (4.58)$$

We normalize to the Kuhn length to ensure that the partition function is dimensionless. Note that the total free energy associated with a knotting event with m knots (i.e. the free energy including all knot-sizes) is given by

$$F_\mathrm{tot}(m, R_\mathrm{b}) = -\log Z(m, R_\mathrm{b}) \quad (4.59)$$

The partition function for finding zero knots $Z(0, R_\mathrm{b}) = 1$. We can define a total knot partition function:

$$Z_\mathrm{tot}(R_\mathrm{b}) = \sum_{i=1}^{n_k} Z(i, R_\mathrm{b}) \quad (4.60)$$

where n_k is the maximum number of knots observed in an event. The probability of finding a state with m knots is then:

$$P(m, R_\mathrm{b}) = Z(m, R_\mathrm{b}) / Z_\mathrm{tot}(R_\mathrm{b}) \quad (4.61)$$

The probability of finding a state with *any* number of knots (i.e. a state with at least one knot) is:

$$P_{\text{all knots}}(R_{\text{b}}) = \frac{Z_{\text{tot}}(R_{\text{b}})}{1 + Z_{\text{tot}}(R_{\text{b}})}. \tag{4.62}$$

This model can be readily solved via direct numerical integration of Eq. 4.58 with $z_m(L_{k1}, \cdots, L_{km}, R_{\text{b}})$ determined by Eq. 4.33, 4.34, 4.35, 4.38, and 4.39 (determine knot free energy contributions), Eq. 4.41-4.43 (determine knot free energy profile) and lastly Eq. 4.51, Eq. 4.54 and Eq. 4.55 (knot formation partitions functions for knots of fixed size). In practice, this integration is performed in Matlab by simply evaluating the partition function z_m for a finely sampled range of knot-sizes and summing. The prefactors A_{b}, A_{wk} and A_{h} are then determined from least-squares fitting to experimental data on one- and two-knot-formation probability.

4.9 Knot Statistics in the Low Compression Limit

The knot statistics are Poissonian in the limit that the compression is low and β is small. This can be easily shown analytically if we assume that the knots have equal size. The general form of the partition function for m trefoil knots on a profile characterized by R_b can be summarized by:

$$z_m(L_{k1}, \cdots L_{km}, R_b) = \frac{e^{-\sum_{i=1}^{m} f_1(L_{ki}, R_b, 0)} e^{\sum_{i=1}^{m} G_{ki}\beta_i}}{\prod_{i=0}^{m-1} \left(e^{2G_{k,m-i}\sum_{j=0}^{i}\beta_{m-j}} - 1 \right)} \tag{4.63}$$

Assuming that the knots have equal size $G_{k1} = G_{k2} = \ldots = G_{km} \equiv G_k$, $\beta_1 = \beta_2 = \ldots = \beta_m \equiv \beta$ and $f_1(L_{k1}, R_b, 0) = f_1(L_{k2}, R_b, 0) = \ldots = f_1(L_{km}, R_b, 0) \equiv f_1$, so we can rewrite equation 4.63 as,

$$z_m = \frac{e^{-mf_1 + mG_k\beta}}{\prod_{i=1}^{m} (e^{2iG_k\beta} - 1)}. \tag{4.64}$$

Equation 4.64 gives the partition sum for the formation of m identical knots on a compressed DNA molecule. In the limit of low compression, hence small free energy slope β, we can approximate equation 4.64 using Taylor expansion,

$$z_m \approx \frac{e^{-mf_1+mG_k\beta}}{(2G_k\beta)^m m!} = \frac{1}{m!}\left[\frac{e^{-f_1+G_k\beta}}{2G_k\beta}\right]^m = \frac{1}{m!}z_1(L_k,R_b)^m \tag{4.65}$$

The probability of forming m knots is then

$$P(m) = z_m/\sum_{j=0}^{\infty} z_j = \frac{1}{m!}z_1^m e^{-z_1} \tag{4.66}$$

which has the form of a Poisson distribution. In the inset to Fig. 3–5 we show numerically that partition functions that include all knot sizes also show Poisson behavior for low compression.

Experimentally Observed Baseline Knotting Probability

According to [187], there is a finite probability that a long DNA molecule becomes knotted in bulk through thermal fluctuations. In our study, on the other hand, the experimentally observed baseline knotting level is negligible (Figure 3–4). Note that we are working with a low salt buffer (10 mM) tris which gives rise to a high DNA effective width w =17 nm. A high effective width is known to suppress knotting in bulk [41]. Simulations investigating knotting along nanochannel extended chains, which predict a high knotting probability, use low effective widths (=2 nm). Moreover, if knots could form on nanochannel extended chains with moderate or even low probability we expect that we would have seen them, yet we have never observed knots to form spontaneously on nanochannel extended DNA molecules (working in buffers $<$ 100 mM). While we do not expect the equilibrium knotting probability to be exactly zero, even for a high effective width [11], there are additional *kinetic* limitations preventing formation of knots along a nanochannel confined chain in the absence of flow. For example, the

free energy cost for a back-fold to form is high in a nanochannel; it is possible that this makes knot-formation along an extended chain highly unfavorable (i.e. give rise to a high free energy barrier for knotting for an equilibrium extended chain, even if the equilibrium knotting probability is not zero). Of course, this barrier will also be a strong function of the effective width, so it is possible that knots could form spontaneously in high salt conditions.

4.10 Comparison to Experiment

Our knot formation theory agrees well with the experimental measurements of knotting probability (Fig. 3–4). In particular, we capture the increasing knot formation probability with increasing pressure (decreasing edge extension) and explain the non-monotonic behavior of the single knot formation probability, suggesting that our model provides an explanation of our experimental findings. Figure 3–5 compares the probability measurements obtained from the experimental data, the probability distribution predicted by the Poissonian model and our fitted model for different ranges of barrier extension R_b. At low pressures, the Poissonian model describes the experimental results very well. At relatively higher pressures (smaller R_b), on the other hand, Poissonian model fails to explain the experimentally obtained probabilities. We argue that the high compression limit where the Poisson model fails corresponds to knots forming in a limited range of positions close to the barrier at X=0. Simultaneous least-square fitting to the experimental one- and two-knot formation probabilities yields: $A_{\mathrm{b}} = 1.43 \pm 0.05$, $A_{\mathrm{wk}} = 0.98 \pm 0.12$ and $A_h = 1.12 \pm 0.07$. The values of determined prefactors are on order of unity, required for self-consistency of the approach.

A feature of our free energy model (Eq. 4.33) is that it explicitly reduces to Dai *et al.*'s knotting free energy in the case where one knot is present and $V = 0$ (e.g. so that

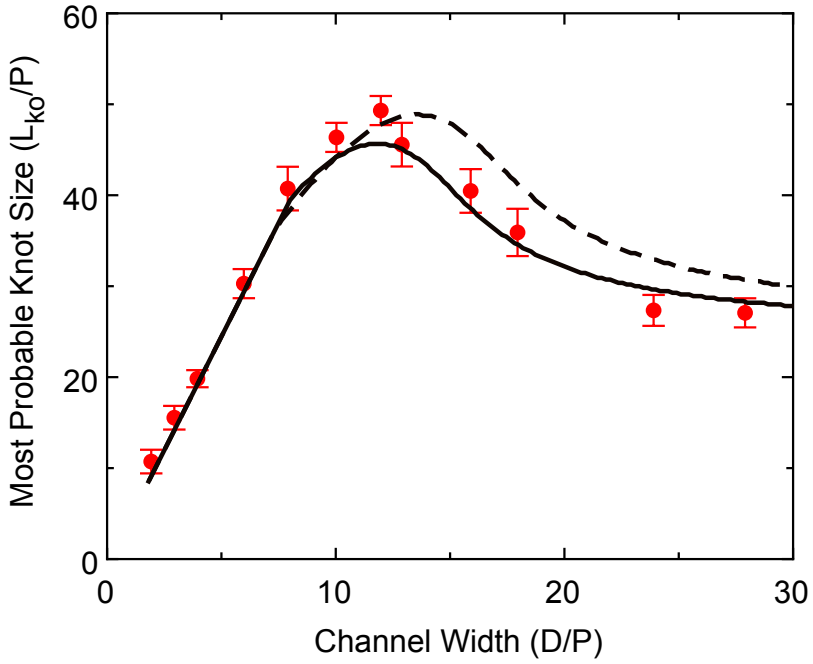

Figure 4–7: Red circles show simulation results for most probable knot size from Dai *et al.* [11]. Bold curve shows prediction for most probable knot size found using Eq. 4.33 with $R = 1$ and $f_{\rm wuk}$ evaluated using Eq. 4.35 (L_b^o determined via classic de Gennes theory). Dashed curve shows prediction for most probable knot size found using Eq. 4.33 with $R = 1$ and $f_{\rm wuk}$ evaluated using Eq. 4.35 (L_b^o determined using the extended de Gennes theory). We have set $P = 0.4w$ to match the parameter values used in Dai *et al.*'s simulation. We have also increased the value of $A_{\rm wuk}$ by 1.5 to ensure that the free energy of the unknotted polymer includes the effect of the spring terms in Eq. 4.35 (in no-flow equilibrium, when $R = 1$, the quadratic spring term only has the effect of adjusting the scaling prefactor).

$R = 1$ everywhere along the chain). The most probable knot size is then determined by minimizing $f_1(L_k, R = 1)$. We can explore if the values of $A_{\rm b}$ and $A_{\rm wk}$ obtained here can also describe Dai *et al.*'s simulation results for a chain in a no-flow equilibrium, a further test of self-consistency. Figure 4–7 shows Dai *et al.*'s simulated most probable knot size versus channel width compared to the prediction of our free energy model (e.g. value of L_k that minimizes Eq. 4.33, with $R = 1$ and P set to the value used in Dai *et al.*'s simulations, $P = 0.4w$). We also show the model prediction using classic

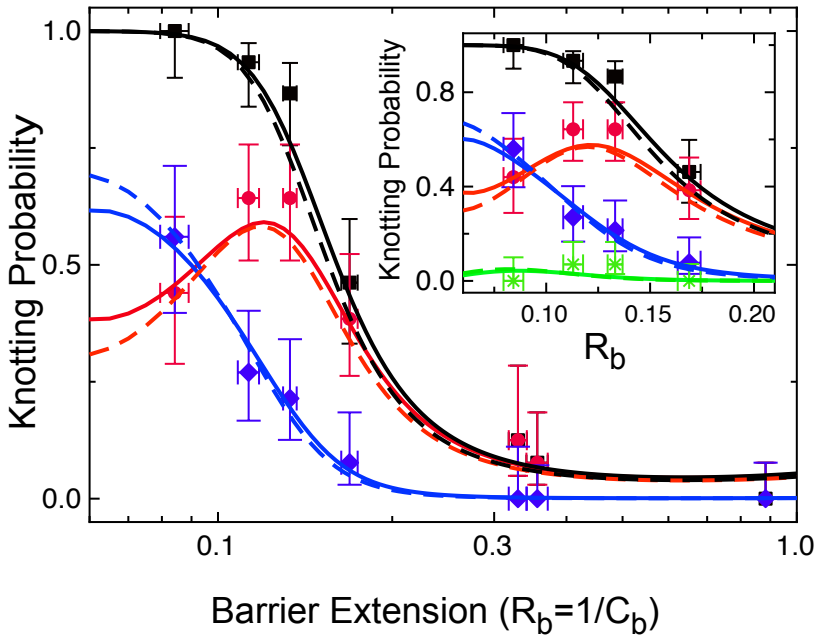

Figure 4–8: The bold curves show the original three parameter fit yielding $A_{\rm b} = 1.43 \pm 0.05$, $A_{\rm wk} = 0.98 \pm 0.12$ and $A_h = 1.12 \pm 0.07$. The dashed curves show a one-parameter fit fixing $A_{\rm wk} = 1$ and $A_h = 1$ and fitting $A_{\rm b}$ (the approach gives $A_{\rm b} = 1.46 \pm 0.01$).

de Gennes bob statistics in Eq. 4.35 to estimate $f_{\rm wuk}$ (classic de Gennes theory, used by Dai *et al.*, is a more appropriate choice for channel widths exceeding around 400-500 nm, covering the range where the most probable knot sizes reach a maximum). The classic de Gennes prediction does not agree exactly, underestimating some of the points near the peak by around 10%, but it reasonably describes the overall trend in the simulated results.

Alternatively, theory suggests fixing $A_{\rm wk} = 1$ and $A_h = 1$ and then performing a *single parameter fit* to fix $A_{\rm b}$. This approach gives agreement of equivalent quality with $A_{\rm b} = 1.46 \pm 0.01$ (see Fig. 4–8 for the comparison between the one parameter and three parameter fits). Note that the value of $A_{\rm b}$ is most critical to the overall agreement: too low and the knotting probability is too high for all R_c, too high and the knotting

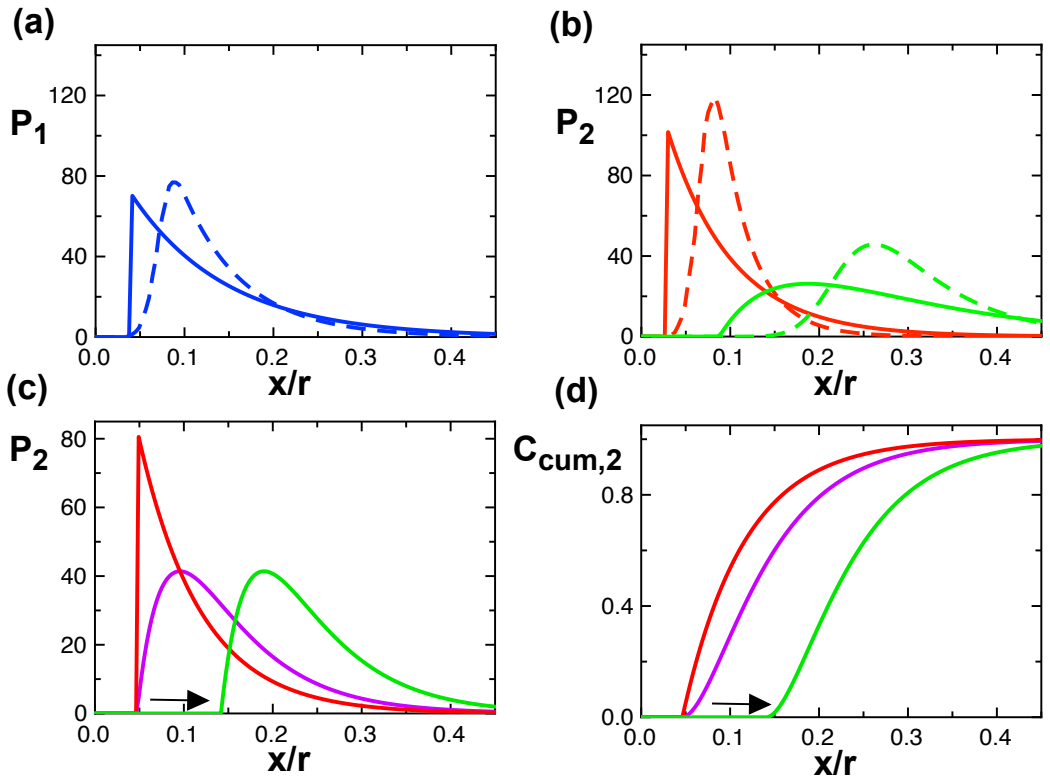

Figure 4–9: (a) The predicted probability distribution of knot position in a one-knot state P_1 for a knot with the most probable knot size (bold) and averaging over all knots sizes (dashed). (b) The predicted probability distribution of knot position in a two-knot state for a lower knot with the most probable knot size (P_{21}, bold, red), averaging over all lower knot sizes ($P_{21}^{\rm av}$, dashed red), the most probable upper knot (P_{22}, bold, green) and averaging over all upper knot sizes ($P_{22}^{\rm av}$, dashed, green). (c) Example predicted probability distributions for knot position along profile for a two-knot state, showing lower knot distribution P_{21} (bold red), upper knot distribution P_{22} (bold green) and the estimated non-interacting distribution P_U (bold purple). (d) Example cumulative probability distributions for knot position along profile for a two-knot state, showing cumulative lower knot distribution C_{21} (bold red), cumulative upper knot distribution C_{22} (bold green) and the estimated cumulative non-interacting distribution C_U (bold purple). The arrows show the shift between the predicted non-interacting upper knot (P_U, C_U purple) and predicted interacting upper knot (P_{22}, C_{22} green). Note that the lower knot distributions and non-interacting distributions in (c) and (d) (bold red, purple curves) have been computed using the most probable size of knots in a one-knot state.

probability is suppressed relative to experiment. We believe that the slightly higher value of A_b relative to unity is required because our model may underestimate the knot confinement free energy. For example, slight compression of knots at high forcing, an effect not taken into account in our model, would lead to additional positive free energy contributions. The value of A_b might need to be increased upon fitting to compensate for the absence of these effects. We feel that these effects would be worthwhile to explore in future theoretical studies of knot-formation on compressed chains.

Note also that the predicted knot sizes are consistent with our single-file ordering hypothesis: two-knots cannot pass in the channel when the ratio of the knot diameter to largest channel dimension $2g_k/D_2 > 0.5$. For the most probable knot size in a one-knot state, our theory predicts $2g_k/D_2 > 0.6$ over the entire range of R_b values (degrees of compressed) used, so our theory self-consistently requires that single-file ordering occurs for multiple knot states.

Our statistical mechanical model also gives predictions for the spatial distributions of knots in steady-state (prior to pressure release). However, caution must be used in comparing our experimental results after pressure release with the predicted distribution in the compressed steady-state. Once the pressure is released, the hydrodynamic flow is no longer present, reducing the slope β of the free energy profile. In addition, during relaxation the concentration profile will flatten out with the profile becoming perfectly uniform in the long-time limit when the no-flow equilibrium is reached. These effects will lead to a dynamic variation in the free energy landscape felt by the knots which in turn will lead to a dynamic variation of the knot distribution. However, while these considerations mean we can't use the distributions following pressure release to draw strict quantitative conclusions about the knot distributions in steady-state, we can draw some qualitative conclusions. The one-knot distribution following pressure

release has an exponential character (see Fig. 4–5(a,c)) with the one-knot probability highest at the slit-barrier. We can infer in steady-state that the knot distributions are also non-uniform with the knots accumulating at the slit-barrier ($x = 0$). Note that the relaxation process must increase entropy and therefore cannot introduce spatial non-uniformity into a profile that was initially uniform (i.e. it is not possible that the non-uniformity we observe was introduced by the relaxation process itself). In particular, we can say that the distributions observed after pressure release represent a lower-limit on the degree of spatial non-uniformity present in steady-state.

The lower knot distribution predicted by our theory, P_{21}, is given by Eq. 4.47; the upper knot distribution predicted, P_{22}, is given by Eq. 4.49 (see Fig. 4–9(b)). If the two knots experience the same free energy profile with identical slope β, note that P_L and P_{21} have the same forms (compare Eq. 4.31 and Eq. 4.47), but that P_U and P_{22} are different. In particular, Eq. 4.49 predicts that P_{22} will be shifted to higher X relative to P_U (see Fig. 4–9(c, d)), with the degree of shifting determined by $2g_k$. In the limit g_k goes to zero, P_U and P_{22} coincide. Thus, we expect that repulsive interactions strong enough to prevent knot crossing will shift the upper knot distribution to higher X relative to P_U, inducing a spatial *segregation* between the distributions. Naturally, while we do not expect quantitative agreement between the steady-state distributions and our measured distributions (due to the complicated dynamics of the relaxation process), any shift observed between the lower and upper knot distributions is indicative of a no-passing condition. While the experimental spatial distribution in fact results from averaging over the distribution of all thermally allowed knot sizes, we find that averaging over all knot sizes has only a moderate effect on the predicted distributions and in fact would shift the upper knot distribution to higher X (see Fig. 4–9(a, b)).

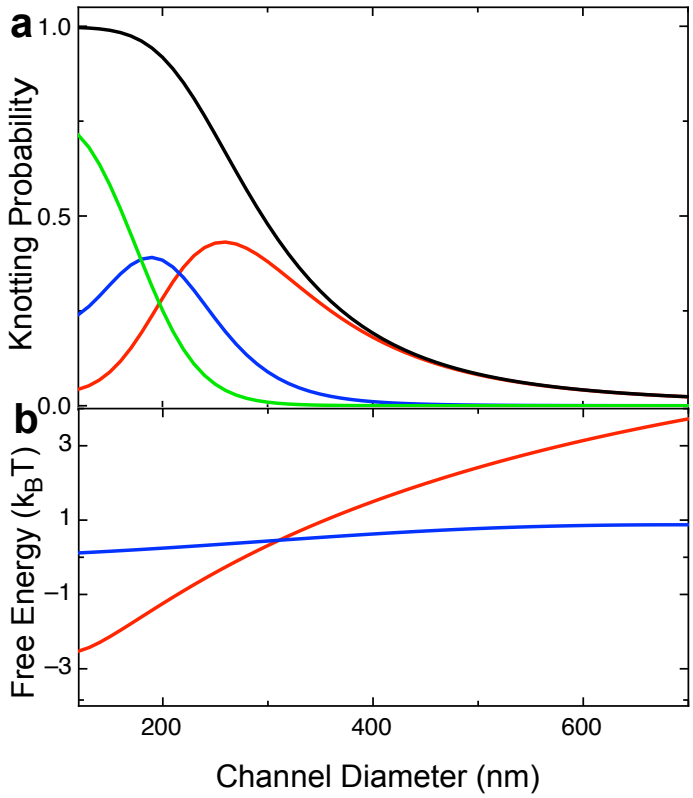

Figure 4–10: (a) Predicted probability as a function of channel dimension of generating an event with one knot (red), two-knots (blue) and three-knot (green) on a profile with $R_b = 0.2$. Note that the large three-knot formation probability would predict that for channel widths below 200 nm, states with even more knots would be generated (although we do not include these explicitly in the calculation). The total probability of generating an event with knots is the black-curve. (b) The free energy of a single-knot state (red) and two-knot interaction free energy (blue) versus channel width. For simplicity, the channels have aspect ratio unity.

4.11 Model Predictions for Varying Channel Diameter

Our model gives predictions for how varying channel dimension should alter knotting probability in the extended de Gennes regime. We find that increasing confinement is predicted to increase the probability of knot generation (see Fig. 4–10(a)). This is unsurprising given that increased confinement, by increasing the f_{wuk} term in the free

energy will tend to enhance the free energy saved upon forming a knot, leading to a decreased single-knot free energy (Fig. 4–10(b), red-curve). Our model also predicts that more multiple knot states would be generated for lower channel width (Fig. 4–10(a)), due to the decreasing knot interaction free energy (Fig. 4–10(b), blue-curve), which falls for channels below 500 nm due to the decreasing knot size with lower channel dimension (see Fig. 4–7, the channel dimension necessarily places an upper limit on the knot size in our model). It would also be interesting to perform analogous experiments in the transition (< 100 nm) and sub-persistence Odijk regimes (< 50 nm). We do know that knots can be formed in channels on order of the persistence length [13], but our model will break down here as we do not expect the free energy scalings to extend to such small channels.

4.12 Experimental Error on Knotting Probability

Each knotting probability measurement (i.e. each data point in Fig. 3–3 (b) and Fig. 3–4) was obtained from around $10 - 15$ individual single-molecule compression-relaxation events. The compression ratio and waiting time shown were determined from averaging over all single molecule events corresponding to a given data point. The horizontal error bars give the corresponding error on the mean over the single-molecule data.

In the current study, we observe three possible outcomes: formation of 0, 1 and 2 knots ($i = 0, 1, 2$). The maximum likelihood estimate for the probability of a knotting state is $\hat{P}_i = n_i/N$ where n_i is the number of molecules in the given knotting state and N is the total number of events corresponding to the data point in question. The maximum likelihood estimate of total knotting probability is $\hat{P}_{\text{total}} = (n_1 + n_2)/N$. We report the maximum likelihood estimate of knotting probability in Figure 3–3 (b) and

Figure 3–4. Our vertical error-bars are defined so that the true probability value lies within the bar with confidence $1-\alpha$ with $\alpha = 0.32$ (for a standard one-sigma error). We argue that binomial statistics are appropriate for our measurements: we either find that the measurement falls within the specific category or outside the category (single knot present or not, two knots present or not, knots present or absent). We then estimate the confidence interval using the Wilson score interval [174]. Let $z^2 \equiv z^2_{\alpha/2}$, where $z_{\alpha/2}$ is the critical value of the normal distribution for error level α. The true knotting probability P_i then lies within the following interval with confidence $1-\alpha$:

$$P_i \in (P_i' - \sigma_n, P_i' + \sigma_n) \tag{4.67}$$

where

$$\hat{P}_i' \equiv \frac{z^2 + 2n_i}{2(N+z^2)} \tag{4.68}$$

$$\sigma_n \equiv \frac{z}{2(N+z^2)} \left(\frac{z^2 + 4n_i(N-n_i)}{N} \right)^{1/2} \tag{4.69}$$

For a one-sigma error, $z_{\alpha/2} = 1$. Note that the confidence intervals are not distributed symmetrically around the maximum likelihood estimate $\hat{P}_i$. We thus use asymmetric error bars to report error on knotting probability, with the top and bottom error bars defined via $\sigma_t = \hat{P}_i' + \sigma_n - \hat{P}_i$ and $\sigma_b = \hat{P}_i - (\hat{P}_i' - \sigma_n)$, respectively.

CHAPTER 5
Complex Knotting Dynamics on Hydrodynamically Compressed Nanochannel-Extended DNA

In the knot factory introduced in chapter 3, a free energy model was proposed for knot formation on nanochannel-confined DNA molecules based on the assumption that knots interact via hard-core repulsion and maintain their single-file ordering on a linear potential. In addition, knot identification criteria were provided on the relaxing DNA molecules based on the observed intensity of the knotted features. This was mainly due to the fact that the knotted features in biopolymers have compact structures and are optically unresolved, thus the topological information is inaccessible. In this chapter, we employ the method of hydrodynamic compression introduced in chapter 3 to report a series of results obtained from observation of knots on DNA molecules following polymer compression. We observe that compression can impose entanglement on DNA contour inside the *knot region*. The trapped DNA contour is subsequently released following expansion of the knot. In some cases, one apparent knot splits into two distinct knots, representing a level of entanglement existing between the two knots prior to their split-up. Moreover, we confirm that two knots maintain single-file ordering on DNA contour and knots interact with each other via hard-core repulsion. Finally, we show that hydrodynamic flow helps unravel knots formed on nanochannel-confined DNA and argue that *step-wise* knot unraveling unveils information about the knot complexity, introduced as the essential number of crossings in the knot contour.

The contents of this chapter have been prepared as a manuscript ready for submission:

Complex Knotting Dynamics on Hydrodynamically Compressed Nanochannel-Extended DNA

Susan Amin, Ahmed Khorshid and Walter Reisner

5.1 Introduction

In a sufficiently long three-dimensional self-avoiding walk, knot probability approaches unity [30]. This is indicative of high probability of knot existence on long chains in nature. In biological systems, knots have been found on umbilical cords and catheters [4] as well as biopolymers such as capsid DNA [38] and proteins [5]. In nanofluidic systems, knots prevent linearization of DNA in the knot region, obscuring sequence information stored in the knot region [20]. Studies have revealed that knots [15, 44] and entanglements [188] strongly influence the statics and dynamics of polymers. Knots decrease the relaxation time of the stretched molecules in elongational fields [14] and reduce the tensile strength of chains [15, 16]. Moreover, there is interdependence between knot dynamics and the global dynamics of the chain. For instance, relaxation of a stretched knotted DNA molecule leads to the swelling of the knot at a time scale comparable to the global relaxation time of the molecule [46] and knot motion toward the end of the polymer can be controlled by varying the local chain tension using elongational fields [189]. Di Estefano *et al.* [190] find that applying longitudinal DC electric field on a stretched knotted DNA model leads to the knot drift in the direction of the electric force, while AC electric field with sufficiently high frequency suppresses the motion of knots. Furthermore, different methods of knot formation on DNA molecules have been proposed. Knot induction via self-assembly of nucleic acids [17], optical trapping [9], exposure to AC electric field [18] and hydrodynamic compression of DNA molecules [163] are some examples of knotting methods. In our recent study, we reported multiple-knot events induced by hydrodynamic compression

of single DNA molecules in nanochannels [163]. In our proposed model, we showed that the assumption of hard-core repulsion between knots in a linear potential leads to the single-file diffusion of knots and affects the distribution of knots on a chain. Thus, observing how knots interact with one another and with the surrounding chain can give an insight to the complex knot dynamics on nanochannel-confined DNA molecules.

A knot is mathematically well defined on a closed nonintersecting curve, as a self-entanglement that cannot unravel [191]. In open chains, the same definition holds if we imagine that the two ends are connected [4, 21]. Knots are categorized into prime and composite knots. Prime knots are the knot types that cannot be decomposed into simpler knots, while composite knots are made up of two or more prime knots [43]. Knots are also categorized based on their number of essential crossings. In polymers, if the knot is far enough from the ends of the chain, the knot complexity can be determined after closing the chain [192]. Knot complexity plays a significant role in determining the extent of knotting effect on polymer dynamics [44, 193]. Knot complexity can be found by either counting the number of crossings in the two-dimensional projection of the knot [4] or using any of the proposed polynomial knot invariants [194] such as Alexander polynomial [195] and Jones polynomial [196], which assign to each knot type a polynomial with integer coefficients. However, employing these methods requires visual access to the knot topology, while on biopolymers knots tend to form compact structures (e.g. metastable knots) [9, 13, 18, 46, 155]. Therefore all the topological information is contained in a region below the instrumental resolution and is thus inaccessible. *Unknotting number* provides more detailed information regarding the complexity of a microscopic knot. Unknotting number is defined as the minimum number of times a knot should pass its end through a loop to unravel [191].

Here, we employ the method of hydrodynamic compression [163] to produce knots on nanochannel-confined DNA molecules and investigate knot interaction with their surrounding contour. In particular, the entanglement between two knots on a chain as well as single-file ordering of two knots are visualized on position-time kymographs. The events illustrate that entanglement exists between two knots formed via hydrodynamic compression. This entanglement appears as a sharp localized intensity, resembling one larger knot. The knots untangle later in time due to thermal fluctuations, leading to self-reptation and the diffusion of knots, which in turn split the apparent large knot into two smaller separate knots. Furthermore, results on knot-knot interactions confirm the existence of hard-core repulsion between knots, which imposes single-file ordering of knots on confined chains. Finally, using the property of *unknotting number* in different knot types, we eliminate the probability of a knot belonging to a knot category. We show that step-wise unraveling of microscopic knots provides useful information regarding knot complexity, which is otherwise inaccessible, as the microscopic knots form compact structures that cannot be resolved.

5.2 Experimental procedure

The experiments are performed on $1\times1\text{cm}^2$ fused silica chips (HOYA). Nanochannels with blunt ends are patterned at the center of the chips using contact UV and electron beam lithography [126](See Fig. 5–1(a)). A 30-nm deep slit (as measured using profilometry) is etched over the array. The slit changes the blunt ends to barriers that allows the solvent to flow while trapping DNA molecules. Two U-shaped microchannels (1μm deep and 50μm wide) are subsequently patterned on the substrate, connecting the sand-blasted reservoirs to the nanochannel array. The chips are then permanently bonded to fused silica coverslips (Valley Design). The nanochannels have dimensions of $300\times300\text{nm}^2$ measured with SEM. T4 bacteriophage molecules (Nippon

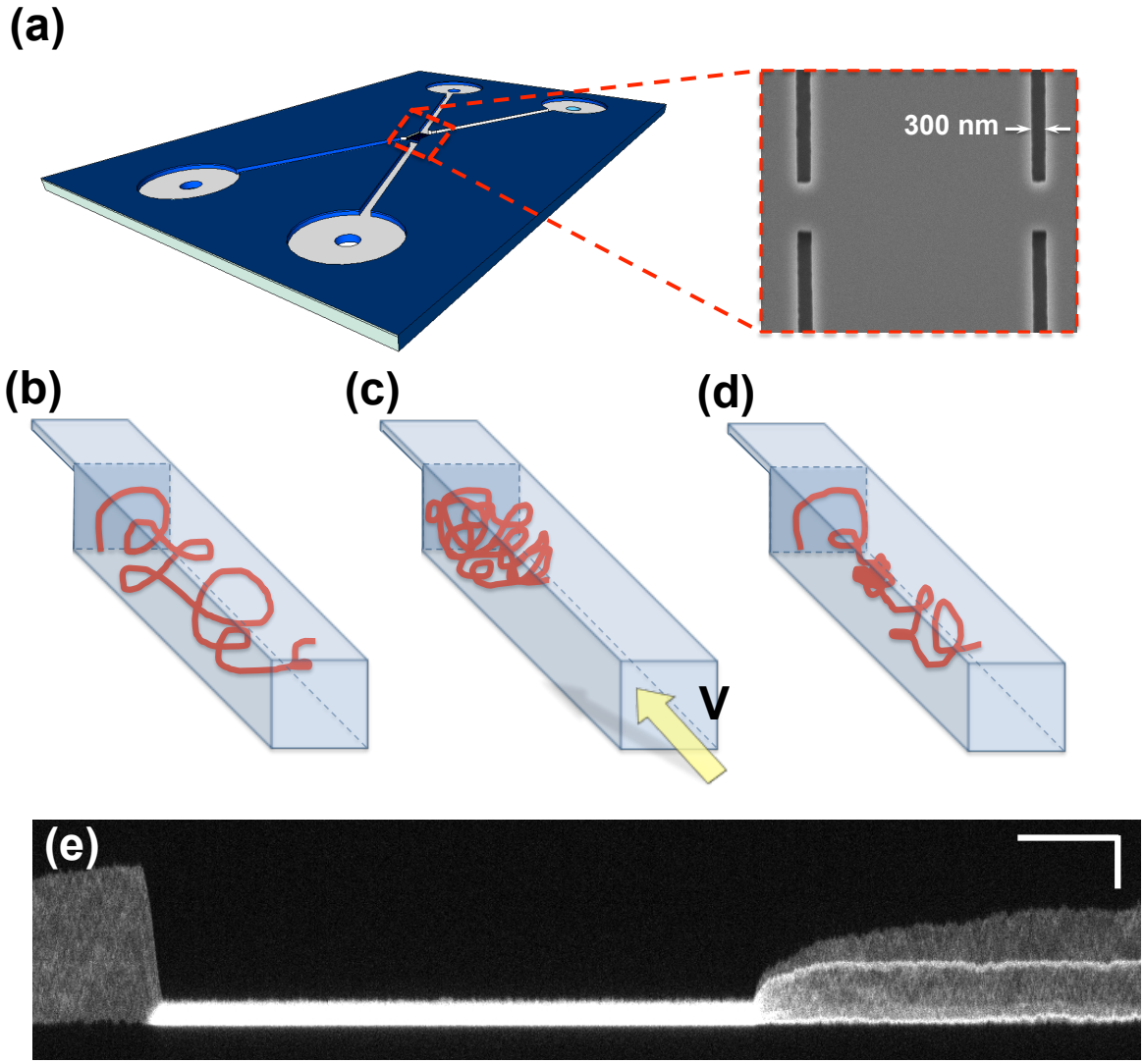

Figure 5–1: Schematic of the experimental set-up. (**a** (*left*))The device is a $1\times1\text{cm}^2$ chip, made of fused silica. The T4 DNA solution is introduced into the microchannels through the sand-blasted holes; (**a** (*right*)) a magnified view of the nanochannels obtained via SEM imaging. (**b-d**) A cartoon showing the steps of knot formation; the molecule is driven into the nanochannel (**b**), compressed via pneumatically-induced hydrodynamic flow with velocity V (**c**) and relaxed (**d**). Knots are present on the DNA molecule. (**e**) A kymograph illustrating knot formation following compression of a DNA molecule. The vertical axis shows the molecule intensity along the nanochannel (scale bar $5\,\mu$m) and the horizontal axis depicts time (scale bar 5 s).

Gene, 166 kbp) are stained with YOYO-1 (Life Technologies) at an intercalation ratio of 10:1. The buffer used in the DNA solution is 10 mM Tris (pH 8.0) and 1.5% β -mercaptoethanol (BME) is added to suppress photobleaching and photonicking. An inverted microscope (Nikon Eclipse Ti-E) with a 100X N.A. 1.5 oil immersion objective is used for visualizing the DNA molecules in the chip. An EMCCD camera (ixon, Andor) is used for imaging with excitation illumination provided by a metal-halide lamp (Xcite).

Pneumatic pressure is applied on the chip through the sand-blasted loading holes to drive T4 DNA into the nanochannels (Fig. 5–1(b)). The DNA molecules are subsequently compressed in a low Reynolds number flow (Re$\sim 10^{-8}$), imposed by pneumatically actuated pressure drop along the nanochannels (5–1(c)). The extension of the DNA molecules at equilibrium state (e.g. prior to compression) is $r_o = 14.6 \pm 0.3\mu$m and the compressed extension at steady state is $r \approx 0.2 r_o$. The polymer is held compressed for *waiting time* $t_w \simeq 30$ sec. The pressure is then released, leading to the relaxation of the molecule (Fig. 5–1(d)). Knots are apparent on the chain with high probability ($P > 0.9$). Each molecule is compressed only once and then driven out of the nanochannel. A new molecule is then introduced to eliminate possible effects of a series of consecutive compression events on knot properties. To study the unknotting process, where knots are not close enough to the end of the DNA molecules, we observe that applying a small hydrodynamic velocity $V \sim 0.5$ μm/sec guides the knot on the chain towards the opposite end of the polymer and accelerates the unraveling process.

5.3 Results and Discussion

Knot-contour interaction

Knots appear as sharply localized high intensity features on the DNA molecules that maintain their size, known as metastable knot size [9, 13, 18, 46, 155]. Figure 5–1(e)

shows the compression-relaxation process of a single DNA molecule. Our observations show that knots usually reach their metastable configuration relatively fast during the first few frames following compression. However, in some instances bright features with relatively constant size during relaxation either shrink or split into two distinct features diffusing on the chain. We have observed two distinct types of phenomena, which we discuss below.

Case 1: Entangled DNA contour trapped in the knot region. Here, wide bright features maintain their size for several seconds following compression followed by a sudden shrinkage in size (Figure 5–2 (a)). We believe that during hydrodynamic compression of the molecule, the entangled contour is forced into the knot region. During molecule relaxation the knot is tightened, trapping the entangled portion inside the knot region (Fig. 5–2(b))). As the knot diffuses along the chain, self-reptation of the knot eventually leads to the ejection of the entangled contour from the knot region, creating a folded structure in the vicinity of the knot (Fig. 5–2(c)). Thermal fluctuation of the molecule inside the nanochannel eventually linearizes the folded region (Fig. 5–2(d)).

Case 2: Entangled knots resembling a larger knot. In the second case, the wide bright feature breaks into two distinct highly localized features (Figure 5–3 (a)). We argue that the knot is in fact a composite knot composed of two simpler entangled knots (Figure 5–3 (b)). Following the pressure release, thermal fluctuation of DNA segments releases the entangled contour and drive the two knots apart. Due to the resolution limits, the two knots are not discernible during the first few seconds, however later after the DNA relaxes (Fig. 5–3(a)), they split up and form two distinct bright localized features that diffuse independently along the chain (Fig. 5–3 (a,c)). In fact, hydrodynamic compression of DNA molecule induces high level of entanglement between the two knots, which restricts the dynamics of each knot to the other. Thus,

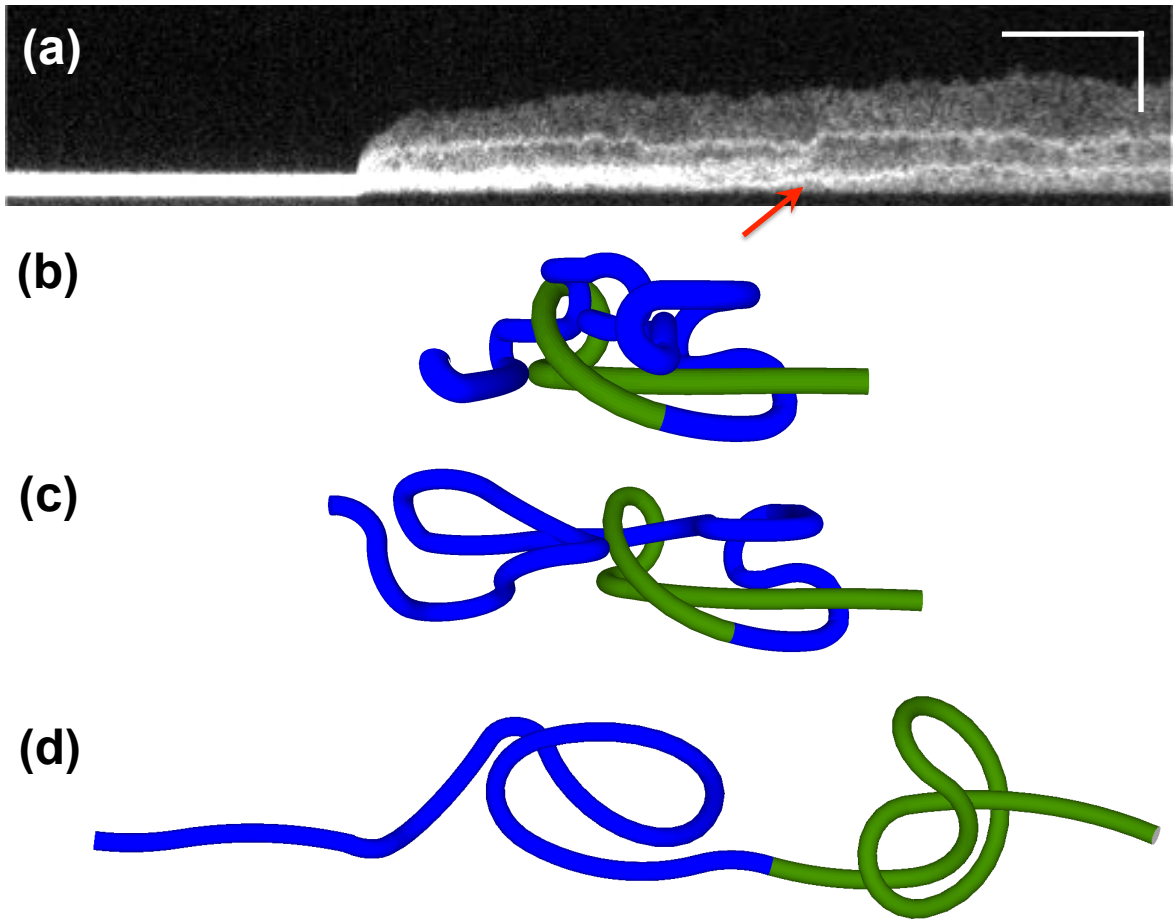

Figure 5–2: (**a**) An example of a kymograph illustrating the relaxation process in a case where a large knot shrinks into a smaller knot (depicted with red arrow). The vertical axis shows the intensity along the nanochannel (scale bar 10 μm) and the horizontal axis represents the time (scale bar 2 sec). Notice a second knot on the upper part of the DNA with a relatively fixed size, which exists throughout the event. (**b-d**) Schematic of the interaction between a knot and the highly entangled contour stored within the knot region; (**b**) a part of the chain is entangled and driven into the knot region via compression of the molecule; (**c**) in the process of DNA relaxation, the knot expels the entangled contour out from the knot region. Due to high level of entanglement, the expelled contour is initially highly folded; (**d**) The released contour eventually unfolds and is linearized in the nanochannel.

for a few seconds following pressure release the two entangled knots appear as a single knot, which maintains its integrity because of high level of entanglement induced

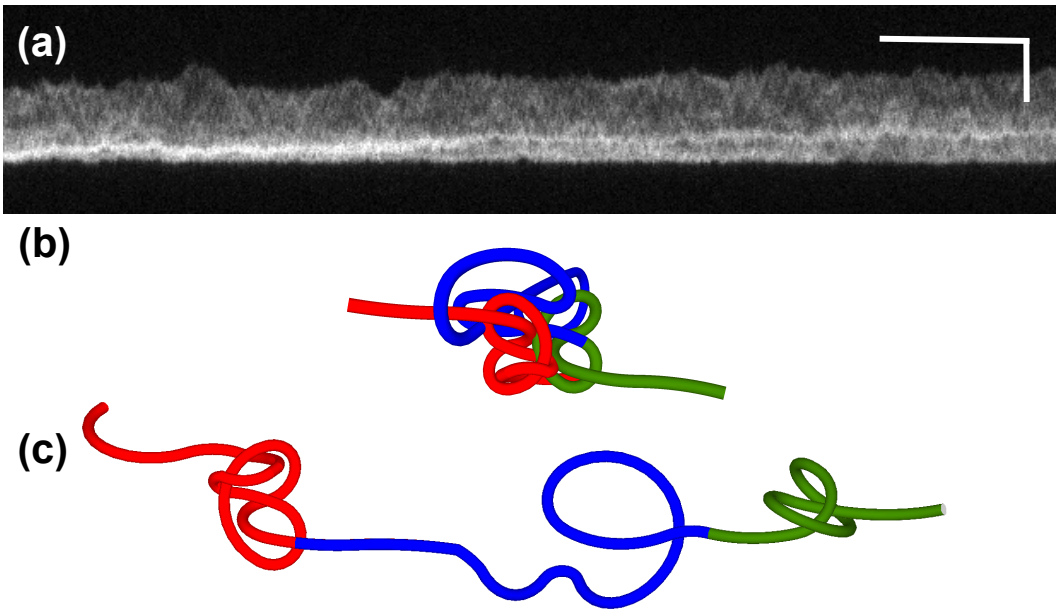

Figure 5–3: (**a**) An example of a kymograph showing the disentanglement of two knots after molecule relaxation. Position (vertical axis, scale bar 10 μm) is plotted versus time (horizontal axis, scale bar 5 sec). A larger knot splits into two smaller knots, one of which unravels later at the molecule's lower free end, indicating that its motion is not restricted by the slit-barrier. During the disentanglement of the knots, the unknotted contour stored between the knots is released gradually. (**b**) A cartoon illustrating the entanglement between two knots induced by hydrodynamic compression. The two knots are shown in red and green. An unknotted part of the chain represented with blue color is trapped between the two entangled knots. (**c**) Thermal fluctuation results in disentanglement of the two knots and the release of the stored contour.

by the hydrodynamic pressure. However, thermal fluctuation of the knots and the surrounding contour will eventually release the contour stored between the two knot regions and split the two knots.

Single-file ordering of knots on the molecule contour. We believe knots formed during hydrodynamic compression have a spatial extent on order of the channel width [163]. This suggests knots cannot pass each other (Fig. 5–4(a)) and must maintain single-file ordering in the linear potential. Studies have shown that in bulk

knots can pass each other via diffusion along the chain contour [178], however this requires swelling of one of the knots to at least the combined size of the two knots added together. As both knots have an extent on order of the channel diameter, this is not possible while nanochannel confinement is maintained. Our observations in the current study confirm the single-file ordering hypothesis on nanochannel-confined molecules. Figure 5–4(b-d) presents several example events where the knots repel each other after approaching a certain distance from one another. In these events, the knot trajectories appear to divert once a certain knot-to-knot separation is reached.

Step-wise knot unraveling. Although unknotting number is not uniquely defined for each knot type, we can categorize different knots based on their unknotting number. Thus, while the number of unknotting steps does not provide a conclusive argument regarding knot topology, unknotting number bears useful information about knot-complexity. By convention, a knot type is shown as C_s, where C denotes the essential number of crossings and s represents the s^{th} knot topology in the sequence of knots with C number of crossings [195]. The unknotting number for a knot is always smaller than half of the knot crossing number C [120] and does not necessarily increase with C or s. For instance, unknotting number in torus knots with the essential number of crossings larger than three is always greater than one. Moreover, composite knots have unknotting numbers of at least two. On the other hand, there exist complex prime knots with essential number of crossings greater than three with unknotting number one. Twist knots belong to this category. For instance, figure eight is a twist knot with four number of crossings, which can unravel after one untying step. Generally, all twist knots have unknotting number equal to one. Thus, if a knot unravels after one unraveling step, it is essentially a prime knot and in the case of crossing number

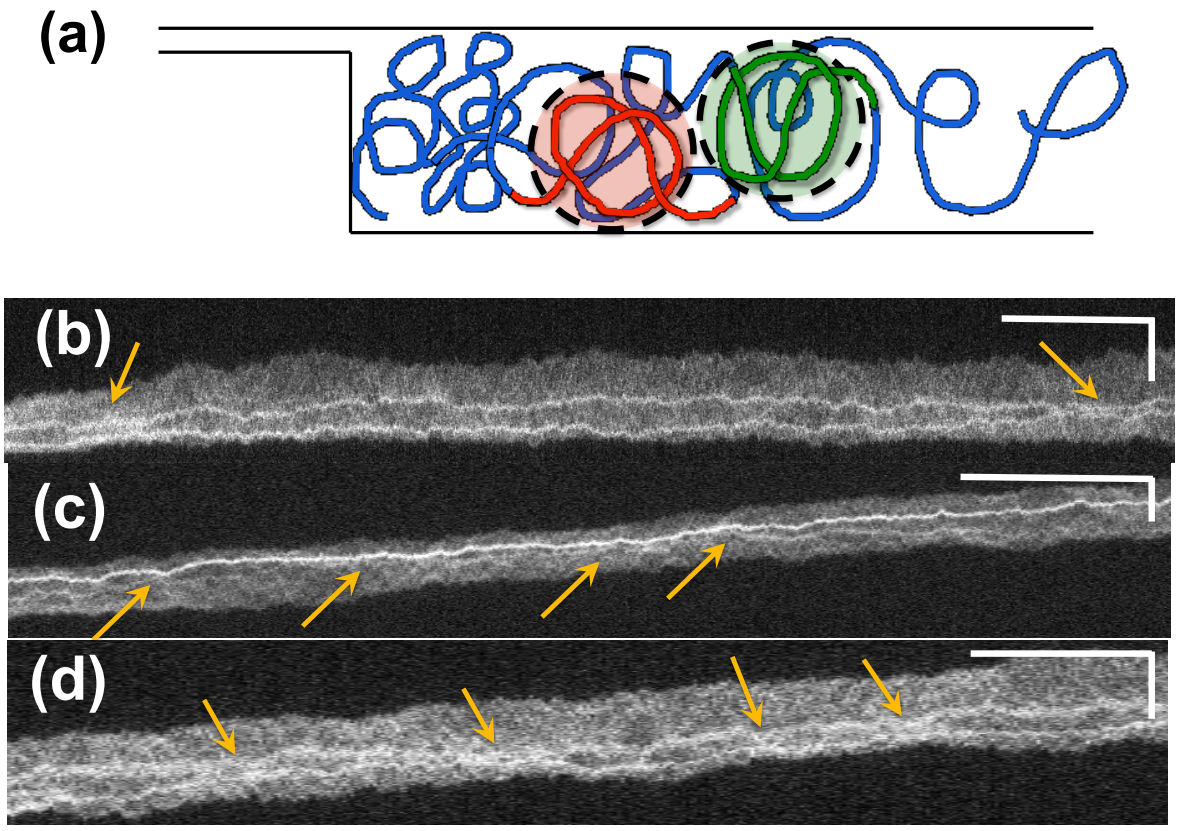

Figure 5–4: (**a**) A cartoon illustrating single-file ordering of knots in nanochannel; The sizes of knots are comparable with channel dimensions, which prevents them from crossing and passing each other. (**b-d**) Examples of kymographs showing the approaching of knots towards each other during and after relaxation. The knots repel each other and maintain their single-file ordering. The orange arrows in all kymographs indicate the times when the two knots are close to each other. The vertical axis in (b)-(d) represent intensity (scale bar 10μ m) and the horizontal axis shows time (scale bar 10 s (b) and 5 s (c,d)).

greater than three, the knot is non-torus. Similarly, if a knot unravels after two steps (Fig. 5–5(e)), we can conclude that the knot is neither trefoil nor figure eight, since both types unravel at exactly one unraveling step. Twist knots with higher complexities, in spite of having unknotting number equal to one can unravel at two or larger number of steps depending on the location on the knot where the unraveling occurs. Figure 5–5 (a-c) illustrates the step-wise unraveling of a torus knot with five essential crossings

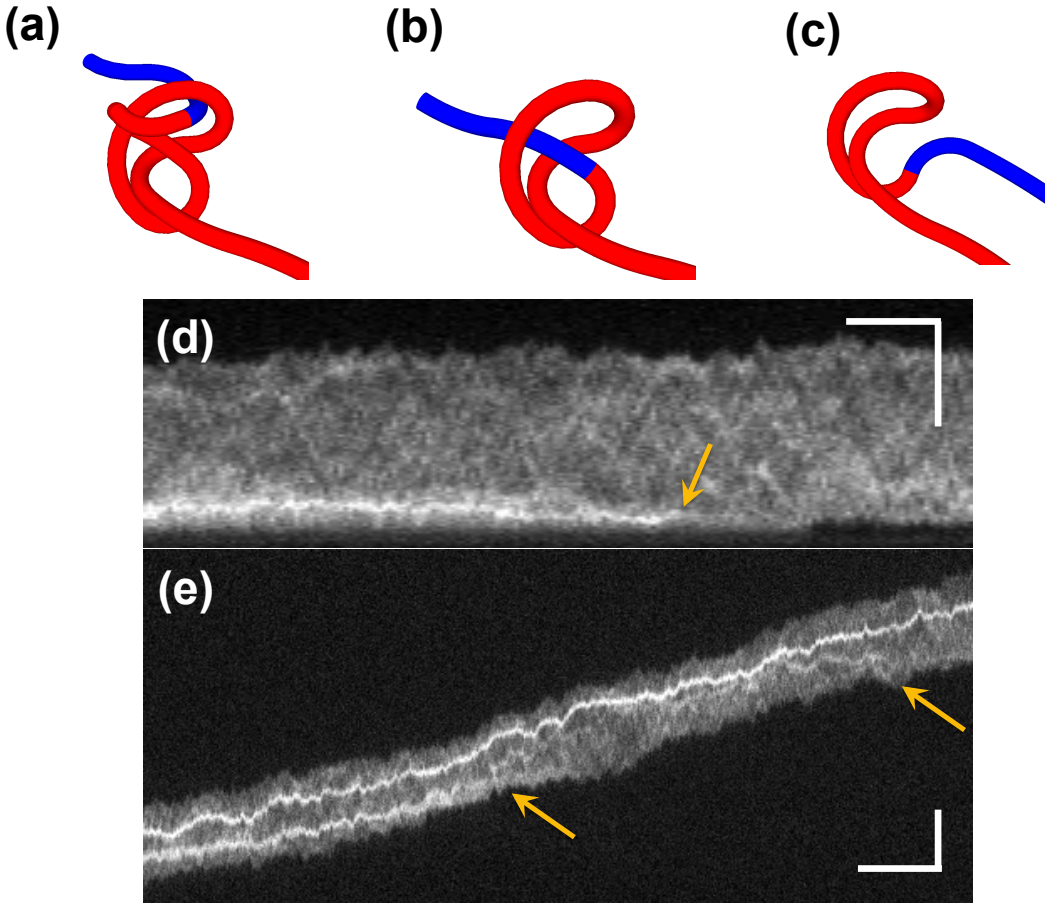

Figure 5–5: (**a-c**) Schematic of the step-wise unraveling process for a complex prime knot; (**a**) the torus knot has five essential crossings. The blue segment shows the end of the knot which is closer to the edge of molecule; (**b**) one of the crossings on the blue-end side unravels and a trefoil knot forms; (**c**) the knot completely unravels. (**d, e**) example kymographs of knot unraveling; (**d**) the unknotting process happens within one unraveling step, indicating that the knot is non-composite. The horizontal and vertical scale bars are 1.5 s and 10μ m, respectively; (**e**) the lower knot unravels in two steps. A hydrodynamic velocity $V = 0.43\mu m/\text{sec}$ is applied to drive the knot towards the lower end of the molecule. The horizontal and vertical scale bars denote 5 s and 10μm, respectively. The orange arrows in (**d** and **e**) show the position of step-wise unraveling on the kymograph.

(Fig. 5–5(a)). The free end of the chain crosses a loop and transforms the knot into a trefoil knot (Fig. 5–5(b)). The trefoil knot subsequently unravels completely after another untying step (Fig. 5–5(c)). The kymographs in Fig. 5–5(d) and Fig. 5–5(e),

show events with one-step and two-step knot unraveling, respectively. In Fig. 5–5(d), the knot unties after one step, which indicates that the knot is neither a composite knot nor a torus knot of complexity greater than three.

We have observed that application of a relatively small hydrodynamic velocity on the order of $V_{\mathrm{h}} \sim 0.5$ to the chain accelerates the knot unraveling process by guiding the direction of knot reptation on the molecule. In Fig. 5–5(e), a small velocity $V_{\mathrm{h}} = 0.43\ \mu m/\,\mathrm{s}$ is applied to the DNA molecule. Figure 5–6 (a,b) presents two knot unraveling events. In figure 5–6 (a), the unraveling happens at $V_{\mathrm{h}} = 0$, while the knot movement in 5–6 (b) is guided by applying a hydrodynamic velocity $V_{\mathrm{h}} \sim 0.44\mu\,\mathrm{m/s}$. Figure 5–6(c) shows that fluctuation in the normalized distance is suppressed in the event where hydrodynamic flow is applied (red circles, figure 5–6 (b)), thus knot unraveling occurs at a faster rate compared to the case where $V_{\mathrm{h}} = 0$ (blue diamonds, figure 5–6(a)).

Upon applying hydrodynamic velocity v_{h}, the molecule starts to move in the hydrodynamic flow direction. We believe that two plausible explanations could describe our findings: (1) Due to the difference in the geometry and width of the knot and the unknotted part of the chain, the drag coefficient for the knot ζ_k is slightly larger than that in the unknotted part ζ_{p}. The knot can be approximated as a ball with drag coefficient $\zeta_{\mathrm{k}} = 6\pi\eta g_k$, where η is the solution viscosity and g_k is the knot diameter. The different drag coefficients lead to two different drag forces $f_{\mathrm{dk}} = -\zeta_{\mathrm{k}} V_{\mathrm{h}}$ and $f_{\mathrm{dp}} = -\zeta_{\mathrm{p}} V_{\mathrm{h}}$ exerted on the knot and the unknotted part, respectively. When the solution is viscous enough, the difference between these two forces becomes significant and suppresses the friction force f_{f} between the knot and the rest of the chain, leading to the relative movement of the knot with respect to the unknotted part of the molecule. (2) As the knot spatial extent is comparable with the channel width, it leads to a

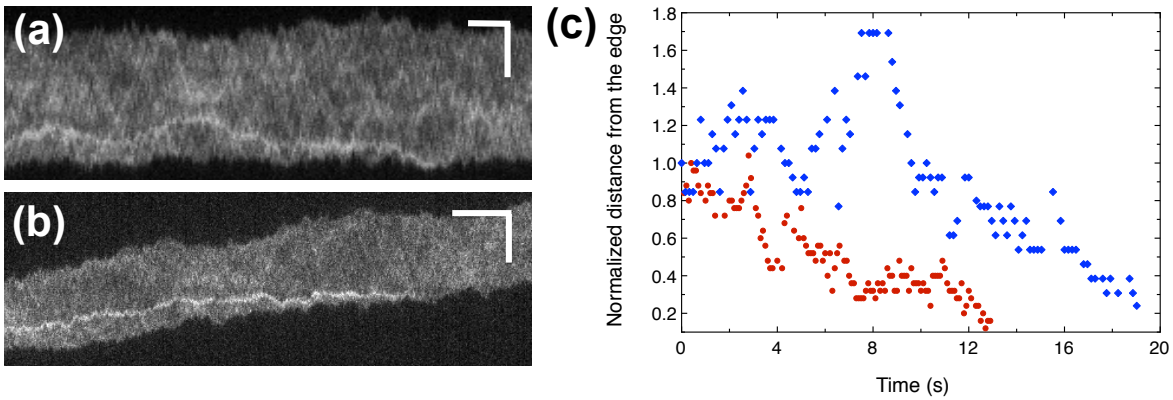

Figure 5–6: (**a-b**) The kymographs present knot unraveling process on chain after DNA relaxation. The vertical axis shows the intensity along the nanochannel (scale bar 5μ m (a) and 10μ m (b)) and the horizontal axis depicts time (scale bar 2 s (a,b)). The hydrodynamic velocity $V_{\rm h} \sim 0.44\mu$ m/s is applied to the molecule in (b) to help faster knot unraveling. The graph in (c) shows the distance between knot and the molecule edge in the process of knot unraveling as a function of time for the events in (a) and (b). The blue diamonds correspond to the kymograph in (a) and the red circles represent the respective values for (b). The distance is normalized to the distance at time $t = 0$.

relatively large friction between the knot and the nanochannel walls compared to that between the knot and the rest of the chain. Therefore, sufficiently small hydrodynamic velocity $V_{\rm h}$ accelerates the self-reptation of the knot on the chain contour leading to the guided relative movement of the knot toward the end of the molecule. For the first explanation to hold, considering that the knot size is much smaller than the channel diameter, the solution must be viscous enough to overcome the friction between the knot and the rest of the chain, which is not the case in the current study ($\eta \approx 1$mPs). Of the two explanations discussed above, we believe that the latter can better explain our findings, as it is also consistent with the observed single-file ordering of multiple knots on the chain.

In conclusion, we show that knots formed via hydrodynamic compression of DNA molecules in nanochannels indeed interact via hard-core repulsion and maintain single-file ordering on the chain contour as predicted by the free energy model proposed in [163]. A study on knotted polymers in bulk [178] has shown that knots can pass each other on polymer contour provided that one of the knots is expanded to at least the total size of the two knots added together. Our observations, on the other hand, confirm that the knots formed via hydrodynamic compression do not pass each other on the nanochannel-confined DNA contour. This finding is indicative of a relatively large size of the knots as predicted in [163]. Moreover, we show that via step-wise unraveling of a knot, we can obtain information regarding the knot complexity, which determines the extent of knotting effect on polymer dynamics [44,193]. Finally, we find that unraveling process can be accelerated by applying a relatively small hydrodynamic flow. While we believe that large knot sizes compared to the channel dimensions can explain this behavior, future theoretical studies and simulations could help clarify the unraveling mechanism.

CHAPTER 6
Flow-Assisted Single-Cell Derived Genomic DNA Purification, Disentanglement and Linearization

Thus far, we have introduced a method of knot fabrication on nanochannel-confined DNA molecules in chapter 3 and in chapter 5 we exploited knot dynamics on the relaxing chains and introduced a step-wise unraveling method to obtain knot topological information, which is otherwise inaccessible. However, there exist instances that entanglements and knots form *in vitro* on long DNA molecules and hinder DNA linearization, particularly in Mbp genomic DNA sequencing applications. In some DNA mapping techniques such as DNA mapping on nanochannel arrays [92] and DNA denaturation mapping [91], DNA linearization is a crucial step. Nevertheless, as long chains tend to form entanglements, when Mbp-long DNA molecules are extracted from cells, they are entangled and cannot be linearized in nanochannels. In this chapter, we introduce a high-throughput method for Mbp-long DNA disentanglement in post corridor-slit structures. In particular, we introduce a platform that performs single-cell trapping and lysis, Mbp-long DNA extraction and purification and high-throughput flow-assisted DNA disentanglement. We show that posts anchor long DNA molecules and post corridors can be used to unravel and stretch DNA coils via hydrodynamic flow. Finally, this chapter complements our knot formation study on DNA molecules by proposing a disentanglement method for unraveling highly entangled long DNA molecules.

The contents of this chapter have been prepared as a manuscript ready for submission:

Flow-Assisted Single-Cell Derived Genomic DNA Purification, Disentanglement and Linearization

Susan Amin, Peter Yao, Ilja Czolkos, Sara Mahshid, Robert Sladek and Walter Reisner

6.1 Introduction

Many lab-on-chip platforms have been developed for DNA sequencing. This is mainly due to the fact that miniaturized platforms eliminate the need for large volumes of reagents as well as long operation times [197] especially in sequencing methods such as polymerase chain reaction (PCR). The genetic heterogeneity in cell population – especially in tumor cells [55–57] – introduces the need for systems capable of analysing DNA molecules extracted from single cells. Different single-cell manipulation methods have been proposed [58, 198], among which microfluidic-based devices are proved to be promising for single-cell studies [65, 66].

According to the theoretical and experimental studies, long molecules can get entangled if their concentration is increased [108, 109]. The complexity of entanglement for n-step self-avoiding walks goes to infinity as n increases [199] and DNA entanglement can become burdensome in the process of long DNA extraction from cells and their purification. Recently, Lam *et al.* [92] have developed a genome mapping technique on nanochannel arrays that requires linearization of DNA in nanochannels. Another coarse-grained DNA mapping technique called *DNA denaturation mapping* [91] provides a sequence-dependent melting probability on long nanochannel-confined DNA molecules. Linearization of DNA molecules in these methods is a crucial step. Entanglements not only hinder DNA linearization, but can also lead to knotted structures with high probability for long molecules [30]. Knots, in turn, inhibit mapping

by concealing the sequence information stored in the knot region [20]. Thus, removing entanglements is crucial for analyzing long DNA molecules. Lam *et al.* [92] use electric field to drive long DNA molecules through a pillar region into nanochannel arrays. While uncoiling DNA molecules using pillars works well for medium-sized DNA molecules (around ~ 50 kb in [92]), very long molecules (Mb in size) are very likely to get anchored between the pillars as suggested in [111, 112].

Here, we introduce a high-throughput lab-on-chip device where: (1) a single cell is trapped at a microchannel barrier via hydrostatic pressure balance and flow suction; (2) the cell is lysed and chromatin is extracted and stored in a chamber; (3) DNA is purified via incubation of chromatin with a protein digestive enzyme at 37°C. Cell lysates are removed from the chamber through the two deep side channels; and (4) the long Mbp entangled DNA molecules are linearized in post corridors using hydrodynamic flow. Our microfluidic-based design provides the power to control the system components (cell and DNA molecules) via hydrostatic balance between different parts of the device. It minimizes DNA fragmentation and provides a high-throughput platform that purifies and linearizes long DNA molecules extracted from a single cell.

6.2 Materials and Methods

Device fabrication and the experimental set-up. The post corridors (100-nm deep, 380-μm long, 2-μm wide) are fabricated on fused silica substrates (HOYA) by UV contact lithography. A 100-nm deep slit is subsequently etched over the post arrays. The device contains a 1-μm deep incubation chamber that connects the post corridors to the 15-μm deep microchannels (50-μm wide) used for cell and lysis solution transport (Figure 6–1). Two 15-μm deep side channels remove the cell debris from the device. In addition, a dense post array is located at the intersection of the incubation chamber and the side channels. The posts anchor long DNA molecules and thus stop

them from leaving the device. Finally, the $2.5 \times 2.5\,\mathrm{cm}^2$ chip are bonded to fused silica coverslips (Valley Design). The chip is filled with phosphate buffer saline (PBS 1X) (Fisher Scientific, CA) and is subsequently mounted on a chuck. The chuck ports are loaded with PBS via 10-ml syringes, actuated and controlled by syringe pumps. The microfluidic set-up is then mounted on an inverted microscope (Nikon Eclipse Ti-E) with a 100X N.A. 1.5 oil immersion objective. An EMCCD camera (ixon, Andor) is used for imaging. A metal-halide lamp (Xcite) is used for excitation illumination.

Cell culture. Lymphoblast cells (10-12 μm in size) are cultured in RPMI 1640 medium supplemented with 15% FBS, 100 units per mL penicillin, 100 μg/mL streptomycin and 4mM L-glutamine (Life Technologies). The cells are incubated with 5% CO_2 at 37°C. The cells are passaged every three days and fresh cell-culture medium is provided.

Cell trapping. The centrifuged and pelleted human lymphoblasts (Quebec Genome Center, CA) are suspended in PBS and incubated with SYTO-24 (Life Technologies, CA) nucleic acid stain at 37°C for 30 minutes to 1 hour. The stained cells are injected into the chip through the deep microchannel. A single cell is then trapped by adjusting the hydrostatic pressure exerted on the cells (through the two deep channels that are in contact with the ambient air) and applying suction at the trap outlet, leading to a hydrodynamic flow towards the trap. Once a fresh cell is trapped, the hydraulic resistance at the cell barrier increases significantly, and thus prevents the other cells from getting trapped. The cells are subsequently washed away from the deep microchannels using PBS at the injection rate of 4ml/hr for 30 minutes.

Cell lysis. RIPA lysis buffer is used to lyse the trapped single cell. The lysis buffer is composed of 1.0% Igepal CA-630 (NP-40), 1.0% soduim deoxycholate, 0.1% sodium dodecyl sulfate (SDS) (all Sigma-Aldrich, CA) dissolved in PBS 1X, pH 7.4.

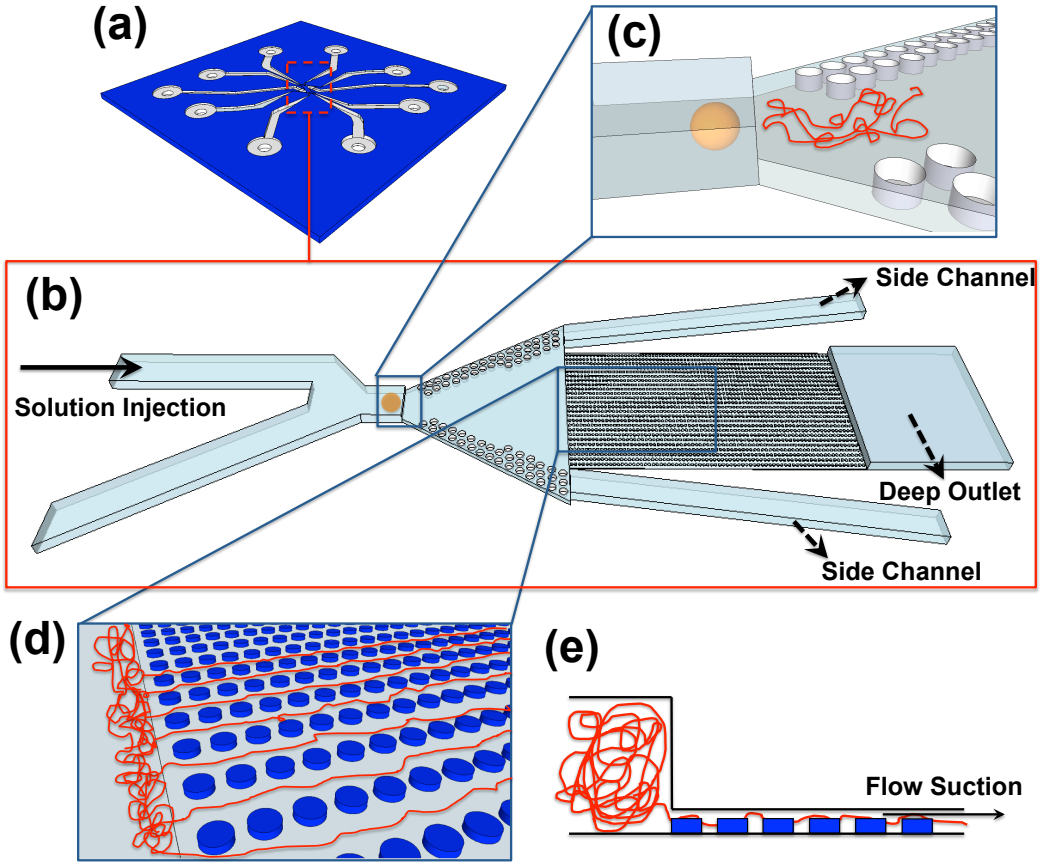

Figure 6–1: A schematic of the microfluidic device. (a) Every $2.5\times2.5\,\mathrm{cm}^2$ chip is composed of two microfluidic devices with ten outlets in total. (b) Each device consists of two inlets (15-μm deep, 50-μm wide) for the injection of cell solution and lysis buffer. A 1-μm deep incubation chamber is connected to the deep channels at their intersection, forming a barrier that traps cell but allows the passage of the DNA molecules following cell lysis (c). The incubation chamber is connected to two side channels through dense post arrays (b). The post arrays anchor long DNA molecules and let the cell debris leave the chamber through the side channels. (d) The long purified DNA molecules are trapped behind an array of post corridors (100-nm deep, 380-μm long, 2-μm wide) and a slit (100-nm deep) spanning the posts. (e) The cross section of the post corridors and the slit. The DNA molecules are disentangled via flow suction.

YOYO-1 (Life Technologies, CA) and Alexa Fluor 568 NHS Ester (Life Technologies, CA) dissolved in DMSO are added to the buffer for staining the DNA molecules and proteins, respectively: 2μL of YOYO-1 and 0.1μg of Alexa Fluor are added to 10mL of

the lysis solution. The solution is then injected into the deep microchannel. The cell lysis process starts in less than 2 minutes. Once the lysis begins, the buffer injection is stopped and the device is incubated with the solution for 90 minutes to allow Alexa Fluor to attach to the proteins in the cell.

Protein digestion and DNA disentanglement. To perform the protein digestion, the concentration of SDS in the lysis buffer is increased to 0.5% and 1mg/mL Pronase (Roche, CA) is added to the solution. The chip temperature is raised to 46°C and the buffer is injected into the device. Subsequently, the flow is stopped and the chip is incubated for 30-45 minutes. After complete protein digestion, the long purified DNA molecules are driven towards the post corridors via flow suction at the post-corridor outlet.

6.3 Results and Discussion

Cell trapping. We use 3D Comsol Multiphysics software to simulate cell trajectories in the cell trapping process. The model is solved in the creeping flow regime with no-slip boundary condition. The simulation predicts the effect of flow suction at the device outlet on the efficiency of cell trapping. The geometry is composed of two deep microchannels (15-μm deep and 50-μm wide). At the intersection of the two channels, there is a 1-μm deep slit, which is a simplified model of the incubation chamber and forms a cell barrier (Figure 6–2 (a)). Simulation results show that due to the large hydraulic resistance at the barrier, cells tend to flow in the microchannels and they do not approach the barrier (Figure 6–2 (b)). When flow suction is applied at the slit, the resistance decreases significantly and the flow is partially diverted towards the barrier (Figure 6–2 (c)). The simulation results are in good agreement with our experimental observations. Figure 6–2 (d) shows a trapped cell behind the barrier. Due to the

large hydraulic resistance, the other cells do not approach the barrier and flow in the microchannels.

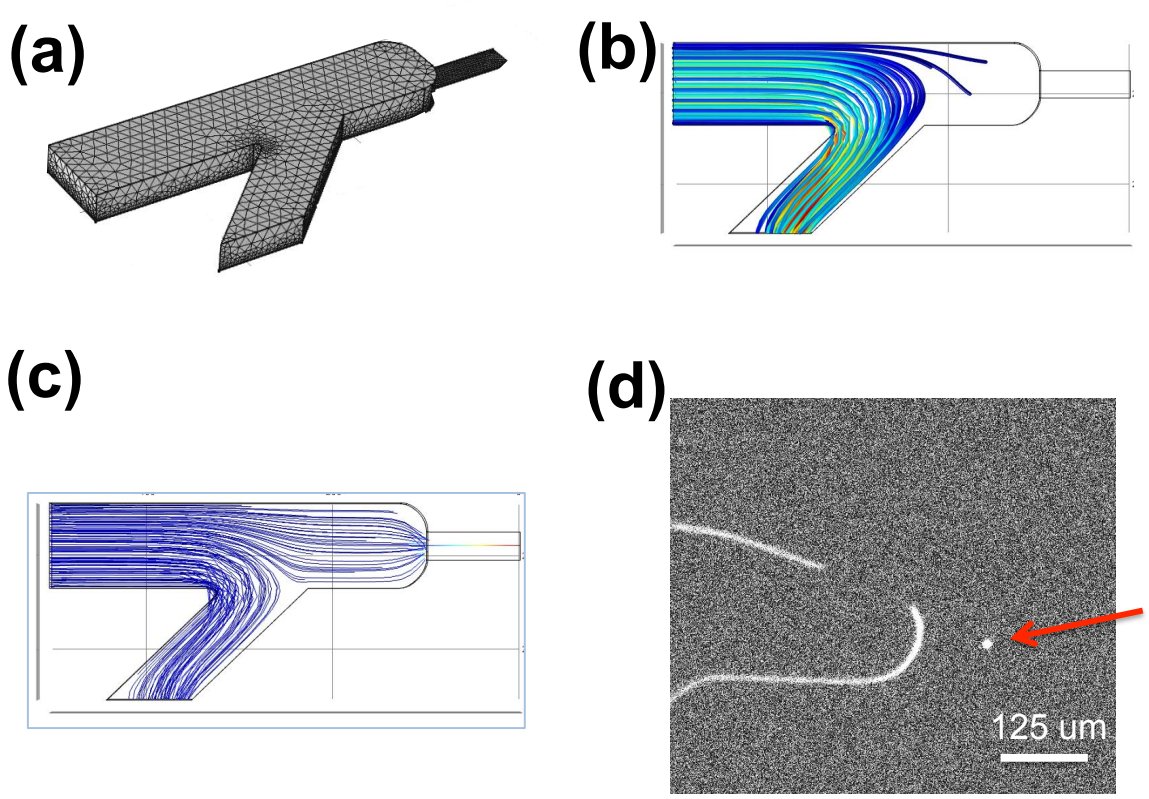

Figure 6–2: A simplified model of the deep microchannels and the cell barrier is simulated using Comsol Software. (a) The deep channels are 15-μm deep and 50-μm wide and the slit is 1-μm deep. (b) In the absence of flow suction, due to the large hydraulic resistance the cells flow in the microchannels and do not enter the cell-trapping region. (c) When flow suction is applied, the flow deviates partially towards the trap. (d) A cell is trapped (shown with a red arrow), while the other cells flow in the deep microchannels. Note that the flowing cells appear as extended objects due to their high speed.

Cell lysis and protein digestion. Cell lysis starts in less that 2 minutes following RIPA injection (Figure 6–3). During the cell lysis process, proteins are stained with Alexa Fluor. Figure 6–4 (a,b) illustrates the co-visualization of DNA molecules (green) and proteins (red) during the cell lysis. After 90 minutes, the protein digestion

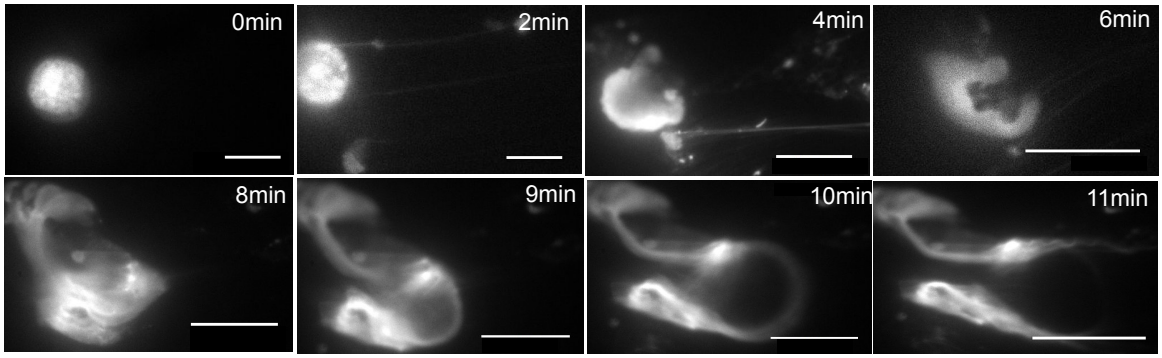

Figure 6–3: A time sequence of cell lysis with RIPA buffer is illustrated. The lysis starts in less than 2 minutes following RIPA injection. The scale bars are 20μm in all panels.

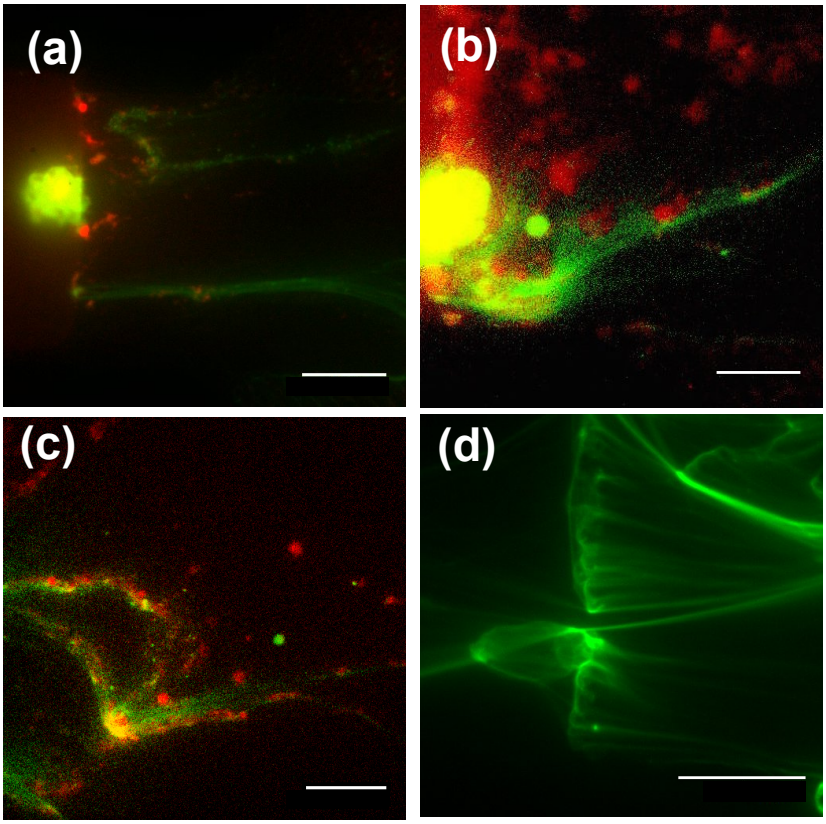

Figure 6–4: Co-visualization of DNA molecules and proteins. (a,b) Proteins are stained and visualized during the cell lysis process. The protein digestion starts after 90 minutes (c) and lasts for 45 mintues. (d) The DNA molecules are completely purified following incubation with Pronase. The scale bars are all 20μm.

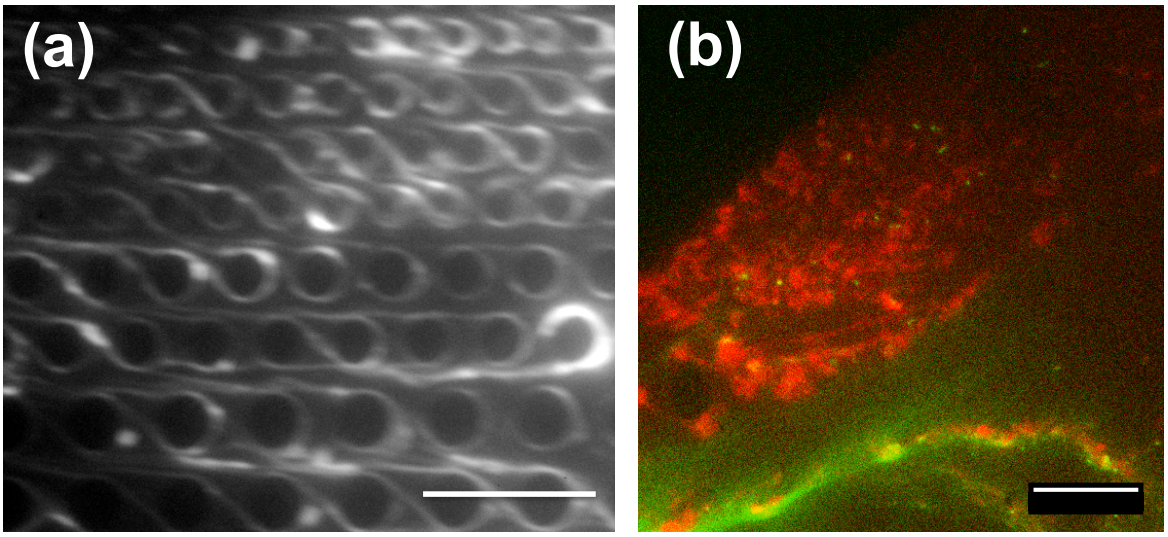

Figure 6–5: DNA anchoring using posts. (a) Long Mbp genomic DNA molecules are pinned by the posts. (b) Proteins (shown in red) fill the side channel, while the DNA molecules (shown in green) do not leave the incubation chamber. The scale bars are 20μm.

process starts. Figure 6–4 (c) depicts the lysed contents of a single cell prior to protein digestion. After 45 minutes of incubation with Pronase, the DNA molecules are purified and the protein is completely digested (Figure 6–4 (d)). Note that during the experiment the shutter is kept closed in between visualization intervals (mostly every 10 minutes) to avoid the photobleaching effects.

The dense post arrays at the inlet of the two side channels in our device are designed to anchor and stop DNA molecules from leaving the chamber. Post arrays are known as interfaces between micro- and nano-scale features [110]. They decrease the local entropic barrier for entering the nanochannels and at the same time help with the pre-stretching of DNA molecules. We argue that while this method works well with relatively short molecules (on order of kbp), it does not have the same effect with long (Mbp) DNA molecules. The relatively small spacing between the posts compared to the radius of gyration of Mbp-long DNA molecules leads to the anchoring of DNA

strands and their fragmentation if they are subjected to high pressure. Figure 6–5 (a) illustrates Mbp genomic DNA molecules extracted from a single cell, pinned in an array of posts. We have used this property of post arrays to minimize the DNA escape from the incubation chamber. Figure 6–5 (b) illustrates the side channel filled with proteins (shown in red), while the DNA molecules remain inside the incubation chamber (shown in green).

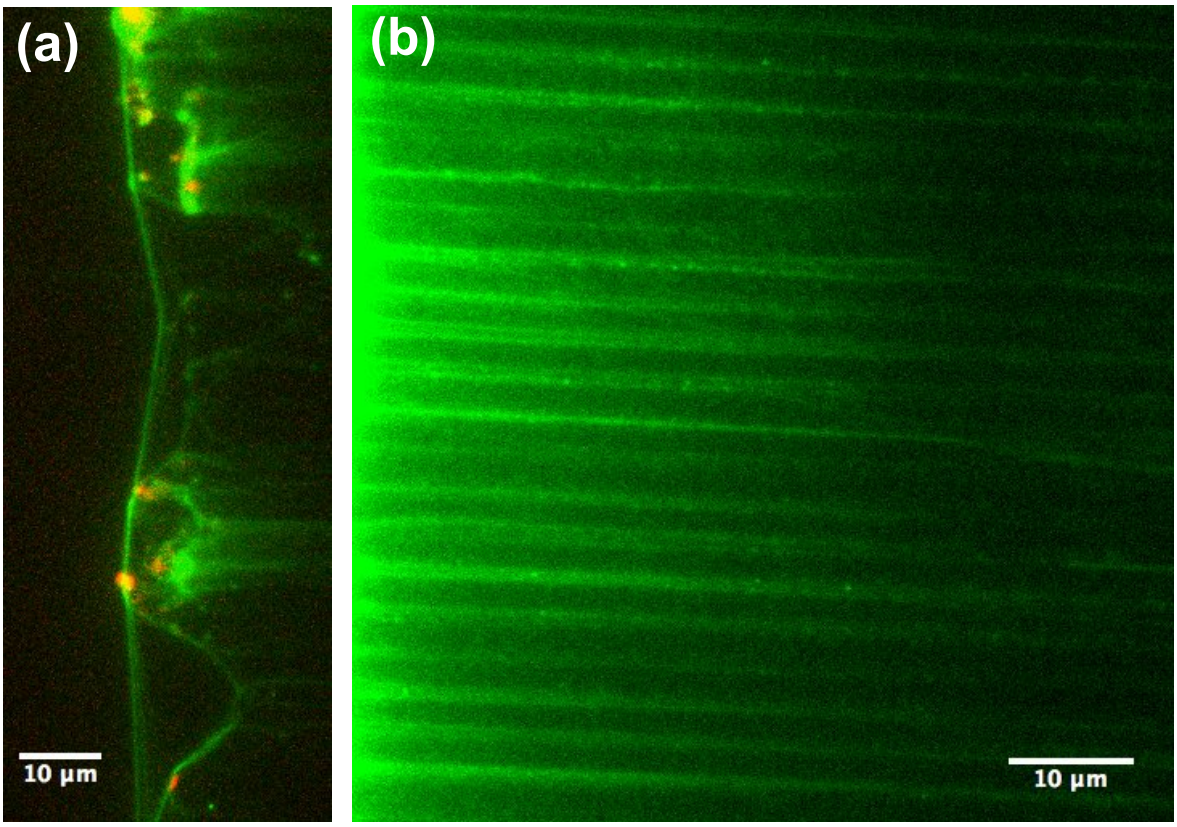

Figure 6–6: High-throughput Mbp-long DNA disentanglement in post corridors. (a) The long genomic DNA molecules are pinned at the entrance of the post corridors. Using hydrodynamic flow, the polymers unravel and enter the corridors. The red spots in the figure indicate that a small amount of protein is not fully digested. (b) High-throughput flow-assisted DNA linearization in the post corridors.

DNA disentanglement. Upon applying suction to the outlet, the purified DNA molecules flow towards the post corridors. Due to their large contour lengths and

given the long enough incubation and purification time, the DNA polymers are highly entangled. The post corridor-slit structure provides an efficient way of DNA disentanglement and the flow-assisted linearization of the molecules. The entangled DNA coils are pinned at the entrance of the post corridors. The suction applied to the outlet creates the required hydrodynamic flow to unravel and stretch DNA molecules (each larger than 1Mbp in size). Figure 6–6 presents the flow-assisted high-throughput stretching of Mbp-long DNA polymers. Several red spots are apparent at the entrance of the post corridors (Figure 6–6 (a)), indicative of some non-dissolved proteins. However, the stretched DNA molecules (Figure 6–6 (b)) are pure and do not contain proteins. Note that the spacing between the posts in the post corridor along the corridor longitudinal axis, reduces the hydraulic resistance, and thus leads to a high-throughput DNA linearization.

In conclusion, we show that our microfluidic device provides a platform for single-cell trapping, lysis, Mbp genomic DNA extraction and purification and flow-assisted DNA disentanglement. A very important feature of our system is its ability to keep the DNA molecules intact and minimize the DNA escape from the device. This is done via a dense array of posts located at the entrance to the two deep side channels. The posts anchor the long DNA molecules while allow the cell lysate to leave the chamber. In addition, we provide an efficient approach for DNA purification and visualization of proteins during the protein digestion. Finally, we disentangle and stretch the DNA molecules in post corridors using hydrodynamic flow. We show that parallel post corridors reduce the hydraulic resistance and thus provide a high-throughput system for DNA linearization. The post corridor-slit structure can be further used as an interface between micro- and nanochannels. Our DNA linearization technique can be

used in the single-cell derived DNA mapping approaches, such as DNA denaturation mapping [91].

CHAPTER 7
Conclusion

In this dissertation, we introduce a knot factory on-chip based on hydrodynamic compression of single DNA molecules in nanochannels. In our proposed system, all parameters are well controlled and knots are produced with high probability. In particular, we show that knots form via compression of single DNA molecules at equilibrium under moderate confinement condition. Moreover, we propose a free energy model based on scaling arguments which can describe our experimental findings. Our model suggests that knotting probability increases with pressure and waiting time at compressed state. We show that knots interact via hard-core repulsion and maintain single-file ordering along the contour. Knot interactions most likely prevent the formation of multiple knot states (we do not observe more than three knots on T4 DNA molecules). Our model predicts that knotting probability obeys Poisson statistics at low compression. Deviation of knotting probability from Poisson statistics at high compression can be explained by knot interactions and is confirmed by our experimental data. Moreover, our model provides predictions for the effect of varying channel dimension in the extended de Gennes regime on the knotting probability. We predict that stronger confinement leads to a larger knotting probability and an increase in the number of multiple knot states.

The most distinct feature of our proposed knotting technique is its reproducibility and that all parameters are well controlled. Our knot factory provides a highly

parallel environment, where knots can form on DNA molecules in parallel nanochannels. Moreover, in our proposed system, knots form on compressed DNA molecules at inhomogeneous equilibrium state, enabling modeling of knot production with free energy models using scaling arguments. We argue that in our proposed model agitation of DNA segments required for knot formation is provided by thermal fluctuations assisted by hydrodynamic flow.

Our knotting approach does not provide detailed information about knot complexity due to the knot compactness and thus unresolved image. We show that *step-wise* knot unraveling can be employed to obtain information regarding knot complexities. Moreover, we observe that application of small hydrodynamic flow can help move the knot toward the end of the molecule and accelerate the unknotting process.

Our proposed knotting technique can be studied in the case of different channel diameters and different regimes. Knotting probability can be investigated for different ionic strengths and DNA contour lengths as well as for different types of polymers. Moreover, the knot factory can help form knots with high probability in a well-controlled system for further studies of knot dynamics. In addition, reptation of knotted polymers through the slit barriers with different depths can be explored. Brownian simulation of knot dynamics during relaxation of DNA molecules can help clarify the evolution of knot extension following pressure release. Simulations can also help understand the single-file ordering of knots observed in the experiments and predicted by the theory.

In our second study, we introduce a microfluidic platform, where long Mbp genomic DNA molecules are extracted from single cells. The DNA is subsequently purified, disentangled and linearized in nano-post corridors using hydrodynamic flow. In our proposed system several steps are performed in the following order:

1. A single cell is trapped via the balance of hydrostatic pressure and flow suction at the device outlet.
2. The single cell is lysed after removing the other cells from the microchannels. A protein-specific dye is added to the lysis buffer for the protein visualization purposes.
3. A protein-digestive enzyme is added to the system for protein digestion and DNA purification.
4. Using flow suction at the device outlet, the purified and coiled DNA molecules that are pinned at the post-corridor entrance, unravel and get stretched in the corridors.

In our proposed microfluidic system, we show that post arrays anchor long DNA molecules and do not let them pass through. We use this property in our device to prevent the long DNA molecules from leaving the chamber. This feature helps the system remove the cell lysates, while keeping the DNA molecules inside the chamber. Finally, the proposed post-corridor set-up decreases the hydraulic resistance in the system, leading to a high-throughput DNA linearization in the corridors.

Our microfluidic device can be further utilized in the denaturation mapping experiments. Nanochannel arrays can be added to the platform to confine DNA molecules following their linearization in the post corridors. As long DNA molecules tend to form entanglements in microfluidic systems, the proposed flow-assisted technique can be used to disentangle and linearize DNA molecules, and the post corridors can be employed as an interface between micro- and nano-features in DNA mapping techniques on nanochannels. In addition, new materials and agents can be employed to reduce the experiment run time. Moreover, different configurations of posts as well as different post and slit depths can be used for improving the efficiency of the system. Finally, the

cell trapping technique can be parallelized by introducing more cell-barrier structures to the platform.

Nomenclature

α	Slope of the steady-state ramp profile for a single chain compressed against the slit-barrier
$\beta(L_k, R_\mathrm{b})$	Slope of the free energy profile for a molecule with edge extension R_b
Δp	Pressure drop across the nanofluidic channel
ΔV	Streaming potential
δ	Offset in the channel dimension due to the presence of a knot
η	Solution viscosity
γ	A numerical factor relating the radius of gyration g_k to the offset in the channel dimension δ
$\hat{P}_i$	Maximum likelihood estimate for the probability of a knotting state with i knots
$\hat{P}_\mathrm{total}$	Maximum likelihood estimate of total knotting probability
λ	Transition rate ratio k_{21}/k_{12}
Π	Osmotic pressure
σ	Width of the point spread function f_PSF
σ_c	Buffer conductivity
σ_n	Wilson score interval
σ_r	Standard deviation in the extension at compressed state
$\xi_\parallel$	Blob extent for a molecule in an extended de Gennes regime at no-flow equilibrium
ζ	Friction factor per unit length
ζ_f	Friction factor at molecule free edge

ζ_k Knot friction factor

$C(X)$ Normalized chain concentration at position X ($\equiv c(X)/c_o$)

$c(x,t)$ The average concentration profile over channel cross section at position x and time t. This quantity reduces to $c(x)$ in steady state.

C_f Normalized chain concentration at the molecule free edge ($\equiv c_f/c_o$)

c_f Chain concentration at the molecule free edge

C_L Cumulative lower knot distribution (in two-knot events) assuming knots do not interact (i.e. non-interacting distribution)

c_o The concentration of the chain in the absence of flow ('no-flow' equilibrium concentration)

C_U Cumulative upper knot distribution (in two-knot events) assuming knots do not interact (i.e. non-interacting distribution)

C_{b} Normalized chain concentration at the barrier edge ($\equiv c_{\mathrm{b}}/c_o$)

c_{b} Chain concentration at the barrier edge

$C_{\mathrm{cum},ij}(X_j)$ Cumulative distribution of the j^{th} knot in i-knot event

$C_{\mathrm{cum}}(x)$ The cumulative knot distribution for one-knot event

D'_{eff} Channel effective diameter, taking into account the reduction in the channel size due to the presence of a knot

D_1 Horizontal dimension of the nanochannel cross section

D_2 Vertical dimension of the nanochannel cross section

D_c Cooperative diffusion constant

D_{av} The geometrical average of nanochannel cross section

D_{eff} Channel cross-section effective diameter

$f_b(L_k)$ Free energy cost of formation of a knot of contour L_k in bulk

$f_h(L_k, V, X)$ Free energy required to move knot away from the barrier ($X = 0$) against hydrodynamic flow

$f_i(L_{k1}, ..., L_{ki}, R_\mathrm{b}, X_1, ..., X_i)$ Free energy of an i-knot state event with the knots on normalized positions on the chain X_1 to X_i

$F_m(L_{k1}, \cdots, L_{km}, R_\mathrm{b})$ Free energy for forming m knots of size $L_{k1}, \cdots, L_{km}$ along a compressed molecule with edge extension R_b

$f_\mathrm{int}(X_2 - X_1)$ Free energy of interaction between two knots at normalized positions X_1 and X_2

f_PSF Point spread function

$F_\mathrm{tot}(m, R_\mathrm{b})$ Free energy for forming m knots including all knot sizes for a compressed molecule with edge extension R_b

$f_{ij}(X_j)$ Free energy of knot j at position X_j in the i-knot event

$f_{wk}(L_k, R)$ Free energy of confining a knot of contour L_k between the channel walls at a position on the molecule concentration profile with local extension R.

$f_{wuk}(L_k, R)$ Free energy of confining an unknotted chain of contour L_k between the channel walls at a position on the molecule concentration profile with local extension R.

G_k Normalized knot gyration radius ($\equiv g_k/r_o$)

g_k Knot gyration radius

$I(x)$ Fluorescence intensity profile of molecule at position x on the chain

I_o Intensity level of the extended molecule at no-flow equilibrium

I_s Buffer ionic strength

J Total segmental current

J_c Convective segmental current along the channel

J_D Cooperative diffusion current along the channel

k_B Boltzmann constant

k_{ij} Transition rate from a state with i knots to a state with j knots

L_b Contour per blob in the no-flow equilibrium chain

L_c Channel length

L_k Knot contour length

n_k Maximum number of knots observed in an event

$n_{\max}$ The number of statistically independent sites at which a knot can form along the profile

P Persistence length

$P(i, R_{\rm b})$ The probability of finding a state with i knots along a compressed molecule with barrier edge extension $R_{\rm b}$.

$P^{\rm av}$ Predicted non-interacting single-knot distribution averaged over all knot sizes

P_i Probability of formation of i knots

$P_i(L_{k1}, ..., L_{ki}, R_{\rm b}, X_1, ..., X_i)$ i-knot probability distribution for finding i knots at position $X_1, ..., X_i$ along a profile with edge extension $R_{\rm b}$.

$P_L(x)$ Probability of knot distribution for the lower knot (closer to the barrier) at position x in 2-knot event assuming knots do not interact (i.e. non-interacting distribution)

$P_L^{\rm av}$ Predicted non-interacting lower knot distribution averaged over all knot sizes

P_o High-salt persistence length

$P_U(x)$ Probability of knot distribution for the upper knot (closer to the free edge) at position x in 2-knot event assuming knots do not interact (i.e. non-interacting distribution)

$P_U^{\rm av}$ Predicted non-interacting upper knot distribution averaged over all knot sizes

$P_{\rm all\ knots}$ The probability of finding a state with any number (non-zero) of knots

$P_{\text{diff}}(\delta x)$ Probability distribution for δx (the distance between two knots on a chain)

P_{tot} Total probability of knot formation

$P_{ij}(X_j)$ The probability distribution for the position X_j of knot j in the i-knot space

r Molecule extension in compressed state

$R(X)$ Local extension at normalized position X on the chain ($\equiv 1/C(X)$)

R_c Normalized chain extension ($= r/r_o$)

r_o Molecule extension in the absence of flow ('no-flow' equilibrium extension)

R_{b} Local extension at the barrier ($\equiv 1/C_{\text{b}} = 1/C(0)$)

R_{ch} Electrical resistance of the nanochannel

r_{relax} Extension of the relaxing molecule measured 2 seconds after pressure release

S_{str} Streaming conductance

T Temperature

t_w Waiting time in compressed state

V Molecule free edge speed during transient compression: a measure of the buffer flow speed

w Effective width of the chain

X Normalized position on the chain ($\equiv x/r_o$)

$x_f(t)$ The position of the molecule free edge at time t

$Z(m, R_{\text{b}})$ The partition function for a system with m knots

$z_m(L_{k1}, \cdots, L_{km}, R_{\text{b}})$ Partition sum for forming m knots of size $L_{k1}, \cdots, L_{km}$ along a compressed molecule with edge extension R_{b}

$z_{\alpha/2}$ The critical value of the normal distribution for error level α

$Z_{\text{tot}}(R_{\text{b}})$ Total knot partition function for a compressed molecule with barrier extension R_{b}

References

[1] JH Przytycki. Knots: From combinatorics of knot diagrams to combinatorial topology based on knots. 2007. *arXiv preprint math/0703096*.

[2] John Christopher Turner. *History and science of knots*, volume 11. World Scientific, 1996.

[3] John Romilly Allen. *Celtic art in pagan and Christian times*. Methuen & Company, 1904.

[4] Dario Meluzzi, Douglas E Smith, and Gaurav Arya. Biophysics of knotting. *Annual review of biophysics*, 39:349–366, 2010.

[5] William R Taylor and Kuang Lin. Protein knots: a tangled problem. *Nature*, 421(6918):25–25, 2003.

[6] Chengzhi Liang and Kurt Mislow. Topological features of protein structures: knots and links. *Journal of the American Chemical Society*, 117(15):4201–4213, 1995.

[7] Gurvan Michel, Véronique Sauvé, Robert Larocque, Yunge Li, Allan Matte, and Miroslaw Cygler. The structure of the rlmb 23s rrna methyltransferase reveals a new methyltransferase fold with a unique knot. *Structure*, 10(10):1303–1315, 2002.

[8] Leroy F Liu, John L Davis, and Richard Calendar. Novel topologically knotted dna from bacteriophage p4 capsids: studies with dna topoisomerases. *Nucleic acids research*, 9(16):3979–3989, 1981.

[9] Xiaoyan R Bao, Heun Jin Lee, and Stephen R Quake. Behavior of complex knots in single dna molecules. *Physical review letters*, 91(26):265506, 2003.

[10] Javier Arsuaga, Mariel Vázquez, Sonia Trigueros, Joaquim Roca, et al. Knotting probability of dna molecules confined in restricted volumes: Dna knotting in phage capsids. *Proceedings of the National Academy of Sciences*, 99(8):5373–5377, 2002.

[11] Liang Dai, C Benjamin Renner, and Patrick S Doyle. Metastable knots in confined semiflexible chains. 2015.

[12] Cristian Micheletti and Enzo Orlandini. Knotting and metric scaling properties of dna confined in nano-channels: a monte carlo study. *Soft Matter*, 8(42):10959–10968, 2012.

[13] Ralf Metzler, Walter Reisner, Robert Riehn, Robert Austin, JO Tegenfeldt, and Igor M Sokolov. Diffusion mechanisms of localised knots along a polymer. *EPL (Europhysics Letters)*, 76(4):696, 2006.

[14] Beatrice W Soh, Vivek Narsimhan, Alexander R Klotz, and Patrick S Doyle. Knots modify the coil–stretch transition in linear dna polymers. *Soft matter*, 14(9):1689–1698, 2018.

[15] Yasuharu Arai, Ryohei Yasuda, Ken-ichirou Akashi, Yoshie Harada, Hidetake Miyata, Kazuhiko Kinosita, and Hiroyasu Itoh. Tying a molecular knot with optical tweezers. *Nature*, 399(6735):446–448, 1999.

[16] A Marco Saitta, Paul D Soper, E Wasserman, and Michael L Klein. Influence of a knot on the strength of a polymer strand. *Nature*, 399(6731):46–48, 1999.

[17] Nadrian C Seeman. Nucleic acid nanostructures and topology. *Angewandte Chemie International Edition*, 37(23):3220–3238, 1998.

[18] Jing Tang, Ning Du, and Patrick S Doyle. Compression and self-entanglement of single dna molecules under uniform electric field. *Proceedings of the National Academy of Sciences*, 108(39):16153–16158, 2011.

[19] Colin Conrad Adams. *The knot book: an elementary introduction to the mathematical theory of knots*. American Mathematical Soc., 2004.

[20] Jeffrey G Reifenberger, Kevin D Dorfman, and Han Cao. Topological events in single molecules of e. coli dna confined in nanochannels. *Analyst*, 140(14):4887–4894, 2015.

[21] William R Taylor. A deeply knotted protein structure and how it might fold. *Nature*, 406(6798):916, 2000.

[22] Behzad Eftekhar and Andrew Hunn. Ventriculoperitoneal shunt blockage due to spontaneous knot formation in the peritoneal catheter. 2008.

[23] Alain Goriely. Knotted umbilical cords. *See Ref*, 19:109–26, 2005.

[24] J Hickford, R Jones, S Courrech Du Pont, and J Eggers. Knotting probability of a shaken ball-chain. *Physical Review E*, 74(5):052101, 2006.

[25] Kleanthes Koniaris and M Muthukumar. Knottedness in ring polymers. *Physical review letters*, 66(17):2211, 1991.

[26] Dorian M Raymer and Douglas E Smith. Spontaneous knotting of an agitated string. *Proceedings of the National Academy of Sciences*, 104(42):16432–16437, 2007.

[27] Piotr Pieranski, Sandor Kasas, Giovanni Dietler, Jacques Dubochet, and Andrzej Stasiak. Localization of breakage points in knotted strings. *New Journal of Physics*, 3(1):10, 2001.

[28] Eli Ben-Naim, ZA Daya, Peter Vorobieff, and Robert E Ecke. Knots and random walks in vibrated granular chains. *Physical review letters*, 86(8):1414, 2001.

[29] Andrew Belmonte, Michael J Shelley, Shaden T Eldakar, and Chris H Wiggins. Dynamic patterns and self-knotting of a driven hanging chain. *Physical review letters*, 87(11):114301, 2001.

[30] DeWitt Sumners and SG Whittington. Knots in self-avoiding walks. *Journal of Physics A: Mathematical and General*, 21(7):1689, 1988.

[31] Clifford Warren Ashley. *The Ashley book of knots*, volume 13. Doubleday Books, 1944.

[32] Andrzej Stasiak, Akos Dobay, Jacques Dubochet, and Giovanni Dietler. Knotted fishing line, covalent bonds, and breaking points. *Science*, 286(5437):11–11, 1999.

[33] Harry L Frisch and Edel Wasserman. Chemical topology1. *Journal of the American Chemical Society*, 83(18):3789–3795, 1961.

[34] M Delbruck. Mathematical problems in the biological sciences. In *Proc. Symp. Appl. Math.*, volume 14, page 55, 1962.

[35] Leroy F Liu, Richard E Depew, and James C Wang. Knotted single-stranded dna rings: A novel topological isomer of circular single-stranded dna formed by treatment with escherichia coli ω protein. *Journal of molecular biology*, 106(2):439–452, 1976.

[36] José M Sogo, Andrzej Stasiak, Marıa Luisa Martınez-Robles, Dora B Krimer, Pablo Hernández, and Jorge B Schvartzman. Formation of knots in partially replicated dna molecules. *Journal of molecular biology*, 286(3):637–643, 1999.

[37] L Olavarrieta, ML Martínez-Robles, P Hernández, DB Krimer, and JB Schvartzman. Knotting dynamics during dna replication. *Molecular microbiology*, 46(3):699–707, 2002.

[38] Leroy F Liu, Luke Perkocha, Richard Calendar, and James C Wang. Knotted dna from bacteriophage capsids. *Proceedings of the National Academy of Sciences*, 78(9):5498–5502, 1981.

[39] Alexander Bucka and Andrzej Stasiak. Construction and electrophoretic migration of single-stranded dna knots and catenanes. *Nucleic acids research*, 30(6):e24–e24, 2002.

[40] JPJ Michels and FW Wiegel. On the topology of a polymer ring. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 403, pages 269–284. The Royal Society, 1986.

[41] Valentin V Rybenkov, Nicholas R Cozzarelli, and Alexander V Vologodskii. Probability of dna knotting and the effective diameter of the dna double helix. *Proceedings of the National Academy of Sciences*, 90(11):5307–5311, 1993.

[42] Enzo Orlandini and Cristian Micheletti. Knotting of linear dna in nano-slits and nano-channels: a numerical study. *Journal of biological physics*, 39(2):267–275, 2013.

[43] L Sumners De Witt and Nicholas R Cozzarelli. *New scientific applications of geometry and topology*, volume 45. American Mathematical Soc., 1992.

[44] Stephen R Quake. Topological effects of knots in polymers. *Physical review letters*, 73(24):3317, 1994.

[45] Steven Wasserman and Nicholas R Cozzarelli. Biochemical topology: applications to dna recombination and replication. *Science*, 232:951–961, 1986.

[46] Alexander R Klotz, Vivek Narsimhan, Beatrice W Soh, and Patrick S Doyle. Dynamics of dna knots during chain relaxation. *Macromolecules*, 2017.

[47] Jong Wook Hong and Stephen R Quake. Integrated nanoliter systems. *Nature biotechnology*, 21(10):1179, 2003.

[48] Douglas B Weibel, Maarten Kruithof, Scott Potenta, Samuel K Sia, Andrew Lee, and George M Whitesides. Torque-actuated valves for microfluidics. *Analytical chemistry*, 77(15):4726–4733, 2005.

[49] Nam-Trung Nguyen and Zhigang Wu. Micromixersa review. *Journal of micromechanics and microengineering*, 15(2):R1, 2004.

[50] Axel Günther, Manish Jhunjhunwala, Martina Thalmann, Martin A Schmidt, and Klavs F Jensen. Micromixing of miscible liquids in segmented gas- liquid flow. *Langmuir*, 21(4):1547–1555, 2005.

[51] Piotr Garstecki, Michael A Fischbach, and George M Whitesides. Design for mixing using bubbles in branched microfluidic channels. *Applied Physics Letters*, 86(24):244108, 2005.

[52] Daniel J Laser and Juan G Santiago. A review of micropumps. *Journal of micromechanics and microengineering*, 14(6):R35, 2004.

[53] Aditya Aryasomayajula, Pouriya Bayat, Pouya Rezai, and P Ravi Selvaganapathy. Microfluidic devices and their applications. In *Springer Handbook of Nanotechnology*, pages 487–536. Springer, 2017.

[54] George M Whitesides. The origins and the future of microfluidics. *Nature*, 442(7101):368, 2006.

[55] So Yeon Park, Mithat Gönen, Hee Jung Kim, Franziska Michor, and Kornelia Polyak. Cellular and genetic diversity in the progression of in situ human breast carcinomas to an invasive phenotype. *The Journal of clinical investigation*, 120(2):636–644, 2010.

[56] Lurdes Torres, Franclim R Ribeiro, Nikos Pandis, Johan A Andersen, Sverre Heim, and Manuel R Teixeira. Intratumor genomic heterogeneity in breast cancer with clonal divergence between primary carcinomas and lymph node metastases. *Breast cancer research and treatment*, 102(2):143–155, 2007.

[57] Fulvia Farabegoli, Donatella Santini, Claudio Ceccarelli, Mario Taffurelli, Domenico Marrano, and Nicola Baldini. Clone heterogeneity in diploid and aneuploid breast carcinomas as detected by fish. *Cytometry Part A*, 46(1):50–56, 2001.

[58] Christopher E Sims and Nancy L Allbritton. Analysis of single mammalian cells on-chip. *Lab on a Chip*, 7(4):423–440, 2007.

[59] Yuichi Taniguchi, Paul J Choi, Gene-Wei Li, Huiyi Chen, Mohan Babu, Jeremy Hearn, Andrew Emili, and X Sunney Xie. Quantifying e. coli proteome and transcriptome with single-molecule sensitivity in single cells. *science*, 329(5991):533–538, 2010.

[60] Nicholas Navin, Jude Kendall, Jennifer Troge, Peter Andrews, Linda Rodgers, Jeanne McIndoo, Kerry Cook, Asya Stepansky, Dan Levy, Diane Esposito, et al. Tumour evolution inferred by single-cell sequencing. *Nature*, 472(7341):90, 2011.

[61] H Christina Fan, Jianbin Wang, Anastasia Potanina, and Stephen R Quake. Whole-genome molecular haplotyping of single cells. *Nature biotechnology*, 29(1):51, 2011.

[62] Michael Gundry, Wenge Li, Shahina Bano Maqbool, and Jan Vijg. Direct, genome-wide assessment of dna mutations in single cells. *Nucleic acids research*, 40(5):2032–2040, 2011.

[63] Yong Hou, Luting Song, Ping Zhu, Bo Zhang, Ye Tao, Xun Xu, Fuqiang Li, Kui Wu, Jie Liang, Di Shao, et al. Single-cell exome sequencing and monoclonal evolution of a jak2-negative myeloproliferative neoplasm. *Cell*, 148(5):873–885, 2012.

[64] Jianbin Wang, H Christina Fan, Barry Behr, and Stephen R Quake. Genome-wide single-cell analysis of recombination activity and de novo mutation rates in human sperm. *Cell*, 150(2):402–412, 2012.

[65] Darwin R Reyes, Dimitri Iossifidis, Pierre-Alain Auroux, and Andreas Manz. Micro total analysis systems. 1. introduction, theory, and technology. *Analytical chemistry*, 74(12):2623–2636, 2002.

[66] Ling Lin, Xuexia Lin, Luyao Lin, Qiang Feng, Takehiko Kitamori, Jin-Ming Lin, and Jiashu Sun. Integrated microfluidic platform with multiple functions to probe tumor–endothelial cell interaction. *Analytical chemistry*, 89(18):10037–10044, 2017.

[67] Alison M Skelley, Oktay Kirak, Heikyung Suh, Rudolf Jaenisch, and Joel Voldman. Microfluidic control of cell pairing and fusion. *Nature methods*, 6(2):147, 2009.

[68] Chuansen Liu, Jiangjiang Liu, Dan Gao, Mingyu Ding, and Jin-Ming Lin. Fabrication of microwell arrays based on two-dimensional ordered polystyrene microspheres for high-throughput single-cell analysis. *Analytical chemistry*, 82(22):9418–9424, 2010.

[69] Qiushui Chen, Jing Wu, Yandong Zhang, Zhen Lin, and Jin-Ming Lin. Targeted isolation and analysis of single tumor cells with aptamer-encoded microwell array on microfluidic device. *Lab on a Chip*, 12(24):5180–5185, 2012.

[70] Stefan Kobel, Ana Valero, Jonas Latt, Philippe Renaud, and Matthias Lutolf. Optimization of microfluidic single cell trapping for long-term on-chip culture. *Lab on a Chip*, 10(7):857–863, 2010.

[71] Jean-Philippe Frimat, Marco Becker, Ya-Yu Chiang, Ulrich Marggraf, Dirk Janasek, Jan G Hengstler, Joachim Franzke, and Jonathan West. A microfluidic array with cellular valving for single cell co-culture. *Lab on a Chip*, 11(2):231–237, 2011.

[72] Di Jin, Bin Deng, JX Li, W Cai, L Tu, J Chen, Q Wu, and WH Wang. A microfluidic device enabling high-efficiency single cell trapping. *Biomicrofluidics*, 9(1):014101, 2015.

[73] Samia Menad, Laure Franqueville, Naoufel Haddour, François Buret, and Marie Frénéa-Robin. ndep-driven cell patterning and bottom-up construction of cell aggregates using a new bioelectronic chip. *Acta biomaterialia*, 17:107–114, 2015.

[74] Shideh Kabiri Ameri, Pramod K Singh, Mehmet R Dokmeci, Ali Khademhosseini, Qiaobing Xu, and Sameer R Sonkusale. All electronic approach for high-throughput cell trapping and lysis with electrical impedance monitoring. *Biosensors and Bioelectronics*, 54:462–467, 2014.

[75] Carlo Liberale, G Cojoc, F Bragheri, P Minzioni, G Perozziello, R La Rocca, L Ferrara, V Rajamanickam, E Di Fabrizio, and I Cristiani. Integrated microfluidic device for single-cell trapping and spectroscopy. *Scientific reports*, 3:1258, 2013.

[76] Arthur Ashkin, James M Dziedzic, JE Bjorkholm, and Steven Chu. Observation of a single-beam gradient force optical trap for dielectric particles. *Optics letters*, 11(5):288–290, 1986.

[77] Yanhua Wu, Dong Sun, and Wenhao Huang. Mechanical force characterization in manipulating live cells with optical tweezers. *Journal of biomechanics*, 44(4):741–746, 2011.

[78] AhRan Kang, JiSoo Park, Jongil Ju, Gi Seok Jeong, and Sang-Hoon Lee. Cell encapsulation via microtechnologies. *Biomaterials*, 35(9):2651–2663, 2014.

[79] Hamed Amini, Wonhee Lee, and Dino Di Carlo. Inertial microfluidic physics. *Lab on a Chip*, 14(15):2739–2761, 2014.

[80] Ling Lin, Qinghua Chen, and Jiashu Sun. Micro/nanofluidics-enabled single-cell biochemical analysis. *TrAC Trends in Analytical Chemistry*, 2017.

[81] Sung-Sik Yun, Sang Youl Yoon, Min-Kyung Song, Sin-Hyeog Im, Sohee Kim, Jong-Hyun Lee, and Sung Yang. Handheld mechanical cell lysis chip with ultrasharp silicon nano-blade arrays for rapid intracellular protein extraction. *Lab on a Chip*, 10(11):1442–1446, 2010.

[82] Jung Kim, Jung Woo Hong, Dong Pyo Kim, Jennifer H Shin, and Inkyu Park. Nanowire-integrated microfluidic devices for facile and reagent-free mechanical cell lysis. *Lab on a chip*, 12(16):2914–2921, 2012.

[83] Guillaume Mernier, Niccolò Piacentini, Thomas Braschler, Nicolas Demierre, and Philippe Renaud. Continuous-flow electrical lysis device with integrated control by dielectrophoretic cell sorting. *Lab on a Chip*, 10(16):2077–2082, 2010.

[84] Nima Jokilaakso, Eric Salm, Aaron Chen, Larry Millet, Carlos Duarte Guevara, Brian Dorvel, Bobby Reddy, Amelie Eriksson Karlstrom, Yu Chen, Hongmiao Ji, et al. Ultra-localized single cell electroporation using silicon nanowires. *Lab on a Chip*, 13(3):336–339, 2013.

[85] Min-Sheng Hung and Ya-Tun Chang. Single cell lysis and dna extending using electroporation microfluidic device. *BioChip journal*, 6(1):84–90, 2012.

[86] Yasuhiro Sasuga, Tomoyuki Iwasawa, Kayoko Terada, Yoshihiro Oe, Hiroyuki Sorimachi, Osamu Ohara, and Yoshie Harada. Single-cell chemical lysis method for analyses of intracellular molecules using an array of picoliter-scale microwells. *Analytical chemistry*, 80(23):9141–9149, 2008.

[87] Chun-Ping Jen, Ju-Hsiu Hsiao, and Nikolay A Maslov. Single-cell chemical lysis on microfluidic chips with arrays of microwells. *Sensors*, 12(1):347–358, 2011.

[88] Yu Hou, Wei Fan, Liying Yan, Rong Li, Ying Lian, Jin Huang, Jinsen Li, Liya Xu, Fuchou Tang, X Sunney Xie, et al. Genome analyses of single human oocytes. *Cell*, 155(7):1492–1506, 2013.

[89] Jeff Gole, Athurva Gore, Andrew Richards, Yu-Jui Chiu, Ho-Lim Fung, Diane Bushman, Hsin-I Chiang, Jerold Chun, Yu-Hwa Lo, and Kun Zhang. Massively parallel polymerase cloning and genome sequencing of single cells using nanoliter microwells. *Nature biotechnology*, 31(12):1126, 2013.

[90] Dmitry Pushkarev, Norma F Neff, and Stephen R Quake. Single-molecule sequencing of an individual human genome. *Nature biotechnology*, 27(9):847, 2009.

[91] Walter Reisner, Niels B Larsen, Asli Silahtaroglu, Anders Kristensen, Niels Tommerup, Jonas O Tegenfeldt, and Henrik Flyvbjerg. Single-molecule denaturation mapping of dna in nanofluidic channels. *Proceedings of the National Academy of Sciences*, 107(30):13294–13299, 2010.

[92] Ernest T Lam, Alex Hastie, Chin Lin, Dean Ehrlich, Somes K Das, Michael D Austin, Paru Deshpande, Han Cao, Niranjan Nagarajan, Ming Xiao, et al. Genome mapping on nanochannel arrays for structural variation analysis and sequence assembly. *Nature biotechnology*, 30(8):771, 2012.

[93] Mark JP Chaisson, John Huddleston, Megan Y Dennis, Peter H Sudmant, Maika Malig, Fereydoun Hormozdiari, Francesca Antonacci, Urvashi Surti, Richard

Sandstrom, Matthew Boitano, et al. Resolving the complexity of the human genome using single-molecule sequencing. *Nature*, 517(7536):608, 2015.

[94] Chenghang Zong, Sijia Lu, Alec R Chapman, and X Sunney Xie. Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science*, 338(6114):1622–1626, 2012.

[95] Stephen Fodor, Richard P Rava, Xiaohua C Huang, Ann C Pease, Christopher P Holmes, and Cynthia L Adams. Multiplexed biochemical assays with biological chips. *Nature*, 364:555–556, 1993.

[96] Adam T Woolley and Richard A Mathies. Ultra-high-speed dna sequencing using capillary electrophoresis chips. *Analytical chemistry*, 67(20):3676–3680, 1995.

[97] Shaorong Liu, Yining Shi, William W Ja, and Richard A Mathies. Optimization of high-speed dna sequencing on microfabricated capillary electrophoresis channels. *Analytical Chemistry*, 71(3):566–573, 1999.

[98] Dieter Schmalzing, Norman Tsao, Lance Koutny, Dan Chisholm, Alok Srivastava, Aram Adourian, Lauren Linton, Paul McEwan, Paul Matsudaira, and Daniel Ehrlich. Toward real-world sequencing by microdevice electrophoresis. *Genome research*, 9(9):853–858, 1999.

[99] Wolfgang Goetzinger, Lev Kotler, Emanuel Carrilho, Marie C Ruiz-Martinez, Oscar Salas-Solano, and Barry L Karger. Characterization of high molecular mass linear polyacrylamide powder prepared by emulsion polymerization as a replaceable polymer matrix for dna sequencing by capillary electrophoresis. *Electrophoresis*, 19(2):242–248, 1998.

[100] Shaorong Liu, Hongji Ren, Qiufeng Gao, David J Roach, Robert T Loder, Thomas M Armstrong, Qinglu Mao, Iuliu Blaga, David L Barker, and Stevan B Jovanovich. Automated parallel dna sequencing on multiple channel microchips. *Proceedings of the National Academy of Sciences*, 97(10):5369–5374, 2000.

[101] Robert G Blazej, Palani Kumaresan, and Richard A Mathies. Microfabricated bioprocessor for integrated nanoliter-scale sanger dna sequencing. *Proceedings of the National Academy of Sciences*, 103(19):7240–7245, 2006.

[102] Raymond Mariella. Sample preparation: the weak link in microfluidics-based biodetection. *Biomedical microdevices*, 10(6):777, 2008.

[103] Michael G Roper, Christopher J Easley, Lindsay A Legendre, Joseph AC Humphrey, and James P Landers. Infrared temperature control system for a

completely noncontact polymerase chain reaction in microfluidic chips. *Analytical chemistry*, 79(4):1294–1300, 2007.

[104] John Eid, Adrian Fehr, Jeremy Gray, Khai Luong, John Lyle, Geoff Otto, Paul Peluso, David Rank, Primo Baybayan, Brad Bettman, et al. Real-time dna sequencing from single polymerase molecules. *Science*, 323(5910):133–138, 2009.

[105] Mark J Chaisson and Glenn Tesler. Mapping single molecule sequencing reads using basic local alignment with successive refinement (blasr): application and theory. *BMC bioinformatics*, 13(1):238, 2012.

[106] David Gordon, John Huddleston, Mark JP Chaisson, Christopher M Hill, Zev N Kronenberg, Katherine M Munson, Maika Malig, Archana Raja, Ian Fiddes, LaDeana W Hillier, et al. Long-read sequence assembly of the gorilla genome. *Science*, 352(6281):aae0344, 2016.

[107] Derek M Bickhart, Benjamin D Rosen, Sergey Koren, Brian L Sayre, Alex R Hastie, Saki Chan, Joyce Lee, Ernest T Lam, Ivan Liachko, Shawn T Sullivan, et al. Single-molecule sequencing and chromatin conformation capture enable de novo reference assembly of the domestic goat genome. *Nature genetics*, 49(4):643, 2017.

[108] Pierre-Gilles De Gennes. *Scaling concepts in polymer physics*. Cornell university press, 1979.

[109] Masao Doi and Samuel Frederick Edwards. *The theory of polymer dynamics*, volume 73. oxford university press, 1988.

[110] Han Cao, Jonas O Tegenfeldt, Robert H Austin, and Stephen Y Chou. Gradient nanostructures for interfacing microfluidics and nanofluidics. *Applied Physics Letters*, 81(16):3058–3060, 2002.

[111] Pierre-Giles de Gennes. Reptation of a polymer chain in the presence of fixed obstacles. *The journal of chemical physics*, 55(2):572–579, 1971.

[112] AR Khokhlov and SK Nechaev. Polymer chain in an array of obstacles. *Physics Letters A*, 112(3-4):156–160, 1985.

[113] C Me Gordon. Some aspects of classical knot theory. In *Knot theory*, pages 1–60. Springer, 1978.

[114] Guy Waldo Dunnington. Carl friedrich gauss, titan of science: a study of his life and work. Exposition Press, 1955.

[115] H Schubert. Sitzungsber. *Heidelberger Akad. Wiss., Math.-Naturwiss. Klasse, 3rd Abhandlung*, 1949.

[116] Alexei B Sossinsky. *Knots: mathematics with a twist*. Harvard University Press, 2002.

[117] Charles Livingston. Knot theory, washington, dc: Math. *Assoc. Amer*, 1993.

[118] Dale Rolfsen. *Knots and links*, volume 346. American Mathematical Soc., 1976.

[119] Martin G Scharlemann. Unknotting number one knots are prime. *Inventiones mathematicae*, 82(1):37–55, 1985.

[120] Kouki Taniyama. Unknotting numbers of diagrams of a given nontrivial knot are unbounded. *Journal of Knot Theory and Its Ramifications*, 18(08):1049–1063, 2009.

[121] Nicole M Baker, Rakhi Rajan, and Alfonso Mondragón. Structural studies of type i topoisomerases. *Nucleic acids research*, 37(3):693–701, 2008.

[122] John L Nitiss. Dna topoisomerase ii and its growing repertoire of biological functions. *Nature Reviews Cancer*, 9(5):327–337, 2009.

[123] Allyn J Schoeffler and James M Berger. Dna topoisomerases: harnessing and constraining energy to govern chromosome topology. *Quarterly reviews of biophysics*, 41(1):41–101, 2008.

[124] Masao Doi. *Introduction to polymer physics*. Oxford university press, 1996.

[125] Michael Rubinstein and Ralph H Colby. *Polymer physics*, volume 23. Oxford University Press New York, 2003.

[126] Walter Reisner, Jonas N Pedersen, and Robert H Austin. Dna confinement in nanochannels: physics and biological applications. *Reports on Progress in Physics*, 75(10):106601, 2012.

[127] Michael E Fisher. Shape of a self-avoiding walk or polymer chain. *The Journal of Chemical Physics*, 44(2):616–622, 1966.

[128] Jonas O Tegenfeldt, Christelle Prinz, Han Cao, Steven Chou, Walter W Reisner, Robert Riehn, Yan Mei Wang, Edward C Cox, James C Sturm, Pascal Silberzan, et al. The dynamics of genomic-length dna molecules in 100-nm channels. *Proceedings of the National Academy of Sciences of the United States of America*, 101(30):10979–10983, 2004.

[129] Walter Reisner, Keith J Morton, Robert Riehn, Yan Mei Wang, Zhaoning Yu, Michael Rosen, James C Sturm, Stephen Y Chou, Erwin Frey, and Robert H Austin. Statics and dynamics of single dna molecules confined in nanochannels. *Physical Review Letters*, 94(19):196101, 2005.

[130] M Daoud and PG De Gennes. Statistics of macromolecular solutions trapped in small pores. *Journal de Physique*, 38(1):85–93, 1977.

[131] Theo Odijk. Scaling theory of dna confined in nanochannels and nanoslits. *Physical Review E*, 77(6):060901, 2008.

[132] Theo Odijk. The statistics and dynamics of confined or entangled stiff polymers. *Macromolecules*, 16(8):1340–1344, 1983.

[133] Theo Odijk. Similarity applied to the statistics of confined stiff polymers. *Macromolecules*, 17(3):502–503, 1984.

[134] Theo Odijk. Dna confined in nanochannels: Hairpin tightening by entropic depletion. *The Journal of chemical physics*, 125(20):204904, 2006.

[135] Peter Cifra, Zuzana Benková, and Tomas Bleha. Chain extension of dna confined in channels. *The Journal of Physical Chemistry B*, 113(7):1843–1851, 2009.

[136] Peter Cifra. Channel confinement of flexible and semiflexible macromolecules. *The Journal of chemical physics*, 131(22):224903, 2009.

[137] Suckjoon Jun, Devarajan Thirumalai, and Bae-Yeun Ha. Compression and stretching of a self-avoiding chain in cylindrical nanopores. *Physical review letters*, 101(13):138101, 2008.

[138] Andrew D Bates and Anthony Maxwell. *DNA topology*. Oxford University Press, USA, 2005.

[139] Hays S Rye, Stephen Yue, David E Wemmer, Mark A Quesada, Richard P Haugland, Richard A Mathies, and Alexander N Glazer. Stable fluorescent complexes of double-stranded dna with bis-intercalating asymmetric cyanine dyes: properties and applications. *Nucleic acids research*, 20(11):2803–2812, 1992.

[140] Chandrashekhar U. Murade, Vinod Subramaniam, Cees Otto, and Martin L. Bennink. Force spectroscopy and fluorescence microscopy of dsdnayoyo-1 complexes: implications for the structure of dsdna in the overstretching region. *Nucleic Acids Research*, 38(10):3423–3431, 2010.

[141] Katrin Gnther, Michael Mertig, and Ralf Seidel. Mechanical and structural properties of yoyo-1 complexed dna. *Nucleic Acids Research*, 38(19):6526–6532, 2010.

[142] Damini Gupta, Julian Sheats, Abhiram Muralidhar, Jeremy J Miller, Derek E Huang, Sara Mahshid, Kevin D Dorfman, and Walter Reisner. Mixed confinement regimes during equilibrium confinement spectroscopy of dna. *The Journal of chemical physics*, 140(21):214901, 2014.

[143] LD Landau. Lifshitz. em 1967: Statistical physics.

[144] Christoph G Baumann, Steven B Smith, Victor A Bloomfield, and Carlos Bustamante. Ionic effects on the elasticity of single dna molecules. *Proceedings of the National Academy of Sciences*, 94(12):6185–6190, 1997.

[145] Alexander R Klotz, Lyndon Duong, Mikhail Mamaev, Hendrick W de Haan, Jeff ZY Chen, and Walter W Reisner. Measuring the confinement free energy and effective width of single polymer chains via single-molecule tetris. *Macromolecules*, 48(14):5028–5033, 2015.

[146] Dirk Stigter. Interactions of highly charged colloidal cylinders with applications to double-stranded dna. *Biopolymers*, 16(7):1435–1448, 1977.

[147] JT Mannion, CH Reccius, JD Cross, and HG Craighead. Conformational analysis of single dna molecules undergoing entropically induced motion in nanochannels. *Biophysical journal*, 90(12):4538–4545, 2006.

[148] Christian Hermann Reccius, John Thomas Mannion, Joshua David Cross, and HG Craighead. Compression and free expansion of single dna molecules in nanochannels. *Physical review letters*, 95(26):268101, 2005.

[149] Ute Zettl, Sebastian T Hoffmann, Felix Koberling, Georg Krausch, Jorg Enderlein, Ludger Harnau, and Matthias Ballauff. Self-diffusion and cooperative diffusion in semidilute polymer solutions as measured by fluorescence correlation spectroscopy. *Macromolecules*, 42(24):9537–9547, 2009.

[150] Masao Doi. *Soft matter physics*. Oxford University Press, 2013.

[151] Ahmed Khorshid, Philip Zimny, David Tétreault-La Roche, Geremia Massarelli, Takahiro Sakaue, and Walter Reisner. Dynamic compression of single nanochannel confined dna via a nanodozer assay. *Physical review letters*, 113(26):268104, 2014.

[152] Ahmed Khorshid, Susan Amin, Zhiyue Zhang, Takahiro Sakaue, and Walter W Reisner. Nonequilibrium dynamics of nanochannel confined dna. *Macromolecules*, 49(5):1933–1940, 2016.

[153] Alexander Y Grosberg and Yitzhak Rabin. Metastable tight knots in a wormlike polymer. *Physical review letters*, 99(21):217801, 2007.

[154] Liang Dai, C Benjamin Renner, and Patrick S Doyle. Metastable tight knots in semiflexible chains. *Macromolecules*, 47(17):6135–6140, 2014.

[155] Erika Ercolini, Francesco Valle, Jozef Adamcik, Guillaume Witz, Ralf Metzler, Paolo De Los Rios, Joaquim Roca, and Giovanni Dietler. Fractal dimension and localization of dna knots. *Physical review letters*, 98(5):058102, 2007.

[156] Vsevolod Katritch, Wilma K Olson, Alexander Vologodskii, Jacques Dubochet, and Andrzej Stasiak. Tightness of random knotting. *Physical Review E*, 61(5):5545, 2000.

[157] Liang Dai, C Benjamin Renner, and Patrick S Doyle. Origin of metastable knots in single flexible chains. *Physical review letters*, 114(3):037801, 2015.

[158] Paul G Dommersnes, Yacov Kantor, and Mehran Kardar. Knots in charged polymers. *Physical Review E*, 66(3):031802, 2002.

[159] Emmanuel Guitter and Enzo Orlandini. Monte carlo results for projected self-avoiding polygons: a two-dimensional model for knotted polymers. *Journal of Physics A: Mathematical and General*, 32(8):1359, 1999.

[160] Wolfram Mobius, Erwin Frey, and Ulrich Gerland. Spontaneous unknotting of a polymer confined in a nanochannel. *Nano letters*, 8(12):4518–4522, 2008.

[161] Luca Tubiana, Angelo Rosa, Filippo Fragiacomo, and Cristian Micheletti. Spontaneous knotting and unknotting of flexible linear polymers: Equilibrium and kinetic aspects. *Macromolecules*, 46(9):3669–3678, 2013.

[162] Xiaozhong Zheng and Alexander Vologodskii. Tightness of knots in a polymer chain. *Physical Review E*, 81(4):041806, 2010.

[163] Susan Amin, Ahmed Khorshid, Lili Zeng, Philip Zimny, and Walter Reisner. A nanofluidic knot factory based on compression of single dna in nanochannels. *Nature Communications*, 9(1):1506, 2018.

[164] Giuseppe D'Adamo and Cristian Micheletti. Molecular crowding increases knots abundance in linear polymers. *Macromolecules*, 48(11):6337–6346, 2015.

[165] Chihiro Nakajima and Takahiro Sakaue. Localization and size distribution of a polymer knot confined in a channel. *Soft-Matter*, 9:3140–3146, 2013.

[166] Cristian Micheletti and Enzo Orlandini. Knotting and unknotting dynamics of dna strands in nanochannels. *ACS Macro Letters*, 3(2):876–880, 2014.

[167] Calin Plesa, Verschueren Daniel, Sergii Pud, Jaco van der Torre, Justus W Ruitenberg, Magnus P. Jonsson, Alexander Y Grosberg, Yitzhak Rabin, and Cess Dekker. Direct observation of dna knots using a solid-state pore. *Nature Nanotechnology*, 11:1093–1097, 2016.

[168] S. L. Levy, J. T. Mannion, J. Cheng, C. H. Reccius, and H. G. Craighead. Entropic unfolding of dna molecules in nanofluidic channels. *Nano Lett.*, 8:3839–3844, 2008.

[169] William R Taylor. Protein knots and fold complexity: some new twists. *Computational biology and chemistry*, 31(3):151–162, 2007.

[170] Henrik Bruus. *Theoretical Microfluidics*. Oxford, 2007.

[171] Hiong Yap Gan, Yee Cheong Lam, Nam Trung Nguyen, Kam Chiu Tam, and Chun Yang. Efficient mixing of viscoelastic fluids in a microchannel at low reynolds number. *Microfluid Nanofluid*, 3:101–108, 2007.

[172] Lorenzo Bertini, Alberto de Sole, Davide Gabrielli, Giovanni Jona-Lasinio, and Claudio Landim. Macroscopic fluctuation theory. *Reviews of Modern Physics*, 87(11):593–635, 2015.

[173] Douglas R. Tree, Yanwei Wang, and Kevin D Dorfman. Modeling the relaxation time of dna confined in a nanochannel. *Biomicrofluidics*, 7:054118, 2013.

[174] Robert G Newcombe. *Confidence intervals for proportions and related measures of effect size*. CRC Press, 2012.

[175] J. F. C. Kingman. *Poisson Processes*. Clarendon Press, 1992.

[176] Jinsheng Zhou, Yanqian Wang, Laurent D. Menard, Sergey Panyukov, Michael Rubinstein, and Michael Ramsey. Enhanced nanochannel translocation and localization of genomic dna molecules using three-dimensional nanofunnels. *Nature Communications*, 8:807, 2017.

[177] Zubair Azad, Maedeh Roushan, and Robert Riehn. Dna brushing shoulders: Targeted looping and scanning of large dna strands. *Nano Letters*, 15:5641–5646, 2015.

[178] Benjamin Trefz, Jonathan Siebert, and Peter Virnau. How molecular knots can pass through each other. *Proceedings of the National Academy of Sciences*, 111(22):7948–7951, 2014.

[179] Daniel Reith, Peter Cifra, Andrzej Stasiak, and Peter Virnau. Effective stiffening of dna due to nematic ordering causes dna molecules packed in phage capsids to preferentially form torus knots. *Nucleic acids research*, 40(11):5129–5137, 2012.

[180] Peter Virnau, Florian C Rieger, and Daniel Reith. Influence of chain stiffness on knottedness in single polymers. *Biochemical Society Transactions*, 41(2):528–532, 2013.

[181] Aiquen Huang, Walter Reisner, and Aniket Bhattacharya. Dynamics of dna squeezed inside a nanochannel via a sliding gasket. *Polymers*, 8:352, 2016.

[182] Giovanni Jona-Lasinio. Thermodynamics of stationary states. *Journal of Statistical Mechanics: Theory and Experiment*, 2:P02004, 2014.

[183] Frank H. J. van der Heyden, Derek Stein, and Cees Dekker. Streaming currents in a single nanofluidic channel. *Physical Review Letters*, 95:116104, 2005.

[184] Frank H. J. van der Heyden, Douwe Jan Bonthuis, Derek Stein, Christine Meyer, and Cees Dekker. Power generation by pressure-driven transport of ions in nanofluidic channels. *Nano Letters*, 7:1022–1025, 2007.

[185] Derek Stein, H. J. van der Heyden, Wiepke J. A. Koopmas, and C. Dekker. Pressure-driven transport of confined dna polymers in fluidic channels. *Proced. Nat. Acad. Sci. USA*, 3:15853–15858, 2006.

[186] H. D. Vollmer. Two particle model for the diffusion of interacting particles in periodic potentials. *Z. Physik B*, 33:103–109, 1979.

[187] Stanley Y Shaw and James C Wang. Knotting of a dna chain during ring closure. *Science*, 260(5107):533–536, 1993.

[188] Alexei E Likhtman and Tom CB McLeish. Quantitative theory for linear dynamics of linear entangled polymers. *Macromolecules*, 35(16):6332–6343, 2002.

[189] Alexander R Klotz, Beatrice W Soh, and Patrick S Doyle. Motion of knots in dna stretched by elongational fields. *Physical review letters*, 120(18):188003, 2018.

[190] Marco Di Stefano, Luca Tubiana, Massimiliano Di Ventra, and Cristian Micheletti. Driving knots on dna with ac/dc electric fields: topological friction and memory effects. *Soft Matter*, 10(34):6491–6498, 2014.

[191] CC Adams and CC Brown. The knot book: An elementary introduction to the mathematical theory of knots. *Nature*, 371(6498):568–568, 1994.

[192] Alexander Vologodskii. Monte carlo simulation of dna topological properties. *Topology in molecular biology*, pages 23–41, 2007.

[193] Frank B Dean, Andrzej Stasiak, Theo Koller, and Nicolas R Cozzarelli. Duplex dna knots produced by escherichia coli topoisomerase i. structure and requirements for formation. *Journal of Biological Chemistry*, 260(8):4975–4983, 1985.

[194] FY Wu. Knot theory and statistical mechanics. *Reviews of modern physics*, 64(4):1099, 1992.

[195] James W Alexander. Topological invariants of knots and links. *Transactions of the American Mathematical Society*, 30(2):275–306, 1928.

[196] Vaughan FR Jones. A polynomial invariant for knots via von neumann algebras. *Bulletin of the American Mathematical Society*, 12(1):103–111, 1985.

[197] P Abgrall and AM Gue. Lab-on-chip technologies: making a microfluidic network and coupling it into a complete microsystema review. *Journal of Micromechanics and Microengineering*, 17(5):R15, 2007.

[198] Sara Mahshid, Mohammed Jalal Ahamed, Daniel Berard, Susan Amin, Robert Sladek, Sabrina R Leslie, and Walter Reisner. Development of a platform for single cell genomics using convex lens-induced confinement. *Lab on a Chip*, 15(14):3013–3020, 2015.

[199] EJ Janse Van Rensburg, DAW Sumners, E Wasserman, and SG Whittington. Entanglement complexity of self-avoiding walks. *Journal of Physics A: Mathematical and General*, 25(24):6557, 1992.