Studying Self-Entangled DNA at the Single Molecule Level

by

C. Benjamin Renner

B.S. Chemical Engineering, University of Tennessee (2009),
B.S. Mathematics, University of Tennessee (2010),
M.Eng. Chemical Engineering Practice, Massachusetts Institute of Technology (2015).

Submitted to the Department of Chemical Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Chemical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2015

Author Signature redacted

Department of Chemical Engineering
August 7, 2015

Certified by Signature redacted

Patrick S. Doyle
Robert T. Haslam (1911) Professor of Chemical Engineering
Thesis Supervisor

Accepted by Signature redacted

Richard D. Braatz
Chairman, Department Committee on Graduate Students

Studying Self-Entangled DNA at the Single Molecule Level

by
C. Benjamin Renner

Submitted to the Department of Chemical Engineering
on August 7, 2015, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Chemical Engineering

Abstract

Knots seem to be found every time one encounters long, stringy objects. At the macroscopic scale, knots are seen every day in shoelaces, tangled hair, or woven clothing, yet they also present themselves at the microscopic scale in long polymer molecules. Knots can be found often in DNA packaged within the viral capsid, occasionally in proteins, and during the transcription and replication of genomic DNA. Biological knots are similarly thought to change the dynamics of viral ejection, protein digestion, and translocation of biomolecules through nanopores. Despite the prevalence of knots in important biological polymers, to date, the physics of knots is only partially understood.

DNA has become a well-accepted model system for investigating the physics of single polymer molecules due to its tremendous biological significance and useful experimental properties. Recent advances in microscopy and nanofabrication have enabled the real-time manipulation and imaging of single DNA molecules, facilitating fundamental studies concerning the physics of individual polymers. Leveraging these experimental techniques, this thesis aims to explore the changes knots can impart on the static and dynamic properties of single DNA molecules.

We first demonstrate a mechanism for the previously observed phenomenon of the compression and self-knotting of a single DNA molecule in the presence of an electric field. We then use this mechanism to study the process of stretching complex DNA knots in an extensional field. These knots dramatically alter the way DNA stretches in two ways: an initially arrested state and a subsequently slowed stretching phase. Our work consists of the first experimental support of these phenomena, originally predicted by simulation and theory.

We then develop theoretical arguments, shown to agree with simulation results, for the physics that govern the distribution of sizes of knots that stochastically occur on DNA molecules, and more broadly, all semiflexible polymers. We then extend our theory to the case where the entire DNA molecule is confined and elongated within a channel. Here, the complex non-monotonic behavior of the sizes of knots agrees with our modified theory.

We finally present the results of dynamical simulations where knots on polymers interact with flows or forces. We first examine the behavior of a knot along a polymer extended by extensional flow. The flow may cause a knot to be swept off a polymer molecule, and the motion of a knot is consistent with a model. Different families of knots display different rates of motion, and we explain this difference with a simple topological mechanism. We then turn to examine the case of knots jamming on a polymer molecule extended with high tensile forces. A simple energy barrier hopping argument qualitatively explains the observed slowdown in dynamics of knots. We use these results

to reexamine the problem of DNA knots jamming during nanopore translocation, and our results establish the potential for using knots to slow and control the rate of translocation by a ratcheting mechanism.

The impact of this thesis is threefold. First, we have demonstrated a novel experimental platform capable of interrogating DNA knots, likely the most efficient of its kind. Second, we have established a theoretical framework for the size and probability of knotting in single molecules capable of directing experiments where these properties need to be controlled. Finally, we have shown how knotted topologies can be manipulated by external flows or forces, which have applications involving preconditioning molecules to unknotted states or the jamming of knotted molecules in nanopores.

Thesis Supervisor: Patrick S. Doyle
Title: Robert T. Haslam (1911) Professor of Chemical Engineering

Acknowledgments

One of the biggest misconceptions about the PhD journey is that it is a solitary one. I have been incredibly fortunate for the productive collaborations, personal and professional support, and lasting friendships that have advanced and sustained me in the pursuit of this goal. A number of important and special people have contributed immensely to this thesis in these ways, and I will do my best to thank them here.

First, I would like to thank my scientific mentors. Foremost is my advisor, Professor Patrick Doyle. Throughout my PhD, he has been a constant source of guidance and direction. His style for approaching problems, "thinking deeply" and striving to "understand the fundamental physics," has left an indelible mark on my own scientific style. I am thankful for his positive attitude, patience, and unwavering faith in me as these qualities made possible overcoming the difficulties of research. I also thank the members of my thesis committee, Professor Alfredo Alexander-Katz and Professor Bradley Olsen. They went above and beyond in providing additional scientific perspectives as well as encouraging me in my research.

I have also been fortunate to directly collaborate with many talented people during the course of research. I first want to thank Ning Du for her collaboration on my first paper. I owe a huge debt of gratitude to Liang Dai and thank him deeply. Our almost effortless collaborations substantially increased the quality and quantity of my PhD work, and I have also enjoyed our personal discussions and adventures at conferences. Finally, I thank Vivek Narsimhan with whom I have also collaborated closely during the final year of my PhD.

Beyond direct collaborations, I have been very fortunate to work along side a number of wonderful people in the Doyle research group. I thank Rathi Srinivas for our many lunches, mocha Mondays, and conversations. I thank Burak Eral, coworker and roommate, whose ability to put people at ease and joyful attitude greatly enriched my time at MIT. I thank Harry An for being an excellent lab social chair and for the fun times he organized. I thank William Uspal for his computer related support and many enjoyable discussions. I thank Jason Rich for his encouragement and perspectives as I was joining the group and he was leaving. I thank all members of the "First Bay": Jeremy Jones and Alona Birjiniuk, Andrew Fiore, and Li-Chiun Cheng, for maintaining the cleanest labspace in the group. I thank my UROP, Ha-young Kang, for her dedication and perseverance. I thank Priyadarshi Panda, Matt Helgeson, Ki Wan Bong, Stephen Chapin, Ankur Gupta, Jiseok Lee, Jae Jung Kim, Hyundo Lee, Lynna Chen, Seung Goo Lee, Gaelle Le Goff, Hyewon Lee, Sarah Shapiro, and anyone else I may have forgotten for all being excellent labmates.

A number of people have been important on supporting me on a personal level during my time at MIT. I have had the privilege to develop many friendships during the course of my PhD, and I am very grateful for them. In particular, I would like to thank Stephen Morton, my roommate for nearly my entire PhD, for the many fun times we've enjoyed. My girlfriend, Connie Gao, has been a constant companion in the last half of my PhD, greatly enriching my life. She has brought great happiness to my life, and I cherish the time we spend together. Finally, I thank my family: my sister, Amy, parents, Omer and Nancy Renner. They were constantly supportive, positive, and encouraging, even when I fall short. Their unconditional love has sustained me during this journey.

Table of Contents

List of Figures

CHAPTER 1

Introduction

This thesis encompasses a series of studies concerning the biophysics of single DNA molecules containing self-entanglements, more commonly referred to as knots. This Chapter supplies the motivation and context necessary for the understanding of the broad designs of this thesis.

1.1 DNA Biophysics

1.1.1 Motivation

DNA is widely known to be an essential molecule in all living things as it encodes and stores genetic information. A DNA molecule consists of two complementary strands, interwoven in the famous double-helix structure [2]. Each strand consists of a repetition four distinct nucleotides: adenine (A), guanine (G), thymine (T), and cytosine (T). The specific sequence of nucleotides in the DNA of an organism has enormous biological implications; it differentiates organisms into distinct species, controls the susceptibility to a variety of diseases and treatments, and effectuates a wide variety of physical traits. As a result, tremendous efforts have been made to develop representative maps (sequences) of the human genome [3] as well as other species [4, 5].

Due to its overarching importance as a biological molecule, the biophysical properties of DNA are of considerable scientific interest. The physics of DNA are intertwined with nearly all techniques used to determine genomic sequences, such as the separation of DNA fragments through sieving medium, the manipulation of DNA in micro and nanofluidic devices, and the translocation of DNA

through nanoscopic pores. Beyond the laboratory sequencing of genomes, transcription, nuclear packing, and a host of other *in vivo* processes are intimately connected to the physical behavior of the DNA molecule.

DNA is a polymer, as it is made up of a series of nucleotides. Due to the ubiquity of polymers in modern life, the physics of polymer molecules has been, is, and will continue to be of tremendous scientific interest. DNA has widely been used "model polymer" in studies that inform the general body of knowledge of polymer physics. The reasons of the status of DNA as a model polymer are numerous and practical: DNA is commercially available in monodisperse samples of varying lengths; DNA can be stained with a wide variety of fluorescent intercalating dyes for visualization [6, 7]; DNA has size and timescales of motion that are resolvable with common fluorescence microscopy setups; and DNA can easily be transported in microfluidic devices with electric fields or hydrodynamic flows.

In short, DNA is a tremendously important biological molecule. The biophysical properties of DNA are a vital to the sequencing DNA molecules, understanding how DNA behaves in the cell, and can be used to infer more general behavior about polymer molecules.

1.1.2 Previous Work

For these reasons, numerous investigators have set about interrogating DNA with well-defined forces, flows, and applied fields [8] as these studies simultaneously address questions of fundamental polymer physics [9] and direct linear analysis for DNA mapping [10]. Magnetic tweezers [11] and optical traps [12, 13] have been used to apply precise forces to single molecules revealing quantitative information such as force-extension behavior [14] and overstretched structure [15]. In hydrodynamic flows, DNA has been used to experimentally test a wide range of theoretical predictions [9]. Similarly, a focus of research in our group has been the design of microfluidic devices to linearize single electrophoresing DNA by collisions with posts [16], translation through hyperbolic contractions [17], and stretching in T-junctions [18] and cross-slot channels [19, 20].

Nanofabrication techniques have allowed for the construction of precise confining geometries. Tube-like (1D) and slit-like (2D) confinement have been used in numerous single molecule studies as they serve as canonical examples for nanopores and many conventional lab-on-a-chip devices, respectively [21]. Furthermore, long-standing scaling theories for single molecules in tubes [22, 23] and slits [24–27] have been rigorously interrogated with single molecule DNA experiments. As introduction of a confining dimension restricts the conformational space of a polymer, confining geometries are useful for controlling molecular configurations. For example, strong nanoconfinement in tubes [22, 28] as well as entropy-driven extension at micro and nano interfaces in slits [29], have been shown to be able to nearly fully extend DNA.

1.2 Self-Entanglements

With an exposition of the rationale and results of biophysical interrogation of DNA in place, we will now proceed to explain why self-entanglements within DNA are of particular interest in this thesis.

1.2.1 Motivation

Self-entanglements of a string, often called knots, are pervasive. Knots have been among the most ancient tools, used by humans long before discovery of the wheel of the wheel [30]. The hoisting of sails, suturing of wounds, and weaving of clothing all involve knots, and the ensuing technological advances in exploration, medicine, and modesty are but a few of the numerous human acheivements enabled by knots.

In many respects, polymer molecules are microscopic strings, and like strings, entanglements can occur between segments along polymers. Concentrated solutions of polymers have tremendous industrial importance, and their physics are dominated by the intermolecular entanglements between molecules. These entanglements introduce new length and time scales into the solution, which in turn affect rheological and material properties. The tube model of Edwards [31] and reptation theory of de Gennes [32] were among the first successful in describing the dynamical properties of these systems.

While the physics of intermolecular entanglements is a mature field of study [33], intramolecular entanglements, *i.e.* the self-entanglements germane to this thesis, are much less well understood. Just as knots are either intentionally tied [34] or randomly occur [35] in macroscopic strings, these types of self-entanglements can be found in cellular DNA. These DNA self-entanglements are manipulated by topoisomerase enzymes, so-named due to their ability to change topology, and this process plays a vital role in DNA transcription and replication [36, 37]. Underscoring the importance of the topology of DNA in basic cellular functioning, poisoning or inhibiting topoisomerases is the mechanism by which many successful chemotherapeutics effect apoptosis [38], and researchers are now beginning to experimentally measure the drastic mechanical changes of the genome resulting from these drugs [39].

Beyond DNA in the cell, knots are found in a remarkably wide range of biophysical systems [40]. Knots have been observed with high frequency in the DNA confined to the tight spaces of viral capsids [41, 42]. Knots occur in proteins [43–45], albeit more rarely than found in similar collapsed polymers [46]. To date, out of all native protein conformations, only enzymes contain knots, yet knots are extremely abundant in denatured proteins [47]. Only up to three out of thousands of examined RNA structures are possibly knotted [48]. These observations have raised intriguing questions concerning connections between topology and biological function of molecules.

In short, self-entanglements are found in a wide range polymers, and their presence can lead to unusual behavior. A greater understanding of the physics of self-entangled DNA can begin to clarify these observations and inform technological applications.

1.2.2 Previous Work

More generally, long polymers almost always contain knots [49], so in some sense, these knotted topologies are unavoidable. This fact has sparked much fundamental scientific inquiry [50], some of which we will explore in greater depth here.

Fundamental theoretical and simulation work on knotted polymers has a rich history. Self-entanglements have been shown to introduce additional long relaxation timescales into polymer rings [51, 52], a qualitatively analogous phenomena to the slowed dynamics of entangled melts. Intuitively, simulations revealed that self-entanglements become common in condensed polymers [53]. Much theoretical [54–56] and simulation [55, 57, 58] work suggests that the expansion of globular polymers to coils in a temperature-jump experiment progresses through long-lived, topo-

logically stabilized arrested states. Furthermore, self-entanglements have been shown to produce counterintuitive phenomena such as imparting effective excluded volume interactions to phantom chains [59, 60] and localization and self-tightening of knots due to entropic effects [61].

More recently, simulations have directly addressed the role of knots in various biological processes. The self-entangled DNA within viral capsids ejects via a ratcheting mechanism [62], and the rate of DNA ejection is greatly influenced by the complexity of the interior self-entanglements [63]. Simulations of the translocation of proteins through nanoscopic pores have shown that knots can greatly slow [64] and possibly jam [65] the this process, and simulations of the translocation of knotted DNA through nanopores demonstrate a similar result [66].

Experiments capable pf testing the results of theory and simulation have been limited. Trefoil [67] and pentafoil [68] knots have been created through highly specific chemical syntheses. Knots have been tied in actin filaments [69] and DNA[70] with optical tweezers, yet this procedure is time consuming and laborious. Other studies have tracked the dyanmics of knots formed via collisions with nanochannel walls [71] or in agarose gels [72], yet formation is stochastic in these studies. Recently, electric fields have been used to compress single DNA [73, 74], providing a simple and efficient means for generating potentially complex single molecular entanglements, a technique is foundational to the experimental work in this thesis.

Notably, knots in linear DNA molecules have specific applications to current approaches in genomic sequencing. Direct linear analysis [10] is technique where DNA is extended and relative positions of sequences are inferred via the relative positions of sequence-specific fluorescent probes. In certain devices used in this technology, knotted chains have likely been observed, and the knots appear as false artificial deletions, slightly complicating data analysis [75]. Recently, nanopores, both biological and synthetic [76], have attracted attention for their ability to interrogate the properties of small molecules. In particular, these devices have shown promise in sequencing DNA molecules [77]. Knots have been experimentally detected with solid state nanopores [78], and the jamming of knotted DNA [66] may play a key role in a reduction in translocation velocity, facilitating genomic sequencing via this platform.

1.3 Objectives

The overall goal of this thesis is to gain fundamental insight into the physics of self-entangled single DNA molecules. To this end, we employ DNA as a model system, due to its biological relevance and physical features, to perform single molecule experiments that directly investigate these topological interactions. These experiments are complemented by theoretical and simulation results on knotted DNA molecules that lend further insight into specific connections between topology and the static, dynamic, and out of equilibrium properties of DNA molecules.

Given this outline, this work of this thesis is broadly structured into the following themes:

1. Experimentally characterizing the mechanisms of entanglement and self-knotting of single DNA molecules and the affects these entanglements can have on out of equilibrium dynamics.

2. Developing models to explain the statistical physics of knots along polymers.

3. Understanding direct connections between dynamics and topology of a knot through computer simulations.

1.4 Overview of Results

In Chapter 2, we first present an overview of the basics of knot theory that are needed for understanding later Chapters of this thesis. Chapter 3 presents experiments that explore a mechanism for compression and self-entanglement of DNA under electric fields. A physical model, adapted from more concentrated suspensions of charged colloids, is shown to explain the observed compression and self-entanglement of DNA under moderate electric fields. In Chapter 4, this mechanism is used to precondition DNA into knotted states, which are then stretched by extensional fields in a microfluidic chip. These experiments reveal that topological entanglements substantially slow the rate at which a molecule may be stretched, and these results are captured semi-quantitatively by a model. Chapter 5 investigates the physics behind the sizes of knots that may occur randomly in dilute, unconfined semiflexible polymers such as DNA. A theoretical model, which includes the semiflexible bending energy of the chain and finite width excluded volume interactions, is presented and is shown to quantitatively reproduce the results of computer simulations of knots along polymers. In Chapter 6 this model is extended to incorporate the effect of confining the polymer into rectangular channel. The length scale of external confinement is shown to drastically and non-monotonically alter the sizes of knots that may form, yet quantitative agreement between the extended model and simulations allow for a physical understanding of this effect. Chapter 7 explores the dynamics of a knot along a DNA molecule extended strongly by an extensional flow. It is demonstrated that the flow can "untie" the knot from the DNA, and the motion of the knot is shown to follow a 1D diffusion-convection equation. Torus knots are shown to untie more readily than all other observed knots, and this discrepancy is reconciled via a simple topological mechanism. Chapter 8 explores the slowdown in dynamics of knots in chains held under high tensions, *i.e.* jamming. Chapter 9 more broadly examines the scientific impact of this thesis and provides commentary on future research directions arising from the natural extension of the ideas of this thesis.

CHAPTER 2

Some Notes on Knot Theory

Knot theory is a well developed and active branch of mathematics. Many of these mathematical approaches can be applied to physical systems like polymers, fibers, and strings where physical knots may occur. In this Chapter, several concepts from knot theory that are useful for understanding simple polymeric knots are presented, and these ideas are essential to results in Chapters 5, 6, 7, and 8. The focus of this Chapter will be of providing a concise and practical overview in lieu of strict mathematical rigor. The reader interested in the latter is invited to consult any of the numerous introductory mathematical texts on the topic, such as references [79, 80].

2.1 Knot Theory

To a topologist, a *knot* is a circle embedded in $\mathbf{R}^3$, *i.e.* a closed curve in $\mathbf{R}^3$ without self-intersection. Knots are traditionally visualized in two-dimensional projections called *knot diagrams*. In a knot diagram, each crossing of the knot consists of an over strand and an under strand. The under strand is illustratively broken at the crossing to orient the two strands. Diagrams of many knots are shown in Fig. 2.1.

Two representations of "simplest" knot, called the *unknot* or *trivial knot* are shown in Fig. 2.1a. While it seems intuitive, one might ask: why do these two diagrams represent the same knot? One way to answer this question is to introduce the three distinct *Reidemeister moves*, shown in Fig. 2.1b. Given two knots (or two diagrams of knots), the knots are said to be *equivalent* if and

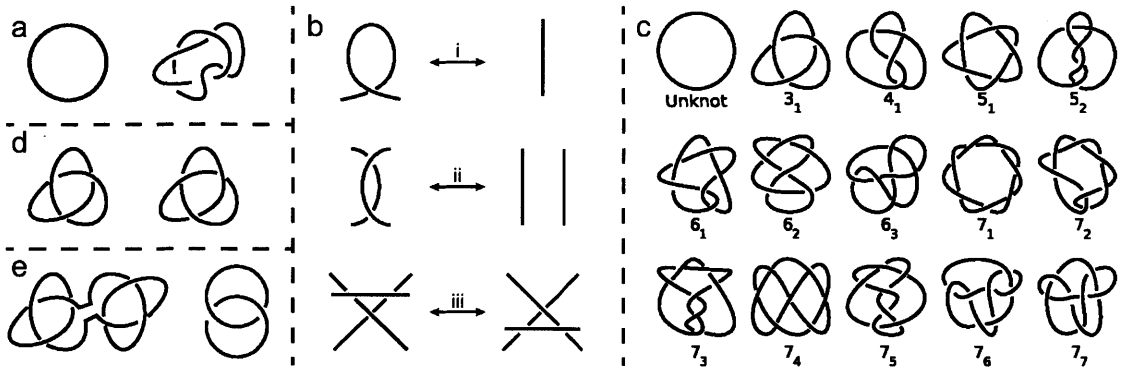

Fig. 2.1: *(a) Knot diagrams of the unknot. (b) Illustration of type (i), (ii), and (iii) Reidemeister moves. (c)A table of the unknot and prime knots of up to seven crossings. Knots are labeled with Alexander-Briggs notation. (d) Knot diagrams for the (left) left-handed and (right) right-handed trefoil knot. (e) (left) The knot sum of two same-handed trefoil knots yields the "granny" knot. (right) A diagram of simplest non-trivial link, the Hopf link. The table of prime knots in (c) as well as the knots in (d) and (e) are modified from https://commons.wikimedia.org/wiki/File:Knot_table.svg*

only if they can be transformed to one another via a finite series of Reidemeister moves. To put this concept in a physical perspective, if one imagines a knot diagram as a representations of a closed string, a series of Reidemeister moves can describe any possible physical manipulations of this string (except breaking strands).

Now with the notion of equivalence of knots in place, we will introduce a number of non-equivalent knots in Fig. 2.1c. These knots are each labeled with a number X_Y; this labeling is referred to as Alexander-Briggs notation. Here, the X represents the *crossing number* of the knot, which is the smallest number of crossings in any diagram of a knot. The value Y is assigned is rather arbitrary (although certain types of knots tend to be listed first), but is necessary to delineate between non-equivalent knots that have the same crossing number, such as the 5_1 and 5_2 knots.

If a knot and its mirror image are not equivalent, that knot is said to be *chiral*. This is illustrated in Fig. 2.1d with the "simplest" non-trivial knot, the 3_1 knot, also called the trefoil knot. The left-handed and right-handed versions of the trefoil knot are each pictured. As one cannot convert between the two via Reidemeister moves, the trefoil knot is chiral. A knot is said to be *achiral* or *amphichiral* that knot and its mirror image are equivalent. The 4_1 knot is a simple example of an achiral knot.

We will now introduce a few final concepts that are relevant to the topology of polymer chains yet outside the purview of this thesis. The closely related notions of the a *knot sum*, *composite knot*, and *prime knot* are depicted on the left of Fig. 2.1e. A *knot sum* of two knots K_1 and K_2 is performed as follows. Take the diagrams of the knots and place them side by side. Chose a direction of motion along the curve of each knot (this is called orienting a knot diagram). Remove a small arc from each diagram and connect the two diagrams together with a consistent final

orientation. The knot that ensues from completing this operation between knots K_1 and K_2, is denoted $K_1\#K_2 = K_2\#K_1$. One trivial property of the knot sum is $K_1\#K_2 = K_1$, if K_2 is the unknot.

A composite knot can be represented as the sum of two non-trivial knots. The $3_1\#3_1$ "granny" knot is shown on the left in Fig. 2.1e. It is the sum of two trefoil knots with the same chirality. Note that the sum of two trefoils of opposite chiralities produces a different knot: the $3_1\#3_1^*$ "square" knot.

Prime knots, analogous in some sense to prime numbers, cannot be represented by the sum of two non-trivial knots. In this thesis, only prime knots are studied, yet in polymer systems, composite knots become increasingly commong as the length of the polymer increases. In the systems under examination in this thesis, however, composite knots tend to "factor" into segregated prime components along the chain, and thus, the exclusive treatment of prime knots can inform these situations as well.

Finally, links, the discrete unions of disjoint knots, are not investigated in this thesis. The simplest non-trivial link of two or more curves, the Hopf link, is depicted on the right in Fig. 2.1e. Here, two unknots are linked together. While interesting and even relevant to the topology of circular genomes [81], this thesis deals with single DNA or polymers, and two or more curves are outside the scope of this work.

2.1.1 Knot Invariants

A *knot invariant* is a quantity that is assigned the same value for all equivalent knots. Implicit in this definition is that the quantity assigned to a knot need not necessarily be unique. One of the primary motivations for research into knot invariants is the problem of distinguishing knots.

One popular knot invariant has already been introduced: the crossing number. While application of Reidemeister moves to the diagram of a knot may increase or decrease the number of crossings in that diagram, the minimum number of crossings needed to represent the knot is obviously unchanged. An easy example of the nonunique nature of knot invariants is that two non-equivalent knots have crossing number five, 5_1 and 5_2.

Some of the most widely used knot invariants are knot polynomials. Knot polynomials are typically calculated from the diagram of a knot, and the degree and coefficients of the polynomials encode information about the type of knot. The oldest and perhaps most famous of these polynomial invariants is the Alexander polynomial of a knot K, $\Delta_K(t)$, and will be used later on in this thesis. Other popular knot polynomials include the Conway polynomial, Jones polynomial, HOMFLY polynomial, Kauffman polynomial.

No currently known knot polynomial is able to distinguish all knots; moreover, it is not even known for certain that any currently known knot polynomial can distinguish all non-trivial knots from the trivial knot. That said, the combination of the calculation of a polynomial invariant along with additional information about the knot of interest may sometimes uniquely identify the knot. For example, $\Delta_K(t) \neq 1$ from a knot diagram with three crossings precisely indicates a trefoil knot (albeit of unknown chirality). From the practical perspective of this thesis, we are only concerned with relatively simple knots, such as those pictured in Fig. 2.1c. Details on an approach to calculate the Alexander polynomial for a polymer conformation as well as combine it with other information to locate knots can be found in Appendix A.

2.2 Closure Schemes

The previous discussion of useful concepts from knot theory quite clearly dealt with closed curves only. The primary interests of this thesis concern the self-entanglements of linear, *i.e.* open, polymer chains. Fortunately, the cognitive dissonance created by these two facts does have a resolution, which we will now spell out in detail.

The conventional approach, used in this thesis as well as in numerous other published papers by others, is to systematically introduce an artificial path or set of paths that connect(s) the ends of the open chain, resulting in a final closed configuration. This technique is particularly attractive since upon applying such a closure, the rich library of techniques to classify topologies of closed chains, namely counting the number of crossings on the knot diagram and calculating knot invariants, becomes unambiguously applicable to the system. For the purposes of our work, the combination of a closure scheme and calculating a knot invariant (the Alexander polynomial) are very frequently used in an iterated approach to calculate the type and size of a knot (see Appendix A for more details). In this section, we will breifly review some schemes for closing open configurations of polymer chains. The majority of these schemes are discussed in reference [82], and it is an excellent resource for the reader interested in understanding this subject in greater detail.

2.2.1 Direct Closure

Perhaps the simplest and most intuitive way to close an open chain is to add a closing segment that directly bridges between the two ends. This scheme is here referred to as *direct closure* and is something that we have avoided in any of our analyses. While reference [82] has indicated that this technique provides reasonable results for polymers at equilibrium, this approach often yields curious results for polymers that are significantly deformed such as by external force, flow, or confinement. This routine becomes problematic as even though one may have a polymer strongly extended by flow with a seemingly well-defined knot in the middle, directly closing the ends can frequently penetrate the knotted region. This results in a closed configuration that is often an unknot or a different kind of knot that what the extended conformation seems to unambiguously suggest.

2.2.2 Implicit Closure

Another scheme, employed in Chapter 7, is what we will here call *implicit closure*. Rather than generating a closure, this scheme simply assumes that some path exists to connect the ends of the projection of the polymer molecule that does not generate any additional crossings. Since the molecule is never closed, the name as well as inclusion in this section may seem misleading, yet it is analogous to the other schemes as it serves the purpose of assuming a state by which techniques used on closed chains can still be applied to calculate topological properties.

With this approach, the 2D diagram of crossings of the linear molecule is sufficient to calculate the Alexander polynomial via the combinatoric approach outlined in Appendix A. This approach is particularly attractive in cases where a knotted molecule is otherwise strongly extended by some external force, flow, or confining geometry. In these cases, the portions outside the knot rarely cross and the ends of the molecule are typically close to minimal and maximal positions of the molecule along the extending axis. Thus, there typically exists an arc to close the molecule without generating additional crossings.

This approach is useful because of its simplicity and computational expediency. To calculate a knot invariant, rather than iteratively removing segments of the chain, one can iteratively remove crossings. The latter approach is significantly quicker for chains with many more segments than crossings.

This approach is not without significant drawbacks. First, this technique is only reasonable for extended configurations where the ends are likely to be near the "perimeter" of the molecule. Furthermore, unlike all the other approaches outlined here, since this approach does not actually close the chain, it is highly sensitive to the 2D projection one chooses. The only clear constraint on the chosen projection is that the vector normal to the 2D plane of projection must also be orthogonal to the axis of extension. Many such normal vectors exist, and the crossing points for a 2D projection will change depending on which one is chosen.

2.2.3 Stochastic Closure at Infinity

A more complex approach, described in reference [83] and implemented in reference [84], seeks to generate an ensemble of closed configurations from a single open configuration by connecting the ends of the chain to random points on the edge of an "infinitely" large (relative to the size of the chain) sphere and then bridging those ends. This technique is called *stochastic closure at infinity.*

An attractive part of this technique is that there is a built-in encoding of the natural ambiguity of a "knot" on an open polymer — some closures will produce one type of knot, others may produce a different type of knot, and still others may produce an unknot. Over a large number of configurations, one is left with a detailed statistical sense of what types of knots those configurations may be considered to contain.

Unfortunately, this same quality leads to many serious drawbacks for certain applications. First, each configuration must be randomly closed a large number N_c of times. Typically, some knot invariant must then be calculated for each of those closures and repeated again for each independent chain configuration of interest. This routine increases the computational cost of calculating knot properties by approximately a factor of N_c, which can be very computationally inefficient. Additionally, finding the boundaries of a knot are complicated by the fact that each individual closure may produce a different knot type. This problem can be addressed by first identifying the dominant knot type upon closing the chain. The fraction of knots with this topology upon closure can be called f_{knot}. As one removes chain segments and repeats this procedure, the edges of this knot can be defined as the segments which if removed cause $f_{knot} < f_{test}$, where f_{test} is some arbitrary threshold.

2.2.4 Minimally-Interfering Closure

Finally, arguably the most complex yet most robust scheme is *minimally-interfering closure*, introduced in reference [82]. In this approach, two candidate closing pathways are considered for each configuration: (i) direct closure of the ends as in Section 2.2.1 and (ii) extending the each end through the nearest point on the convex hull that encompasses the entire chain and then bridging those two extensions with an arc at infinity. The distance between the two ends of the chain d_{in}, and the sum of the distances from the ends to their nearest points on the convex hull of the chain d_{out} are the respective measures of "interference." Therefore, if $d_{in} \leq d_{out}$, the ends are directly bridged. Otherwise, the closure described in (ii) is applied.

This approach has the advantage of generating a single ring of fixed topology for any given configuration. Furthermore, this approach, like *direct closure* and *stochastic closure at infinity*, creates physically closed loops, and therefore, the results are independent of a chosen 2D projection that is used for later analysis.

The drawbacks of this approach are the added complexity and computational expense of computing the convex hull. Fortunately, efficient routines for calculating the convex hull of a set of points have been developed, such as Qhull [85].

CHAPTER 3

Enhanced electrohydrodynamic collapse of DNA due to dilute polymers

In this Chapter, we examine a mechanism by which DNA self-knots in the presence of an electric field. We notice that adding small, charge neutral polymers to the electrophoresis buffer can cause two wildly divergent behaviors: accelerated DNA collapse and self-knotting or DNA elongation and stretching. These two cases are shown to be functions of the physical properties of the small, charge neutral polymers. By thinking carefully about the ways in which these dilute polymers may collide with DNA, we are able to explain these behaviors. Next, in the case when DNA self-knots, we are able to borrow a mechanism for aggregation of charged colloids under electric fields to show the same mechanism applies to single DNA. These results represent the first physical understanding of DNA knotting by this mechanism, a technique we will exploit in later chapters.

The results in this chapter have been published in reference [86] and are reproduced with permission from AIP Publishing, Copyright 2014. This research was performed in close collaboration with Ning Du, who performed most of the experiments shown here.

3.1 Overview

We experimentally demonstrate that addition of small, charge-neutral polymers to a buffer solution can promote compression of dilute solutions of single electrophoresing DNA. This phenomenon

contrasts with the observed extension of DNA during capillary electrophoresis in dilute solutions of high molecular weight polymers. We propose these discrepancies in micron-scale DNA configurations arise from different nano-scale DNA-polymer collision events, controlled by solute polymer properties. We build upon theories previously proposed for intermolecular DNA aggregation in polymer-free solutions to develop scaling theories that describe trends seen in our data for intramolecular DNA compaction in dilute polymer solutions.

3.2 Introduction

Electric fields are widely used to transport and manipulate DNA in micro/nanodevices [87, 88] with applications in molecular genetics [89, 90], nucleic acid-based diagnostics [91, 92], and fundamental studies of polyelectrolytes [93]. The interplay between electrokinetics and polymer physics in electric field mediated DNA transport has been shown to cause DNA aggregation [94–97], a phenomenon that either hinders the separation of long DNA molecules by capillary electrophoresis, or can be harnessed to concentrate DNA on-chip. Recently, it was demonstrated that electric fields can cause compression of single, large ($\sim$ 100 kbp) DNA molecules in a standard electrophoresis buffer [73, 74]. As shown by Tang et al. [73], a moderate electric field of a few hundred V/cm induces strong coil-to-globule compaction and self-entanglement of DNA. Increasing ionic strength or decreasing DNA molecular weight lessens this effect. Tang et al. developed scaling relations to collapse data that lends support to the postulate that the mechanism driving the intramolecular compression of a single DNA molecule is analogous to what Isambert et al. [96, 97] developed for electric field induced intermolecular DNA aggregation in more concentrated solutions ($\sim$overlap concentration c^*). Other models including curvature condensation [74, 98] were proposed to explain the physical mechanism leading to compression. However, as remarked by Tang et al. [73], the curvature condensation theory neglects electrohydrodynamic effects which are the underlying physics of the mechanism proposed by Isambert et al. [96, 97].

According to Isambert et al. [96, 97], the electrohydrodynamic flow that ultimately leads to DNA aggregation is a consequence of concentration fluctuations of macroions and augmented by the different mobilities of the small salt ions and much larger macroions (*e.g.* DNA or colloids). If this mechanism also explains the compression of single DNA, adding neutrally-charged macromolecules to the system should enhance these compressive flows since they serve to reduce the mobility of the macroion (DNA) without significantly changing the mobility of smaller salt ions. In this manuscript, we experimentally investigate the proposed mechanism by systematically studying the conformations of T4 DNA under electric fields in various dilute neutral polymer solutions of different concentrations and molecular weights. We find that the addition of dextran polymers to the electrophoresis buffer enhances DNA compression at polymer concentrations well below that required for depletion-induced DNA compaction in bulk [99]. We contrast this phenomenon with an observed DNA elongation when the same volume fraction of hydroxypropyl cellulose (HPC) polymers is added to the electrophoresis buffer. We argue that this dissimilarity arises due to differences in DNA-polymer collisions. Following the arguments of Isambert et al. [96, 97], we derive the key dimensionless group that drives the compaction process and show that it is able to collapse the experimental data.

3.3 Experimental Methods

The conformations of dilute (concentration $c \ll c^*$ where c^* is the overlap concentration, see Appendix C for details) fluorescently labeled T4 DNA (165.6 kbp, Nippon gene) in 0.5X Tris-Borate-EDTA (TBE, Sigma-Aldrich) buffer were examined under uniform electric fields in 2 μm tall by 200 μm wide straight microchannels using an inverted fluorescence microscope (IX71, Olympus) combined with an EM-CCD camera (Andor) with an exposure time of 0.01 s and at a rate of 24 frames per second. For convenience in describing the polymer concentrations, we define a scaled concentration $\Phi = c/c^*$. Physically, Φ corresponds to the fraction of the volume pervaded by the polymer coils when envisioned as spheres with a radius equal to the polymer radius of gyration. Unless otherwise noted, all polymer concentrations were at a value of $\Phi = 0.6$ and so can be considered dilute. Dextran polymers are charge-neutral, branched, flexible, and readily dissolved in most solutions. The dextran polymers used in this study have molecular weights of 5 kDa, 80 kDa, 410 kDa and 2000 kDa (Pharmacosmos). The concentration of dextran polymers used is well below the value of $\Phi \sim 3-4$ needed for the smallest dextran polymers ($M_w = 5$k) to induce condensation of unconfined DNA in the absence on an electric field due to depletion interactions [100]. The larger molecular weight dextran polymers require even higher volume fractions to collapse DNA. We also examined the effects of the following charge-neutral linear polymers: hydroxypropyl cellulose (HPC) of molecular weight 100 kDa, 370 kDa, and 1000 kDa from Sigma-Aldrich, and polyvinylpyrrolidone (PVP) of molecular weight 10 kDa and 1000 kDa from Polysciences. A uniform electric field (up to a few hundred V/cm) was applied to the channel using an external DC power source. More detailed experimental information (*e.g.* overlap concentrations and polymer properties) is given in Appendix C.

3.4 Experimental Results and Discussion

The probability distributions, P, of the radius of gyration, R_g, of T4 DNA at equilibrium and under uniform electric fields of 23 and 46 V/cm in a dextran solution ($M_w = 410\text{k}, \Phi = 0.6$) are shown in Fig. 1a. Without an applied electric field, the DNA molecules sample a wide range of expanded configurations. Under an electric field of 46 V/cm, the DNA molecules are significantly more compact and exhibit much smaller size fluctuations (Fig. 1a). As the field strength increases, the molecules become more isotropic, reflected by a decrease in the ratio of the radii of the major and minor axes R_M/R_m of the radius of gyration tensor [27] (Fig. 1b). Results for three dextran solutions ($M_w \leq 410\text{k}$) with the same ionic strength (0.5X TBE) and volume fraction ($\Phi = 0.6$) but different sizes of dextran polymers are shown in Fig. 1c & d. We observe a continuous decrease in $\langle R_g \rangle / \langle R_{g,\text{eq}} \rangle$ (ensemble average R_g normalized by their equilibrium average $\langle R_{g,\text{eq}} \rangle$) and $\langle R_M/R_m \rangle$ of T4 DNA with increasing field strength. Without added dextran polymers, the onset of T4 DNA compression occurs at higher field strengths. The largest dextran polymers among the three sizes ($M_w = 410\text{k}$) reduce the electric field threshold required for DNA compression to 40 V/cm, compared to 150 V/cm in a dextran-free solution (Fig. 1c & d). An enhanced compression of DNA due to solute polymers ($M_W = 410\text{k}$ dextran) was also seen for $\Phi = 0.3$ and 0.6, shown in Fig. S4 in Appendix C. It is important to note that the trend of decreasing compression field thresholds with increasing size of dextran polymers is opposite to that observed in depletion-induced DNA condensation in absence of electric fields [100, 101] where, for a given volume fraction, dextran polymers with a larger molecular weight cause less compression. Thus, depletion effects are not the

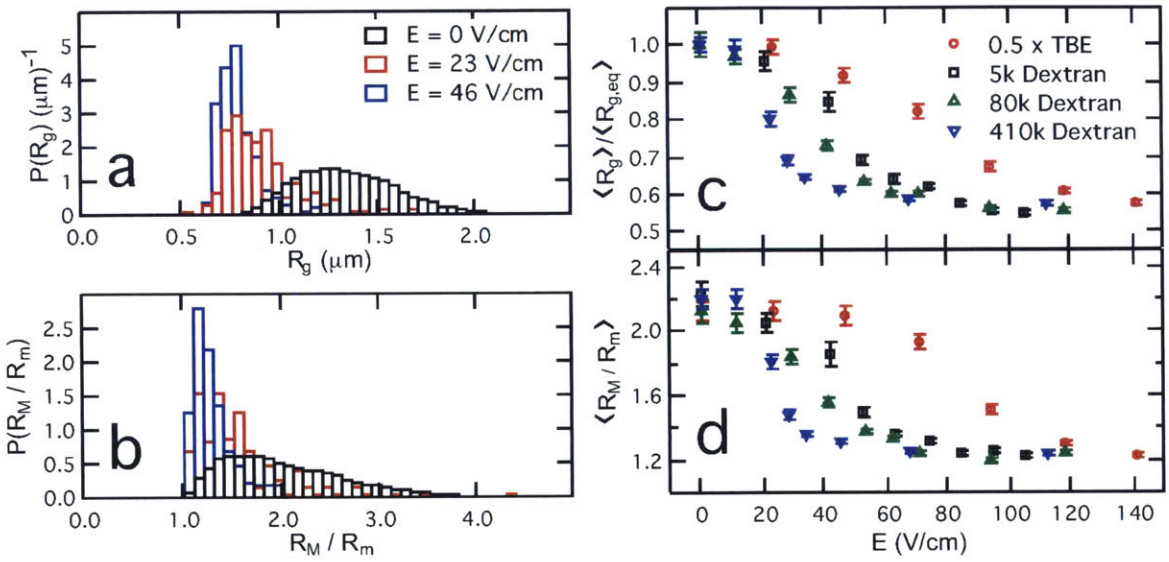

Fig. 3.1: *Compression of T4 DNA in dextran solutions. Probability distributions of (a) R_g, and (b) R_M/R_m of T4 DNA in dextran solutions (M_w = 410k, volume fraction $\Phi = 0.6$, 0.5X TBE) at equilibrium (0 V/cm) and under uniform DC electric fields of $E = 23$ and 46 V/cm. All probability distributions, $P(x)$, are constructed to satisfy the normalization criteria: $\int_{-\infty}^{\infty} P(x)dx = 1$. (c)-(d) Conformations of T4 DNA under uniform DC electric fields in 0.5X TBE and dextran solutions with the same volume fraction ($\Phi = 0.6$) but different molecular weights (M_w = 5k, 80k, and 410k) in 0.5X TBE. (c) Ensemble average radius of gyration R_g of T4 normalized by the equilibrium average $\langle R_{g,\mathrm{eq}} \rangle$, and (d) the corresponding average ratio between the major and minor axes $\langle R_M/R_m \rangle$ as functions of field strength E. If not visible, the error bar (standard error) is smaller than the symbol size.*

driving mechanism behind the polymer-assisted compression of DNA in electric fields.

The results in Fig. 1 are remarkable considering it is widely known that *extension* of DNA is observed during electrophoresis in solutions containing relatively rigid linear polymers such as hydroxyethyl cellulose (HEC) [102, 103] or hydroxypropyl cellulose (HPC) [104]. In fact, the extension of DNA in these dilute polymeric solutions is a result of the mechanism that enables size dependent separation via capillary electrophoresis. HPC, a stiff linear polymer with similar physical properties as HEC, induces significant extension of DNA in an applied field, shown in Fig. 2a by the increase in $\langle R_g \rangle / \langle R_{g,\mathrm{eq}} \rangle$ with increasing field strength. Furthermore, the extension and alignment of DNA with the applied field naturally result in increasingly anisotropic configurations, displayed in Fig. 2b. Additionally, the size of added HPC polymers has a marked effect on DNA conformations. The addition of high M_w HPC polymers (370k and 1000k) results in DNA extension, whereas lower M_w HPC polymers (100k) facilitate DNA compression.

These drastic differences between the conformations of electrophoresing DNA in different dilute polymer solutions indicate that the properties of the solute polymer are critically important to the overall physics. To further explore these effects, we investigate the conformations of DNA electrophoresing in high molecular weight dextran ($M_w = 2000$k) solutions. As opposed to solutions with smaller dextran polymers ($M_w \leq 410$k), the larger 2000k dextran solution causes moderate initial compression at low field strengths followed by substantial extension of T4 DNA at high field strengths (Fig. 2c & d). This observation is consistent with the observed extension of T2 DNA in 2000k dextran solutions at 200 V/cm by Wang and coworkers [105]. The effect of PVP, a linear polymer, on the conformation of electrophoresing DNA is shown in Fig. 2e & f. Compression of T4 DNA is seen in 10k PVP solutions while elongation is seen in 1000k PVP solutions at field strengths larger than 100 V/cm. This is the same qualitative behavior as seen in dextran solutions with low (Fig. 1c & d) and high (Fig. 2c & d) molecular weights. Thus, for a given polymer, increasing molecular weight can result in a crossover between assisting compression to inducing extension of DNA, indicating a change in the physical process that governs DNA conformations.

3.5 Theory for DNA-Polymer Collisions

We propose that the extension of DNA in dilute polymeric solutions results from the forces involved in nanoscopic DNA-polymer collisions. These forces are largely determined by the electrophoretic velocity U of the DNA, the physical properties of the solute polymer (radius of gyration $R_{g,\mathrm{p}}$, diffusivity D_{p}, and persistence length $l_{p,\mathrm{p}}$), and the physical properties of the DNA (bare width $w_{\mathrm{DNA}} \approx 2$ nm, persistence length $l_{p,\mathrm{DNA}} \approx 50$ nm, contour length $L_{c,\mathrm{DNA}} \approx 75$ μm, and diffusivity D_{DNA}). Inspired by literature on DNA-post collisions during electrophoresis, we define a Péclet number [16, 106]

$$\mathrm{Pe_c} = \frac{U R_{g,\mathrm{p}}}{D_{\mathrm{p}}} \tag{3.1}$$

which represents the ratio the time for the solute polymer to diffuse versus the time to convect over the collision length scale which is $R_{g,\mathrm{p}}$ (as $R_{g,\mathrm{p}} > w_{\mathrm{DNA}}$ for all solvent polymers studied). For $\mathrm{Pe_c} \ll 1$, the solute polymer diffuses around the approaching DNA, avoiding collisions; DNA-polymer collisions begin to occur at $\mathrm{Pe_c} \sim 1$, depicted in Fig. 3.3. After collision, the solute polymer becomes entrained with the electrophoresing DNA and imparts a local drag force of $F_{D,\mathrm{p}} = U\zeta_{\mathrm{p}}$ where ζ_{p} is the drag coefficient of the entrained polymer. By comparing this drag force of the entrained polymer to the scales for the elastic spring forces of the polymer $F_{S,\mathrm{p}} \sim \frac{k_B T}{l_{p,\mathrm{p}}}$ and DNA

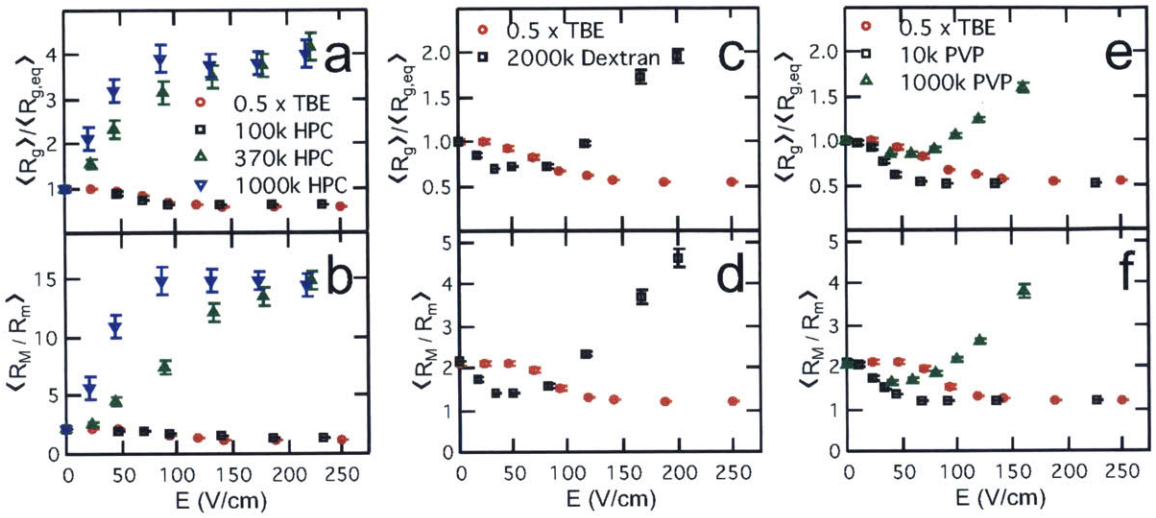

Fig. 3.2: *Conformations of T4 DNA under uniform DC electric fields in HPC solutions of* $M_w = 100$k, *370k, and 1000k (a) & (b), dextran solution of* $M_w = 2000$k *(c) & (d), and PVP solutions of* $M_w = 10$k *and 1000k (e) & (f). All solutions are at* $\Phi = 0.6$ *in 0.5X TBE. If not visible, the error bar (standard error) is smaller than the symbol size.*

$F_{S,\mathrm{DNA}} \sim \frac{k_B T}{l_{p,\mathrm{DNA}}}$ two additional Péclet numbers can be formed,

$$\mathrm{Pe_p} = \frac{U l_{p,\mathrm{p}} \zeta_\mathrm{p}}{k_B T}, \text{ and} \tag{3.2}$$

$$\mathrm{Pe_{DNA}} = \frac{U l_{p,\mathrm{DNA}} \zeta_\mathrm{p}}{k_B T}. \tag{3.3}$$

We note that prior to the collision, the solute polymer is considered to be undeformed, so the drag coefficient can be calculated from the equilibrium diffusivity of the polymer via the Stokes-Einstein relation, $\zeta_\mathrm{p} = \frac{k_B T}{D_\mathrm{p}}$. The Péclet number for the solute polymer can then be written as $\mathrm{Pe_p} = \frac{U l_{p,\mathrm{p}}}{D_\mathrm{p}}$, a more common expression. The magnitudes of $\mathrm{Pe_p}$ and $\mathrm{Pe_{DNA}}$ determine whether a single collision event can cause deformation in the solute polymer or DNA, respectively. However, when considering the conformation of DNA, the magnitude of the force on the DNA due to drag of all (N) entrained polymers must be considered, $F^N_{D,\mathrm{DNA}} = N \times F_{D,\mathrm{p}}$. The number of solute polymers entrained with the DNA contour (N) can be calculated by

$$N = \tau_\mathrm{c} \times k_\mathrm{enc} \times P_\mathrm{c}\left(\mathrm{Pe_c}\right) \tag{3.4}$$

where k_enc is the rate at which DNA encounters solute polymers, $P_\mathrm{c}\left(\mathrm{Pe_c}\right)$ is the probability that the encountered solute polymer actually collides with the DNA contour, and τ_c is the duration time of such a polymer-DNA collision event. We note that P_c is a function of $\mathrm{Pe_c}$ and asymptotically

approaches zero for small $\mathrm{Pe_c}$ and unity for large $\mathrm{Pe_c}$.

The rate at which the electrophoresing DNA encounters solute polymers is

$$k_{\mathrm{enc}} = \frac{3UR_{g,\mathrm{p}}L_{c,\mathrm{DNA}}}{4\pi R_{g,\mathrm{p}}^3}\Phi. \tag{3.5}$$

This expression requires some explanation. The product $R_{g,\mathrm{p}}L_{c,\mathrm{DNA}}$ is the cross-sectional area available for polymer-DNA collisions. Multiplying this term by the electrophoretic velocity gives the rate at which such a collision volume is generated by the electrophoresing DNA. Dividing through by the volume of a solute polymer coil, $4\pi R_{g,\mathrm{p}}^3/3$, gives the rate at which the DNA would encounter a solute polymer for $\Phi = 1$, and multiplying through by Φ renormalizes this rate for the actual volume fraction.

Now, using Eqs. 3.4 & 3.5, $F_{D,\mathrm{DNA}}^N$ can be compared to $F_{S,\mathrm{DNA}}$. Dropping the factor of $4\pi/3$, we arrive at

$$\mathrm{Pe_{eff}} = \mathrm{Pe_{DNA}} \times \left(\tau_{\mathrm{c}} \times \frac{UL_{c,\mathrm{DNA}}}{R_{g,\mathrm{p}}^2}\Phi\right) \times P_{\mathrm{c}}\left(\mathrm{Pe_c}\right), \tag{3.6}$$

the dimensionless group that determines the configuration of the electrophoresing DNA in a dilute solution of neutral polymers.

In the Table 3.1, we use a representative scale for an electrophoretic velocity of $U = 100$ μm/s to present values of these groups for all studied polymer solutions. To consistently define polymer properties for all polymer solutions, we estimate $R_{g,\mathrm{p}} = \left(\frac{3M_w}{4\pi N_A c^*}\right)^{1/3}$ and $D_{\mathrm{p}} = \frac{k_BT}{6\pi\eta R_{h,\mathrm{p}}}$, where $\eta = 1$ cP is the solvent buffer viscosity and $R_{h,\mathrm{p}} = R_{g,\mathrm{p}}/1.56$ in the limit of long, linear chains [107]. For nearly all polymers studied, $\mathrm{Pe_p} \ll 1$, and $\mathrm{Pe_p}$ is still below unity for the two largest HPC solutions. Therefore, for all solutions, we make the assumption that the polymers do not deform during collision, so $\zeta_{\mathrm{p}} = \frac{k_BT}{D_{\mathrm{p}}}$ in all expressions and $\mathrm{Pe_{DNA}} = \frac{l_{p,\mathrm{DNA}}}{l_{p,\mathrm{p}}}\mathrm{Pe_p}$. A further consequence of this fact is that the polymer must disengage from the DNA via diffusive motion, and since $D_{\mathrm{p}} \gg D_{\mathrm{DNA}}$ for all solutions studied, the duration of a polymer-DNA collision event is $\tau_{\mathrm{c}} = \frac{R_{g,\mathrm{p}}^2}{D_{\mathrm{p}}}$. This yields a simplified expression of

$$\mathrm{Pe_{eff}} = \mathrm{Pe_{DNA}} \times \left(\frac{UL_{c,\mathrm{DNA}}}{D_{\mathrm{p}}}\Phi\right) \times P_{\mathrm{c}}\left(\mathrm{Pe_c}\right). \tag{3.7}$$

Inspection of the relative sizes of these dimensionless groups clarifies the physical picture. For 10k PVP, 5k, 80k, and 410k dextran, and 100k HPC polymers, $\mathrm{Pe_c} \ll 1$, so $P_{\mathrm{c}}\left(\mathrm{Pe_c}\right)$ is near zero, and thus $\mathrm{Pe_{eff}}$ becomes small, and the DNA do not extend during electrophoresis. In solutions where extension is observed, we note that $\mathrm{Pe_c} \sim 1$, and thus we expect DNA-polymer collisions to occur. In these cases, we use a conservative lower bound of $P_{\mathrm{c}}\left(\mathrm{Pe_c}\right) = 0.05$ for all solutions to calculate $\mathrm{Pe_{eff}}$. We find that $\mathrm{Pe_{eff}} \gg 1$ for all these solutions, consistent with the experimentally observed extension of DNA. A comparison of the proposed DNA-polymer collision physics for both compacted and extended DNA configurations under an electric field is shown in Fig. 3.3.

It is interesting to note that this analysis does not explicitly consider the structure of the solute polymer, emphe.g. branched dextran polymers vs. linear PVP or HPC polymers. These effects are reflected in the $R_{g,\mathrm{p}}$, $l_{p,\mathrm{p}}$, and ζ_{p}. These differences are sufficient to order the solute polymers in Table 3.1. The structure of the solute polymer (linear vs. branched) may affect the conformational change of DNA beyond what is captured by $R_{g,\mathrm{p}}$, $l_{p,\mathrm{p}}$, and ζ_{p}. This is suspected

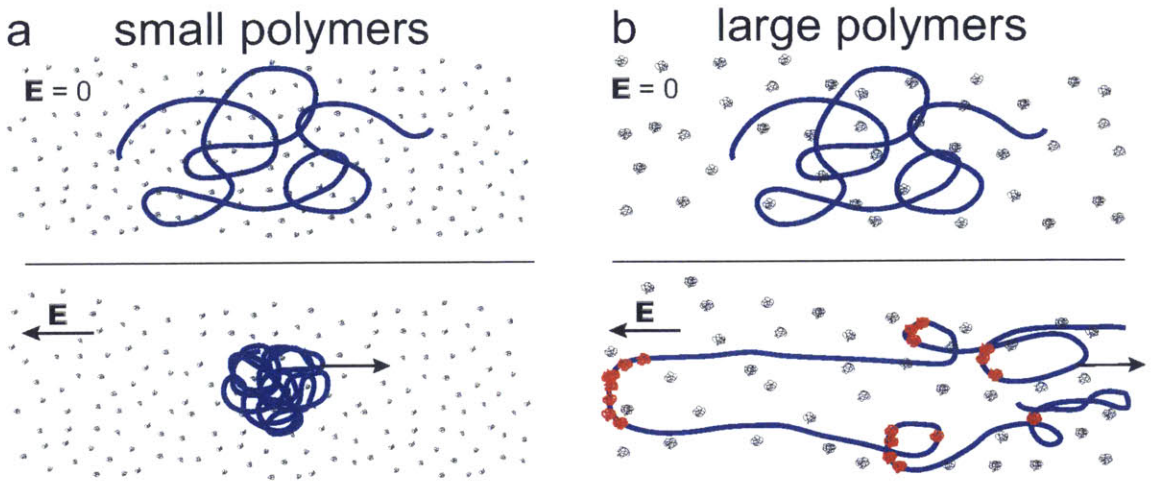

Fig. 3.3: *Schematic of DNA electrophoresing in dilute solutions of polymers. Top: DNA at rest prior to application of the electric field. Bottom: Steady state configurations long after switching on the electric field. (a) DNA with "small" polymers such that* $\mathrm{Pe_c} \ll 1$ *and* $\mathrm{Pe_{eff}} \ll 1$. *(b) DNA with "large" polymers such that* $\mathrm{Pe_c} \sim 1$ *and* $\mathrm{Pe_{eff}} \gg 1$. *Polymers entrained with DNA contour are shown in red.*

to be at most a secondary effect since changing only the molecular weight is able to induce a transition from compressed DNA to extended DNA for both linear (PVP) and branched (dextran) solute polymers.

3.6 Electrohydrodynamical Mechanism of DNA Collapse

We next explore the mechanism by which relatively small polymers may act to assist DNA compression by examining a proposed electrohydrodynamic scaling. In these solutions, collision events are not substantially extending the DNA, and the polymers primarily act to reduce the overall mobility of the DNA.

We will develop our theory by building off the main result from the Isambert and coworkers [96, 97]. The key tenets in their theory are: 1) DNA (macroions) and salt (microions) will have differing electrophoretic mobilities denoted by μ_M and μ_s, respectively, and 2) Brownian motion will give rise to spontaneous fluctuations in the DNA concentration denoted by δc_M. Isambert et al. showed that in the vicinity of these concentration fluctuations there will be an induced hydrodynamic flow that leads to further increase the local DNA concentration and ultimately gives rise to DNA aggregates. The flow field was found to scale as

$$v_h \sim \epsilon\epsilon_0 \frac{\mu_s \delta c_M N_M L_h}{\mu_M c_s \eta} E^2 \tag{3.8}$$

where N_M is the number of charges per DNA, $\epsilon\epsilon_0$ is the solvent dielectric constant, c_s is salt

Table 3.1: *Various Péclet numbers for the polymeric solutions used in this study.*

	Pe_c	Pe_p	Pe_{DNA}	Pe_{eff}
PVP, 10k	0.0019	0.00097	0.037	-
PVP, 1000k	0.33	0.013	0.49	11
Dextran, 5k	0.0020	0.00031	0.038	-
Dextran, 80k	0.021	0.0010	0.12	-
Dextran, 410k	0.086	0.0020	0.25	-
Dextran, 2000k	0.34	0.0040	0.50	11
HPC, 100k	0.042	0.035	0.18	-
HPC, 370k	0.20	0.077	0.39	6.7
HPC, 1000k	0.66	0.14	0.70	22

concentration, η is the solvent viscosity, and L_h is the length scale of the electrohydrodynamic flow. Readers are referred to the original papers of Isambert et al. for a detailed derivation.

To describe the effect of this flow on a single macroion, we derive a scaling for a dimensionless group, the Deborah number, which characterizes the balance between compression due to electrohydrodynamics and the macroion (DNA) entropic tendency to expand back to coiled conformation: $\mathrm{De} \equiv \dot{\varepsilon}\tau$, where $\dot{\varepsilon}$ is the strain rate of the electrohydrodynamic flow, and τ is the longest relaxation time of the macroion in solution. We then have

$$\dot{\varepsilon} \equiv \frac{v_h}{L_h} \sim \epsilon\epsilon_0 \frac{\mu_s \delta c_M N_M}{\mu_M c_s \eta} E^2. \tag{3.9}$$

Since the longest relaxation time $\tau \sim \eta$ when c_s and molecular weight of DNA are kept constant, we find

$$\mathrm{De} \equiv \dot{\varepsilon}\tau \sim \epsilon\epsilon_0 \frac{\mu_s \delta c_M N_M}{\mu_M c_s} E^2. \tag{3.10}$$

We now relate the above scaling to our experimental data. For all dextran solutions, c_s and N_M are held constant, and the dielectric constant $\epsilon\epsilon_0$ [108, 109] and μ_s are approximately constant (see the Appendix C for the μ_s data). As a result, De is only dependent on μ_M, δc_M, and E. As described above, the primary consequence of the dextran polymers is to effectively decrease the overall mobility of the macroion. As shown in Fig. S1 C, DNA molecules migrate much slower in dextran solutions than in 0.5X TBE, *i.e.* 2 (μm/s)/(V/cm). As larger polymer chains produce greater tension on electrophoresing DNA coils, DNA migrate the slowest in 410k dextran solution, and their mobilities exhibit the largest magnitude of fluctuations (Fig. S1 C). This finding agrees with our observation of DNA conformational fluctuations in dextran solutions at low field strength (Fig. 4). During electrophoretic migration, DNA coils are moderately stretched by the dextran polymers and recoil back as these polymers are released from them. Consequently, periodic oscillations in DNA size and mobility are observed. Larger dextran polymers likewise induce more significant fluctuations in DNA volume δV, and since $\delta c_M \sim \delta$V they also give rise to larger magnitudes of segment density fluctuations. As a result, dextran polymers with larger molecular weights promote DNA compression more strongly by reducing DNA mobility and amplifying DNA segment density fluctuations more significantly (see Eq. 3.10).

We finally test if the collapse of DNA in solutions of low molecular weight dextran polymers ($M_w \leq$ 410k) is quantitatively consistent with our proposed model by considering a scaling analysis

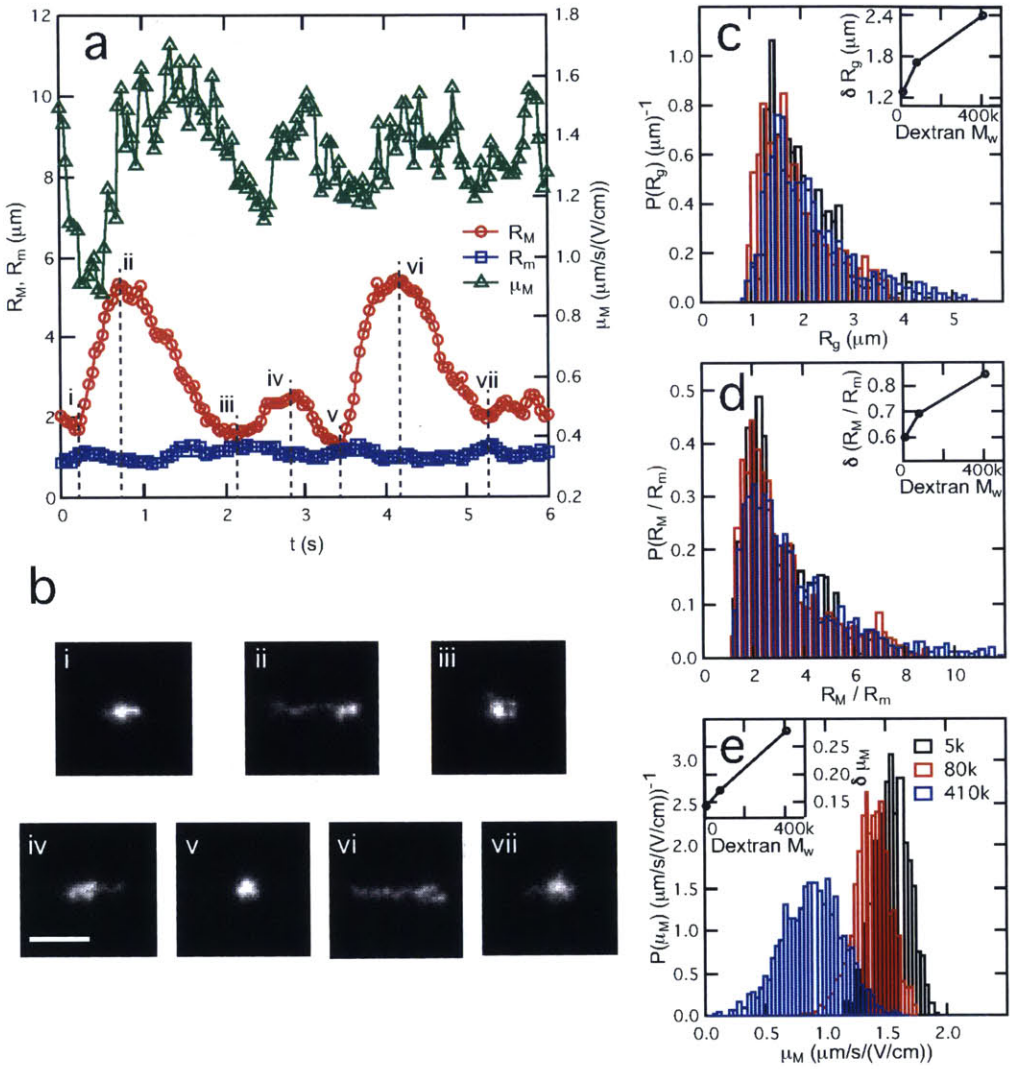

Fig. 3.4: *DNA conformations and mobility fluctuations in dextran solutions at 15 V/cm. (a) Representative oscillation profiles of μ_M, R_M, and R_m of T4 DNA in dextran solutions M_w = 80k, $\Phi = 0.6$, 0.5X TBE. (b) Snapshots of DNA conformations during electrophoresis corresponding to each time points in (a). Scale bar: 5 μm. (c)-(e) Probability distributions of R_g, R_M/R_m, and μ_M of T4 in three dextran solutions with different molecular weights (M_w = 5k, 80k, and 410k), all $\Phi = 0.6$. All probability distributions, $P(x)$, are constructed to satisfy the normalization criteria: $\int_{-\infty}^{\infty} P(x)dx = 1$. Insets of (c)-(e) are the standard deviations δ of R_g, R_M/R_m, and μ_M as a function of the size of dextran polymers, respectively.*

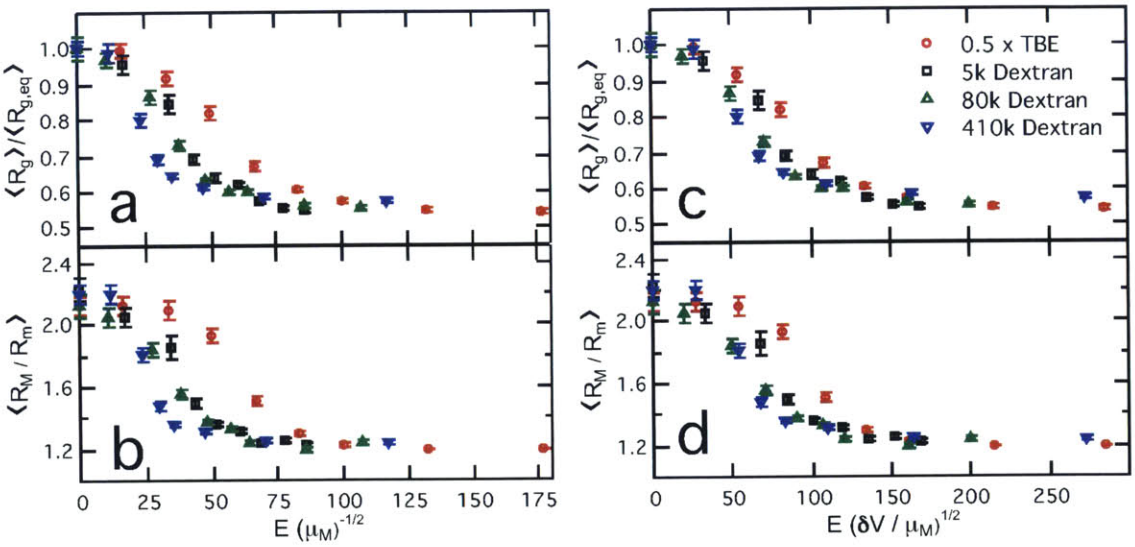

Fig. 3.5: *Replotting the data in Fig. 1c & d versus (a)* $E\mu_M^{-1/2}$*, units:* $(\mathrm{V/cm})\left(\frac{1}{(\mu m/s)/(\mathrm{V/cm})}\right)^{1/2}$*; (b)* $E\left(\frac{\delta \mathrm{V}}{\mu_M}\right)^{1/2}$*, units:* $(\mathrm{V/cm})\left(\frac{\mu m^3}{(\mu m/s)/(\mathrm{V/cm})}\right)^{1/2}$ *results in a data collapse onto a master curve.* μ_M *and* δV *used were measured at at* $E = 15$ V/cm.

to collapse the data. First, we examine the effects of macroion mobility by replotting the data in Fig. 1c & d versus $E/\mu_M^{1/2}$, finding a substantial collapse of data (Fig. 5a). Further collapse of the data is achieved by accounting for differing macroion segment density fluctuations and replotting the data versus $(E(\delta \mathrm{V}/\mu_M)^{1/2})$(Fig. 5b). Here, the fluctuations in macroion segment density δc_M are estimated from the magnitude of fluctuations in macroion volume δV. Assuming the 3D DNA molecule is described by an ellipsoid with R_M and R_m as its major and minor principle axes in the observable 2D-plane, we get $\delta c_M \sim \delta \mathrm{V} \sim \delta R_M R_m^2 + 2\delta R_m R_M R_m$ (see Appendix C). As DNA coils are quickly compressed ($\sim$ 1 s) to globules upon applying a high enough electric field, a change which greatly influences volume fluctuations and mobilities, we use characteristic values of δV and μ_M at a low field strength (15 V/cm) to achieve data collapse in Fig. 3.5. Replotting the data using μ_M as a function of E results in a similar extent of data collapse (Fig. S3 C). The collapse supports our proposed mechanism for polymer-facilitated DNA collapse.

3.7 Conclusions

Our observations demonstrate that the electric-field-induced compression of single DNA molecules is significantly enhanced by the addition of dextran polymers ($M_w \leq$ 410k) to solution. We sharply contrast this behavior to the observed DNA *extension* in HPC solutions, a widely known phenomenon. By varying the size of added polymers, we show that high molecular weight dextran ($M_w \sim$ 2000k) and PVP ($M_w \sim$ 1000k) polymers can also cause similar DNA extension. Through derivation and inspection of the dimensionless groups governing the DNA-polymer collisions, these counterintuitive observations are explained. When the polymers act only to slow the bulk mobility of the DNA, we find that the compression of single isolated DNA molecules is well described by a scaling consistent with the arguments developed by Isambert et al. for intermolecular DNA aggregation that occurred at an overlap concentration.

For DNA separations in dilute polymer solutions, the compact globule conformation of DNA is undesirable, and a lower field strength, higher ionic strength [73], and avoidance of short and flexible polymers as separation medium are expected to prevent the compression of large DNA. Conversely, it is challenging to produce self-compacting DNA vectors for therapeutic gene delivery [110]. To achieve this end, the addition of dextran or similar short flexible polymers to DNA solutions could be used to lower the field threshold for compaction, mitigating potential damage to DNA chains by high electric fields. This enhancement of DNA compression could also be used to systematically induce topological states, such as knots or self-entanglements, in single DNA for fundamental studies in polymer physics [66, 73]. It is reasonable to expect that small polymers can also assist in intermolecular DNA aggregation and hence improve assays that exploit these instabilities for label-free DNA detection in lab-on-chip devices [111]. In future work, it will be interesting to explore how the interplay between nanofluidic-confinement and polymer collisions affects the conformation and dynamics of DNA.

CHAPTER 4

Stretching self-entangled DNA molecules in elongational fields

In this Chapter, we use a mechanism for collapse and self-entanglement of DNA under eletric fields, examined in Chapter 3, to generate self-entangled globules that we then strech. Compared to previous published results, this approach allows us to generate a large statistical set of data on knotted molecules, which we can then examine with external forces. Upon closely inspecting large number of DNA molecules during the process of stretching, we will see these knotted molecules stretch much differently than identical molecules without the presence of knots.

The results in this Chapter have been published in reference [112] and are reproduced with permission from the Royal Society of Chemistry, Copyright 2015.

4.1 Overview

We present experiments of self-entangled DNA molecules stretching under a planar elongational field, and their stretching dynamics are compared to identical molecules without entanglements. Self-entangled molecules stretch in a stage-wise fashion, persisting in an "arrested" state for decades of strain prior to rapidly stretching, slowing down the stretching dynamics by an order of magnitude compared to unentangled molecules. Self-entangled molecules are shown to proceed through a transient state where one or two ends of the molecule are protruding from an entangled, knotted

core. This phenomenon sharply contrasts with the wide array of transient configurations shown here and by others for stretching polymers without entanglements. The rate at which self-entangled molecules stretch through this transient state is demonstrably slower than unentangled molecules, providing the first direct experimental evidence of a topological friction. These experimental observations are shown to be qualitatively and semi-quantitatively reproduced by a dumbbell model with two fitting parameters, the values of which are reasonable in light of previous experiments of knotted DNA.

4.2 Introduction

Advances in nanofabrication, microscopy, and molecular biology have both motivated and enabled the direct observation of the static and dynamic properties of single DNA molecules. These experiments help guide applications such as direct linear analysis [113] or nanopore translocation [77] for sequencing genomes. Experiments on single DNA molecules have a rich history in addressing number of fundamental questions in polymer physics [9, 25, 114]. Optical tweezers have been used to stretch molecules [11], and the data were well described by the theory of Marko and Siggia for semiflexible chains [14]. A wide range of microfluidic devices have been designed to actively manipulate DNA molecules with hydrodynamic flows or electric fields for analysis [106, 115] such as t-junctions [18], cross-slots [19, 116, 117], posts [16, 118, 119], contractions [17, 120, 121], and nano-scale slits [26, 29] and channels [22]. In particular, cross-slot microfluidic devices have been used as a way to stretch molecules without bulky probes for detection of specific DNA sequences [122] or to understand the subsequent relaxation of polymers in slits [123] or collapse of polymers in poor solvents [124]. These devices have also been used to study the transient dynamics of polymer molecules in well-controlled elongational flows/fields, and experiments have revealed surprising configurational diversity [125, 126] and hysteresis in the coil-stretch transition [127].

One of the most dramatic findings of such experiments is that polymer molecules unraveling in elongational flows do so at distinct rates that are largely controlled by a diverse set of transient configurations [125, 126], referred to as *molecular individualism* [128]. These transient configurational classes, dumbbells, half-dumbbells, kinks, hairpins/folds, and coils, were investigated in the simulations of Larson and coworkers [129]. They showed that a bead-spring polymer model with only polymer connectivity, hydrodynamic drag of the solvent, and Brownian fluctuations can recreate the qualitative features of DNA experiments. From an applications perspective, the intrinsic variance in the rate of stretching DNA molecules due to molecular individualism has complicated the design of flow-based stretching devices for DNA analysis. In response, cross-linked gels [130] or post array [131] "preconditioning" devices have been developed to reduce this variability.

More recently, there has been a focus on how the *topology* of a polymer molecule can affect polymer properties [132]. Topological entanglements are found in biological contexts; [40] knots occur in DNA confined to the tight spaces of viral capsids [41, 42] as well as in folded proteins [43]. Simulations have investigated the statistics of knots on polymers in confining geometries at equilibrium [133–135]. In dynamical processes, simulations have indicated knots can significantly slow the ejection of viral DNA [63], slow or jam the sequencing of DNA through nanopores [66], and reduce the rate at which a protein is digested by the proteasome [64]. Theory suggested that topological entanglements can arrest the swelling of polymer globules [54], and simulations supported this idea [55, 57]. More recently, Tang et al. reported an experimental technique for compressing DNA with electric fields and demonstrated an arrested state prior to expanding back

to a swollen coil, which they attributed to self-entanglements [73].

In this work, we present experiments in which DNA molecules are initially preconditioned to a self-entangled state and are then subjected to a planar elongational field. We compare these results with DNA molecules which are stretched in a planar elongational field starting from an equilibrium (unentangled) state. We find substantial differences between these two cases, and we quantify these differences by analyzing their different rates of stretching. Finally, we present a simple model that is shown to semi-quantitatively capture the mean stretching behavior of the self-entangled molecules in our experiments.

4.3 Experimental methods

The experimental buffer consisted of 4% (vol) β-mercaptoethanol (BME, Cabiochem), 0.1% 10kDa polyvinylpyrrolidone (PVP, Polysciences) in 0.5x Tris-Borate-EDTA (TBE, Accugene). T4GT7 DNA (165.6 kbp, Wako) and λ-DNA (48.502 kbp, New England Biolabs) were fluorescently labeled with YOYO-1 intercalating dye (Invitrogen) at a 4:1 base pair to dye ratio in the experimental buffer, leading to a final contour length of 75 μm, 38% larger than bare DNA [136]. This mixture was allowed to stain for 12-48 hours prior to viewing. Cross-slot channels, 1.65 μm in height, were manufactured in PDMS (Sylgard 184, Dow Corning) using soft lithography on a silicone master template (SU8-2 photoresist). Channels were soaked overnight in the experimental buffer at $40°C$ to mitigate permeation-driven flow [130], quickly rinsed with RO water, dried with argon, and sealed to a glass cover slide. Stained DNA solutions were diluted in the experimental buffer 10 to 25-fold for optimal viewing concentrations and loaded in the channel reservoirs. The channel was flushed with buffer for a minimum of 30 minutes prior to collecting data *via* the application of a moderate ($\sim$ 50 V) electric potential at the reservoirs.

A planar elongational field may be used to linearize a charged macromolecule such as DNA [19, 20, 27], and the kinematics of this field are described by the following equation:

$$V_x = \dot{\varepsilon}x;\ V_y = -\dot{\varepsilon}y \tag{4.1}$$

where V_x and V_y are the x and y components of velocity and $\dot{\epsilon}$ is the strain rate of the field. We used a cross-slot device to generate a homogeneous elongational field within a $\sim$ 100 x 100 μm field of view. The strain rate was controlled by varying the voltages applied at the reservoirs. Molecules were trapped at the metastable stagnation point at the center of the field by manually perturbing the potential ($\sim$ 2 V) of the right reservoir. See Appendix D for device layout and strain rate calibration curve.

For a molecule in an elongational flow or field, the relevant dimensionless group is the Deborah number, $De \equiv \dot{\varepsilon}\lambda$, where λ is the longest relaxation time of the polymer molecule. In such fields, a polymer will undergo the coil-stretch transition at $De_c \approx 0.5$, and the critical strain rate for the onset of this transition is $\dot{\varepsilon}_c = \frac{1}{2\lambda}$. The longest relaxation time of DNA was measured as $\lambda = 2.6$ s by fitting the long time decay of the autocorrelation function of orientation angles of DNA molecules at equilibrium [137] (see Appendix D for details).

For molecules with unentangled initial conditions, the molecule was brought to the stagnation point and there allowed to relax for $\sim$ 30 s $> 10\lambda$, allowing the molecule to thoroughly sample its equilibrium configurations. The elongational field was then turned on, and the molecule was stretched. The procedures for generating the initial self-entangled molecular states are more com-

plex and are shown in Fig. 4.1. For a self-entangled initial condition, the molecule was brought to a channel arm and allowed to relax for $\sim$ 30 s. An AC square-wave electric field of strength $E_{\text{rms}} = 200$ V/cm and frequency $f = 10$ Hz was applied for 30 s to compress and self entangle a molecule in a fashion demonstrated by Tang et al. [73] After entanglement, the reservoir potentials were switched to generate an elongational field, and the molecule was quickly (< 5 s) swept into the straining region of the device. The molecule was held in this region until it stretched.

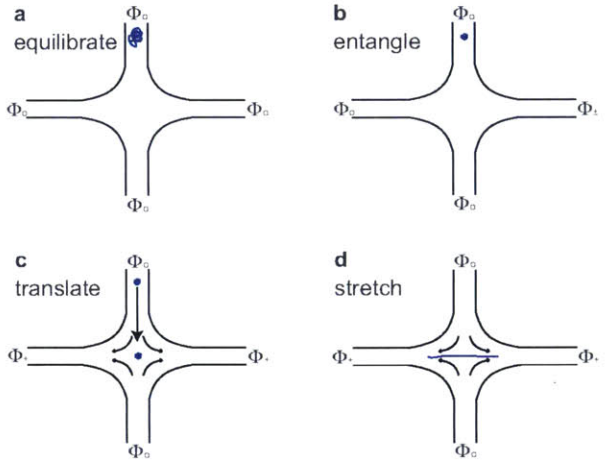

Fig. 4.1: *Schematic for stretching self-entangled DNA. (a) A molecule is brought to an inlet arm and allowed to equilibrate for* $\sim$ *30 s with no applied field. (b) A square-wave AC electric field* ($\Phi_{\pm}$) *of strength* $E_{\text{rms}} = 200$ V/cm *and frequency* $f = 10$ Hz *is turned on for 30 s to compress and self-entangle a molecule in the channel arm. (c) The elongational field is switched on* ($\Phi_{+} > \Phi_{\circ}$)*, and the self-entangled molecule rapidly translates to the stagnation point and is trapped there. (d) The molecule stretches some time after the translation step shown in (c).*

4.4 Results and Discussion

4.4.1 Differences Due to Entanglements

Representative snapshots of the initial, transient, and fully stretched configurations of both initially unentangled and self-entangled molecules are shown in Figure 4.2. Several stark differences between the stretching processes for each type of molecule are immediately apparent. The initial configurations for unentangled molecules are visually diverse and somewhat anisotropic since the molecules are exploring the full configurational space of a polymer in a good solvent prior to the onset of the field. The initial conditions of the self-entangled molecules are highly isotropic "globules," and the important differences between their individual configurations exist at a length scale smaller than can be resolved by fluorescence microscopy. It is important to emphasize that the subsequent differences in stretching behavior of unentangled and self-entangled molecules arise *solely* due to

these differences in the initial conditions. The so-called "half-dumbbell," "kink," "hairpin/fold," and "dumbbell" configurations emerge as transient configurations of unentangled molecules. This phenomenon as well as the corresponding differences in the rates of stretching grouped by transient configurational class is the key concept behind *molecular individualism*. The transient configurations of self-entangled molecules are less diverse, all consisting of a knotted core from which one or two ends of the chain unravel. The differences between initially unentangled and self-entangled molecules even persist in their fully stretched conformations. While absent in the vast majority of unentangled molecules, a region of increased fluorescence (indicated with arrows) can be seen in the stretched conformation of an initially self-entangled molecule. These regions are persistent and suggest that a topological knot has been pulled tight in the chain.

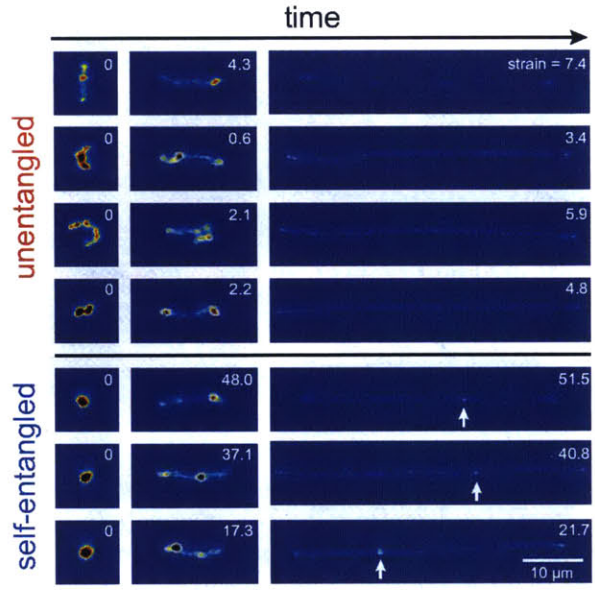

Fig. 4.2: *Snapshots of initially unentangled and self-entangled molecules stretching under an electric field of* $De = 2$. *The white arrows indicate the presence of a persistent, localized knot along the fully stretched contour of the DNA molecule. The white numbers are the accumulated strain experienced in each snapshot.*

Selected trajectories of extension versus strain are plotted for both initially unentangled and self-entangled DNA molecules in Figure 4.3. The trajectories of initially unentangled molecules demonstrate that their stretching dynamics are *progressive* - the molecule orients, aligns, and immediately begins stretching until fully extended. Initially self-entangled molecules exhibit *stage-wise* stretching dynamics - the molecule persists in a low extension arrested state before rapidly stretching. The snapshots corresponding to the bolded trajectory in each plot are shown on the right further illustrate this point. The unentangled molecule immediately begins stretching, and within 2 units of strain, the transient "dumbbell" configuration can be easily seen. The self-entangled molecule persists in its arrested state until a strain of 15. At a strain of 15, the nascent

ends of the molecule can be seen to first clearly protrude from the highly knotted core. This event immediately precedes the onset of a comparably rapid stretching phase of this molecule, indicative of a nucleation phenomenon. Nearly all molecules in experiments follow this general nucleation then stretch behavior, a phenomenon with qualitative similarities to the onset of stretching in collapsed polymers in elongational flows [138, 139].

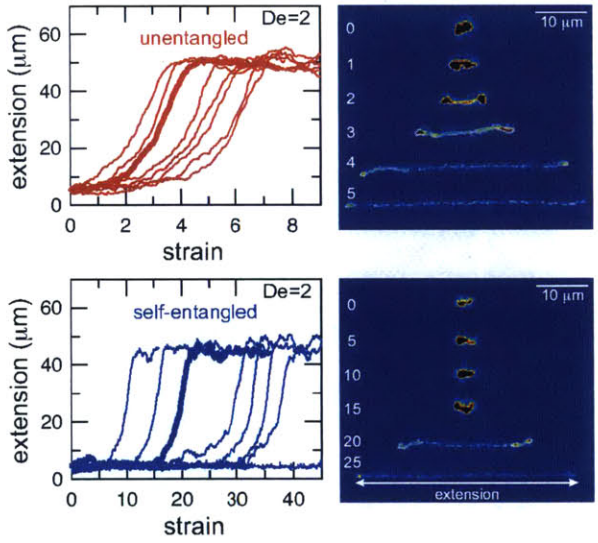

Fig. 4.3: *Extension vs. strain trajectories for initially unentangled (top) and self-entangled (bottom) DNA at* $De = 2$. *The snapshots to the right correspond to the bolded trace in each graph. The white numbers are the accumulated strain experienced in each snapshot. The reported extension is the maximum distance along the extensional axis of two points on the molecule, indicated in the snapshots. Note the different scales of the x-axes.*

In Figure 4.4, stretching trajectories are shown for both unentangled (red) and self-entangled (blue) molecules unraveling in elongational fields of $De = 1$, 2, 2.9, and 5. The immediate difference in time scales associated with stretching the two populations is evident at all field strengths, indicating that even at relatively strong fields of $De = 5$, the presence of topological restrictions can dramatically slow the stretching process. At all field strengths, the trajectories clearly indicate significant dispersity in the times to reach a fully stretched conformation for both initially unentangled and self-entangled molecules. For unentangled molecules, this dispersity is referred to as *molecular individualism* and can be largely associated with the differences in transient configurations. For initially self-entangled molecules, the variation in the time to fully stretch a molecule arises almost entirely from the different lengths of time spent in the arrested state.

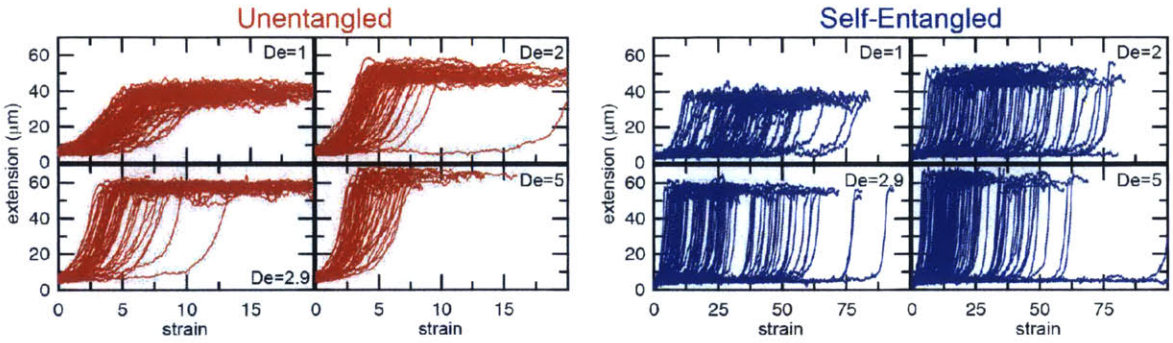

Fig. 4.4: *Extension vs. strain trajectories trajectories for initially unentangled (left) or self-entangled (right) DNA for all Deborah numbers (De = 1, 2, 2.9, and 5) in this study.*

4.4.2 Stage-wise Decomposition of Trajectories

The images and trajectories in Figs. 4.2, 4.3, and 4.4 all indicate a stage-wise stretching of a self-entangled molecule. In general, each trajectory can be decomposed into three distinct stages: a metastable arrested state, a transient stretching phase, and an extended steady state, shown in Fig. 4.5. We now present the method by which we algorithmically defined these stages for all experimental trajectories.

The arrested state is the initial phase of a molecule, and the molecule is considered to remain in this state until its extension permanently passes a lower extension threshold. The duration of the arrested state is deemed the nucleation time, denoted t_{nuc}, since the end of the arrested state corresponds to the nucleation events described in the previous section. To best delineate the arrested state from the stretching phase, the value for the lower extension threshold should be chosen to be the smallest extension where molecules all immediately stretch upon passing the threshold for the last time. A value of 10 μm was found to be suitable for our ensemble of experimental data.

The stretching phase of the molecule begins at the end of the arrested state and continues until the extension of the molecule passes an upper extension threshold. The value of the upper extension threshold is again chosen so that all molecules in the ensemble continue to stretch rapidly. As the steady state extension of DNA is a function of De, this parameter, too, is a function of De. Values for the upper extension threshold of 30, 42, 46, and 50 μm for $De = 1$, 2, 2.9, and 5 were used in the following analysis. Both the upper and lower extension thresholds were chosen empirically to best segregate the phases.

In Figure 4.6, the distributions of nucleation times are shown for both unentangled and self-entangled molecules. The x-axes of these distributions are scaled by the here-called "excess strain" rate. The strain, $\dot{\varepsilon}t$, time scaled by the strain rate, relates the exponential increase in separation between tracers along streamlines in an elongational field, $x(t + \delta t) = x(t)\exp(\dot{\varepsilon}\delta t)$. The excess strain, $(\dot{\varepsilon} - \dot{\varepsilon}_c)t$, time scaled by the strain rate exceeding the onset of the coil-stretch transition, relates the exponential increase in separation between two points connected by a Hookean spring with the same entropic elasticity as the DNA molecule, $x(t + \delta t) = x(t)\exp((\dot{\varepsilon} - \dot{\varepsilon}_c)\delta t)$. This

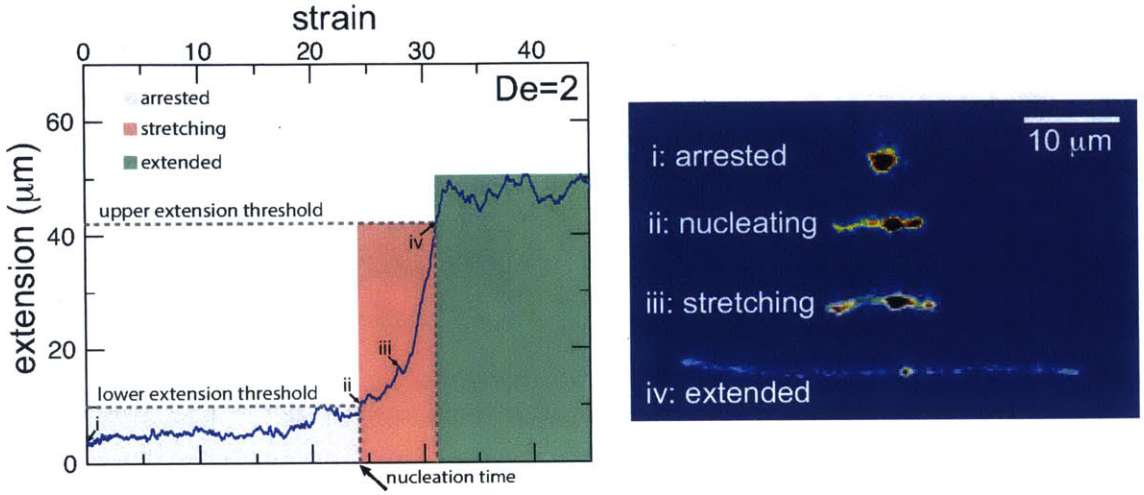

Fig. 4.5: *Experimental trajectories are decomposed into three stages: arrested, stretching, and extended. (top) The molecule is "arrested" until its extension passes and remains above a lower extension threshold. Afterwards, the molecule is "stretching" until its extension passes an upper extension threshold for the first time. A molecule is considered stretched thereafter. The extension thresholds were chosen empirically to best segregate the phases. The lower extension threshold used was* 10 μm. *The upper extension threshold was chosen as* 30, 42, 46, *or* 50 μm *for* $De = 1$, *2, 2.9, and 5, respectively.*

quantity has been shown to reasonably scale the stretching kinetics in experiments [125, 126] and nucleation-type behavior of the coil stretch transition in simulations [140]. In Figure 4.6, the x-axes thus correspond to the amount of excess strain accumulated until the nucleation event occurs at t_{nuc}, *i.e.* $(\dot{\varepsilon} - \dot{\varepsilon}_c)t_{\text{nuc}}$.

For initially unentangled molecules, the "nucleation time," also called the transition time by others [140], represents the time required for a molecule to begin to align, orient, and immediately stretch to the lower extension threshold. Although this process is fundamentally different than the nucleation-type events seen in initially self-entangled molecules, we will use the term "nucleation time," still the time required to pass the 10 μm lower extension threshold, for the sake of consistency. The distributions of nucleation times for unentangled molecules appear to be properly scaled by the excess strain, in agreement with the simulations results of Cifre and de la Torre [140]. For all field strengths, ~1.5 "excess strain" is the characteristic quantity required to nucleate the unentangled molecules, and ~3 'excess strain" is approximately the characteristic width of the distributions. The simulations of Cifre and de la Torre [140] found $(\dot{\varepsilon} - \dot{\varepsilon}_c)\langle t_{\text{trans}}\rangle = 3.23 \pm 0.04$, where $(\dot{\varepsilon} - \dot{\varepsilon}_c)\langle t_{\text{trans}}\rangle$ is analogous to the nucleation time of an unentangled molecule. This result appeared universal for ideal chains or chains with excluded volume as well as chains with and without hydrodynamic interactions. However, the results of these simulations show some disagreement with our result of 1.5, the reason for which is not immediately clear.

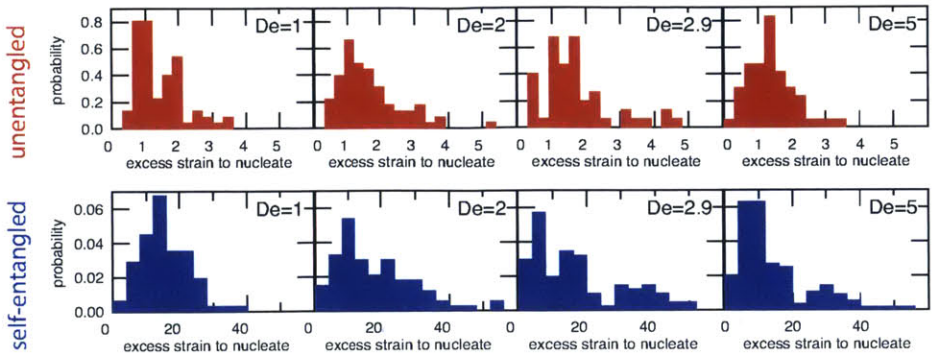

Fig. 4.6: *Probability distributions of the excess strain required to nucleate (begin stretching) a molecule for initially unentangled (left) and self-entangled (right) DNA at all Deborah numbers (De = 1, 2, 2.9, and 5) in this study. The excess strain to nucleate is defined as* $(\dot{\varepsilon} - \dot{\varepsilon}_c)t_{nuc}$. *Note the different scales for the x-axes.*

For initially self-entangled molecules, the nucleation time represents the time for the ends of the molecule to free themselves from an entangled core a sufficient distance to initiate stretching. The distributions of nucleation times for initially self-entangled molecules are similarly scaled by the excess strain rate, and interestingly, this quantity appears to be able to reasonably scale these distributions as well, although the physical reason is not as clear. For all field strengths, ~15 excess strain is required to stretch these tangled globules and the width of the distributions is ~30 excess strain. These quantities are an order of magnitude larger than their counterparts in initially unentangled molecules, underscoring the fundamentally different physics in the arrested state that is present in the self-entangled molecules.

We now turn our attention to the stretching phase. The distributions of nucleation times in Fig. 4.6 are useful in quantifying the differences between the initially unentangled and self-entangled populations, but the main result, the existence of a the arrested state in self-entangled molecules, is already apparent when viewing the images in Fig. 4.2 or the trajectories in Figs. 4.3 & 4.4. Less obvious, however, are the differences in the rapid stretching phase between the two populations. Using the decomposition criteria described in Fig. 4.5, the stretching phase of each molecule can be separated from the prior nucleation process. This decomposition affords the ability to independently analyze this phase.

In Figure 4.7a, trajectories of the stretching phase for initially self-entangled molecules are shown at a field strength of $De = 2$. From the ensemble of trajectories, a master stretching curve was created by sorting the extensions of the molecules into 2 μm bins and calculating the average strain of all trajectories as they pass through each bin. Such a master curve is shown superimposed over the underlying trajectories in Fig. 4.7a.

The process of generating master stretching curves was repeated for all initial conditions and field strengths. These curves are plotted versus excess strain in Fig. 4.7b. The affect of varying the lower and upper extension thresholds is shown in Appendix D. For both initially unentangled and

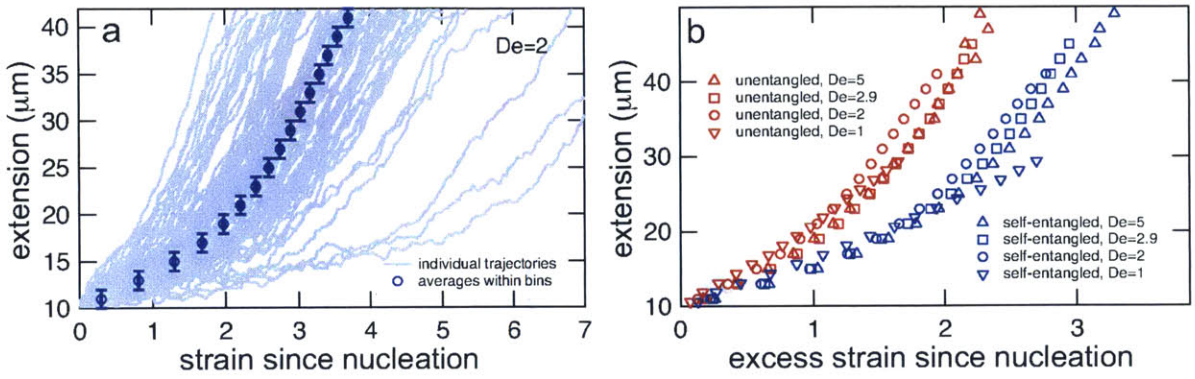

Fig. 4.7: *(a) Generating mean stretching curves from stretching trajectories. All trajectories were binned into 2 μm bins, and the mean strain within the extension bin was computed by equally weighting the mean strain within each bin for each individual trajectory. Error bars represent the bin widths (y-axis) and 95% confidence intervals around the mean (x-axis)(b) Mean stretching curves for initially unentangled and self-entangled molecules at all Deborah numbers (De = 1, 2, 2.9, and 5) in this study. These mean extension curves collapse into two polulations when plotted vs. excess strain, defined as* $(\dot{\varepsilon} - \dot{\varepsilon}_c)t$.

self-entangled molecules, the master curves collapse when scaled by the excess strain, consistent with previous experiments of unentangled molecules in extensional flows [126]. Interestingly, initially self-entangled molecules stretch $\sim 50\%$ more slowly than initially unentangled molecules. Initially self-entangled molecules appear to have a localized, knotted core that shrinks and is tightened during stretching, as can be seen in Figs. 4.2, 4.3b, and 4.5. However, such a structure is very rare in stretching unentangled molecules. It follows that the difference in the rate of stretching between initially unentangled and self-entangled molecules corresponds to this topological difference. In other words, pulling the strands through a tangled core introduces and additional "topological friction" into the stretching process. The concept that knots can introduce an effective topological friction has been seen in computer simulations of knots slowing (or even jamming) the translocation of DNA through nanopores [66], the ejection of viral capsids [63], and the pulling of knotted proteins through the proteasome [64]. *To our knowledge, the results in Fig. 4.7b are the first direct experimental observation that is suggestive of the phenomena of topological friction due to a knotted domain at the molecular scale.*

After exiting the stretching window, a molecule will remain in an extended state as long as the elongational field is applied. Once extended, initially self-entangled molecules frequently have a localized area of increased fluorescence along the contour, like those shown in Fig. 4.2. These regions are likely localized knots. We estimated the amount of contour stored in the knots by comparing the ensemble average steady state extensions, $\langle X_{ss} \rangle$ of initially unentangled and self-entangled molecules, shown in Figure 4.8(a). The trend of increasing molecular extension with increasing field strength is expected and readily seen. For all field strengths, the initially self-entangled molecules exhibit lower $\langle X_{ss} \rangle$ due to the finite amount of contour length stored within localized knots on the

stretched chains. The difference in the steady state extension between initially unentangled and self-entangled ensembles is the excess knot length, $\langle L_{\text{knot,excess}} \rangle = \langle X_{ss,\text{unentangled}} \rangle - \langle X_{ss,\text{entangled}} \rangle$, shown in Figure 4.8(b). This quantity was first introduced as a length reduction due to a knot in a perfectly tightened string [141] and has been used to characterize the size of knots in the experiments of Bao et al. [70] In Figure 4.8 (b), $\langle L_{\text{knot,excess}} \rangle$ is plotted versus De. A gradual decrease of $\langle L_{\text{knot,excess}} \rangle$ is seen as De increases, consistent with the notion of tightening a knot. Over the range of field strengths in this study, $\langle L_{\text{knot,excess}} \rangle$ varies approximately from 1 to 3 μm. These data can be compared to the excess knot lengths found by Bao et al. [70] of ~0.25 - 0.55 μm found for the 3_1, 4_1, 5_1, 5_2, and 7_1 knots held at tension of 0.1 - 2 pN, suggesting that the interior knots may be quite complex in our experiments. The large amount of extra knot length we observe in our experiments may also be an important factor in our ability to measure a strong signature of topological friction in Figure 4.7b. Overall, the data in Figure 4.8 verify the notion of a topological knot containing some finite amount of chain contour.

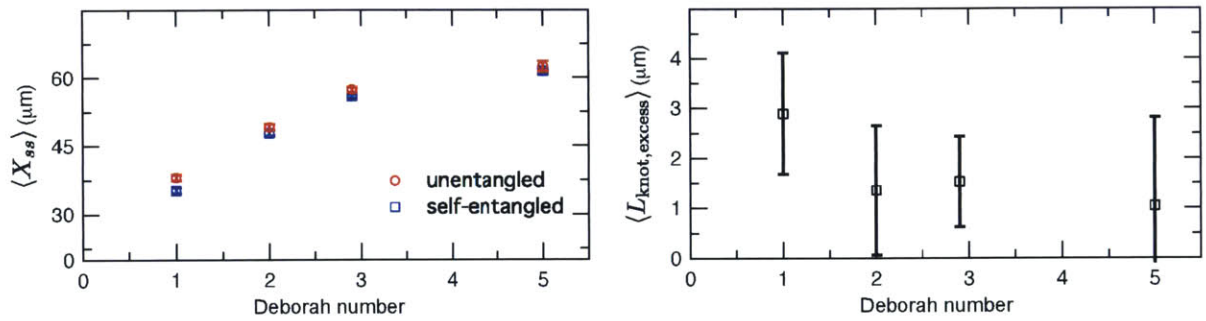

Fig. 4.8: *(a) Ensemble average steady state extension of initially unentangled and initially self-entangled DNA molecules for all Deborah numbers (De = 1, 2, 2.9, and 5) in this study. Error bars represent 95% confidence intervals about the mean. (b) Ensemble average excess knot length, $\langle L_{\text{knot,excess}} \rangle$, plotted for all Deborah numbers (De = 1, 2, 2.9, and 5) in this study. Error bars represent 95% confidence intervals about the mean.*

If the applied field is switched off, the molecule will relax back to the equilibrium coil. The trajectories of relaxing molecules are shown in Fig. 4.9. On average, the molecules that begin with interior bright spots relax more quickly than molecules without. This increase in the rate of relaxation contrasts with the slowing of stretching in initially self-entangled globules, shown in Fig. 4.7. Both observations, however, can be qualitatively explained by the electrostatic interactions within the vicinity of the knot. Stretching a knotted chain requires the knot to be pulled "tight," and the electrostatic repulsions in the vicinity of the knot will act to slow the stretching process. Similarly, when tension is released from a tightened knot, the electrostatic interactions seek to swell the knotted core. Swelling can only be achieved by retraction of the ends of the molecule, and this will increase the observed rate at which the extension of the molecule relaxes. These observations further reinforce the notion that localized knots are frequently present upon stretching self-entangled molecules.

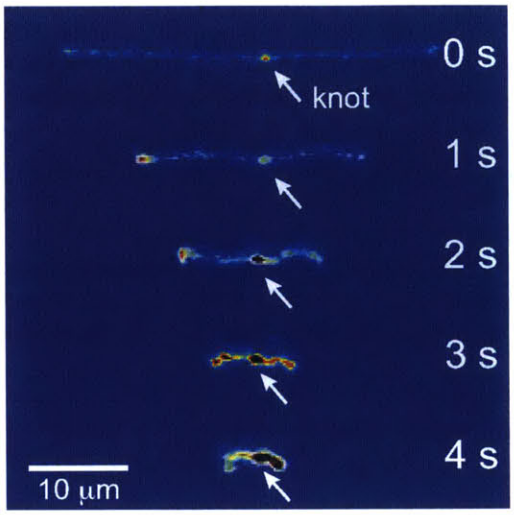

Fig. 4.9: *Relaxation of a stretched DNA molecule. Selected snapshots for an initially unentangled and self-entangled molecule relaxing after the shutoff of the field.*

4.4.3 Modeling Stretching Dynamics of Entangled DNA

The preceding discussion has emphasized key differences in stretching of initially self-entangled and unentangled DNA, both qualitative (initial, transient, and stretched configurations) and quantitative (nucleation times and stretching rates) in nature. Here, we lay out a physical framework that is consistent with these observations, and show that a simple dumbbell model motivated by this framework can semi-quantitatively capture the mean nucleation times and stretching curves for the experimental data.

Consider a highly self-entangled blob. The topological entanglements can be thought to form a transient network structure. In order to stretch such a blob, this network must be resolved by transporting contour from a centrally entangled region to an unentangled region near the ends of the molecule. We suggest that this transport can occur *via* two mechanisms: (i) diffusive release of entanglements *via* cooperative motions of the DNA ends and entangled loops and (ii) convective transport of contour to the ends *via* pulling contour out of the central entangled region. This process will result in the gradual "softening" of an initially inextensible blob until soft enough to be extended by the flow, which could reproduce the initial arrested state.

We will seek to model this process by the stretching of a single non-Brownian dumbbell. The contour of the molecule is partitioned into both "free" and "entangled" regions such that $L_f + L_e = L_c$, where L_f, L_e, and L_c are the contour lengths of the free portion, entangled portion, and entire molecule, respectively. We will consider only the free portion of the molecule as being able to directly interact with the external field. The beads of the dumbbell, therefore, have drag coefficients that vary as $\zeta_f = \zeta\left(\frac{L_f}{L_c}\right)$. Here, $\zeta = 4H\lambda$ is the total drag on the DNA molecule at equilibrium and $H = \frac{3k_BT}{2l_pL_c}$ is spring constant of the total molecule [27].

Given the concept of a varying length of extensible contour, the Deborah number, $De = \dot{\varepsilon}\lambda$, cannot describe the transient dynamics of an entangled strand; the relaxation time, λ is that of unentangled state. We therefore define a "free" Deborah number as

$$De_f = De\left(\frac{L_f}{L_c}\right)^2 \tag{4.2}$$

as an estimate for the effective dumbbell of containing L_f length of contour, assuming a freely-draining dumbbell. For a fixed $\dot{\varepsilon}$, De is constant while De_f changes in time as L_f changes from a small fraction of L_c to nearly the entire contour length of the molecule as contour is progressively freed. Given an initially highly entangled state, $L_f \approx 0$, De_f will progress from $De_f \approx 0$ at time $t = 0$ to $De_f = De$ as time $t \to \infty$. The dumbbell will persist in a low extension (coiled) state until $De_f \approx 0.5$, giving rise to an apparent arrested state.

We now turn to the specifics of the model that give rise to the evolution equation of the total extension of the molecule. The initially entangled state is considered to have an extension between the ends of $X_{\text{knot}} = 1.6$ μm (experimental diameter of gyration), and we assume that the extension in this region does not change. We consider the free contour on each end to be tethered to this inextensible core. The cumulative extension of both ends from their tethers is R, and the total extension of the molecule is $R + X_{\text{knot}}$. By neglecting inertia, the evolution equation for the dumbbell becomes

$$\frac{d(R + X_{\text{knot}})}{dt} = \left[(R + X_{\text{knot}}) - \frac{0.5}{De_f} f\left(\frac{R}{L_f}\right) R\right]\dot{\varepsilon} \tag{4.3}$$

where

$$f\left(\frac{R}{L_f}\right) = \left(\frac{L_f}{6R}\right)\left[\left(1 - \frac{R}{L_f}\right)^{-2} + \frac{4R}{L_f} - 1\right] \tag{4.4}$$

is the dimensionless force law for a wormlike chain of fractional extension $\frac{R}{L_f}$. The first term in the right hand side of Equation 4.3 corresponds to the affine deformation of two points separated by $R + X_{\text{knot}}$, and the second term slows this deformation rate due to a nonlinear spring of fractional extension $\frac{R}{L_f}$.

In order to solve Eq. 4.3, an expression for $L_f(t)$ is needed. We consider two factors can free entangled contour from the knotted core: (i) cooperative diffusion of the ends and knot and (ii) "pulling" of the ends out of the globule by the external flow. While both factors (particularly (i)) are quite complex in reality, we considered the simplest estimates of each for the purpose of this model, described as follows.

We estimate the rate of diffusive release of entangled contour to the free ends as

$$L_{f,\text{diff}}(t) = \sqrt{D_{\text{knot}} t} \quad \text{while} \quad De_f < 0.5, \tag{4.5}$$

where D_{knot} is diffusion coefficient for release of contour, t is the residency time in the elongational field. The constraint on De_f prevents the diffusive release of contour once the dumbbell stretches and the entangled globule is localized. In this scenario ($De_f < 0.5$), the ends have been extended and cannot easily further relax knots *via* diffusion.

We estimate the rate at which contour is pulled out of the entangled center due to the field is

$$L_{f,\text{pull}}(t) = \frac{1}{\delta}\int_0^t F_{\text{spring}}dt', \tag{4.6}$$

where δ is a topological friction coefficient and F_{spring} is the time-dependent stretching force experienced by the dumbbell.

The total contour that has been freed from the initially entangled globule is then written as

$$L_f(t) = \min\left(L_{f,\text{diff}}(t) + L_{f,\text{pull}}(t),\ L_c - X_{\text{knot}}\right), \tag{4.7}$$

where X_{knot} now represents the contour remaining in the knot on a fully stretched strand, which is taken to be approximately the initial diameter of gyration of the entangled globule $X_{\text{knot}} = 1.6\ \mu\text{m}$.

Eq. 4.7 was solved simultaneously with Eqs. 4.6 and 4.5 *via* a midstepping Euler integration scheme. The resulting stretching curves are shown in Figure 4.10a. These curves qualitatively reproduce three key experimental findings. 1) Each curve persists in an arrested state for some time. 2) The duration of the arrested state decreases as De increases. 3) The arrested state is followed by a rapid stretching phase.

The values of $\delta = 0.5\zeta$ and $D_{\text{knot}} = 22\ \mu\text{m}^2/s$ were chosen to best match the ensemble average nucleation times and the master stretching curves from experiments. The results from the model are compared to the experimental results for average nucleation times and master stretching curves in Figs. 4.10b and c. The model does predict a slowed rate of stretching, seen in Fig. 4.10b. It captures the slowed rate of stretching well at high De, but fails to capture the experimental curve at $De = 1$. This discrepancy is likely due to the documented difficulty in a simple dumbbell model in capturing polymer dynamics over ranges of De (regardless of entanglements); additional complexity such as configurationally dependent drag is often needed to capture this behavior [142]. The model and experiment find good agreement for the average nucleation time over the entire range of field strengths, $De = 1$, 2, 2.9, and 5, seen in Fig. 4.10c. This quantity depends only on the rate of diffusive release of the contour coupled with convective release due to flow on a globule of low and nearly constant extension, further supporting the notion that configurationally dependent drag causes the discrepancy in Fig. 4.10b.

Given the simplicity of the dumbbell model, the results in Figure 4.10 are encouraging. We now turn to discuss the limitations of this modeling approach. Foremost, by seeking to capture average properties, this model ignores the wealth of information encoded in the *distributions* of nucleation times or stretching rates. This issue seems likely intractable by any model of such simplicity. Fine-scaled simulations by Larson and others [129] were the first successful attempt to recreate the qualitative configurational classes seen in the molecular individualism experiments [125, 126]. We expect that even finer-grained simulations that preserve the topology of individual entanglements will be necessary to make further progress into our experimental work. Also, the assumptions of constant diffusion coefficients and friction coefficients for the diffusive and convective transport of contour out of the entangled structure are oversimplifications. Both the rate of disentangling and the friction of pulling an end through an entangled core should be functions of the entangled state, but the functional dependence of each is likely complex and not presently clear.

We now seek to put the values of $\delta = 0.5\zeta$ and $D_{\text{knot}} = 22\ \mu\text{m}^2/s$ in proper context. With similar experimental conditions, Tang et al. [73] found the relaxation of an initially entangled molecule to the equilibrium state to proceed *via* a two stage process with average durations of 19.7 s and

20s, giving a total disentanglement time of $\sim 40s$. From our model, the value of $D_{\text{knot}} = 22$ μm/s would give a characteristic time of $\tau_{\text{unknot}} \approx \frac{L_c^2}{D_{\text{knot}}} = 256$ s. While this discrepancy appears large at first, it can be understood in context of experimental results. The 40 s disentanglement time of Tang et al. [73] corresponds to the average time when the fluctuations and size of the molecule are indistinguishable from the equilibrium state under no external forces. In Fig. 4.10, the experimental nucleation time for $De = 1$ is ~ 80 s. These results suggest that in the process of disentangling a molecule under no field, there exists a period of time where measurements of the molecule approach their equilibrium values, yet entanglements significant enough to affect an out of equilibrium process remain. This observation underscores the well-established fact that dynamical measurements, such as stretching a molecule in an elongational field, can provide a wealth of information beyond that provided by experiments at equilibrium.

The value of $\delta = 0.5\zeta$ means that pulling contour out of the entangled glob experiences an additional friction of approximately half the total friction coefficient of the molecule. This value appears eminently reasonable given the cooperative nature of pulling segments through a topological knot. Our value of δ can be estimated with units as $\delta = 0.5\zeta = 2\lambda\frac{3k_BT}{2l_pL_c} = 8.6$ pN/(mm/s). In experiments where knots of specific topologies were tied with optical tweezers, Bao et al. [70] were able to directly measure the friction coefficients of knots *via* the diffusivities of the knot along an extended contour. They found frictions of $\sim 3 - 8$ pN/(mm/s) for the 3_1, 4_1, 5_1, 5_2, and 7_1 knots, and it is interesting to note the similarity between our results and their measured friction coefficients.

4.5 Conclusions

We have presented a systematic study comparing the rates at which initially unentangled and self-entangled DNA molecules stretch while in elongational fields. These experiments revealed two striking qualitative differences between their stretching pathways. For initially self-entangled molecules, there exists and initial topologically stabilized arrested state leading to nucleation-type behavior after which molecules stretch at a slowed rate due to interior knots. The herein observed slowed rate of stretching represents the first direct experimental evidence of a topological friction in polymer molecules, seen recently in computer simulations of driven knots [143]. We systematically characterized the nucleation times and stretching rates over a variety of field strengths. Guided by this experimental data, we developed a simple dumbbell model which consists of a gradually softening spring due to diffusive and convective release of contour from the knot. This model is able to semi-quantitatively capture both the average nucleation time and stretching dynamics of an initially self-entangled molecule, reinforcing the key physics in play. Looking forward, we anticipate this approach will stimulate future work on expanding the model in more detail, possibly through the use of fine-grained dynamical computer simulations. From an experimental point of view, we hope our work motivates further experimental studies of knotted polymers, such as knotted DNA in nano-scale confining geometries or knotted DNA driven through nanopores.

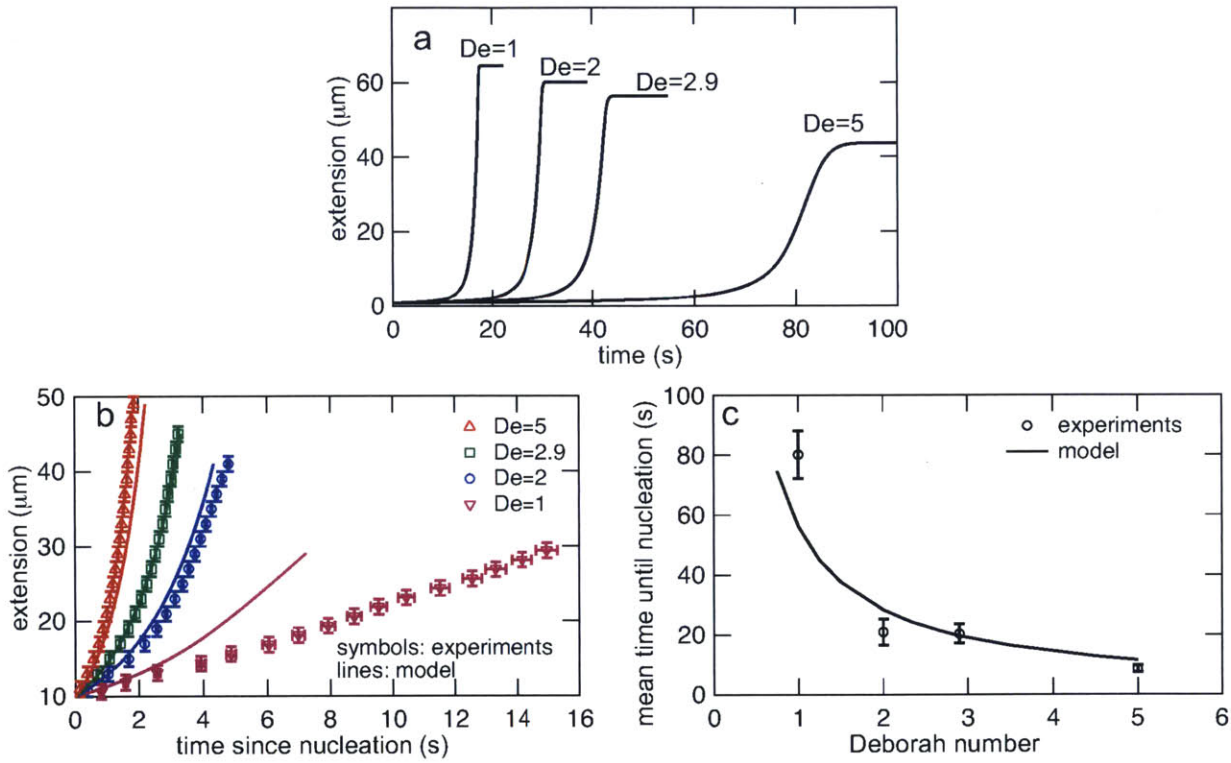

Fig. 4.10: *(a) Stretching curves for generated by the model all Deborah numbers ($De = 1$, 2, 2.9, and 5) in this study with $\delta = 0.5\zeta$ and $D_{\text{knot}} = 22\ \mu\text{m}^2/s$. Note the presence of an arrested state followed by a rapid stretching phase and finally a fully extended state. (b) Comparison of the rates of stretching for curves in the model vs. the experimental data from Fig. 4.7b. Error bars represent the bin widths (y-axis) and 95% confidence intervals around the mean (x-axis). (c) Comparison of the nucleation times from the model vs. the mean nucleation time from the data in Fig. 4.6. Error bars represent 95% confidence intervals around the mean (y-axis).*

CHAPTER 5

Metastable tight knots in semiflexible chains

In this Chapter, we turn our attention to a fundamental question: what controls the size of a knot that randomly forms on DNA? This question was first addressed by Grosberg and Rabin [61]. They presented an elegant idea that two competing effects, the desire of a DNA molecule to resist bending of its contour and the need to maximize the entropy of its configuration, results in a *metastable* size of a knot on a long DNA molecule. Inspired by this work, we here extend their theory to account for a finite length of excluded volume interactions that exists in real molecules, and we validate their original theory and our modified theory with the results of extensive computer simulations. These results represent the first supported theory for the physics of randomly occuring knots in DNA molecules, and as we will see in the next Chapter, this platform can be extended to include other external effects.

The results in this Chapter have been published in reference [144] and are reproduced with permission from the ACS Publications, Copyright 2014. This work would not have been possible without close collaboration with Liang Dai, whose expertise in simulation and theory greatly expedited and enhanced the final work.

5.1 Overview

Knotted structures can spontaneously occur in polymers, such as DNA and proteins, and the formation of knots affects biological functions, mechanical strength and rheological properties. In this work, we calculate the equilibrium size distribution of trefoil knots in linear DNA using off-lattice simulations. We observe metastable knots on DNA, as predicted by Grosberg and Rabin. Furthermore, we extend their theory to incorporate the finite width of chains and show an agreement between our simulations and the modified theory for real chains. Our results suggest localized knots spontaneously occur in long DNA and the contour length in the knot ranges from 600 to 1800 nm.

5.2 Introduction

Knotted structures can occur in polymers, such as DNA [73, 145] and proteins [146]. The investigation of knots in mathematics has a rich history, and recently research into these structures is increasingly motivated by biological relevance and practical concerns such as DNA sequencing [66]. Much like the well known tangling of jumbled strings [35], linear polymers can spontaneously form knotted structures[147], and their topolgies can be determined via procedural closing schemes[82]. Knots (most frequently simple ones) are present in the DNA contained in viral capsids and influence the rate of ejection of the viral genome [41, 148], and a number of knotted protein structures have been identified[45, 146]. Beyond their biological importance, knots in polymers can affect their mechanical strength [149], rheological properties [150], and may play a role in crystallization [151], cause jamming during nanopore DNA sequencing [66], or be used for controlled drug delivery [152].

For these reasons, there is a desire to elucidate the size and probabilities of knots in polymer molecules. Accordingly, simulations addressing this issue have been performed for numerous cases: linear [53, 147] and circular [153] chains, ideal [153] and self-avoiding [147] chains, flexible [53, 147, 153] and semiflexible [154–157] chains, lattice [154, 158], and off-lattice [154, 156, 157] models, good [147] and bad solvents [53, 159], as well as in free space [154, 156, 157], in confinement [134, 135, 160] and under tension [156, 157]. An intriguing finding from simulations is that the cores of knots very often localize at small portion of chain [53, 153, 158, 160]. For example, Katritch et al. found the most probable size of trefoil knot on an unconfined circular ideal chain is only seven segments [153]. The localization of polymer knots has also observed in experiments [161], and several theories have been developed to explain this behavior [61, 162]. The theory of Grosberg and Rabin [61] employed the idea, similar to the tube theory for polymer melts, that the chain in the area of the knot forms a self-consistent confining tube. The bending energy and confinement free energy within a knot tend to swell and shrink the knot respectively, leading to a localized, metastable knot that is likely to untie by diffusing to the end of the molecule. Beyond qualitatively predicting knot localization, this theory is able to predict the size of such a metastable knot in a polymer, but the theory has yet to be rigorously tested against simulation or experiment.

In the current study, we validate the Grosberg-Rabin theory [61] for metastable knots with computer simulations of long, linear wormlike polymers. We extend the theory to incorporate the finite thickness of polymer chains and similarly validate it with computer simulations. Our results demonstrate a quantitative agreement between simulation and theory and provide testable predictions for sizes of knots in dsDNA, which has been used in previous experimental work on knotting [41, 70, 148, 163].

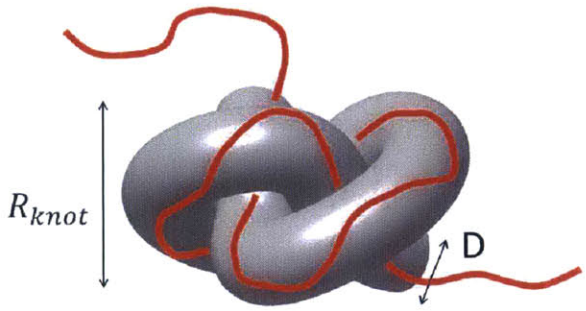

Fig. 5.1: *Illustration of a trefoil knot on an open chain (red). The sub-chain with a contour length of L_{knot} in the knot core is confined in a virtual tube (grey) with a diameter of D.*

5.3 Theory and Simulation

5.3.1 Theory of Knots in Semiflexible Chains

We first recall the Grosberg-Rabin theory for a knot wormlike polymer of zero thickness [61]. The central tenet of this theory is that the polymer contour in the knotted region is self-confined by a virtual tube (Figure 5.1). The tube can be imagined as a tight knot formed by pulling both ends of a rope with diameter D. The free energy cost to form a knot contains contributions due to both bending and confining the polymer contour in the knotted region. First, the bending energy F_{bend} in the knot core scales as $L_{knot}R_{knot}^{-2}$, where and L_{knot} and R_{knot} are the contour length and the size of the knot core, respectively. Second, the free energy of confining the polymer within the knot core is given as $F_{conf} \sim L_{knot}D^{-2/3}L_p^{-1/3}$ after applying the Odijk scaling [164], where L_p is the persistence length. For a tight knot, the quantities L_{knot}, R_{knot} and D are proportional to each other. Accordingly, we replace R_{knot}, and D by L_{knot} after ignoring numerical coefficients. Note that L_{knot} can be measured in simulations, while D is difficult to determine directly. Accordingly, Grosberg and Rabin obtain the total free energy $F_{knot} = F_{bend} + F_{conf}$

$$\frac{F_{knot}}{k_BT} = k_1\left(\frac{L_{knot}}{L_p}\right)^{-1} + k_2\left(\frac{L_{knot}}{L_p}\right)^{1/3}, \tag{5.1}$$

where k_BT is the thermal energy, and k_1 and k_2 are prefactors that take account of all numerical coefficients ignored in the derivation. The first and second terms in the above equation tend to swell and shrink the knot size, respectively. The competition of these terms leads to a local minimum of free energy, which corresponds to a metastable knot.

Next, we adapt the Grosberg-Rabin theory to real chains with finite thicknesses. For a chain with an effective width w confined in a tube with a diameter D, the effective diameter of the confining tube becomes $D_{eff} = D - w$ due to the repulsion between the chain and tube walls. As a result, the confinement free energy becomes $F_{conf} \sim L_{knot}(D-w)^{-2/3}L_p^{-1/3}$. Similar to the

previous study [61], we define p as the ratio of L_{ideal} to D

$$p \equiv L_{ideal}/D \approx L_{knot}/D, \tag{5.2}$$

where L_{ideal} is a topological property defined for a maximally inflated knot[141].

We then obtain $F_{conf} \sim L_{knot}(pD - pw)^{-2/3}L_p^{-1/3} = L_{knot}(L_{knot} - pw)^{-2/3}L_p^{-1/3}$. The free energy cost of forming a knot on a real chain becomes

$$\frac{F_{knot}}{k_BT} = k_1\left(\frac{L_{knot}}{L_p}\right)^{-1} + k_2L_{knot}(L_{knot} - pw)^{-2/3}{L_p}^{-1/3}. \tag{5.3}$$

It is easy to see that when $w = 0$, Eq. 5.3 returns to Eq. 5.1. The size of metastable knots for real chains can be calculated by minimizing F with respect to L_{knot}. This calculation requires the values of p, k_1, and k_2. We estimate these values for trefoil knots as follows. The value of p is estimated to be 12.4, which is calculated for the tightest trefoil knot with both ends in a line [141]. To estimate k_1, we assume the bending in the core of trefoil knots is uniform and the total bending angle is θ_{total}. Then, the bending energy follows $(1/2)\theta_{total}^2L_p/L_{knot}$ and $k_1 = (1/2)\theta_{total}^2$. To form any knot, the total bending angle should be larger than 2π, so k_1 should be larger than 19.7. To estimate k_2, we apply the scaling of confinement free energy with a prefactor [165] $F_{conf} \approx 2.4L_{knot}(D - w)^{-2/3}L_p^{-1/3}$, leading to an estimate of $k_2 \approx 2.4p^{-2/3} \approx 12.9$. We will show the estimations of p and k_1 are close to the fitted values from simulation results while the estimation of k_2 is much larger than the fitted value.

5.3.2 Simulations of Knots in Semiflexible Chains

To examine the theory, we performed simulations of long open semiflexible chains and analyzed the knotted conformations. The polymer chain is modeled as a string of touching beads.[166] The diameter of each bead equals the effective chain width w, and the contour length L is thus $L = (N - 1)w$, where N is the number of beads. There are only two interactions between beads: the pairwise hardcore repulsion between beads and the bending energy $E_{bend}(\theta)/k_BT = (1/2)(L_p/w)\theta^2$, with bending angles θ, to reproduce the persistence length L_p. We use the Pruned-enriched Rosenbluth method (PERM) algorithm [167] to generate polymer configurations and then analyze the topologies of these configurations after closing the both ends. In the following paragraphs, we will describe how we implement PERM algorithm, how we close the both ends of a linear chain, and how we determine the size of knot core if the chain is knotted.

The PERM algorithm generates chain configurations by a growth process.[167, 168] The growth starts from a bead at the origin $(0, 0, 0)$. In each step of growth, a new bead (the i-th bead) is placed at the end of the chain. The orientation of the newly added bead follows a Boltzmann distribution of bending energies $\exp[-E_{bend}(\theta_i)/k_BT]$, where θ_i is the bending angle formed by (i-2)-th, (i-1)-th and i-th beads. If the new bead overlaps with any other bead, then this chain dies. In the simulation, each chain is grown in a batch of N_c chains. After each step of adding a bead, a few chains may die. The surviving chains are duplicated, which is the so-called enrichment. Without enrichment, more and more chains die as the length of chain grows, and we cannot obtain enough sampling of long chains. The idea of the PERM algorithm is to duplicate the surviving chains and reduce the statistical weight of these surviving chains such that these chains are not overrepresented in the final analysis of conformational statistics. We implement the enrichment

scheme as follows. In each simulation, we grow a batch of N_c chains. The initial weights of these chains are $W_{j,i} = 1/N_c$ for $i = 1$, where i denotes the chain length and j denotes the index of chain. Suppose N_d chains die after adding the i-th bead. Then, we randomly pick N_d chains from $(N_c - N_d)$ surviving chains and duplicate these N_d chains so that the number of chains in our ensemble is restored to N_c for the next chain growth step. Considering that all surviving chains have the same probability $q = N_d/(N_c - N_d)$ to be duplicated, the weights of all chains are reduced by a factor $1/(1+q) = (N_c - N_d)/N_c$. That is, $W_{j,i+1} = W_{j,i}(N_c - N_d)/N_c$ after adding (i+1)-th bead to the j-th chain. When the chain length (number of beads) reaches the desired chain length, the growth stops, and the configurations of $(N_c - N_d)$ surviving chains and the associated weights $W_{j,L}$ are used to calculate the probability of the trefoil knot $f = (\Sigma_{j=1}^{j=N_c-N_d}[A(j)W_{j,L}])/(\Sigma_{j=1}^{j=N_c-N_d}W_{j,L})$, where $A(j)$ equals one when the j-th chain is a trefoil knot, and equals zero when the j-th chain is not a trefoil knot.

To determine the topology of an open chain, the chain must first be closed. In the current study, we employ the minimally interfering closure scheme[82]. After closing an open chain, we calculate the Alexander polynomial to identify the topology.

To identify the core of a trefoil knot, we cut the maximum number of beads from each end of the chain while the trefoil knot remains. The two boundaries of knot core are determined sequentially. First, we keep one end unchanged and cut beads from the other end to determine the boundary. After obtaining this boundary, we cut beads to determine the other boundary. We find that the size of knot core may depend on which boundary is determined first. For consistency, we calculated the knot size both ways and used the smaller value as the final size of knot core because the core of knot corresponds to the smallest portion of the chain that forms the knot.

5.4 Results and Discussion

Figure 5.2a shows the probability of forming a trefoil knot as a function of the rescaled knot size for a wormlike chain of zero thickness. The contour length is $L = 400L_p$. The probability $f(x)$ is normalized such that $\int_0^{L/L_p} f(x)dx = f_{total}$, where $x = L_{knot}/L_p$ and $f_{total} \approx 0.23$ is the total probability of trefoil knot. The probability exhibits a peak at a certain knot size $L^*_{knot} \approx 12L_p$. The value L^*_{knot} is insensitive to increasing the contour length (see Appendix E) since the knots are localized. The most probable size of a trefoil knot in a previous simulation of a circular freely jointed ideal chain by Katritch et al. [153] is seven segments, each corresponding to $2L_p$. Thus, their finding of a knot size of $14L_p$ is close to the $12L_p$ reported here.

In Figure 5.2b, the probability is converted to the potential of mean force (PMF) using $F = \ln(f) - F_0$, where $F_0 = \min(-\ln(f))$ to offset the minimum. The PMF curve exhibits a local minimum in a potential well, corresponding to a metastable knot. This minimum is a local minimum because the globule minimum corresponds to the unknotted state. As can be seen from our simulation results, the potential well is relatively broad, and the depth is $\sim 1\ k_BT$ up to $L_{knot}/L_p \approx 50$, which is over 3 times the size of the knot with the minimum free energy. This observation indicates the metastable minimum is relatively broad, and long chains (several times pL_p — a characteristic size of a tight knot) are required to observe it. In the simulations of knots by Zheng and Vologodskii [169], shallow or even flat potential wells are observed for the 7_1 and 10_{151} knots (both much larger than the trefoil knot) under a slight stretching force. The authors concluded no metastable knot size exists for these topologies, attributing the slight minimum they observe to the applied stretching force. The apparent difference between their results and the present work may arise

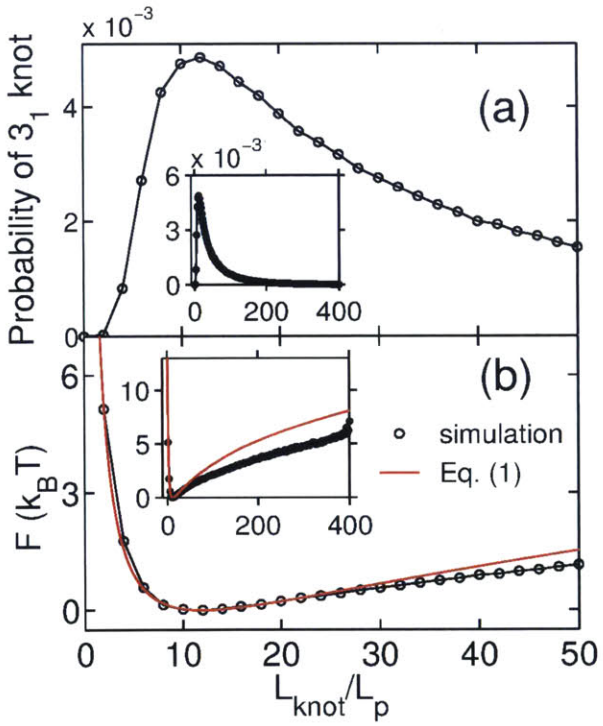

Fig. 5.2: *(a) The probability of a wormlike chain containing a trefoil knot with* $L = 400L_p$*. (b) The potential of mean force as a function of knot size. The line of best fit is shown in red:* $y = 17.06x^{-1} + 1.86x^{1/3} - 5.69$*. Both insets show curves over wider ranges.*

due to the very shallow nature of the potential well — if a metastable knot size exists for these large knots, extremely long chains (several multiples of pL_p) would be needed to visualize the well. For the long chains in our simulations, however, the depth of the potential well we observe is ~ 6 k_BT for $L_{knot}/L_p = 400$, which suggests that the spontaneous unknotting process of a knot on a long and thin chain is more likely through the diffusion of knot along the chain than through the swelling of knot towards the full contour length [71, 170].

We use Eq. 5.1 to fit the PMF curve in the range $L_{knot}/L_p \leq 28$ and determine the prefactors $k_1 = 17.06$ and $k_2 = 1.86$. The fitted line (red) deviates from simulation results for $L_{knot}/L_p > 30$. This is expected from the theory because the theory assumes the chain is strongly confined in a virtual tube, and the confinement free energy follows the Odijk scaling $F_{conf} \sim D^{-2/3}$, where D is the diameter of the virtual tube. This assumption becomes invalid when $D \equiv L_{knot}/p > 2L_p$. If we use the estimation $p = 12.4$ as discussed above, then the Grosberg-Rabin theory is supposed to be inapplicable for $L_{knot} > 2pL_p = 24.8L_p$, which is consistent with the results in Figure 5.2b. The assumption of Odijk-like confinement, however, is always valid for small knots. As a result, the fit of the theory matches the simulation data well through the smallest observed knots, $L_{knot} = 2L_p$.

The fitted value of $k_1 = 17.06$ is slightly less than the estimate for the lower bound of 19.7, as described above. This discrepancy likely arises from the fact that the bending energy in the unknotted subchains is ignored when we estimate the total bending energy cost of forming a knot. To re-examine the bending energy term, $17.06L_p/L_{knot}$, we directly analyzed the increase in bending energy due to forming a knot from the chain configurations in simulations. This analysis confirms that the term $17.06L_p/L_{knot}$ indeed roughly captures the bending energy increase in knots (see Appendix E). On the other hand, for the confinement free energy, the fited value of $k_2 = 1.86$ is much less than the estimate of $k_2 = 12.9$. This deviation may result from the assumption that the virtual tube of the knot (Figure 5.1) has hard, solid walls. In reality, the walls of the virtual tube are made of fluctuating sections of the same chain, and this "softer" confining tube results in the smaller prefactor from the simulation data.

Figure 5.3a shows the probability of forming a trefoil knot as a function of the rescaled knot size for real chains with different chain widths. As the chain width increases, the total probability of forming a trefoil knot monotonically decreases. The most probable knot size, L^*_{knot} also increases with increasing chain width, as expected. We analyze the most probable knot size L^*_{knot} rather than the mean knot size $\langle L_{knot} \rangle$ because L^*_{knot} is insensitive to the contour length and is a local quantity (see Appendix E). On the other hand, the mean size $\langle L_{knot} \rangle$ strongly depends on the contour length [53, 147, 171] due to a long tail in the distribution of knot sizes. The inset of Figure 5.3a shows the total probability of forming a trefoil knot as a function of the rescaled chain width (circle symbols). The probability decreases from 22.8% to 0.57% as w/L_p increases from 0 to 0.5. We also plot the total probability (square symbols) for $L_{knot}/L_p \leq 100$ because some natural ambiguity arises in classifying knots on open chains when the size of the knot approaches the total contour length. Using this criterion, the probability decreases from 18.8% to 0.33%. Furthermore, this only modest reduction in probability when the size of the knot is restricted to a small fraction of the total contour underscores the notion that these knots are localized.

Figure 5.3b shows the potential of mean force as a function of the rescaled knot size for $L_{knot}/L_p \leq 100$. When $w/L_p \leq 0.4$, the depth of potential well is larger than 1 k_BT. Extending the curves to $L_{knot}/L_p = 400$, the depths further increase (see Appendix E). We note that the potential well becomes shallower as the chain width increases, indicating that knots on thick chains, while less probable, will exhibit much more variation in size.

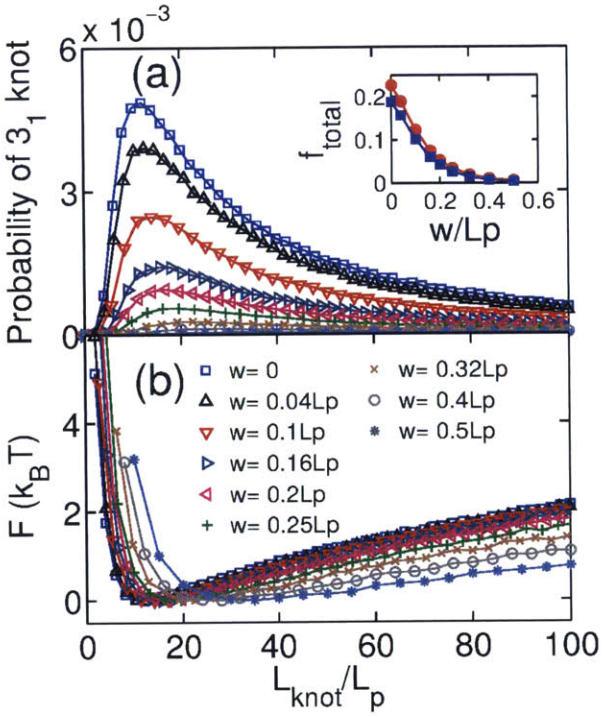

Fig. 5.3: *(a) The probability of forming a trefoil knot as a function of the rescaled knot size in real chains with different chain widths. The contour lengths are fixed as* $L = 400L_p$. *The circle and square symbols in the inset show the total probability for* $L_{knot} \leq 400L_p$ *and* $L_{knot} \leq 100L_p$, *respectively. (b) The potential of mean force (PMF) as a function of the rescaled knot size for different chain widths. The curves are shifted such that the* F *minimum is zero.*

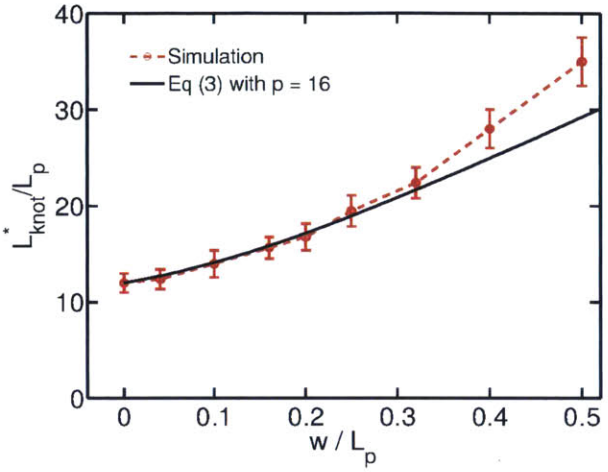

Fig. 5.4: *The most probable size of a trefoil knot as a function of the rescaled chain width. The solid line is calculated from Eq. 5.3 with* $k_1 = 17.06$, $k_2 = 1.86$, *and* $p = 16$.

The dependence of L^*_{knot} on w is plotted in Figure 5.4 and compared with theoretical predictions from Eq. 5.3. The values of k_1 and k_2 are already determined from the fit in Figure 5.2 for the wormlike chain of zero thickness, so only p is used as a fitting parameter. Recall that Eq. 5.3 is supposed to be applicable only when the knot is small. So we manually tune p to match simulation results and theoretical predictions when L_{knot} is small, and we find $p = 16$ is suitable for $w/L_p \leq 0.25$. Furthermore, $p = 16$ is a reasonable value for the ratio of L_{knot} to the tube diameter D, considering that p is estimated as 12.4 for the tightest trefoil knot [141], which is the lower bound of p value. For large knots, the Odijk scaling $F \sim D^{-2/3}$ overestimates the confinement free energy. Recall that the overestimation in confinement free energy tends to shrink the knot, and leads to the observed discrepancy between the simulation results and the theoretical prediction for large knots on thick chains. Similar analysis was performed for the figure eight (4_1) knot (see Appendix E), yielding a fitted value of $p = 38$ for 4_1 knot.

After obtaining the fitted value of $p = 16$ for the trefoil knot, we calculate the potentials of mean force using Eq. 5.3 for $w = 0.1L_p$ and $w = 0.2L_p$ and compare with the ones obtained from simulations, as shown in Figure 5.5. Near the minima, the theoretical predictions agree with the simulation results.

From the metastable knot sizes in Figure 5.4 and the fitted value of $p = 16$, we can infer that the diameter of the virtual tube $D = L_{knot}/p$ varies from $0.75L_p$ to $2.2L_p$ as the chain width increases from 0 to $0.5L_p$. A similar virtual tube also exists in entangled polymers, and the effective potential in the virtual tube has been calculated by Zhou and Larson. [172] If we set 1 k_BT as a cutoff in the tube potential, the tube diameter is roughly 8 L_p, while the effective chain width w roughly equals L_p in their calculation [172]. An important difference in these scenarios is that in the case of a knot, one chain is confined by itself, while in the case of entanglement, one chain is usually confined by other chains.

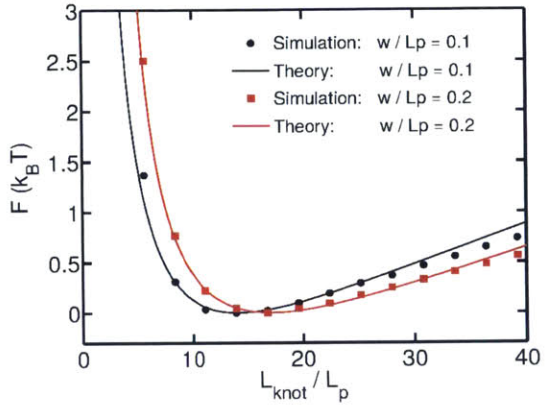

Fig. 5.5: *Comparison of potential of mean force calculated from simulations and the theory for two chain widths $w=0.1L_p$ and $w=0.2L_p$.*

We now turn our attention to "real" polymers of non-zero thickness. The total probability of forming a trefoil knot as a function of the rescaled chain width in the inset of Figure 5.3a. The simulation parameters we employed in this study are directly comparable to λ-DNA, which is widely used in experiments [22, 173]. The contour length in simulations is 400 L_p, which is close to the case of λ-DNA, considering that the YOYO-1 intercalated λ-DNA has a contour length about 22 μm and a persistence length about 50 nm [174]. The range of chain width in simulation is from 0 to $0.5L_p$, which is within the range of the effective width of DNA at experimental ionic strengths [175]. These simulations predict the knotting probability of λ-DNA can be around 10% for high salt conditions. The contour length of the metastable trefoil knot typically range from 600 to 1800 nm, while the knot radius of gyration ranges from 65 to 138 nm (see Appendix E). Note that the knotting probability of a circular 10-kb P4-DNA, much shorter than 48.5-kb λ-DNA, has been shown experimentally to be 4% in high salt solution [176].

As discussed above, the modified Grosberg-Rabin theory is not applicable for larger knots because the Odijk scaling becomes invalid when $D \gg L_p$. For these more swollen knots, the virtual tube is wide, and the confinement free energy is proportional to the number of blobs $N_{blob} = L_{knot}/D = p$ as discussed by Grosberg and Rabin. In this case, the confinement free energy is independent of L_{knot} and thus no longer tends to shrink the knot core. Accordingly, Grosberg and Rabin suggest that the force driving the chain towards tight knots disappears when the tube is wider than L_p. However, our simulation results in Figure 5.3a show that the PMF curve keeps increasing with the increasing size of knot core. Other factors may need to be considered for the spontaneous tightening of large knots.

5.5 Conclusions

We have extended the theory of Grosberg and Rabin to calculate the size and distribution of knots on a real, semiflexible polymer molecule and validated the predictions of this theory with computer simulations. These tight metastable knots exist on semiflexible chains due to the competition of

bending energy and confinement free energy, which tend to swell and shrink knot size, respectively. The metastable knots are trapped by potential wells with depth of several k_BT. Our simulation results should be universal because the model contains only two parameters: L_p and w, or a single dimensionless parameter w/L_p. Looking forward, we expect that the Grosberg-Rabin theory can be further modified for chains confined in channels. Finally, our simulations are directly comparable to λ-DNA at experimentally relevant ionic strengths, and we hope that future experiments will allow for further testing of these ideas.

CHAPTER 6

Metastable knots in confined semiflexible chains

In the previous Chapter, we established a validated theory for the sizes of knots in DNA molecules. Here, we we ask and answer the technologically relevant question: what happens to the sizes of those knots when the DNA molecule is squeezed into a channel? We show this effect can be easily incorporated as a modification of the previously validated theory, yet the addition of the external channel dramatically changes the sizes of knots that may be found. As in the previous Chapter, we show this theory agrees with the results of extensive computer simulations.

The results in this Chapter have been published in reference [177], and are reproduced with permission from the ACS Publications, Copyright 2015. This work would not have been possible without close collaboration with Liang Dai, whose expertise in simulation and theory greatly expedited and enhanced the final work.

6.1 Overview

We study the size distribution of spontaneous knots on semiflexible chains confined in square channels using Monte Carlo simulations. The most probable knot size L^*_{knot} is found to vary non-monotonically with the channel size. For knotted polymers confined in channels larger than the size of a knot in bulk, our analysis reveals that the metastable knot size in weak confinement is

larger than the knot size in absence of confinement because the confinement free energy gained by shrinking the knot is lessened when the chain experiences the confinement of a channel. In the case of strong confinement, the metastable knot size is smaller than the knot size in the absence of confinement, because the segments in the knot core experiences more confinement free energy, and the channels pushes the segments out of the knot core. We demonstrate that a simple theory can capture this non-monotonic behavior and quantitatively explain the metastable knot size as a function of the channel size. These results may have implications for tuning the channel size to either generate or screen knots.

6.2 Introduction

Long linear polymers both tie and untie knots through the random motions of the chain [147]. These types of knots have been shown to affect the mechanical [149, 178], rheological [150], and structural properties of polymer molecules. In biological contexts, knots have been found in proteins [146] and viral capsid DNA [42]. Experiments and simulations have begun to address the structure of a knot along a polymer in a good solvent— some experiments [161] and simulations [153] have suggested the localization of a knot whereas other simulations have suggested the opposite [155, 169]. For semiflexible polymers like DNA, Grosberg and Rabin posited that a competition between bending and confinement energies within a knotted region of the chain results in a localized, metastable knot for long, thin polymers [61]. In a Chapter 5, we extended the Grosberg-Rabin (GR) theory to incorporate the effect of a finite chain width and validated the theory with computer simulations.

Advances in microfabrication have spurred fundamental experimental research into the static and dynamic properties of double-stranded DNA (dsDNA) in microfluidic confinement [23, 115]. Computer simulations have systematically investigated the scaling regimes of a confined polymer [166, 168, 179], and experiments have added support to these arguments for both static and dynamic properties [22, 24, 26, 27, 180]. More recently, the topological properties of polymers in confinement have been investigated. Simulation results have indicated that the probability of forming a knot can depend non-monotonically on the degree of confinement in slits [133, 135] or channels [134, 160]. The most direct experimental measurements of knot sizes have been on actin [69] and DNA [70] held under tension by optical tweezers. Knotted DNA molecules have been observed in nanochannels [71], and confinement could provide a passive means of visualizing the size of knots in polymers. Beyond these fundamental studies, nanochannels have been used to extend dsDNA for the optical mapping of genomes [181–183], and knots may have been observed to interfere with this process by reducing apparent separation between sites [184].

In the current study, we analyze how confining a knotted polymer in a channel may affect the size of the knot. We have extended the modified GR theory to incorporate the effect of confinement in a channel. We find that the modified theory predicts a non-monotonic variation of the knot size with the confining dimension. We tested this theory with detailed Monte Carlo simulations to sample trefoil knot configurations in the channel geometry. Our results demonstrate the modified GR theory quantitatively agrees with the metastable knot size from simulations, and we provide predictions for the sizes of channels where the theory can be tested with dsDNA experiments.

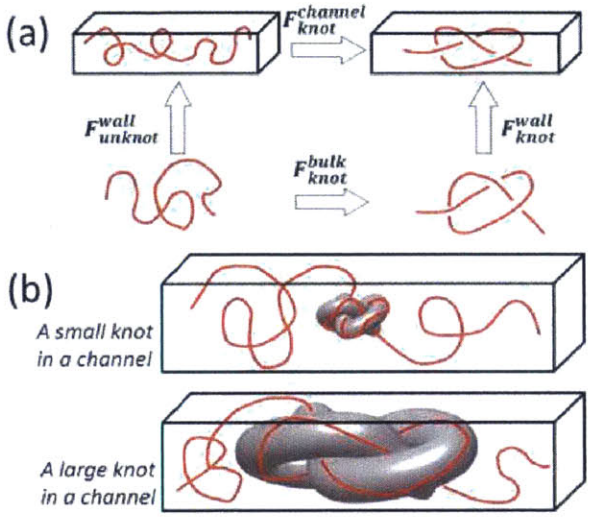

Fig. 6.1: *(a) A diagram to show the free energy difference between four states: an unknotted subchain in bulk, a knotted subchain in bulk, an unknotted subchain in a channel, and a knotted subchain in a channel. (b) Illustrations of small (top) and large (bottom) knots confined in a channel.*

6.3 Theory and Simulation

6.3.1 Theory for Metastable Knots

Let us first recall the modified GR theory for semiflexible polymers with finite thickness in bulk (no confinement). The free energy cost of knot formation in bulk follows

$$F^{bulk}_{knot}(L_{knot}) = k_1(L_{knot}/L_p)^{-1} + k_2 L_{knot}(L_{knot} - pw)^{-2/3} L_p^{-1/3} \tag{6.1}$$

where L_{knot} is the contour length of knot, L_p is the persistence length, and w is the chain width. The numerical coefficients $k_1 = 17.06$, $k_2 = 1.86$, and $p = 16$ were determined for trefoil knots in the previous study. The first term represents the bending energy due to the increased curvature of the contour within the knot, and the second term corresponds to the confinement free energy within the knot resulting from topological constraints. Note that Eq. 6.1 only contains the terms involving L_{knot} because our primary purpose is to calculate the size distribution of knot from the free energy.

For semiflexible chains in square channels, we calculate the free energy cost of knot formation using the diagram of states, illustrated in Fig. 6.1a, is:

$$F^{channel}_{knot}(L_{knot}) = F^{bulk}_{knot}(L_{knot}) + F^{wall}_{knot}(L_{knot}) - F^{wall}_{unknot}(L_{knot}), \tag{6.2}$$

where $F^{wall}_{knot}(L_{knot})$ corresponds to the free energy cost of confining a knot while the contour length of the knot is maintained as L_{knot}, and $F^{wall}_{unknot}(L_{knot})$ corresponds to the confinement free energy

of an unknotted chain with contour length L_{knot}. Note that in the definition of $F_{knot}^{wall}(L_{knot})$, we ignore the confinement free energy experienced by the unknotted portion $(L - L_{knot})$ in a knotted chain because this contribution would cancel with an identical term in F_{unknot}^{wall}. For convenience, we define an excess free energy

$$F_{excess}(L_{knot}) \equiv F_{knot}^{wall}(L_{knot}) - F_{unknot}^{wall}(L_{knot}). \tag{6.3}$$

Eq. 6.2 then becomes

$$F_{knot}^{channel} = F_{knot}^{bulk} + F_{excess}. \tag{6.4}$$

Rewriting Eq. 6.2 in this way enables a simple physical interpretation of how the confining geometry interacts with the knotted subchain. When the knotted region of the chain experiences more confinement free energy due to the channel walls than the unknotted region of the chain chain, *i.e.* $F_{excess} > 0$, forming a knot is more difficult inside the channels than in bulk, and the formation of the knot becomes less likely. Conversely, in situations were the unknotted region of the chain experiences more confinement free energy due to the walls than the knotted region of the chain, *i.e.* $F_{excess} > 0$, forming a knot becomes more likely within the confining geometry.

Now, we turn to calculate F_{excess}. The term $F_{unknot}^{wall}(L_{knot})$ can be approximated as the confinement free energy of equilibrium chains in channels $F_{chain}(L_{knot})$, which is ensemble-average of knotted and unknotted chains. This approximation is based on the fact that knots usually occur in small portions of a long chain. The term $F_{chain}(L_{knot})$ depends on the strength of confinement, such as Odijk regime or de Gennes regime. In the current study, we mainly deal with the classic de Gennes regime, leading to

$$F_{unknot}^{wall}(L_{knot}) = 5.0D^{-5/3}w^{1/3}L_p^{1/3}L_{knot}, \tag{6.5}$$

where the prefactor 5.0 was determined in a previous study [168]. The term $F_{knot}^{wall}(L_{knot})$ depends on the size of the knot relative to that of the channel. When the knot size is much smaller than the channel size, the knot is weakly deformed and can be considered as a ball with an effective diameter b_{knot}. The term b_{knot} is approximately proportional to L_{knot}, so we write $b_{knot} = \alpha L_{knot}$, where α is a numerical coefficient less than 1. The confinement free energy of this ball is approximated as the one for a bead on a flexible chain of identical balls [185]. Then, we have $F_{knot}^{wall}(L_{knot}) \approx b_{knot}^{5/3}(D - b_{knot})^{-5/3}$. In the expression, $(D - b_{knot})^{-5/3}$ is used in favor of instead of $D^{-5/3}$ because the center of knot-ball is restricted in a channel of size $(D - b_{knot})$ due to the finite thickness of the ball. After introducing a numerical coefficient β, we obtain

$$F_{knot}^{wall}(L_{knot}) = \beta L_{knot}^{5/3}(D - \alpha L_{knot})^{-5/3}. \tag{6.6}$$

When the knot size is much larger than the channel size, we can consider the different portions of knot are confined in many sub-channels with channel size less than D, as discussed by Nakajima and Sakaue [160] and shown in Figure 6.1b. In such a situation, the contour within the knot experience a larger confinement free energy penalty that outside the knot. In the current study, we examine channels of size $\geq 2L_p$, and we have shown in a previous study that the metastable radius of gyration of a trefoil knot is approximately $2L_p$ in bulk. Accordingly, we will use Eq. 6.5 to calculate the metastable knot size in channels since our simulations occur in the regime small knot regime, $R_{g,knot} < D$. Later in this manuscript, we will verify that the studied knots are indeed in

this small knot regime.

6.3.2 Polymer Model and Simulation Methods

We performed Monte Carlo simulations of polymers confined in square channels to test the preceding theoretical predictions. Polymer chains were modeled as a string of touching beads. The diameter of each bead corresponds to the effective chain width w, and the contour length L follows as $L = (N-1)w$, where N is the number of beads. The present model contains only three interaction energies: pairwise hardcore repulsion between beads, hardcore repulsions between the beads and the channel walls, and a harmonic bending energy $E_{bend}(\theta)/k_BT = 1/2(L_p/w)\theta^2$ used to reproduce the persistence length L_p.

Configurations of chains were sampled with a standard Monte Carlo procedure. Each Monte Carlo cycle consisted of alternating crankshaft or reptation moves. For a given set of parameters, we typically performed 10^11 cycles and sampled the chain configuration every 1000 cycles. For each parameter set, the number of knotted chains was approximately 10^6 or 10^7, resulting in knotting probabilities between 0.01 and 0.1 for our simulation conditions.

Since the topology of an open chain is not well defined, the chains must be systematically closed in order to determine its topology. In the current study, the closing loop is generated by minimally interfering closure scheme [82]. After closing the chain, the topology of the chain is determined by calculating the Alexander polynomial [186]. Further details about the determination of knot size are presented in Chapter 5 as well as Ref. [144]. Note that the knots of open chains might be ambiguous when the knot size is close to the entire chain. However, the knots can be easily distinguished from unknots, when the knot size is much less than the entire chain. In the current study, the contour lengths of metastable knots are usually one order of magnitude less than the entire chain, and hence these knots are well defined.

6.4 Results and Discussion

The probability distributions of trefoil knot sizes for semiflexible chains confined in channels of varying sizes as well as in bulk are shown in Figure 6.2. The contour length is fixed as $L = 400L_p$ and the chain width is fixed as $w = 0.4L_p$. For each curve, the probability $f(x)$ is normalized such that $\int_0^{L/L_p} f(x)dx = f_{total}$, where $x = L_{knot}/L_p$ and f_{total} is the total probability of forming a trefoil knot on the chain. For all channel sizes as well as the bulk case, the probability distribution attains a maximum at L^*_{knot}, indicating a metastable knot. The shapes and magnitudes of the probability distributions of knot size are markedly influenced by confinement, and the metastable knot size L^*_{knot} depends on the channel size.

The total probability of forming a trefoil knot as a function of the channel size is shown in the inset of Figure 6.2. As the degree of confinement is increased, f_total increases from the bulk value 0.011 to the peak value 0.043 at $D = 8L_p$. A further reduction of the channel size causes a sharp decrease in f_{total} to zero. A similar non-monotonic behavior in knotting probability was also observed in a previous study of circular chains in slits [135].

The metastable knot size L^*_{knot} also varies non-monotonically with D, shown in Figure 6.3. As D decreases, L^*_{knot} increases from the bulk value of $27.2L_p$ to the peak value of $49.6L_p$ at $D = 12L_p$, decreasing rapidly thereafter. The metastable knot sizes shown here are much less than the contour length of the chain $L = 400L_p$; accordingly, the end effects due to finite chain length are weak.

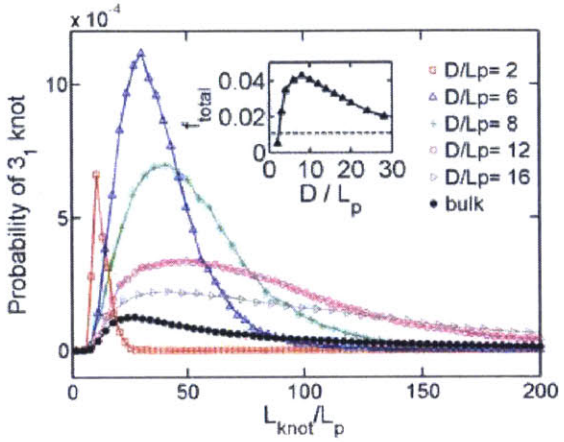

Fig. 6.2: *Probability distributions of the sizes of trefoil knots for different confining channel widths, D. For all curves, the contour length is fixed as $L = 400L_p$, and the chain width is fixed as $w = 0.4L_p$. The inset shows the total probability of trefoil knots as a function of channel size. The dashed line in the inset shows the probability of trefoil knots in bulk.*

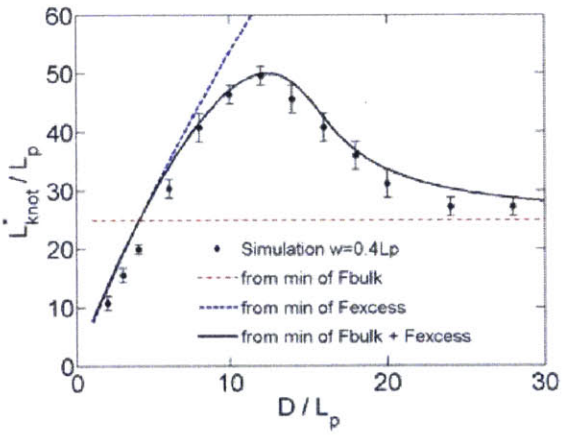

Fig. 6.3: *The most probable size of a trefoil knot as a function of the channel size. The solid line corresponds to the minimization of free energy in Eq. 6.2 with respect to L_{knot}. The contour length is fixed as $L = 400L_p$, and the chain width is fixed as $w = 0.4L_p$.*

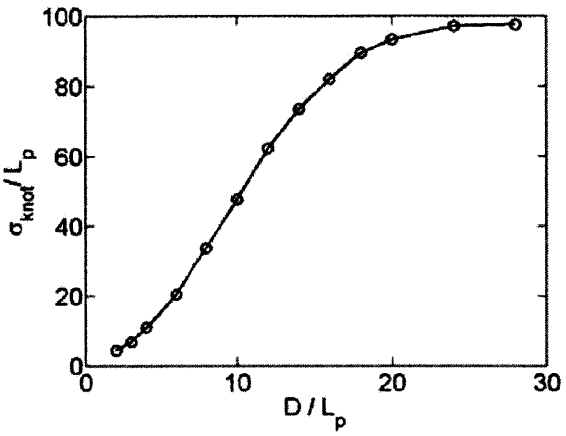

Fig. 6.4: *The standard deviation of knot size as a function of the channel size. The contour length is fixed as* $L = 400L_p$, *and the chain width is fixed as* $w = 0.4L_p$.

To rule out possibility that the non-monotonic change of L^*_{knot} versus D is caused by entering the weak confinement regime, we identified the de Gennes regime as $4 \leq D/L_p \leq 20$ by analyzing the confinement free energy of the chain (see Appendix F). The peak of L^*_{knot} at $D/L_p \approx 12$ is located well inside the de Gennes regime, so a transition in the effective confinement regimes is not responsible for the non-monotonic change of L^*_{knot} versus D.

The standard deviation of knot size, σ_{knot}, as a function of the channel size is shown in Figure 6.4. The fluctuations in the size of the knots monotonically decrease as the channel size is decreased. This result directly reflects the tightening of the distributions of knot sizes, shown in, Figure 6.2. In other words, confinement acts to stabilize and more sharply localize the knot on a chain. The data in Figure 6.4 agree qualitatively with experimental observation of small fluctuations in knot size during the diffusion and escape of a knot along a DNA molecule confined in a channel [71].

The average deformation ratio of trefoil knots as a function of the channel size is shown in Figure 6.5a. The aspect ratio is calculated as $2\langle R_x \rangle / (\langle R_y \rangle + \langle R_z \rangle)$, where $\langle R_x \rangle$, $\langle R_y \rangle$, and $\langle R_z \rangle$ are the radii of gyration in each direction. The x-axis corresponds to the axis of the channel. In the absence of confinement, knots are isotropic, and the average aspect ratio is one. Even in weak confinement, noticeable anisotropy is present in the knotted regions. This result may occur due to deformation of knots located near the channel walls. The aspect ratio increases monotonically as the channel size decreases. When $D \leq 6L_p$, the average aspect ratio increases more sharply. The extensions of knots along the direction of the channel axis are shown in Figure 6.5b. The extension is defined as the maximum span of the knot projected in the direction of the channel axis. We chose to plot the most probable value of extension $L^*_{||}$ rather than the mean value of extension $\langle L_{||} \rangle$ because $L^*_{||}$ is insensitive the contour length of the chain. The knot extension is smaller than the channel size for $D > 4L_p$

Next, we compare the metastable knot size in simulations with the theoretical prediction from Eq. 6.1, shown by the black curve in Figure 6.3a. By choosing numerical coefficients $\alpha = 0.1$ and

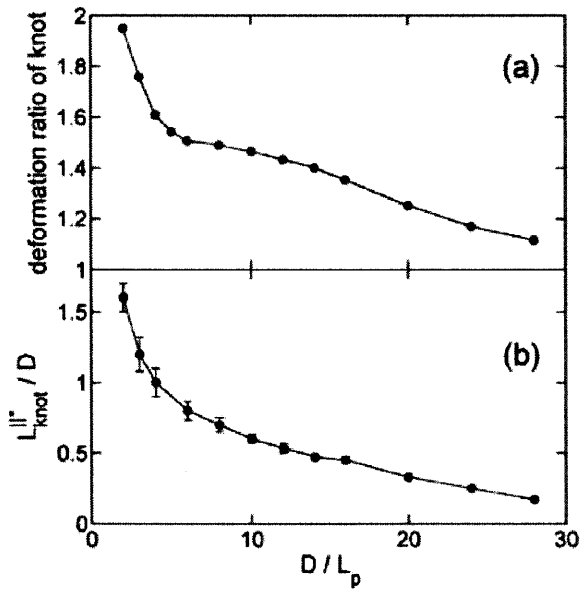

Fig. 6.5: *The aspect ratio of knots as a function of the channel size. The aspect ratio is calculated as* $2\langle R_x\rangle / (\langle R_y\rangle + \langle R_z\rangle)$, *where* $\langle R_x\rangle$, $\langle R_y\rangle$, *and* $\langle R_z\rangle$ *are the radii of gyration in each direction. The x-axis corresponds to the axis of the channel. (b) The most probable extension of trefoil knot as a function of the channel size. The contour length is fixed as*$L = 400L_p$, *and the chain width is fixed as* $w = 0.4L_p$.

$\beta = 0.02$, our theory is able to quantitatively reproduce the dependence of L^*_{knot} on D with two prefactors, shown in Fig. 6.3. The fitted value of $\alpha = 0.1$ suggests that the effective diameter of knot-ball, b_{knot}, is roughly $0.1L_{knot}$, which is reasonable. After substituting $b_{knot} = 0.1L_{knot}$ into Eq. 6.3 with $\beta = 0.02$, we have $F^w_{knot}all = 0.93b^{5/3}_{knot}(D - b_{knot})^{-5/3}$, where the prefactor 0.93 is close to unity.

In our theory, the metastable knot size is determined by the minimum of $F^{bulk}_{knot}(L_{knot})$ and $F_{excess}(L_{knot})$ for a given channel size. To further examine our theory, we compare F^{bulk}_{knot} and F_{excess} calculated from simulations and theory. The free energy of knots in bulk F^{bulk}_{knot} as a function of knot size is shown in Figure 6.6a. The metastable knot size in bulk, L^{bulk*}_{knot}, predicted by Eq. 6.1 is $27.2L_p$, slightly larger than $24.9L_p$ observed in the simulations. Note that, the calculation of $F^{bulk}_{knot}(L_{knot})$ by Eq. 6.1 is no longer valid for very large knots, because Eq. 6.1 is derived based on the Odijk scaling for self-confinement free energy of knots in bulk, and the Odijk scaling is only valid in strong confinement, *i.e.* the tight knots [61, 144].

The values of F_{excess} extracted from simulations of knots are shown in Figure 6.6b. The value of F_{excess} is extracted using $F_{excess} \equiv F^{wall}_{knot} - F^{wall}_{nonknot} = F^{bulk}_{knot} - F^{channel}_{knot} = \log(f_{channel}) - \log(f_{bulk})$, where $f_{channel}$ and f_{bulk} are the knotting probabilities as a function of L_{knot}, shown in Figure 6.2. Recall that $F^{wall}_{knot} - F^{wall}_{nonknot} = F^{channel}_{bulk} - F^{channel}_{knot}$ is based on the diagram in Figure 6.1a. In addition, applying $F^{bulk}_{knot} = -\log(f_{bulk})$ and $F^{channel}_{knot} = -\log(f_{channel})$is accurate only when the probability of forming a trefoil knot is small and other knotting probabilities are negligible in comparison. This requirement is satisfied since the probability of forming a trefoil knot is small ($< 5\%$) and the probabilities of other knots are always less than 1%.

The theoretical predictions of F_{excess} calculated using Eqs. 6.3–6.5 with the fitted coefficients $\alpha = 0.1$ and $\beta = 0.02$ (determined in the fit of L^*_{knot} in Figure 6.3) are shown in Figure 6.6c. Our theory qualitatively captures three major features of F_{excess}. First, $F_{excess}(L_{knot})$ exhibits a minimum. Second, the value of L_{knot} at the minimum of F_{excess} decreases with decreasing D. Third, the potential well becomes deeper with decreasing D. Quantitatively, our theoretical predictions of F_{excess} deviate from the ones extracted from simulation, particularly for large L_{knot}. This deviation is caused by two major assumptions in our theory: (i) the knot core is treated as an isotropic ball (ii) the effective diameter b_{knot} of the ball is proportional to L_{knot}. The first assumption is not valid for small channels, because the knot is significantly deformed as shown in Figure 6.5a. The second assumption is not valid for large knots, as shown by data of knot radius versus L_{knot} in Appendix F.

The insights from Figure 6.6 enable a deeper understanding of the dependence of L^*_{knot} on D in Figure 6.3. The metastable knot size corresponds to the minimum of $F^{channel}_{knot} = F^{bulk}_{knot} + F_{excess}$. The location of F^{bulk}_{knot} minimum is independent of D (red line Figure 6.3a). The value of L_{knot} which minimizes F_{excess} becomes smaller with the decreasing D (blue line Figure 6.3a). In wide channels, the confinement effect is weak, $F^{channel}_{knot}$ is dominated by F^{bulk}_{knot}, and thus L^*_{knot} is close to the bulk value L^{bulk*}_{knot}. In narrow channels, $F^{channel}_{knot}$ is dominated by F_{excess}, and L^*_{knot} approaches the minimum of F_{excess}. In between these extremes, the competition between F^{bulk}_{knot} and F_{excess} causes the non-monotonic change in L^*_{knot}. Interestingly, although our theoretical predictions of F^{bulk}_{knot} and F_{excess} in Figure 6.6 are not precise, the prediction of L^*_{knot} is significantly more accurate probably due to cancellation of errors. As shown by Figure 6.6a, Eq. 6.1 overestimates the shrinking force $-\partial F^{bulk}_{knot}/\partial L_{knot}$ of knots when the knot size is larger than the metastable knot size in bulk. On the other hand, as shown by two y-axes in Figures 6.6b and 6.6c, Eq. 6.3 overestimates the swelling forces $-\partial F_{excess}/\partial L_{knot}$.

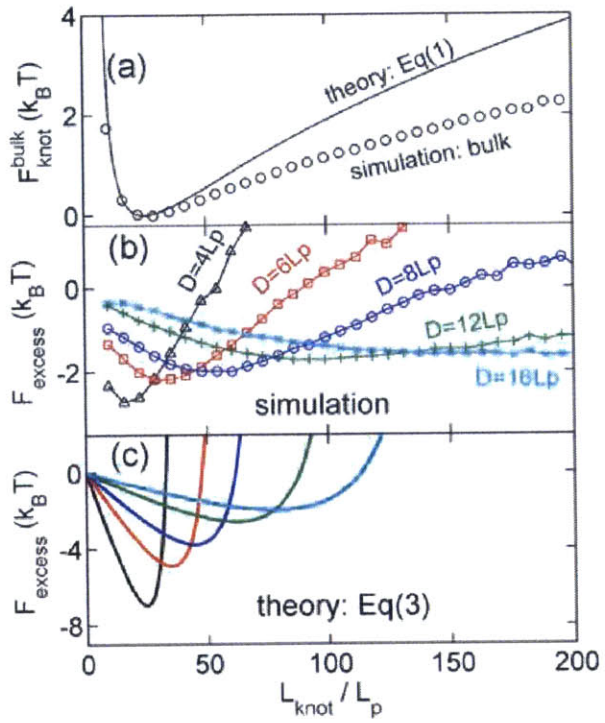

Fig. 6.6: *(a) The potential of mean force as a function of knot size for a semiflexible chain in bulk. (b) The difference in confinement free energy, F_{excess}, between a knotted and unknotted subchain as a function of the knot size calculated from simulation. (c) F_{excess} calculated from theory. The contour length is fixed as $L = 400L_p$, and the chain width is fixed as $w = 0.4L_p$.*

The physical reason why the shrinking a knot can reduce the confinement free energy of the knot is that a subchain of L_{knot} is compacted in the knotted region. For the knotted and unknotted states, the confinement free energies scale as $(D - \alpha L_{knot})^{-5/3} L_{knot}^{5/3}$ and $D^{-5/3} w^{1/3} L_p^{1/3} L_{knot}$ respectively. The ratio is $L_{knot}^{2/3}/(w^{1/3} L_p^{1/3})$, if we approximate $(D - \alpha L_{knot}) \approx D$. This ratio approaches zero as the knot size approaches zero. As the knot size increases, $L_{knot}^{2/3}/(w^{1/3} L_p^{1/3})$ eventually becomes larger than 1. Furthermore, as the knot size increase, the term $D/(D - \alpha L_{knot})$ increases to an infinitely large value, which corresponds to a large ball without transitional freedom in the cross-section of the channel. The knotted region, however, is a soft ball rather than a hard ball. Thus the entropy gained from deformability becomes the dominant effect.

The results in Figures 6.2–6.6 are for confined semiflexible chains with the chain width $w = 0.4L_p$ and contour length $L = 400L_p$. To investigate the effect of changes in chain width on the agreement between simulation and theory, we performed simulations for confined semiflexible chains with chain widths $w = 0.2L_p$ and $w = 0.1L_p$. The metastable knot sizes for three different chain widths are compared to our theoretical predictions in Figure 6.7. Recall that our theory requires two

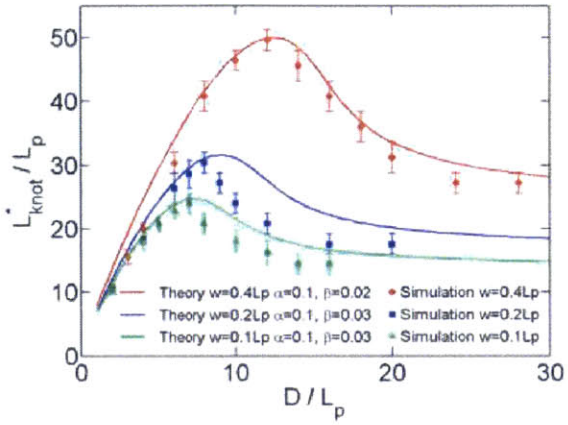

Fig. 6.7: *The metastable knot size as a function of channel size for different chain widths. The symbols and lines correspond to simulation results and theoretical predictions, respectively. The contour lengths are fixed as* $L = 400L_p$.

numerical coefficients α and β. For $w = 0.4L_p$, we use the coefficients $\alpha = 0.1$ and $\beta = 0.02$ to attain agreement between theory and simulation. For $w = 0.2L_p$ and $w = 0.1L_p$, we maintain $\alpha = 0.1$ and change β to 0.03 in order to match our theory to simulation results. The coefficient α is used to relate the effective diameter of the knot, b_{knot}, and the contour length of the knot, L_{knot}. Our simulation results show that the relationship between b_{knot} versus L_{knot} is insensitive to the chain width (See Appendix F), and thus the value of α is independent of the chain width. The reason why we change β for different chain widths is as follows. The theoretical equation for F_{knot}^{bulk} (Eq. 6.1) is fairly accurate for thin chains (See Appendix F and Ref. [144]) but significantly overestimates the force which shrinks a knot for thick chains (Figure 6.6a). To compensate this overestimation, we need to similarly overestimate the swelling force due to F_{excess}. Therefore, the value of β used in Eq. 6.5 is larger for thicker chains.

Our theory is able to capture three major trends concerning the dependence of L_{knot}^* on w and D, shown in Figure 6.7. First, as w decreases, the peak value decreases. Second, as w decreases, the peak in L_{knot}^* shifts to smaller channel sizes. Third, for all w, the curves of L_{knot} versus D converge at very small channel sizes.

6.5 Conclusions

We have used computer simulations to show how confinement in a tube may alter the sizes and probability of formation of knots in polymer molecules. We put these simulation results in context of a theory that explains the physics of the process: the competition between bending energy within the knot and confinement energy within the knot and due to the walls determines the knot size. The metastable knot size in weak confinement is larger than the one in bulk because the segments in the knot core experiences less confinement free energy than the ones in the unknotted subchains and it behaves like that the channel pushes segments into knots to reduce the confinement free

energy. On the other hand, the metastable knot size in strong confinement is smaller than the one in bulk, because the segments in the knot core experiences more confinement free energy, and the channels pushes the segments out of the knot core.

CHAPTER 7

Untying Knotted DNA with Elongational Flows

In this Chapter, we shift our focus to studying the detailed *dynamics* of knots on DNA molecules. In particular, we look at the motion of a knot on a DNA molecule that is stretched out by an elongational flow. We show this type of flow will cause a knot to be swept off the DNA and "untie," and we are able to describe different regiemes of the motion of the knot through a simple mathematical model. Interestingly, we see, as others have seen in quite different simulations, that torus knots, a specific family of knots, move more quickly than other types of knots. We explain this difference by obseving that these torus knots can spin while moving along the DNA while other knots are unable to do so.

The results in this chapter have been published in reference [187] and are reproduced with permission from the ACS Publications, Copyright 2014.

7.1 Overview

We present Brownian dynamics simulations of initially knotted double-stranded DNA molecules untying in elongational flows. We show that the motions of the knots are governed by a diffusion-convection equation by deriving scalings that collapse the simulation data. When being convected, all knots displace non-affinely, and their rates of translation along the chain are topologically

dictated. We discover that torus knots "corkscrew" when driven by flow whereas non-torus knots do not. We show that a simple mechanism can explain a coupling between this rotation and the translation of a knot, explaining observed differences in knot translation rates. These types of knots are encountered in nanoscale manipulation of DNA, occur in biology at multiple length scales (DNA to umbilical cords), and are ubiquitous in daily life (*e.g.* hair). These results may have broad impact in manipulations of such knots via flows, with applications to genomic sequencing and polymer processing.

7.2 Introduction

Knots are commonly encountered and manipulated in everyday experiences such as tying one's shoelaces or untangling spontaneously knotted strings [35]. Formally defined only for closed rings, the topologies of "open" knots (referred to hereafter simply as knots) are often unambiguous (*e.g.* shoelaces and neckties) and can be closed and algorithmically defined [53, 83, 84]. At microscopic scales, chromosomal knots are modified by topoisomerases during cell division [36] and are thought to participate in gene regulation [188]. Knots are found in proteins [43, 189] and viral capsid DNA [41, 42], likely with yet to be fully understood functions. It has been mathematically proven that knots become asymptotically likely as the length of a polymer increases [49], a fact that explains their ubiquity.

Due to the emerging significance of knotted polymers, a growing body of simulation literature is devoted to their study [40, 46]. For instance, while the topology of a ring is fixed, an open polymer can spontaneously form and untie knots [147]. The probability of forming such knots can be non-monotonic when the polymer is confined in slits[133, 135] or tubes[134], and increasing the stiffness of a polymer can similarly influence the knotting probability in unintuitive ways [156, 157]. Such knots can substantially affect the mechanical properties [69, 149] and rheological behavior [150] of a polymer, and the probability of forming knots has been used to infer the effective diameter of DNA molecules[176]. Recently, simulations have shown dramatic slowing of processes wherein a knot is *driven* along a chain such as entropic ejection of DNA from a viral capsid [62] and the translocation of single-stranded DNA (ssDNA) [66] and polypeptides [64] through pores.

Common nanofluidic experiments have led to the spontaneous formation of knots in DNA by collision with channel defects [71] or the application of moderate electric fields [73, 86] during electrophoresis. More broadly, the growing library of methods to manipulate DNA molecules in nanofluidic devices has enabled fundamental research about single polymer molecules [25, 190]. These studies inform important applications such as genomic sequencing via nanopore translocation [77] or direct linear analysis [113]. Thus, (un)tying knots in polymers is of interest in its own right. To this end, knots have been intentionally tied with optical tweezers in actin filaments [69] and double-stranded DNA (dsDNA) [70]. Impressively, simulations reproduced the sizes and diffusion coefficients of dsDNA knots within a factor of two [170].

In this Letter, we use simulations to investigate the transport of a knot on a dsDNA molecule that has been extended by an elongational flow. We show that such flows cause the knot to be driven off the chain and untied, and we elucidate the relevant length and time scales for this process by examining the diffusion-convection equation. We observe that knots of different topologies translocate at different rates when strongly driven by the external flow. We show the different rates of knot translations are explained by a rotational mode of motion, available to torus knots [79], that facilitates the translation of the knot, providing unique mechanistic insight.

7.3 Simulation Methods

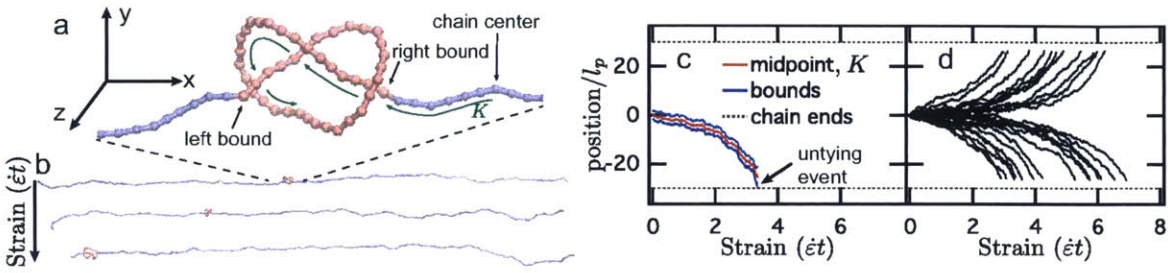

Fig. 7.1: *(a) Knotted (red) and unknotted (blue) regions of DNA extended by elongational flow (Wi=16) for the* 3_1 *knot. The distance along the contour to the midpoint of the knot,* K*, is shown in green. (b) Simulation snapshots of an initially centered* 3_1 *knot untying from DNA at flow strength* $\mathrm{Wi} = 16$*. (c) The midpoint and bounds of the knot pictured in (b) are plotted versus strain. (d) Untying trajectories of 25 initially centered* 3_1 *knots at flow strength* $\mathrm{Wi} = 16$*.*

We have used a Brownian dynamics approach to simulate dsDNA, which has been extensively parameterized by others[170, 191, 192]. The dsDNA molecule is represented by a fine-grained bead-spring model with $n_p = 5$ stiff bonds per persistence length, $l_p = 50$ nm. Screened Debye-Hückel interactions are used to model the long-range electrostatics of DNA-DNA interactions, and all simulations used an ionic strength of $I = 10$ mM, leading to a Debye length $\kappa^{-1} \approx 3$ nm and a effective chain diameter of $w \approx 16$ nm[176, 193, 194]. These values were chosen to be representative of common low-salt conditions in microfluidic dsDNA experiments. An external planar elongational flow of the form $\mathbf{u}(\mathbf{r}_i) = \dot{\varepsilon}\,(\hat{\mathbf{x}} - \hat{\mathbf{y}}) \cdot \mathbf{r}_i$ was considered, where $\dot{\varepsilon}$ is the strain rate and $\mathbf{r}_i$ is the position of the ith bead. We neglected inter-bead hydrodynamic interactions in this work, so the drag force on the ith bead is simply $\mathbf{F}_i^{\mathrm{d}} = \zeta_b \left[\mathbf{u}(\mathbf{r}_i) - \frac{d\mathbf{r}_i}{dt}\right]$ where ζ_b is the drag coefficient of a single bead. The Weissenberg number, $\mathrm{Wi} = \dot{\varepsilon}\lambda$, is the appropriate dimensionless group for such flows. The DNA longest relaxation time, λ, was determined by fitting the long-time decay of the squared end-to-end distance of an initially stretched chain to a single exponential. See Appendix G for additional simulation details, including considerations of how the addition of intra-bead hydrodynamic interactions and DNA torsional rigidity could alter the results of the simulations.

7.4 Results and Discussion

In Figure 7.1a, a snapshot of the 3_1 knot (Alexander-Briggs notation) on an extended DNA molecule at flow strength $\mathrm{Wi} = 16$ is shown. The "knotted" region, identified algorithmically (see Appendix G), is shown in red and the unknotted region in blue. The knot, initially centered in the internal chain coordinates, is ultimately driven to the end of the chain where it unties, shown in the snapshots in Fig. 7.1b.

Following previous studies [195], the position of a knot is tracked in an "internal" chain coordinate by measuring the length of chain contour between the central index of the chain and midpoint index of the knot, a length we denote as K. This scalar is signed; a negative sign indicates a position nearer the first index of the chain than the last, a positive sign the converse. The trajectory corresponding to the snapshots in Fig. 7.1b is plotted in Fig. 7.1c. A knot is considered to have "untied" itself when it's boundary reaches the first or last index of the chain. The escape trajectories of 25 initially centered 3_1 knots are shown in Fig. 7.1d. Knots escape to the left and right with equal probability, yet after moving a finite distance from the center, the directionality of escape is fixed. Furthermore, it is *extremely* difficult for a molecule that is extended by such a flow to form new knots. Thus, these results show the ability to guide the topology of a polymer to an unknotted state, which may be useful in aforementioned applications.

We next explore the physics behind these observations. In a statistical sense, the evolution of knot trajectories corresponds to the evolution of an underlying probability density function Ψ of knot positions. This idea can be cast in the form of a diffusion-convection equation for a knot:

$$\frac{\partial \Psi}{\partial t} = \nabla \cdot (D \nabla \Psi) - \nabla \cdot (\mathbf{u} \Psi) \tag{7.1}$$

where D is the diffusivity of the knot.

We consider the 1-D case of a knot with constant diffusivity confined to the stretching direction of the applied elongational flow, $\mathrm{u}_x = \dot{\varepsilon} x$. We propose this scenario is the simplest approximation of the knot motion along a chain extended by an elongational flow (See Appendix G). For this simplified system, the diffusion-convection equation becomes:

$$\frac{\partial \Psi}{\partial t} = D \nabla^2 \Psi - \dot{\varepsilon} \left(x \frac{\partial \Psi}{\partial x} - \Psi \right) \tag{7.2}$$

Examination of Eq. 7.2 reveals the fundamental scale for time to be $\dot{\varepsilon}^{-1}$. A length scale, $l = \sqrt{D \dot{\varepsilon}^{-1}}$, emerges where the probability flux contributions due to diffusion and flow are balanced. To the first order, the motion of the knot will be diffusive for $|x| \ll l$ and follow the deterministic path of the flow for $|x| \gg l$. We independently measured knot diffusivities from chains held at the ends by constant tension (see Appendix G). We found the diffusivity of the knots in this study to be topologically dependent but not strongly influenced by tension (and thus, Wi, see Appendix G). Thus, the scalings depend on knot topology alone.

We simulated the process of a knot, initially centered in the internal chain coordinates ($K(t = 0) = 0$), escaping from a chain extended by elongational flow for a variety of topologies and flow strengths. In Figure 2a, the mean squared displacements of knot position are plotted versus time. The scalings from the diffusion-convection equation, $\dot{\varepsilon}^{-1}$ for time, and $l^2 = D\dot{\varepsilon}^{-1}$ for length are used to rescale the data, plotted in Fig. 7.2b. The collapse of the data is good, especially given the simplistic approximations implicit in the scalings. When $K/l \ll 1$, a slope of 1 agrees with the collapsed simulation averages, indicating a diffusion-dominated regime. When $K/l \gg 1$, the trajectories expectedly diverge from this slope, indicating the predicted dominance of the external flow, a convection-dominated regime. These results show an uncomplicated model can capture the dynamics of simple knots on chains in elongational flows. Furthermore, these results elucidate the fundamental length and time scales of interest in a practical application — untying knotted molecules with an elongational flow.

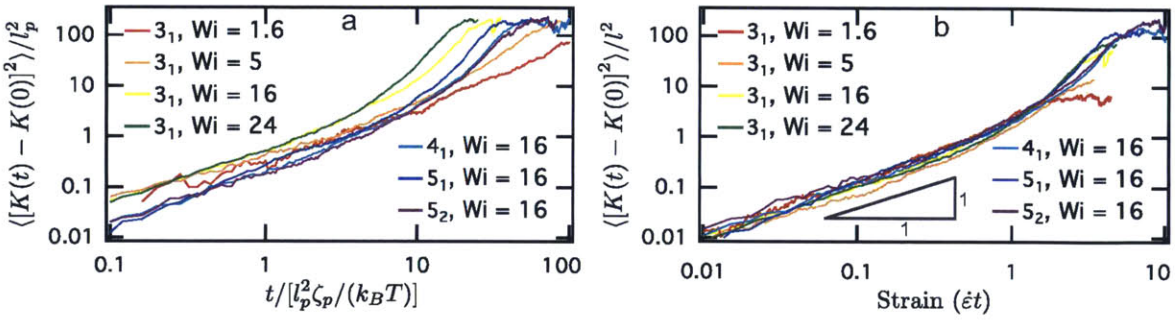

Fig. 7.2: *Mean squared displacements of initially centered knots ($K(t = 0) = 0$) in elongational flows. (a) Knot displacement plotted versus dimensionless time ($\zeta_p = n_p\zeta_b$). (b) Scaled knot displacement plotted versus strain. The triangle represents the slope of the diffusion-dominated mean squared displacement.*

We next explore the convection-dominated regime, *i.e.* $K \gg l$, in greater detail for the following topologies: 3_1, 4_1, 5_1, 5_2, 6_1, 6_3, 7_1, 9_1, 10_{28}, 11_1, and 15n165258 [1]. The mean displacements of these knots are plotted in Fig. 7.3a. The slope of any given trajectory is less than that of affine deformation, indicating that all knots lag the applied flow. The degree to which they lag the flow, remarkably, is strongly influenced by topological class rather than size or self-diffusion coefficient. For instance, the 4_1 knot, despite being the second simplest knot, moves much slower than the larger 7_1 knot. The non-torus knots have remarkably similar rates of convection despite drastically different sizes and topologies. Moreover, the $(2n + 1)_1$ knots, all torus knots, convect faster than every other knot studied. As the crossing number of the torus knots increases, the mean displacements seem to asymptotically approach that of the non-torus knots.

These findings indicate the torus knots access an additional collective mode of motion while being driven along the chain. This idea is akin to topologically controlled breathing modes that have been described in knots under tension [196] but is distinct in that we propose driving the knot along the chain preferentially activates this dominant mode. To visualize the relative motion of knot segments, we examined the displacements of the knot midpoint from the knot center of mass in two axes orthogonal to the stretching direction, ΔK_y and ΔK_z, where $\Delta K_y = \left(\mathbf{r}_{(R+L)/2} - \frac{1}{R-L+1}\sum_{i=L}^{R}\mathbf{r}_i\right)\cdot\hat{\mathbf{y}}$ and L and R denote the indices of the left and right bounds of the knot. The projections of dimensionless versions of these quantities, $\Delta\tilde{K}_y = \Delta K_y/l_p$ and $\Delta\tilde{K}_z = \Delta K_z/l_p$, are plotted for the 3_1 and 4_1 knots in Figs. 7.3b and c, respectively. The central segment of the 3_1 knot follows a circular trajectory around the center of mass of the knot, but the 4_1 knot shows no discernable pattern. The rotational modes of motion become evident in the plots in Figs. 7.3(d-i). The 3_1, 5_1, and 7_1 knots demonstrate sustained "corkscrew" rotations of the central segment where, as the knot is driven down the chain by the elongational flow, a secondary global rotation of the knot develops. These sustained rotations are clearly specific to the knot, not the flow; the flow, itself, is irrotational. These trajectories sharply contrast to those of the 4_1 knot where no sustained rotation of the central segment can be seen. The 5_2 and 6_1 knots demonstrate half rotations of the central segment, but full rotations are absent. It is here that we

make a critical mechanistic observation: all knots move via self-reptation of segments, yet the torus knot topologies permit that a global, sustained *rotation* of the knot facilitates an additional *translation* of that knot down the chain. The non-torus knots cannot access these means of transport due to topological self interference.

We tested this mechanistic insight by predicting the chirality of knot rotations from simplistic sketches for the 3_1 knot. The 3_1 knot exists in right and left-handed chiralities. The sketches in Fig. 4 show that for a knot of given chirality, globally rotating the knot will produce a translational motion of the knot along the extended chain. Conversely, when a 3_1 knot of given chirality is being convected by an elongational flow, the chirality of rotation that coincides with the direction of translation can be predicted from the sketches. The colored trajectories in Fig. 7.4 show the displacements of the central segments from the center of mass of knots of given chirality (right or left-handed) being convected by an elongational flow off the right or left of the chain. The mechanism in the sketches is in complete agreement with the visualized simulation results. We have found this agreement between knot chirality, direction of knot translation, and knot rotation to hold in all observed simulation trajectories (~ 50), *i.e.* knot rotation always occurs as predicted by the sketches. These results confirm our idea that the translational motion of torus knots is facilitated by knot rotation. Further, this mechanism clearly extends to all of the $(2n+1)_1$ family of torus knots; the addition of extra loops in the knot does nothing to prohibit the coupling of knot rotation and translation. The increased mobility of torus knots appears to be quite general, having been demonstrated in macroscopic shaking chain experiments [197], simulations of DNA ejection from viral capsids [63], and simulations of tensioned electrophoresing DNA[143]. We postulate our mechanism plausibly explains these results.

7.5 Conclusions

We have demonstrated the ability of elongational flows to untie knotted dsDNA molecules with computer simulations. Through scaling analysis, we revealed a critical length scale along the molecule that separates diffusive from convective transport of a knot in these flows. We have shown that a subset of knots, torus knots, can move linearly via sustained global rotation, which increases the speed at which they release from the chain. This represents an important mechanistic insight into the motion of driven knots, and since the mechanism is solely a function of topology, we expect it will apply in many other driven processes (*e.g.* nanopore sequencing or viral ejection of DNA) and for other polymers. We speculate that for longer DNA or for higher flow strengths, self-jamming of knots may be observed as has been seen in tensioned knots [64], and we hope future work will address this notion. Simulations show knots can jam nanopore translocation [66], and experiments suggest jamming of knots during ultrafiltration of plasmid DNA [198]. Practically, our results could guide development of microfluidic devices that precondition molecules to unknotted states for such applications. Finally, since the model employed is parameterized to dsDNA at experimentally realizable conditions, we hope future experiments will test our results.

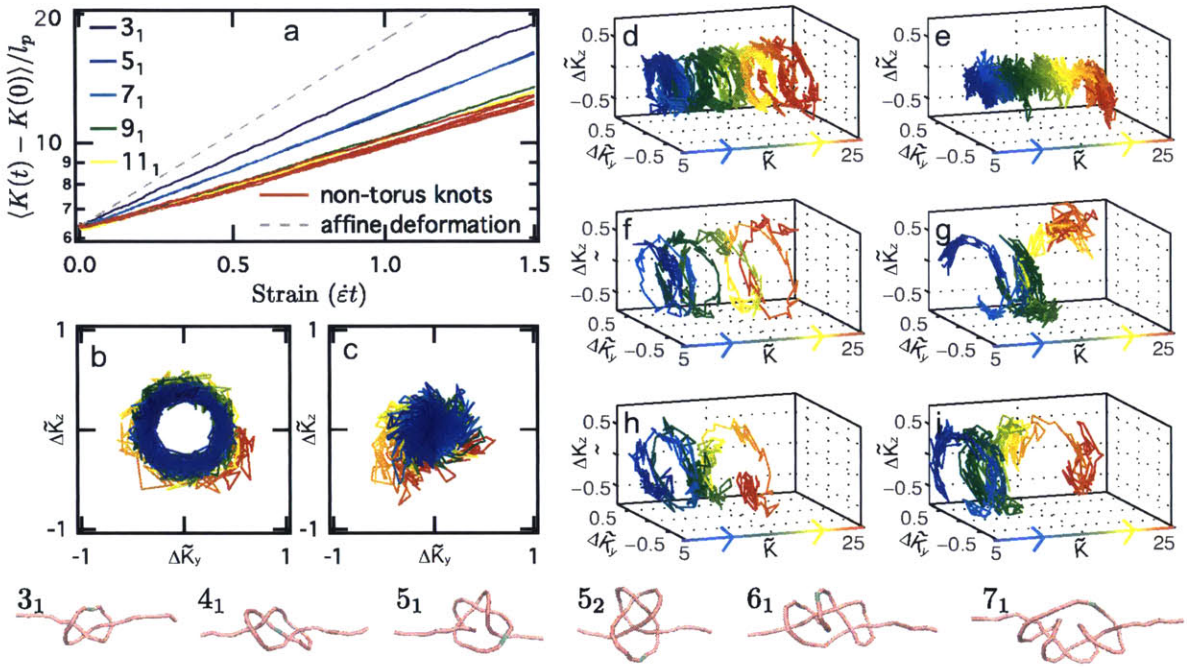

Fig. 7.3: *(a) Average knot displacements are plotted vs. strain for various knots at flow strength* $\mathrm{Wi} = 16$ *for knots initialized off-center,* $K(t=0) = 6l_p \gg l$ *for all topologies. The non-torus knots consist of the following:* 4_1, 5_2, 6_1, 6_3, 10_{28}, *and 15n165258 [1]. Displacements from the knot center of mass of the central segment of the knot made dimensionless by* l_p *(*$\Delta\tilde{K}_y$ *&* $\Delta\tilde{K}_z$*) for the* 3_1 *(b) and* 4_1 *(c) knots in the plane orthogonal to the extensional axis; color changes from blue to red as the knot moves off the chain. Displacements from the knot center of mass of the central segment of the knot (*$\Delta\tilde{K}_y$ *&* $\Delta\tilde{K}_z$*) in the plane orthogonal to the extensional axis plotted versus knot position (*$\tilde{K} = K/l_p$*) for the* 3_1 *(d),* 4_1 *(e),* 5_1 *(f),* 5_2 *(g),* 7_1 *(h), and* 6_1 *(i) knots. Color changes from blue to red as the knot moves off the chain. Bottom: Snapshots of knots in (d-i) with central segments highlighted in green.*

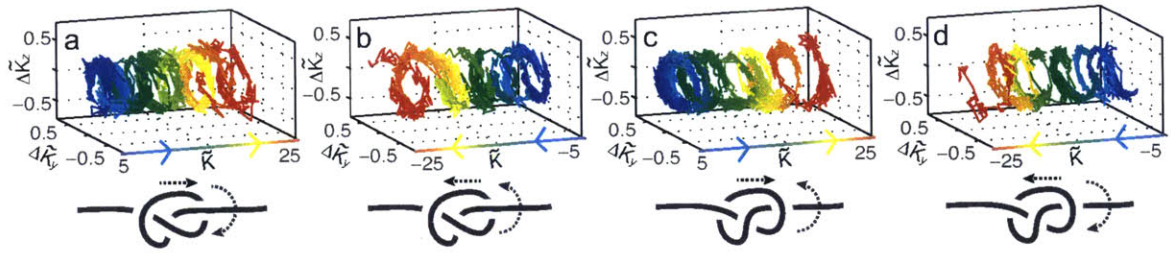

Fig. 7.4: *Displacements from the knot center of mass of the central segment of the knot ($\Delta\tilde{K}_y$ & $\Delta\tilde{K}_z$) in the plane orthogonal to the extensional axis plotted versus knot position ($\tilde{K}$) for the 3_1 knot. Color changes from blue to red as the knot moves off the chain. The right-handed (a & b) and left-handed 3_1 knots (c & d) rotate with opposite chiralities as they move off the right (a & c) and left (b & d) of the chain. All knots were initialized $K(t=0) = \pm 6l_p$, $6l_p \gg l$ with flow strength* $\mathrm{Wi} = 16$.

CHAPTER 8

Jamming of Tensioned Knots

The results in this Chapter are in preparation for publication and presented in exploratory form. This work was performed in close collaboration with Vivek Narsimhan, who ran the simulations shown here.

We performed Brownian dynamics simulations of a linear, flexible chain containing knots at high tensions. The chains under examination consisted of 100 to 300 spherical monomers of diameter b. Connectivity between adjacent monomers was enforced with rigid rods (rather than stiff springs) of length b, *i.e.* touching beads. Harmonic excluded volume interactions between beads started at a separation between beads $r < b$, and the magnitude of these interactions was chosen so as to apply the repulsive force needed to bring the overlapping beads back to contact within the next time step. The equations of motion for the chain were integrated using a semi-implicit predictor corrector scheme. The routines for initializing different knotted topologies and tracking the positions were used in Chapter 7 and are discussed in greater detail in Appendix A

A schematic illustrating the overarching idea of our simulations is shown in Fig. 8.1. In Fig. 8.1a, a stretched chain with a simple knot tied far from the chain ends is shown. In Fig. 8.1b, and c, trajectories of the evolution of the position of knots at two different tensile forces are shown. These trajectories show a progressively discretized motion of the knot as the tension on the chain is increased. At high forces, we are able to fit a piece-wise constant to the motion of the jumps, shown in Fig. 8.1d. The sizes of these jumps correspond to roughly the bead diameter and the distribution of jump frequencies can be described as a Poisson process.

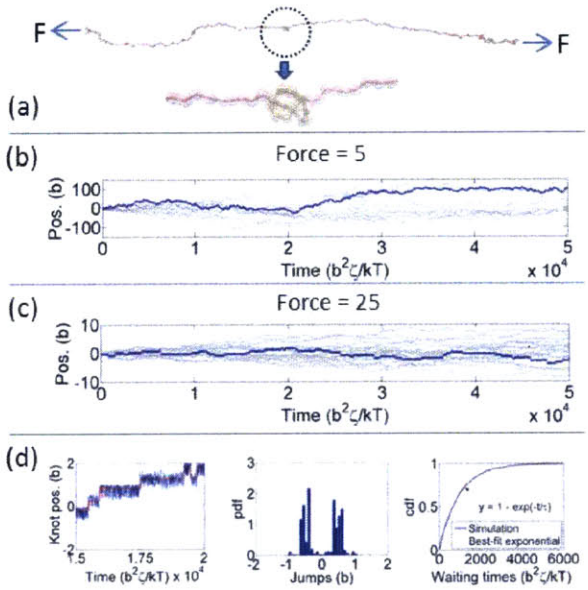

Fig. 8.1: *(a) Schematic of knotted polymer under tension. (b, c) Position of a knot along the chain as a function of time for force $Fb/kT = 5$ and 25. Twenty runs are shown with one run highlighted in blue. (d) Left: Best piece-wise constant fit for trajectory for force $Fb/kT = 25$. Center and right: distribution of the jumps and waiting times for this fit. The motion is well described as a Poisson process.*

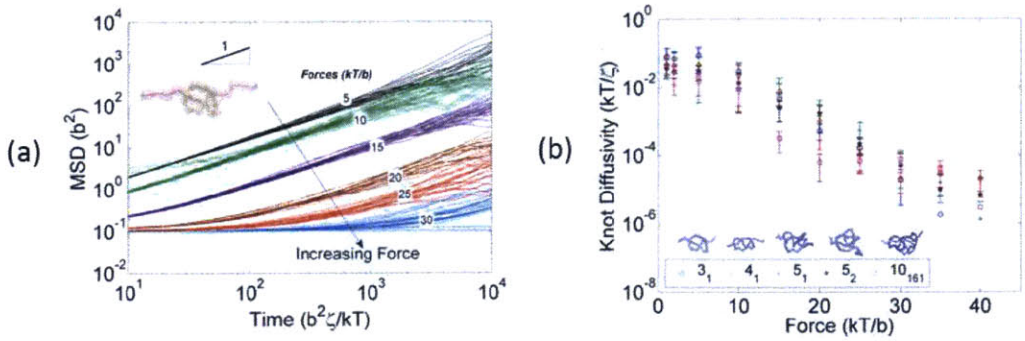

Fig. 8.2: *(a) Mean-squared displacement vs. time for the* 3_1 *knot at various tensions (*Fb/kT *ranges between 5 and 30). (b) Knot diffusivities vs. applied force for various types of knots (*3_1, 4_1, 5_1, 5_2, *and* 10_161*). The diffusivity decreases exponentially beyond a critical force (onset of jamming).*

We next seek to understand in greater detail how the long-time diffusivities of knots change as a function of tension for "high" tensile forces, $Fb/kT \geq 5$. The lowest of these forces corresponds to the onset of "jamming." In Fig. 8.2a, we show how the ensemble averaged mean-squared displacements of the 3_1 knot vary with applied tension. Increasing tension clearly serves to reduce the mobility of these knots. At long enough times, the mean-squared displacements become linear. From this linear region, a long-time diffusivity can be extracted. In Fig. 8.2b, the diffusivities for a number of different knots as a function of applied tension are plotted. Jamming begins to occur upon the application of some critical tensile force, $Fb/kT \sim 5$. At higher forces, the diffusivites decrease exponentially.

This Arrhenius decrease in the diffusivities of knots can be explained via a simple mechanism. At high forces, nearly all the force in the chain is directed normal to crossings of segments, trying to untie the knot. This is resisted by the excluded volume interactions. As we use spheres to model our chain, the excluded volume interactions have a roughness length scale, which is proportional to the diameter of the beads, b. This roughness couples with the applied tension to trap segments at crossings in energy wells, $E_{well} \sim Fb$. *We postulate the knot moves by these segments hopping out of these wells due to thermal fluctuations.* Given this precept, the rate at which a knot can randomly take a step left or right is $r \sim \exp(-E_{well}/kT)$, where the attempt frequency which is independent of applied force. As the diffusivity is proportional to this rate, we recover $D_{knot} \sim \exp(-E_{well}/(kT))$, explaining the exponential slowdown in the jammed regime.

As we have now presented the first systematic investigation of the jamming of tensioned knots, we turn to use this knowledge in a technological application, namely, translocation of knotted polymers through nanopores. The jamming of knotted DNA translocating through a pore was first phenomenologically observed in the simulations of Rosa and coworkers [66]. We simulated translocation of our flexible polymer through a tight pore (in an otherwise infitie wall) with depth $10b$. The beads within the pore experienced a force of f, and so a total force of approximately $10f$ acted on the chains. Excluded volume interactions with the pore geometry were treaded the same as bead-bead interactions.

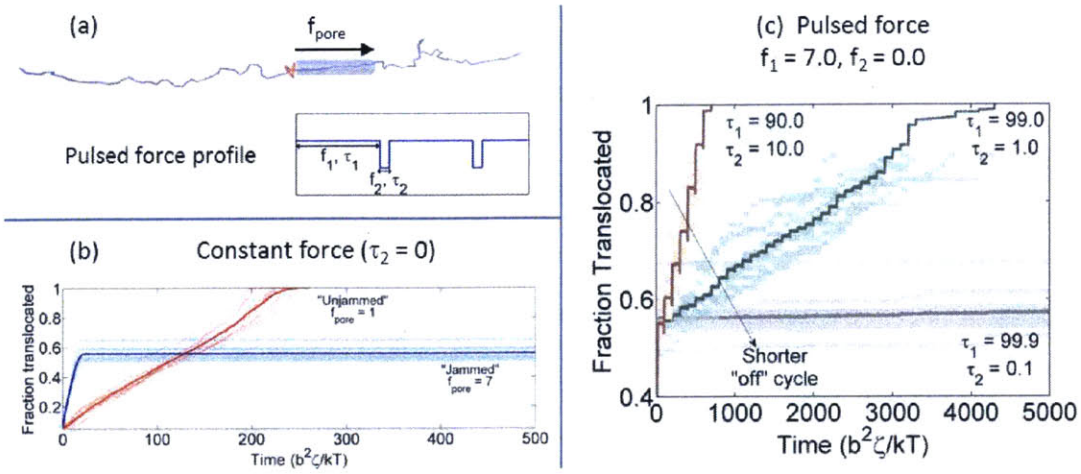

Fig. 8.3: *(a) Schematic of a knotted polymer translocating through a pore. Inset shows the force profile on each bead inside the pore. (b) Polymer translocation for the case of constant force. Translocation is halted by a "jammed" knot at large forces. (c) Polymer translocation with a pulsed force field. By varying the "off" rate, the rate at which polymer moves through the pore can be controlled. For (b) and (c), faint lines are results of 10 replicas. Solid lines represent the ensemble averages.*

We propose that by cycling between two forces, f_1 and f_2, with respective durations, τ_1 and τ_2, jamming of knots at the pore can be used to translocate series of a finite number of segments through the pore, partitioned by jamming events. This procedure would therefore lend control to the rate of polymer translocation through the pore. In Fig. 8.3a, we provide a snapshot of this setup as well as the pulsed force profile previously described. In Fig. 8.3b, we show, similar to reference [66], that sufficient forces (*e.g.* $f = 7$) indeed cause jamming of the knot and the cessation of translocation whereas at low forces (*e.g.* $f = 1$), the knotted polymer continuosly translocates.

In Fig. 8.3c, we show translocation results for a knotted polymer when the force in the nanopore is cycled between an "on" force $f_1 = 7.0$, sufficient to cause jamming, and an "off" force $f_2 = 0$. Fixing the duration of τ_1, the rate of translocation can be controlled by varying τ_2, the time in the "off" cycle. For longer off cycles of $\tau_2 = 10$ and $\tau_2 = 1.0$, translocation proceeds in a step-wise fashion every cycle. Reducing the "off" portion of the cycle sufficiently ($\tau = 0.1$), translocation no longer proceeds in this manner and is essentially always jammed, with rare hopping events that do not occur every cycle.

We can understand these results with the mechanism shown in Fig. 8.4. Upon close inspection of the "off" portion translocation cycles, the swelling and diffusion of the knot can be seen. Upon switching to the "on" portion of the cycle, additional segments are pulled through the knot until it finally tightens and re-jams at the orifice of the poor. This process repeats itself every cycle.

This mechanism breaks down for sufficiently short "off" cycles. As previously discussed, at high applied tensions, the segments in the knot get locked into local energy wells, the mechanism that

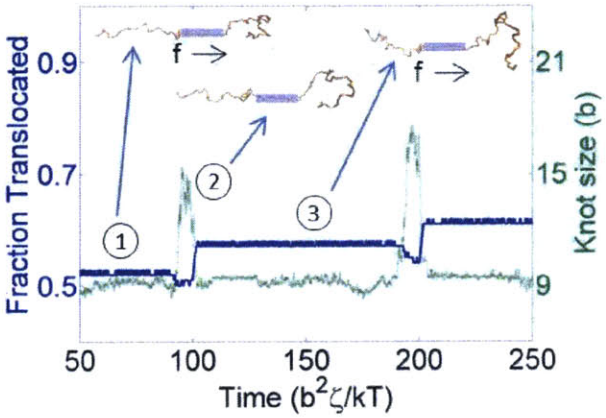

Fig. 8.4: *Mechanism of polymer "racheting." We show polymer translocating through the pore with a pulsed force field with* $f_1 = 7.0$, $f_2 = 0$, $\tau_1 = 90$, $\tau_2 = 10$ *(*cf. *Fig. 8.3). The knot swells and diffuses away from the pore when the force is in the "off" state, allowing the polymer to ratchet through a finite number of segments upon reapplication of the field.*

explains the exponential decrease in diffusivity shown in Fig. 8.2b. When the field is switched off, segements of the knot to diffuse out of these locally locked conformations during τ_2 in order for them to advance upon reapplication of the translocating force. The timescale for this diffusion is the time required for a bead to diffuse over the roughness length scale, previously argued to be proportional to b. Thus, $\tau_{unlock} \sim b^2\zeta/(kT)$, which is the scale already used for simulation time. This argument explains the break between stair-step translocation observed when $\tau_2 = 10 > 1$ and $\tau_2 = 1.0$ and total jamming when $\tau_2 = 0.1 < 1$, providing additional evidence for our mechanistic explanation of the jamming process.

While jammed polymer knots have been seen in previous simulations, this study represents the first mechanistic understanding of this phenomenon. Furthermore, we have shown how our insights can be leveraged into controlling the problem of polymer translocation through a nanopore provided we can introduce a knot into the polymer configuration either through stochastic means (see Chapters 5 and 6) or through a mechanism previously discovered (see Chapter 3).

CHAPTER 9

Outlook

9.1 Conclusions and Impact

The body of work presented in this thesis represents a series of steps forward in understanding the effects that knots can have on the equilibrium and out of equilibrium properties of DNA and polymers more broadly. If the current trend in this area of research continues, the results in this thesis should develop a substantial position in the body of scientific literature.

The major experimental result of this thesis concerns the first single molecule experiments revealing how knotted topologies can impact the process of stretching. Importantly, we verified experimentally two phenomena suggested by simulations and theory: knots can lead to metastable arrested states and knots can slow driven dynamical processes. Additionally, we have explained a mechanism for making these DNA knots, establishing an experimental platform for creating and interrogating single DNA molecules with knots in microfluidic devices. We expect this platform to be used to great success in our group as well as others to perform further experiments on knotted DNA.

The major theoretical result of this thesis involves understanding the statistical physics that govern the size of randomly occurring knots in semiflexible polymers. The original theoretical framework of this work was developed by Grosberg and Rabin to describe the size of knots in thin semiflexible polymers in bulk; we modified this theory to account for excluded volume interactions and showed agreement between the theory and results of simulations. We then extended the theory to a previously uninvestigated case, the size of knots on polymers confined in channels. We showed

that the confinement of the channel can remarkably change the size of the knot in a non-monotonic fashion, yet these strange results of simulations agreed with a simple modification of the theory for molecules in bulk. The first of these studies represents a significant leap forward in understanding the physical forces and energies that govern knots on polymer molecules, establishing a simple theoretical framework. The second study shows how this framework can be extended to other systems with surprising results. Due to its simplicity and the high level of agreement we have shown with simulations, we expect this theoretical model to be extended by others in many directions, such as polymers in slit-like confinement and polymers with weak attractive interactions. DNA knots have been recently measured with solid state nanopores [78], and we expect these and future experiments to quickly proceed and allow for the testing of our results.

Finally, this thesis demonstrates an advance in understanding the motion of knots on polymers acted upon by external forces. We demonstrated extensional flows can cause knots to be swept off a molecule, and these results have implications in preconditioning DNA to unknotted states for applications such as direct linear analysis or nanopore sequencing of genomes. In this study, we also showed the family of torus knots can couple rotation and translation in a way that other knot families cannot, and this explains recent results in the literature where torus knots move more easily along driven chains or eject more quickly from viral capsids. We suspect these results may inspire future work on how "topological families," such as twist or pretzel knots, may access different modes of motion, particularly when driven by forces. In our study to understand the jamming of tensioned polymer knots, we have shed light on a topic that is the subject of increasing simulation and experimental interest, describing a simple model combines the roughness in local energy landscape with the applied tension to explain the slowed dynamics. At the present time, it appears feasible to directly test this work by tying knots in DNA with optical tweezers and stretching the strands at high tensions. Furthermore, these results have direct implications to DNA nanopore translocation, and we anticipate others performing experiments involving jamming of long strands containing knots in the near future.

9.2 Future Work

Original research often begins by setting out to answer an unanswered question, succeeding in answering a different question, and raising dozens more questions to be answered along the way. In this section, I will give my perspectives on a few of the questions that this thesis work raises. In particular, I will focus on three extensions of this thesis that I believe our research group is particularly well-positioned to investigate. Two of these leverage the experimental platform used in this thesis for creating and investigating knotted DNA. The third proposes a natural extension of the theory we have developed for the statistical physics of knots and a possible means to compare with experiment. Given the abundance of theoretical and simulation results concerning polymer knots and the relative paucity of experimental work on this topic, experiments on single knotted molecules are inherently valuable, and I hope they proceed apace.

9.2.1 Relaxation of Stretched Knotted DNA

Direct observation of the ways knots can change the dynamics of single DNA molecules has been of obvious interest in this thesis. Particularly, the experiments in Chapter 4 provided the first direct experimental evidence of topological friction within a single molecule. Fig. 4.9 showed the

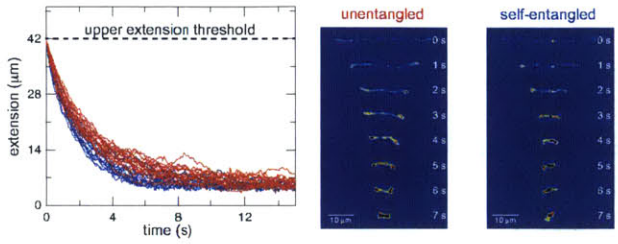

Fig. 9.1: *(left) Trajectories of extension vs. time for ensembles of initially unentangled and self-entangled molecules. The trajectories were all initialized when the extension of the molecule first fell below 42 μm. Representative snapshots of a unknotted and knotted DNA molecule relaxing from an initially stretched configuration.*

relaxation process from stretch of a single knotted compared to a single DNA molecule without a knot. We noticed that molecules relaxed more quickly with interior entanglements and rationalized this due to electrostatic interactions which caused knots to swell and thus the chain ends to retreat more quickly.

Here, the results of a series of DNA relaxation experiments are shown in Fig. 9.1. All molecules were prepared with a similar experimental protocol as in Chapter 4. The molecules were initially stretched at $De \approx 5$. The trajectories are plotted with time initialized at 0 seconds for each trajectory at the first frame where the total extension of that molecule passed below an extension threshold, 42 μm. The molecules with internal knots quite clearly relax more quickly than those without such entanglements. Moreover, this difference in relaxation rates between the two populations is large enough to likely produce a strong enough experimental signal for a more fine-grained analysis of this process.

Many forward-looking ideas spring from these results. The swelling of these relaxing knots may prove another way to directly measure the excluded volume interactions of DNA, a topic of recent interest in nanofluidic devices [199]. While these experiments lack precise topological resolution (one does not know exactly what type of knot resides on the chain), the total contour entrapped in the knot can likely be inferred through careful processing of the experimental images. Thus, the evolution of this quantity could provide more direct experimental investigation into dynamics of knots in single molecules. Both of these directions should be strongly influenced by excluded volume interactions, which can be easily changed by manipulating the ionic strength of the buffer solution.

9.2.2 Dynamics of Knotted DNA in Confinement

Our research group has long been at the forefront of examining the static and dynamic properties of DNA in micro and nanofluidic confining geometries. In particular, many experiments in our group have been performed in nanoscopic slits [19, 20, 26, 27, 123, 137, 200–202]. Here, we show some preliminary results of the dynamics of knotted DNA these nanoscopic slits and discuss ideas for the extension and completion of such a project.

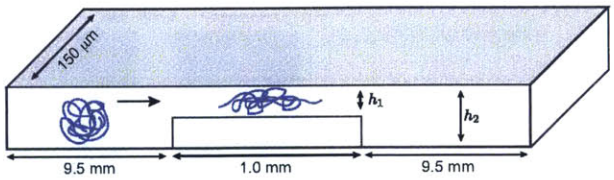

Fig. 9.2: *Schematic of DNA knotted DNA in nanoslit devices. The devices consists of two symmetric slits of height h_2, each connected to a distant fluid resevoir (not shown), and separated by a more constricted nanoslit of height h_1.*

In our group, we have access to a series of nanoslits of varying dimensions. A schematic of the slits used in the forthcoming results appears in Fig. 9.2. These slits offer two different levels of confinement, indicated by h_1 and h_2, referred to as a nano and microslit region, respectively.

DNA molecules were stained and prepared in electrophoresis buffers consistent with the experimental protocol in Chapter 4, with the notable omission of PVP from the electrophoresis buffer. Molecules were introduced into the device via the application of electric potentials across the slits. Knotted molecules were initially compressed with electric fields of ~200 V/cm in the microslit region of the device and then quickly translated into the nanoslit region, indicated by the arrow in Fig. 9.2.

In Figure 9.3, traces of the radius of gyration, R_g, of one unknotted and one knotted molecule are plotted. The knotted molecule exhibits a substantial decrease in mean size (dashed lines) compared to the unknotted molecule. Furthermore, the the magnitude of fluctuations of the R_g is significantly smaller for the knotted molecule compared to the unknotted molecule. The presence of a knotted core, a persistent region of unusually high contour density, can be seen in the snapshots of the knotted molecule. The snapshots of an unknotted molecule, by comparison, rarely have conformations containing a similarly highly localized density of contour; the conformations of increased contour density are always transient in the unknotted case.

The observations of a reduction in molecular size and fluctuations due to knots can be explained as direct consequences of the storage of a substantial amount of the contour of the molecule within the knotted core. The knotted core traps contour, increasing the local segment density due to topological interactions and preventing the molecule from extending as much as it otherwise would. Stored contour similarly reduces the contour available to the "free ends" (a concept discussed in Chapter 4), which are primarily responsible for the fluctuations in the size of the molecule.

The knotted state appears stable (no evidence of unknotting) over the duration of the trajectory and snapshots presented in Figure 9.3. Knots created at similar conditions in T4 DNA were shown to begin to unravel in 20 s on average[73] in the absence of confining geometries. Previous work has shown the relaxation time of λ-DNA to increase by a factor of roughly 5 when a molecule is confined in a 160 nm slit compared to the same conditions without confinement[27]. The trajectory presented here shows no signs of unknotting in 80 seconds. Other movies of knotted molecules (not shown) show stable knots consistently for 200 seconds or more with few unknotting events observed. These observations raise a more subtle point that may prove interesting for future investigation: the dynamics of unknotting in confinement are possibly fundamentally different than unknotting in bulk. This phenomenon (if indeed substantiated in future work) could be qualitatively explained as

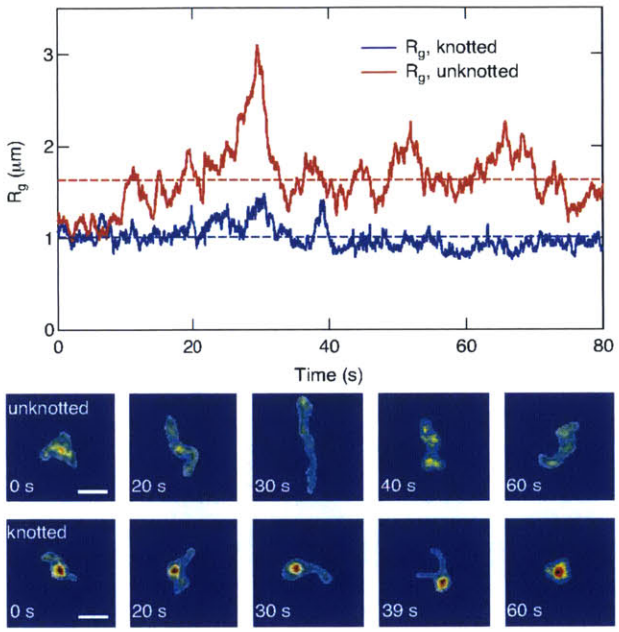

Fig. 9.3: *Top: The radius of gyration (R_g) vs. time plotted for an unentangled molecule and a knotted molecule confined in a 160 nm nanoslit. The dashed lines show the respective mean R_g for each trajectory. Bottom: Selected snapshots for each molecule. The scale bar is 3 μm. The scales for intensity of each molecule are the same.*

follows. In bulk, collective motions of the chain ends and the knotted region allow relatively rapid knotting whereas in a confining geometry, the ends are more extended from this knotted core. As a result, a self-reptation of the knotted core to an end of the molecule, likely a slow process for the complex knots seen in these experiments, must account for a greater fraction of unknotting events.

Looking forward, these results raise many interesting questions for investigation. How do the knots change the diffusivity of the molecule? While the size of the knotted molecule R_g is reduced, more segments are in close proximity to the slit walls, introducing additional hydrodynamic drag. Is the mechanism of unknotting different in slits compared to the unconfined case? These questions and more should prove to be fertile grounds for future research.

Appendix A

Tips and Tricks

This Appendix details some useful procedures and routines that were used in generating the results found in the previous chapters.

A.1 Alexander Polynomial Calculation

Here, the combinatoric approach to calculating the Alexander polynomial ($\Delta_K(t)$ of a knot, K) is outlined. As this section was originally motivated by DNA knotting, it follows closely the procedure laid out in Reference [186], although it can be found in many introductory texts on knot theory. A more recent description of this approach is found in reference [132]. The reader who has examined both references [186] and [132] will notice a slight difference in their construction of the Alexander matrix; both constructions produce the identical final results.

Prior to beginning this calculation, for an open chain, one must employ a closure scheme, such as those detailed in Section 2.2. After the appropriate scheme is chosen and applied, we have the coordinates of a closed loop. In order to proceed with our calculation, we must take a chain with 3D coordinates project them down to a 2D plane. For a chain with a knot, there will be at least three crossings in this 2D plane and often many more, and we will call the number of crossings in the chosen projection N_{cross}. (Obviously, if there are 2 or fewer crossings, the chain is the unknot and this procedure may be skipped entirely.) Looking locally, each crossing appears as the intersection of two lines. The out-of-plane dimension can be used to determine which line passes above and

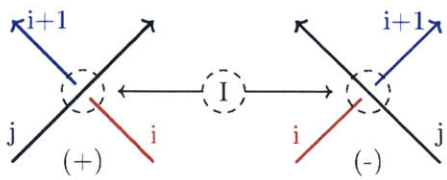

Fig. A.1: *Illustrations of positive (+) and negative (-) crossings. Here, j, i, and $i+1$ correspond to the generators involved in the Ith crossing.*

which line passes below at the crossing. By identifying the undergoing and over going lines at each crossing, we will have constructed the *2D knot diagram.*

Next, we start at an arbitrary point on the chain and begin moving along chain indices in either direction. As we approach a crossing, if the segment we are following is the over going segment, we proceed. If not, we label the crossing as 1. We repeat this procedure until we have labeled all N_{cross} crossings. From this labeling, we can similarly break the ring into N_{cross} continuous curves, called generators, that are "broken" by under crossings. Specifically, the generator $i+1$ is the segment of the curve that lies between crossings I and $I+1$. Note that due to the circular nature of our system, indices are only defined between 1 and N_{cross}. Thus, the closure of indices via $N_{cross}+1 \rightarrow 1$ and $0 \rightarrow N_{cross}$ applies.

Now, we have broken up the chain into a series of crossings and generators. For a given crossing I, we know it separates generators i and $i+1$, and there is some other generator, say j, that is the over going generator of this crossing, shown in Figure A.1. The orientations of these generators can be used to assign one of two types to the given crossing: (+) or (-). These types of crossings are also depicted in A.1. One can imagine many ways, such as the sign of the cross product of the two lines, to differentiate between these two types. We can now go through the diagram and assign a type of (+) or (-) to each crossing.

With all of this bookkeeping complete, we are ready to construct the Alexander matrix, $A(t)$. The ith row of contains entries corresponding to the Ith crossing and its over going generator, j. All entries are zero, except:

---

if $i = j$ or $i = j+1$ **then**
  $A_{ii} = -1$ and $A_{ii+1} = 1$
else
  if the crossing type is (+) **then**
    $A_{ii} = 1$, $A_{ii+1} = -t$, and $A_{ij} = t-1$
  else (the crossing type is (-))
    $A_{ii} = -t$, $A_{ii+1} = 1$, and $A_{ij} = t-1$
  end if
end if

Upon assembling $A(t)$, we can compute any $(N_{cross}-1)$th order minor, *e.g.* by taking the determinant of $A(1:N_{cross}, 1:N_{cross})$. The determinant of this minor is $\pm t^m \Delta_K(t)$, *i.e.* the desired Alexander polynomial multiplied through by $\pm t^m$, where m is a whole number.

To obtain $\Delta_K(t)$, one must factor out t^m. This can be done by multiplying through by t^{-m} where m is chosen such that $t^{-m}\det A(1 : N_{cross}, 1 : N_{cross})$ produces a polynomial where the lowest degree of any term is zero, *i.e.* it can be written in the form $t^{-m}\det A(1 : N_{cross}, 1 : N_{cross}) = c_n t^n + cn - 1t^{n-1} + ... + c_1 t + c_0$ with $c_0 \neq 0$. Finally, multiplication through by ± 1 to produce a positive constant term removes all remaining ambiguity and yields the desired Alexander polynomial, $\Delta_K(t)$.

In practice, one usually implements these schemes on computers. Here, taking *algebraic* determinants is computationally expensive. Depending on the application, evaluating $\Delta_K(t)$ at a few carefully chosen values of t reveals essentially the same information as evaluating the algebraic form of the Alexander polynomial, especially for "simple" knots with fewer than 10 crossings. For the simplest applications, such as finding the boundaries of a knot where the topology is known *a priori*, evaluating $\Delta_K(t = -1)$ is often sufficient. The particular case of $\Delta_K(t = -1)$, note that for a given crossing, the elements of A are identical for both (+) and (-) types.

The Alexander polynomial has a number of useful properties, a few which have had particular relevance to this thesis are:

- For all knots, K, $\Delta_K(t = 1) = \pm 1$. (Note with the convention described above, the $\pm$ is removed.) This property can be used to easily screen for errors in calculating the $\Delta_K(t)$.
- For the unknot, $\Delta_K(t) = 1$. It follows that $\Delta_K(t) \neq 1$ indicates a non-trivial topology.
- For some non-trivial knots, K, $\Delta_K(t) = 1$. Therefore, $\Delta_K(t) = 1$ is a necessary but not a sufficient for identifying the unknot. The simplest of these knots K have 11 crossings (11n42 and 11n34). Therefore, $\Delta_K(t) = 1$ and K presented with 10 or fewer crossings is sufficient identify the unknot.
- For a composite knot $K\#K'$, $\Delta_{K\#K'}(t) = \Delta_K(t)\Delta_{K'}(t)$.

A.2 Generating Knotted Topologies

Generating chain configurations for knots of desired topologies has been an essential task for the simulation work in Chapters 7 and 8 of this thesis. Depending on the application, two fundamentally different approaches have been used in this work. The two approaches employed, referred to hereafter as *biased* and *unbiased knot construction*, are now discussed.

A.2.1 Biased Knot Construction

Biased knot construction consists of selecting an arbitrary space curve that describes a knot of a desired topology, and thus the configuration is biased based on the specific choice of the space curve. A variety of parametric curves for knots exist. We will now discuss a few families of these curves.

Torus knots, written here as $T(p, q)$, can be described by

$$\begin{aligned} x(t) &= (\cos(qt) + 2)\cos(pt) \\ y(t) &= (\cos(qt) + 2)\sin(pt) \\ z(t) &= -\sin(qt), \end{aligned} \tag{A.1}$$

where $0 \leq t \leq 2\pi$ and p and q are relatively prime integers. This family of knots is named because it encompasses all knots which can be drawn on the surface of a torus. This leads to an easy physical interpretation of p and q. $T(p,q)$ will wind p times through the "hole" in the torus and q times around the "perimeter" of the torus.

For a given torus knot, $T(p,q)$, a number of interesting and useful results have been obtained:

- Both $T(p=\pm 1,q)$ and $T(p,q=\pm 1)$ are the unknot[80].
- $T(p,q)$, $T(q,p)$, and $T(-p,-q)$ are all the same knot[80].
- $T(p,q)$ and $T(p,-q)$ are mirror images[80].
- The Alexander Polynomial of $T(p,q)$ is $\Delta_{T(p,q)}(t) = \frac{(t^{pq}-1)(t-1)}{(t^p-1)(t^q-1)}$[203].

The mirror image property is particularly useful for creating torus knots of desired chiralities.

Harmonic knots, $H(a,b,c)$ can be described by

$$\begin{aligned} x(t) &= T_a(t) \\ y(t) &= T_b(t) \\ z(t) &= T_c(t), \end{aligned} \tag{A.2}$$

where $0 \leq t \leq 2\pi$, a, b, and c are integers, and $T_n(t)$ are the Chebyshev polynomials of the first kind. These polynomials satisfy $T_n(\cos(t)) = \cos(nt)$, which is equivalent to the recurrence relation:

$$\begin{aligned} T_0 &= 1 \\ T_1 &= t \\ T_{n+1} &= 2tT_n(t) - T_{n-1}(t) \end{aligned} \tag{A.3}$$

The values of a, b, and c for some Harmonic knots are found in reference [1].

Obviously, myriad parametric equations for knots exist, these are simply two used in this thesis. Other popular curves for knots include Lissajous and Fourier parameterizations. Lissajous knots have been the subject of sustained interest by mathematicians (likely due to the tractability availed by their symmetries), yet interestingly, the *analogies to supercoiled DNA* originally motivated the first studies [204]. Fourier knots are a natural extension of Lissajous knots, and it has been proven that every knot admits some Fourier parameterization [205]. In addition to equations, a useful compilation knot representations on lattices is available in reference [206].

These parameterizations are used in what we have called biased knot construction. We have used biased knot construction when we want to fix the topology of a molecule *a priori* with the downside of starting from a specific initial condition. This technique is employed in Chapters 7 and 8 to generate initial conditions for a wide range of knot complexities. In the applications in these chapters, the chains were then subjected to high tensions and allowed to sample configurations. In this way topology of the knot is preserved yet the arbitrariness of the initial condition is rapidly lost. Biased knot construction may also be used for generating initial conditions for knots in ring polymers, provided these initial conditions can be subsequently by schemes that can prevent chain-crossings such as the BFACF algorithm for lattice polymers [207], careful choice of moves in an off-lattice Monte Carlo simulation, or dynamical simulations which are fine-grained to approximately the length scale of steric excluded volume interactions.

A.2.2 Unbiased Knot Construction

Unbiased knot construction consists of statistically sampling knots of a given topology with preference only to configurations that are topologically equivalent to the desired knot. The Pruned-Enriched Rosenbluth Method (PERM) as well as conventional Monte Carlo schemes were employed in Chapters 5 and 6 are the tools we used to sample knotted configurations in an unbiased way. For linear chains, these approaches have the advantage of sampling all the conformations in an unbiased way. These large unbiased ensembles were then combed using chain closure and knot detection techniques to find the unbiased subset of knots of a specific topology. The benefit of this approach is that it is capable of generating a large statistical ensemble of knotted conformations. The drawbacks, however, are substantial. Since the subset of chain conformations with a specific knot is small relative to the total ensemble of polymer conformations, the majority of computational time is spent generating polymer conformations that are not part of the desired knot ensemble. For instance, this method is inherently biased to sampling simple knots since they are the most abundant in the ensemble of random chain configurations [147]. Thus, creating very complex knots (crossing number ~ 7 or higher) by these methods is difficult.

A.3 Simplifying Knots

Frequently, it is useful to reduce the number of crossings in a projection of a knot. Especially for configurations resulting from unbiased knot construction, the number of crossings in an arbitrary projection of the knot can greatly exceed the minimum number of crossings in the knot. Since calculation of the Alexander polynomial requires the expensive operation of taking the determinant of an N_{cross} by N_{cross} matrix, simplifying the conformation of the knot without changing its topology is often necessary.

We have employed a simple scheme to this end from reference [132]. Consider successive indices along the chain $A = B - 1 = C - 2$. If the planar region $\triangle ABC$ does not intersect any segments of the chain, index B may be removed and A connected directly to C without generating a self-crossing of the chain. Applying this procedure iteratively over the closed chain will not chain the topology of the chain but will often dramatically reduce the number of segments and crossings in a chosen projection. Note, the length of the chain is not conserved by this procedure, so it should only be used for simplifying a conformation for topological calculation.

This procedure can also be operated in reverse; if A and C are successive indices of a chain, an additional chain index B can be inserted between them without changing the topology of the chain provided that $\triangle ABC$ does not intersect any segments of the chain. For the purposes of this thesis, increasing the apparent complexity of a knot of fixed topology in this manner was of no use.

A.4 Finding Knot Boundaries

Finding the boundaries of knots has been an essential technique for the work in Chapters 5-8 of this thesis. For a chain configuration described by coordinates with indices 1, 2, ..., n_{bead}, we will define this task to be equivalent to the task of finding the indices of the chain $1 \leq k_l < k_r \leq n_{bead}$ such that subchain defined by indices k_l, k_{l+1}, ..., k_{r-1}, k_r has the same topology as the entire chain while the subchains defined by indices k_{l+1}, ..., k_{r-1}, k_r and k_l, k_{l+1}, ..., k_{r-1} both do not have the same topology as the entire chain.

The procedure to finding the "left" (k_l) and "right" (k_r) boundaries of a knot is as follows. One first determines the type of knot the entire chain contains. For chains initialized by biased knot construction, one has *a priori* knowledge of the topology and thus $\Delta_K(t)$. For the statistical ensembles of simple knots of unknown topology generated by unbiased knot construction, a combination of $\Delta_K(t=-1)$, $\Delta_K(t=2)$, and finding an upper bound on the crossing number through knot simplification proved sufficient to discriminate knots.

With this in mind, the full procedure, used in this thesis, for calculating the position of a knot is as follows:

1. If the chain is open, employ a closure scheme on the whole chain.
2. If the topology is unknown, identify the topology.
   - $\Delta_K(t=-1)$, $\Delta_K(t=2)$, and knot simplification.
3. Remove bead indices from the "left" side of the chain
4. Employ a closure scheme on the new subset of indices
5. Calculate the topology of this subset of the chain
6. Repeat steps 3-5 until the index k_l is identified such that the topology, *e.g.* $\Delta_K(t=-1)$, changes from the original topology of the whole chain to a different topology when only k_l is removed. The left bound of the knot is k_l.
7. Repeat steps 3-5 removing indices from the right side of the chain until k_r is similarly identified.

This procedure, while straightforward, is not without problems. First, the determination of k_l and k_r is sensitive to the order in which they are calculated. This is because, in the algorithm shown above, k_r is determined after indices have been deleted from the left end of the chain whereas no indices have been deleted from the right during the calculation of k_l. One approach to reducing the sensitivity is to alternate the side of the chain from which indices are deleted. In Chapters 5 and 6, k_l and k_r were calculated twice for each conformation, once with indices initially removed from the left and then again with indices initially removed from the right. The average of each was then the accepted k_l and k_r.

It should be noted that more rigorous methods of calculating knotted subsections of a chain have been discussed in reference [82]. In particular, the authors discuss two approaches: a bottom up and top down approach. The bottom up approach consists of looking at all subsections of the chain of a given length. This length is increased from the smallest possible length until some subsection of the chain is identified with the same topology as the entire chain. The top down approach consists of looking for the shortest possible subsection of the chain, which cannot be further shortened without deviating from the topology of the entire ring. In this view, the approach employed in this Appendix would be classified as a modified top down search.

Appendix B

Experimental Tips and Tricks

This Appendix describes in greater detail some of the important experimental protocols that were used in generating the results in Chapters 3 and 4.

B.1 General experimental tips

B.1.1 Cutting PDMS devices

- Use a sharp, clean scalpel blade.
- When cutting, attempt to hold the blade orthogonal (not at an angle) to the surface of the PDMS.
- Cut out the molded PDMS by cutting slightly inside perimeter of a molded silicon wafer. Pull off this section of PDMS in one piece and put it, feature side down, on a new, clean petri dish where the cuts for individual devices can be made — don't cut individual devices out on the wafer because it will damage the wafer.
- Mark the underside of the petri dish with dots at the device reservoirs to help guide cutting the reservoirs.
- Cut out reservoirs before cutting around the perimeter of the individual device.

B.1.2 Preventing DNA fragmentation by shear

- When pipetting DNA, always cut off the end of the pipette tip (approximately 3 mm above the end) to increase the diameter of the outlet orifice.
- Discharge DNA solutions from a pipette tip slowly and at a uniform flow rate.
- When mixing DNA solutions, slow stirring of the pipette tip in the receiving container and/or gradual pipetting in/out of the mixing solution can help to disperse DNA molecules. **Never vortex any solution containing DNA.**

B.2 A few protocols

The following protocols have proven useful in performing successful experiments with DNA in microfluidic devices. An excellent resource for additional detailed protocols (*e.g.* staining DNA, attaching quantum dots to DNA, making DNA concatemers, and more) may be found in reference [208].

B.2.1 Molding PDMS

- Vigorously mix Dow-Corning 184 PDMS Sylgard elastomer base and curing agent in a 10:1 weight ratio for over 30 s until air bubbles are incorporated throughout the mixture in a small plastic cup. Typically 80g:8g produces a sufficient volume of PDMS to mold an entire large petri dish. 50g:5g produces enough to mold an area the size of the master silicon wafer.
- Place the cup containing the mixed PDMS into a vacuum chamber for approximately 30 minutes. This step helps remove the air bubbles quickly.
- Once the bubbles have been removed from the mixed PDMS, slowly pour the PDMS onto the desired silicon wafer template. Pour approximately 1 inch above the surface of the wafer; pouring from higher distances can often cause formation of bubbles.
- Place the PDMS coated wafer in an oven set at 65 C overnight to cure.

B.2.2 Cleaning glass coverslips (slides)

- Remove coverslip from box with tweezers. Inspect for obvious dust or defects, and discard the slide and select a new slide if defects are found.
- Holding the slide with tweezers, rinse both sides of the slide with ethanol from a squirt bottle.
- Holding the slide with tweezers, rinse both sides of the slide with filtered water from a squirt bottle.
- Carefully place the slide in a bath of 1M NaOH. Allow the slide to sit in the bath for at least 1 hour. Do not allow the slides to sit for more than several hours; the etching of the slide by the NaOH can make them much more fragile if left too long.
- Carefully remove the slide from the NaOH bath. Rinse the slide with filtered water bottle several times. Notice the water wets the slide differently after etching the surface with NaOH.

- Dry the slide with *clean* gas. (Note: the house compressed air is "dirty" in that it contains some particulates. Typically, ultrapure gas from a cylinder, *e.g.* argon, works well for this.) It often helps to attach a pipette tip to the Tygon tubing attaching to the source of the glas to direct the gas stream. Typically, one wants to hold the slide with tweezers and "push" the liquid tangentially across the surface of the slide to the edge with the stream of gas from the pipette tip. Then, the liquid pools at the edge, it will either shed due to the forces of the gas or evaporation.

- Place a piece of lens paper inside a large petri dish and place the sldie on top. Approximately 6 slides can fit into this setup without overlapping.

- After cleaning all desired slides, seal the edge of the petri dish by stretching a strip of paraffin film around the edge — this procedure aims to prevent airborne particulates from contaminating the glass if not used immediately.

B.3 Calculating buffer ionic strength

The ionic strength of a buffer, I, is a critical experimental parameter in DNA experiments as it can be used to change the level of electrostatic interactions between segments of DNA due to salt screening. Here, we outline the procedure for calculating the ionic strength in the complex experimental buffers employed in our DNA experiments. This procedure was first employed in reference [200], and follows closely with their results.

The ionic strength of a solution of ions i each with concentration c_i and charge z_i is

$$I = \frac{1}{2}\sum_i c_i z_i^2, \tag{B.1}$$

usually expressed in mol/L (M).

Unlike relatively simple ionic buffers (*e.g.* NaCl), in TBE buffer solutions, the equilibria between multiple acid-base dissociations determines the total concentrations of all ions. For an acid A the dissociation equilibrium is described by

$$-\log\left(\frac{a_{\mathrm{H}^+}a_{\mathrm{A}^-}}{a_{\mathrm{HA}}}\right) = \mathrm{pK_a}, \tag{B.2}$$

where a_i is the activity coefficient of species i.

Similarly, the dissociation equilibrium for a base B can be written

$$-\log\left(\frac{a_{\mathrm{B}^+}a_{\mathrm{OH}^-}}{a_{\mathrm{BOH}}}\right) = \mathrm{pK_b}. \tag{B.3}$$

The relevant values for TBE buffers are $\mathrm{pK_a}$ for boric acid (9.24), EDTA (1.99, 2.67, 6.16, 10.26), and BME (9.6) and the $\mathrm{pK_b}$ of Tris base (5.94).

The activity of ion i can be expressed as

$$a_i = \gamma_i \frac{c_i}{c_o}, \tag{B.4}$$

where γ_i is the activity coefficient and $c_o = 1M$ is a reference concentration.

The dissociation reactions are also subject to conservation equations. For example given an acid A which only sheds one hydrogen,

$$[\mathrm{A}]_o = [\mathrm{A}] + [\mathrm{A}^-], \tag{B.5}$$

where $\mathrm{A}]_o$ is the initial added concentration of acid. For reference, 5x TBE has initial concentrations of 0.45 M Tris, 0.45 Boric acid, and 0.01M EDTA.

In salted solutions, the activities of ions increasingly deviate from ideal behavior $\gamma_i = 1$ as the salt concentration increases. These deviations can be modeled in many ways, one of which the Davies equation:

$$\log \gamma_i = -Az_i^2 \left(\frac{\sqrt{I}}{1+\sqrt{I}} - 0.2I \right), \tag{B.6}$$

where $A \approx 0.5085\ \mathrm{M}^{-1/2}$ in aqueous solutions at 25 C and I is in M.

Finally, our aqueous buffers must satisfy

$$\mathrm{pH} + \mathrm{pOH} = -\log(a_{\mathrm{H}^+}) - \log(a_{\mathrm{OH}^-}) - = 14. \tag{B.7}$$

and the condition of charge neutrality

$$\sum_i c_i z_i = 0. \tag{B.8}$$

The above equations can be solved in an iterated scheme until all values of c_i have converged. The value of I and pH follow accordingly.

Appendix C

C.1 Channel and DNA preparation

2 μm tall, 200 μm wide, and 1 mm long straight channels were constructed in polydimethylsiloxane (PDMS, Sylgard 184, Dow Corning) using soft lithography on a silicon master (AZ 5214 image reversal photoresist). The PDMS channels were soaked in 0.5X Tris-Borate-EDTA (TBE, Sigma) buffer at 45 °C overnight to eliminate permeation driven flow through the PDMS[130], rinsed and dried briefly, and sealed to a glass coverslide. The microchannels were flushed for 20 minutes with the experimental buffer via application of electric potentials at the fluid reservoirs before use. The experimental buffer contained 0.5X TBE, 0.05% (wt/vol) 10 kDa polyvinylpyrrolidone (PVP, Polysciences) to dynamically coat the channel walls to suppress electroosmotic flow and prevent interactions between the channel walls and DNA, and 4% (vol.) beta-mercaptoethanol (BME, Sigma) to slow down photobleaching and photo-induced damage to the stained DNA molecules. T4 DNA molecules (165.6 kbp, Nippon gene) were stained (at a DNA concentration of 0.69 μg/ml) with YOYO-1 (Invitrogen) dye at a base pair to dye ratio of 4:1 and allowed to sit at least overnight. DNA samples were diluted 2 to 10-fold immediately before experiments to reach an optimal concentration for observation.

C.2 DNA solution concentration relative to c^*

We estimated the overlap concentration of our DNA solutions using the formula:

$$c^* = \frac{3M_w}{4\pi R_g^3 N_A} \tag{C.1}$$

where M_w is molecular weight, N_A Avogadro's constant, and R_g the radius of gyration. Takings values of $R_g = 1.3$ μm for T4 DNA results in $c^* = 19.4\mu$g/mL. As described above, our solutions were typically at concentrations ranging from $0.07 - 0.35$ μg/mL, which is significantly below c^*.

C.3 DNA conformations under electric fields

Single DNA molecules under uniform electric fields were observed in 2 μm tall straight channels using an Olympus IX 71 microscope with a 60X NA 1.35 oil-immersed objective and an X-Cite 120 light source. A typical experiment involved first applying a weak DC electric field of 5-15 V/cm for 5 minutes to electrophoretically flush out any DNA molecules that previously experienced high electric field and to introduce equilibrium DNA molecules into the microchannel, then switching to the desired field strength DC, waiting for 3 minutes, and starting image acquisition. A Trek (Model 677B) Supply/Amplifier was used as a power supply for application of the electric field. DNA images were collected using an Andor EM-CCD camera and Andor SOLIS X-5220 image software. An exposure time of 0.01 s and a frame rate of 24 frames per second were used for image acquisition. Thus, the conformations of DNA molecules under fast electrophoretic motion (especially at $E > 50$ V/cm for a low viscosity buffer, e.g. 0.5X TBE without added polymers) could not be directly observed. As a result, once a DNA molecule of interest enters the field of view, we switched off the electric field to stop the motion of the molecule, captured the DNA conformation in the first frame where the DNA image is not blurred (within 0.04 s of stopping the electric field), and turned on the field again. The electric field was kept off for less than 0.5 s during this process, ensuring that no significant expansion of the DNA has occurred before the field was switched back on.

C.4 Physical Properties of dextran, PVP, and HPC solutions

For dextran polymers, radii of gyration R_g were calculated with $R_g = 0.066 \times M_w^{0.43}$ with M_w in g/mol and R_g in nm[209]. The overlap concentrations c^* of dextran polymers were calculated from

$$c^* = \frac{3M_w}{4\pi R_g^3 N_A}, \tag{C.2}$$

where M_w and R_g are molecular weight and radius of gyration of the polymer, respectively. N_A is the Avogadro constant. The overlap concentrations c^* of linear polymers, hydroxypropyl cellulose (HPC) and polyvinylpyrrolidone (PVP), were estimated from their intrinsic viscosity $[\eta]$[210]

$$c^* = \frac{1.5}{[\eta]}. \tag{C.3}$$

Table C.1: *Physical properties of dextran polymers*

Property \ M_w (g/mol)	5000	80,000	410,000	2,000,000
c^* (mg/ml)	117	52.2	32.5	20.5
R_g (nm)	2.6	8.5	17.1	33.8
R_h (nm)	1.7	5.4	11.0	21.7
D (μm^2/s)	131.0	40.1	19.9	10.1
l_p (nm)[211]	0.4	0.4	0.4	0.4

Table C.2: *Physical properties of PVP polymers*

Property \ M_w (g/mol)	10,000[212]	1,000,000[213]
$[\eta]$	6.3	142
c^* (mg/ml)	238	10.6
R_g (nm)	2.6	33.4
R_h (nm)	1.6	21.4
D (μm^2/s)	133.4	10.2
l_p (nm)[214]	1.3	1.3

From c^*, the radius of gyration for these polymers was estimated via

$$R_g = \left(\frac{3M_w}{4\pi N_A c^*}\right)^{1/3}. \tag{C.4}$$

For all polymers, hydrodynamic radii were estimated using the long chain limit for linear chains[107]: $R_h = R_g/1.56$. We note that while dextran polymers are not linear, the scaling coefficient approaches 2 for randomly branched polymers with high polydispersity in good solvents. Using a value of 2 to estimate R_h of dextran polymers does not change any of the calculated Péclet numbers enough to alter the conclusions reached. Finally, the diffusivity of a polymer was estimated via the Stokes-Einstein relation

$$D = \frac{k_B T}{6\pi\eta R_h}, \tag{C.5}$$

where k_B is the Boltzmann constant, $T = 298.15$ K is the experimental temperature and $\eta = 1$ cP is the background viscosity of the buffer with no added polymers. The estimated physical properties for dextran, PVP, and HPC polymers are given in Tables S1, S2, and S3, respectively.

C.5 Electrophoretic Mobility of DNA

Movies of DNA electrophoresis were collected with the imaging system described in Section III to obtain the electrophoretic mobility of DNA molecules as a function of electric field strength. The electrophoretic mobility is defined as $\mu_M = \frac{v}{E}$, where v is the migrating velocity of DNA and E is the electric field strength.

Table C.3: *Physical properties of HPC polymers*

Property \ M_w (g/mol)	100,000[215]	370,000[215]	1,000,000[210, 215]
c^* (mg/ml)	23	8.0	3.7
R_g (nm)	12.0	26.4	47.5
R_h (nm)	7.7	16.9	30.4
D (μm^2/s)	28.4	12.9	7.2
l_p (nm)[216]	10	10	10

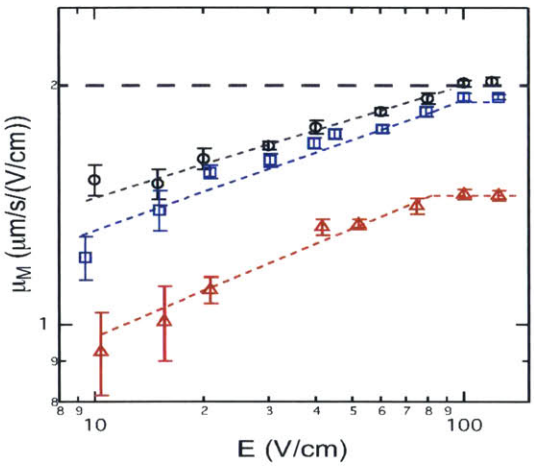

Fig. C.1: *Electrophoretic mobility of T4 DNA in dextran solutions (Φ = 0.6, in 0.5X TBE) as a function of electric field strength E. The dotted straight lines are added to guide the eye. The horizontal dashed line indicates the DNA mobility in 0.5X TBE. If not visible, the error bar (standard error) is smaller than the symbol size.*

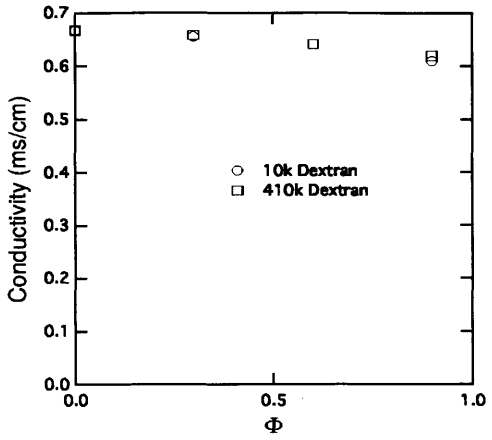

Fig. C.2: *Conductivity of 10k and 410k dextran solutions, 0.5X TBE, at* $T = 22\,^{\circ}\mathrm{C}$.

C.6 Buffer conductivity

The mobility of salt ions is scaled as $\mu_s \sim \Lambda$, where Λ is the conductivity of salt solutions[217]. The conductivity of the buffer was measured using a conductivity meter (Lutron CD-4303). A standard conductivity calibration solution 1413 μs/cm (HANNA instruments) was used for instrument calibration before sample measurements. All the conductivity measurements were carried out at room temperature (22 °C). The conductivity of both dextran solutions of 10k and 410k slightly decrease ($< 10\%$) with increasing volume fraction of dextran polymers.

C.7 Size of T4 DNA at equilibrium

Table C.4: *Radius of gyration of T4 molecules at equilibrium (0 V/cm) in dextran solutions,* M_w *=5k, 80k, and 410k,* $\Phi = 0.6$ *in 0.5X TBE.*

M_w (g/mol)	5k	80k	410k
$\langle R_{g,\mathrm{eq}} \rangle$ μm (mg/ml)	1.29 ± 0.03	1.28 ± 0.03	1.25 ± 0.02

C.8 DNA volume fluctuations under electric fields

To calculate volume fluctuations, the 3D DNA molecule is described by an ellipsoid where R_M and R_m are its major and minor principle axes in the observable 2D-plane and z is the out of plane direction. With $V_M \sim R_M R_m R_z$, using chain rule we get $\delta V_M \sim \delta R_M R_m R_z + \delta R_m R_M R_z + \delta R_z R_M R_m$. Since the main drag is along the travelling direction of the DNA, i.e. the direction of the electric field, R_M exhibits the strongest fluctuation among the three directions. From symmetry,

one has $R_z \sim R_m$, both showing much smaller magnitude of fluctuation. Therefore, we get $\delta V_M \sim \delta R_M R_m^2 + 2\delta R_m R_M R_m$.

C.9 Rescaling the data in Fig. 1c & d using $\mu_M(E)$

The collapse of the data according to the scaling arguments in Section V of the manuscript depends on the electrophoretic mobility of the macroion (DNA), μ_M. This quantity depends weakly on the electrophoretic velocity of DNA, and thus the applied electric field. In Fig. 4 of the manuscript, a characteristic mobility was used to collapse the data for all field strengths. In Figure C.3, the experimental data is collapsed in the same way as Fig. 4 of the manuscript, but $\mu_M(E)$ is used so that the field-strength dependence is included. The quality of the data collapse is essentially indistinguishable from that in Fig. 4 of the manuscript.

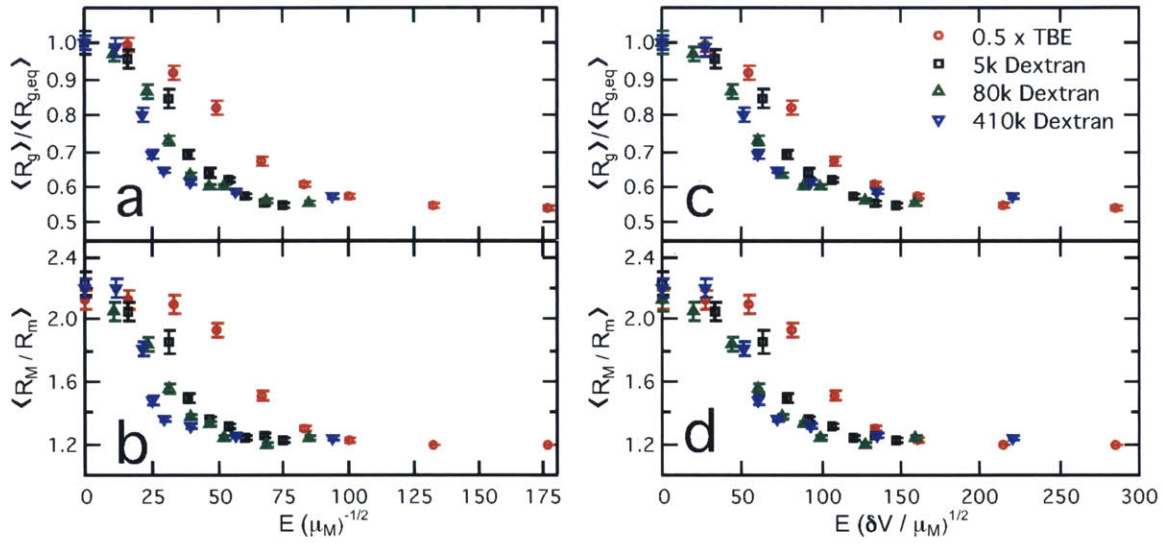

Fig. C.3: *Replotting the data in Fig. 1c & d versus (a)* $E\mu_M^{-1/2}$, *units:* $(\mathrm{V/cm})\left(\frac{1}{(\mu\mathrm{m/s})/(\mathrm{V/cm})}\right)^{1/2}$; *(b)* $E\left(\frac{\delta \mathrm{V}}{\mu_M}\right)^{1/2}$, *units:* $(\mathrm{V/cm})\left(\frac{\mu\mathrm{m}^3}{(\mu\mathrm{m/s})/(\mathrm{V/cm})}\right)^{1/2}$ *using* μ_M *as a function of* E *and* δV *at 15 V/cm results in a similar extent of data collapse as Fig. 5.*

C.10 Compression T4 DNA in dextran solutions with various volume fractions

All of the data shown in the manuscript was acquired at volume fractions of solute polymers of $\Phi = 0.6$. In Figure C.4, an enhancement of compression occurs for $M_w = 410$k dextran solute polymers at $\Phi = 0.3$, 0.6, and 0.9 is shown. In this range of volume fractions, increasing volume fraction increases the compressive effect.

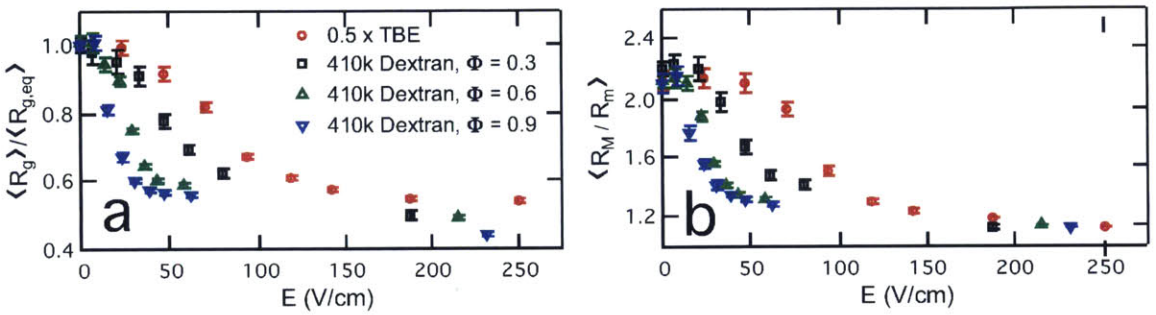

Fig. C.4: *Conformation of T4 DNA in dextran solutions (M_w = 410k, 0.5X TBE) with $\Phi = 0.3$, 0.6, and 0.9.*

Appendix D

D.1 Channel Schematic

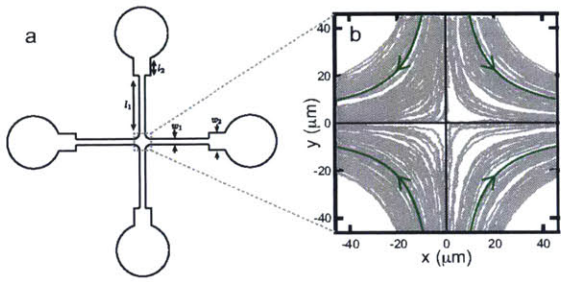

Fig. D.1: *(a) 2-D illustration of the cross-slot channel geometry with the following dimensions:* $l_1 = 3000$ μm, $l_2 = 1000$ μm, $l_H = 100$ μm, $w_1 = 40$ μm, *and* $w_2 = 100$ μm. *DC power supplies are used to establish electric potentials* (Φ_+) *at the left and right reservoirs and ground* ($\Phi_\circ$) *the top and bottom reservoirs, generating a planar elongational electric field at the center of the device. (b) Streamlines (gray) of* λ*-DNA electrophoresing through the center of the device with 30 V applied potentials. The direction and shape* ($y \propto 1/x$) *of the field is indicated by the field lines (green).*

D.2 Relaxation Time of DNA

The relaxation time of DNA, λ, is affected by the contour length length of the molecule as well as several experimental variables such as the slit height [27], buffer ionic strength [200], and buffer viscosity, to name a few. Accordingly, the calculated value of the relaxation time is specific to the experimental setup.

We calculated the rotational relaxation time of T4 DNA by fitting a single exponential decay to the angular autocorrelation function, $C_r(\delta t)$, of an ensemble of molecules, seen in Figure D.2a. This procedure has been frequently employed in past work from this research group [27, 137, 200, 201].

For a given snapshot of a molecule, $\theta(t)$ is defined as the angle between the principle eigenvector of the radius of gyration tensor and the x-axis. This signal is used to construct an autocorrleation function, $C_r(\delta t)$, by

$$C_r(\delta t) = \frac{\langle(\theta(t) - \theta_{avg})(\theta(t+\delta t) - \theta_{avg})\rangle}{\langle(\theta(t) - \theta_{avg})(\theta(t) - \theta_{avg})\rangle} = \frac{\langle\theta(t)\theta(t+\delta t)\rangle}{\langle\theta(t)^2\rangle} \tag{D.1}$$

with $\langle\rangle$ denoting both temporal and ensemble averages and $\theta_{avg} = 0$. The fitting region is $0.3 < C_r(\delta t) <$statistical noise. The upper bound for the fitting region acts to exclude the influence of higher order decay modes on the fit for the longest relaxation time. The lower bound, the statistical noise, is defined as $1/n^{1/2}$ where n is the total number of statistical samples in the data set. Here, $n \approx \frac{t_{\text{total}}}{3\tau}$ where t_{total} is the total amount of video time analyzed for all molecules.

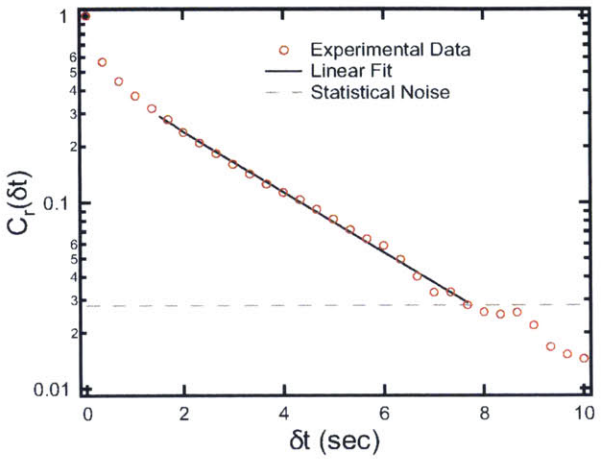

Fig. D.2: *Measurement of the rotational relaxation time of T4 DNA in the buffer and channel described in the manuscript. In the fitting region, the slope of the exponential yields a relaxation time of* $\lambda = 2.6\ s$.

D.3 Strain Rate Calibration

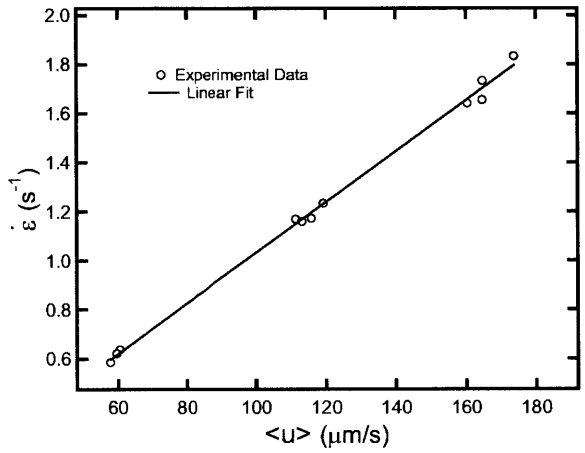

Fig. D.3: *Strain rate calibrated against the average electrophoretic velocity in the cross-slot channel arms. The linear best fit to the experimental data is given by* $\dot{\varepsilon} = (0.0103 \pm 0.002)\mu m^{-1}\langle u \rangle - (0.0016 \pm 0.02)s^{-1}$.

D.4 Effect of Extension Thresholds

Upper and lower extension thresholds were required in order to decompose the trajectories of unentangled and self-entangled molecules into their composite stages: arrested, stretching, and extended. In Figure D.4, we show the effect that changing these cutoffs has on the results shown in Figure 7b.

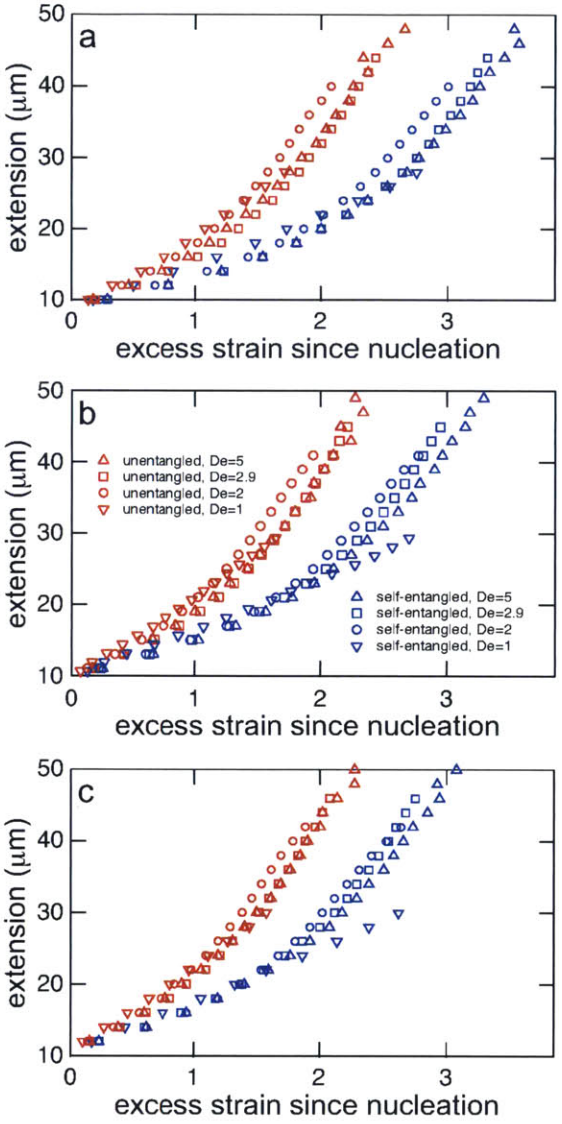

Fig. D.4: *Effect of varying upper and lower extension thresholds on mean stretching curves. The binning and averaging conditions are identical to those employed in Figure 7 of the main text. (a) The mean stretching curves with a lower extension threshold of 9 μm and upper extension thresholds of 29 μm, 41 μm, 45 μm, and 49 μm. (b) The mean stretching curves with a lower extension threshold of 10 μm and upper extension thresholds of 30 μm, 42 μm, 46 μm, and 50 μm. This is simply a copy of Figure 7b. (c) The mean stretching curves with a lower extension threshold of 9 μm and upper extension thresholds of 31 μm, 43 μm, 47 μm, and 51 μm.*

Appendix E

E.1 Bond lengths of the polymer model

In our simulation model, the continuous semiflexible chain is discretized as a string of beads with bond length of L_b. To examine whether such discretization affects the knotting property, we simulate a wormlike chain with $L = 400L_p$ using different bond lengths. Figure E.1 shows the distribution of knot size for different bond lengths. The four curves agree with each other, and the peak locations are identical. These findings strongly suggest that discretization errors do not affect the knotting property significantly.

E.2 The effect of contour length on the metastable knot size

Figure E.2a and E.2b show the rescaled probability of forming a trefoil knot for different contour lengths in the absence and presence of excluded volume interactions. Three contour lengths are used: L/L_p=200, 400, and 800. We rescale the curves for L/L_p=200 and L/L_p=800 such that the peak probabilities of these two curves match the peak probability of the curve for L/L_p=400. After rescaling, we can see that the three curves agree with each other for small knot sizes in both figures. Three curves slowly diverge for larger knot sizes because the longer chain has a longer tail in the distribution of knot sizes. In the current study, we focus on the peak location, *i.e.* the most probable knot size, which is insensitive to the contour length.

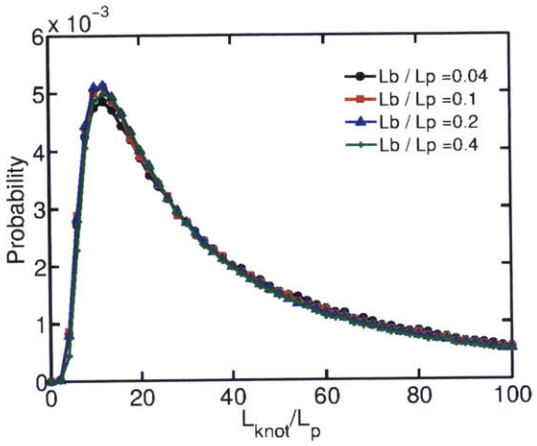

Fig. E.1: *Probability of trefoil knot as a function of rescaled knot size. Different curves correspond to different bond lengths used in simulations.*

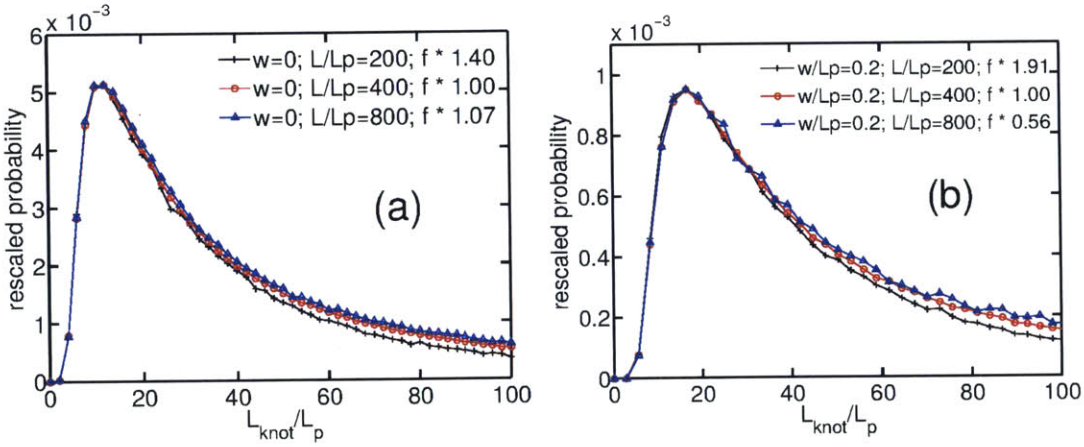

Fig. E.2: *Rescaled probability of forming a trefoil knot for different contour lengths in the absence (a) and presence (b) of excluded volume interaction. The curves for L/L_p=200 and L/L_p=800 are rescaled such that the peak probabilities of these two curves match the peak probability of the curve for L/L_p=400.*

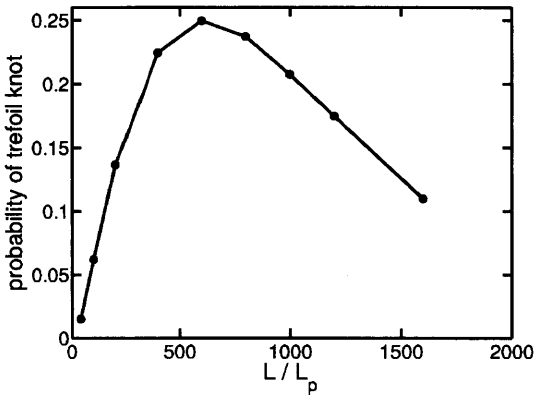

Fig. E.3: *The total probability of forming a trefoil knot as a function of the contour length in the absence of excluded volume interactions.*

The total probability of forming a trefoil knot does not always increase with the contour length. As shown by Figure E.3, the total probability of trefoil knot of a wormlike chain is maximized when the contour length is about 600 L_p. Further increasing the contour length from $600L_p$ reduces the probability of forming a trefoil knot because the chain tends to form more complex knots rather than the simple trefoil knot.

E.3 Positions of knots along chains

To examine how knots are distributed along a chain, we plot the histogram for the location x_{knot} of trefoil knot along a wormlike chain with $L = 400L_p$ (Figure E.4). When we analyze trefoil knots of all sizes (black line), the histogram drops at both sizes because the large knot center is impossible to locate at the ends of chain. For example, the distance from the center of a knot of size $200L_p$ to both ends must be larger than about $100L_p$. In the middle part of the histogram, the curve is nearly flat, which indicates the knot is nearly uniformly distributed along the chain.

When we analyze the trefoil knots with size $L_{knot} \leq 100L_p$, the histogram exhibits two humps with a distance $50L_p$ to ends. Knots appear less frequently at both ends for $x_{knot} \leq 50$ and $x_{knot} \geq 350$ due to the reason discussed in the previous paragraph. Looking into the histogram in the range with $50L_p \leq x_{knot} \leq 350L_p$ we can see that the knot is slightly more likely to exist at the ends. This increase may be caused by the larger conformational entropy of segments at both ends of an open chain.

E.4 The potential of mean force as a function of knot size

Figure E.5 shows the potential of mean force as a function of knot size for different chain widths with a wider range than Figure E.3. The potential barriers keep increasing as the knot size increase. For all of these chain widths, escaping the potential well costs more than 2 k_BT. It means that tight knots on chains with $L/L_p = 400$ are thermally stable when $w/L_p \leq 0.5$.

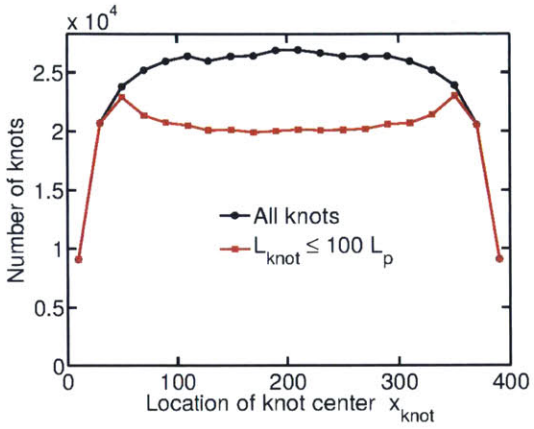

Fig. E.4: *Histogram for the location of the knot center along a wormlike chain with L=400L_p.*

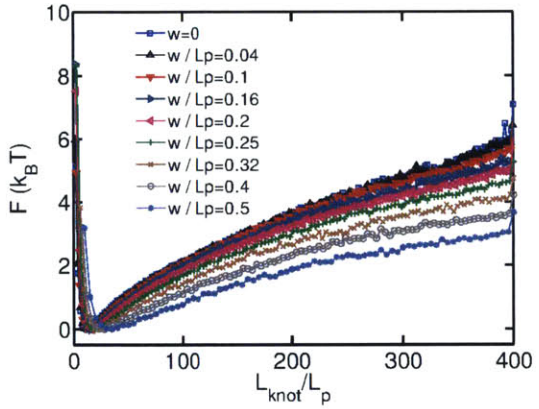

Fig. E.5: *Potential of mean force as a function of knot size for different chain widths.*

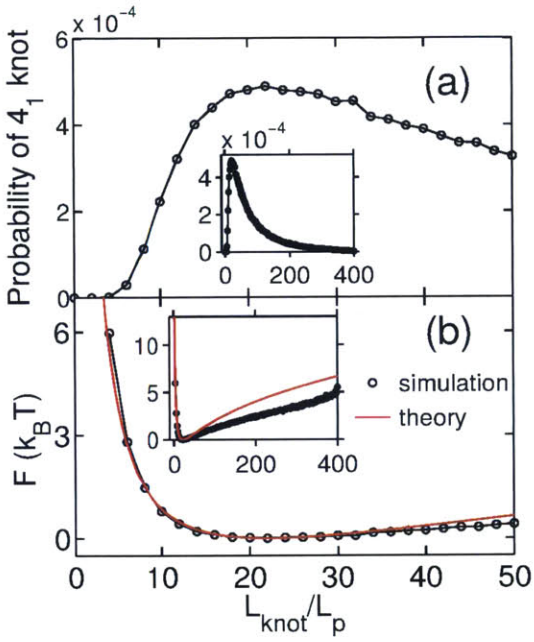

Fig. E.6: *(a) The probability of a wormlike chain containing a figure eight (4_1) knot with $L = 400L_p$. (b) The potential of mean force as a function of knot size. The line of best fit is shown in red: $y = 37.0x^{-1} + 1.8x^{1/3} - 6.7$. Both insets show curves over wider ranges.*

E.5 Simulation results for figure eight (4_1) knots

Figure E.6 shows the simulation results of 4_1 knots in a wormlike chain, similar to Figure E.2, which is for 3_1 knots. The fit to the potential of mean force yields an equation $F/(k_BT) = 37.0(L_{knot}/L_p)^{-1} + 1.8(L_{knot}/L_p)^{1/3} - 6.7$.

Figure E.7 shows the most probable size of a figure eight (4_1) knot as a function of the chain width in the simulations of real chains. The fit to the simulation results yields $p = 38$ for 4_1 knots.

E.6 Increase of bending energy within knots

As described in the main text, the subchain in the knot core experiences more bending than the unknotted subchain with the same contour length. We calculate the increase of bending energy within the knotted region from simulations using two methods. In the first method, we calculate the bending energy of each chain and then group these bending energies into bins E_{bend}^{knot} according to the size of 3_1 knot on chain. Then, we subtract the average bending energy of unknotted chains $\langle E_{bend}^{unknotted} \rangle$ from E_{bend}^{knot} in each bin, as shown by the red curve in Figure E.8a. The statistical errors in the red curve are relatively large due to the large thermal fluctuations in long chains. In the second method, we only calculate the bending energy for the core of 3_1 knot, and then we

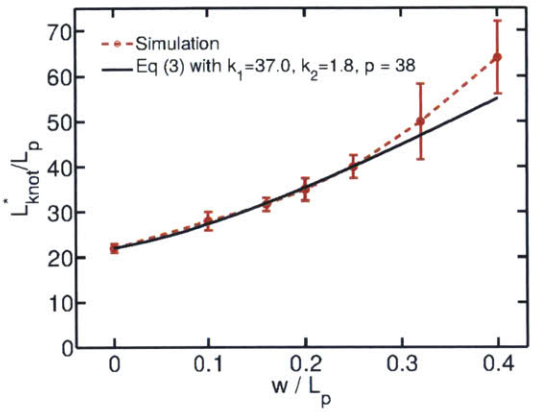

Fig. E.7: *The most probable size of a figure eight (4_1) knot as a function of the rescaled chain width. The solid line is calculated from Eq. (3) with $k_1 = 37.0$, $k_2 = 1.8$, and $p = 38$.*

subtract the bending energy of unknotted chains with the same contour length, as shown by the black curve in Figure E.8a. The bending energies calculated by the first method are larger than the ones calculated by the second method, which suggests that the unknotted region of a knotted chain may experience less bending energy than the random case. It is possible that the segments marginally outside knot core are more likely to be straight than fully random segments. From comparison, we also plot $E_{bend}^{knot} = 17.06L_p/L_{knot}$, which is the first term (bending energy term) in the fitting equation used in Figure E.2b. This equation is close to both red and black curves, which indicates this equation indeed captures the bending energy change due to forming the knot. The same analysis for the 4_1 knot is shown in Figure E.8b, and we observe the same trend as the 3_1 knot.

E.7 The radius of gyration of the knot core

Figure E.9 and E.9b show the probability of the 3_1 knot and the 4_1 knot, respectively, as a function of the radius of gyration of the knot R_{knot}. The probability for the 3_1 knot exhibits a peak at $R_{knot} = 1.3L_p$, and the probability for the 4_1 exhibits a peak at $R_{knot} = 1.9L_p$. Figure E.10 shows the probability of the 3_1 knot for a real chain with $w/L_p = 0.5$, and the peak is located as $R_{knot} = 2.75L_p$.

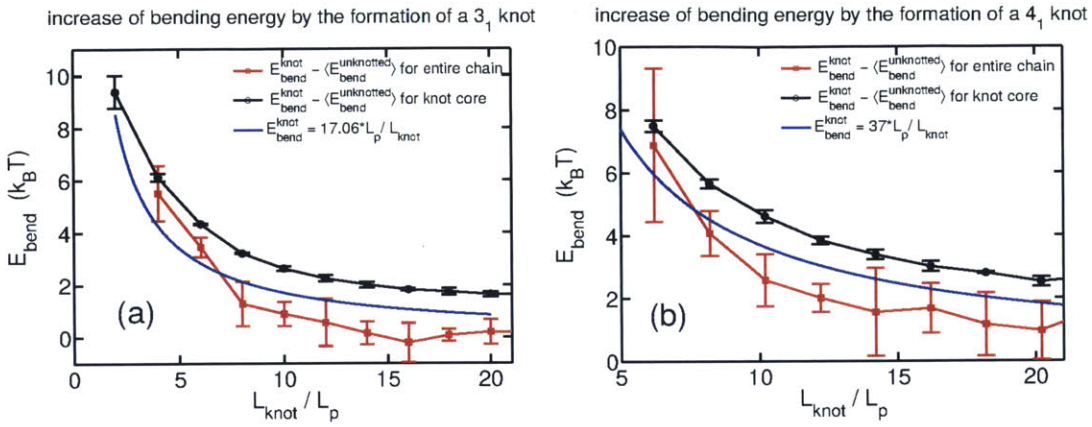

Fig. E.8: *The increase of bending energy by the formation of a* 3_1 *knot (a) or a* 4_1 *knot (b). The red and black curves are calculated from the configurations in the simulation of a wormlike chain with* $L/L_p = 400$. *The blue curve is calculated from the the first term (bending energy term) in the fit equation used in Figure 2(b).*

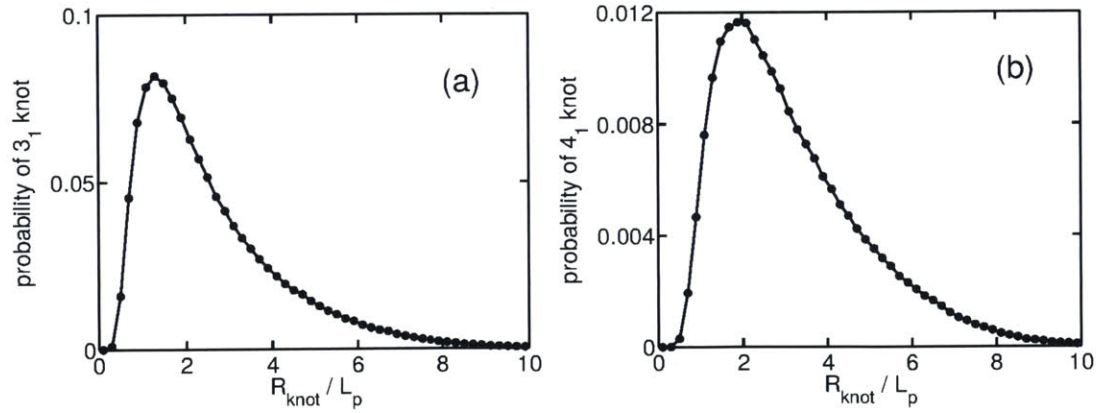

Fig. E.9: *The probability of the* 3_1 *knot (a) and the* 4_1 *knot (b) as a function of the radius of gyration of the knot in the simulation of a wormlike chain with* $L/L_p = 400$.

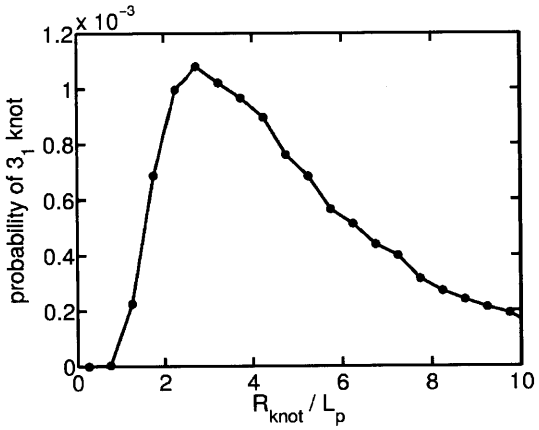

Fig. E.10: *The probability of the* 3_1 *knot as a function of the radius of gyration of the knot in the simulation of a real chain with* $L/L_p = 400$ *and* $w/L_p = 0.5$.

Appendix F

F.1 The confinement free energy

Figure S1 shows the confinement free energy of an equilibrated chain in a square channel as a function of the rescaled channel size. The red line is based on Eq. 6.5 in Chapter 6, which was obtained from the de Gennes regime in our previous study[168]. The simulation results agree with Eq. 6.5 for the channel size in the range $4 < D/L_p < 20$.

F.2 Location of knots along the contour

Figure F.2 shows the distribution of knot locations along the contour. The knots appear to be less frequent at both ends, simply because the centers of knots are separated from the chain ends with a minimum distance of $L_{knot}/2$, which has been discussed in 5 and Ref. [144].

F.3 Radius of gyration of knots in bulk

Figure F.3 shows the average radius of gyration of a knot as a function of the contour length of knot. It is expected that for very tight knots, the radius of gyration is proportional to the contour length[141]. A linear function $y = 0.1x$ is included in Figure F.3 for comparison. The three curves, corresponding to three chain widths, are close to each other. As a result, we approximate that the

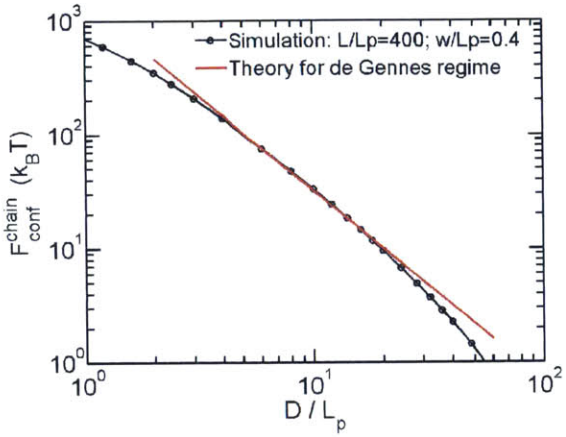

Fig. F.1: *The confinement free energy as a function of the rescaled channel size.*

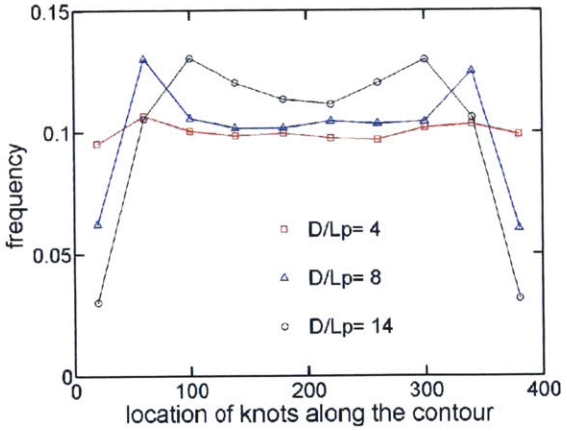

Fig. F.2: *Location of knots along the contour.*

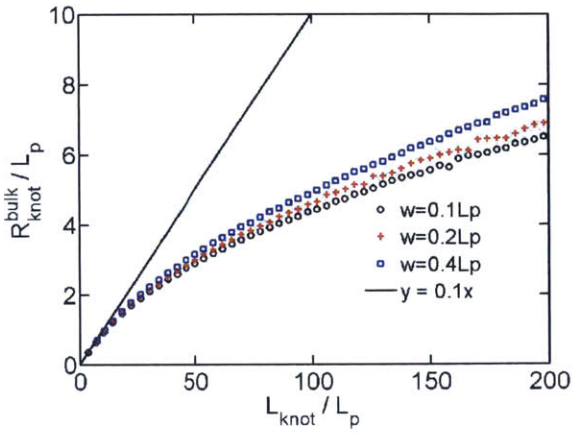

Fig. F.3: *The average radius of gyration of knot as a function of the contour length of knot.*

radius of knot only depends on the contour length inside the knot when we derive the free energy of knots in confinement.

F.4 The mean size of knots

Figure F.4 shows the mean size of knots as a function of the channel size. The mean size of knots exhibits different behavior from the most probable knot size, *i.e.* metastable knot size. The difference is caused by the asymmetry of the size distribution of knots. It is worth mentioning that the metastable knot size is much more insensitive to the contour length of the chain than the mean size of knots.

F.5 The size distribution of trefoil knots for other chain widths

Figure F.5 shows the size distribution of trefoil knots in bulk and channels for the chain width $w = 0.1L_p$. Figure F.6 shows the size distribution of trefoil knots in bulk and channels for the chain width $w = 0.2L_p$.

F.6 Free energy of knots in bulk

Figure F.7 shows the potential of mean force as a function of the knot size for semiflexible chains in bulk. The symbols and lines correspond to simulation results and theoretical predictions.

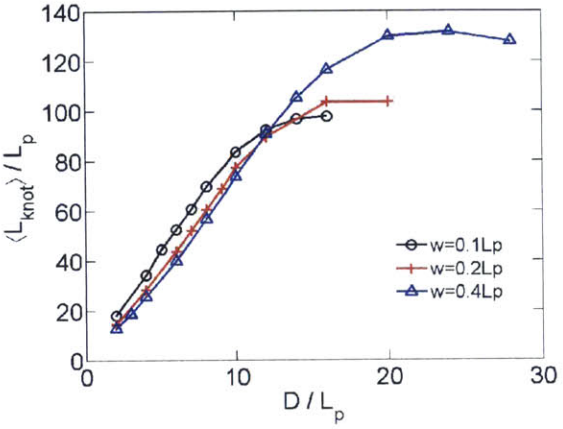

Fig. F.4: *The mean size of knot as a function of the channel size.*

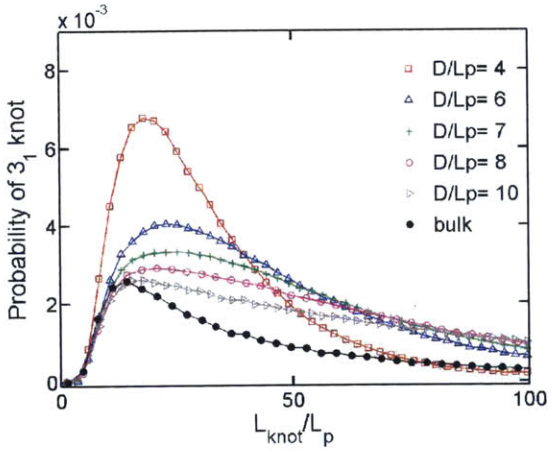

Fig. F.5: *The probability distribution of trefoil knots for chains with* $w = 0.1L_p$ *and* $L = 400L_p$.

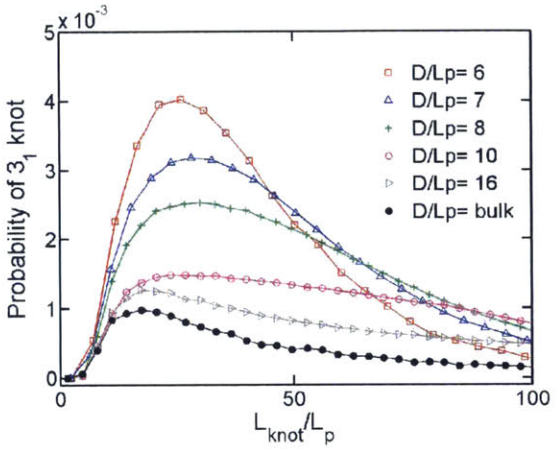

Fig. F.6: *The probability distribution of trefoil knots for chains with* $w = 0.2L_p$ *and* $L = 400L_p$.

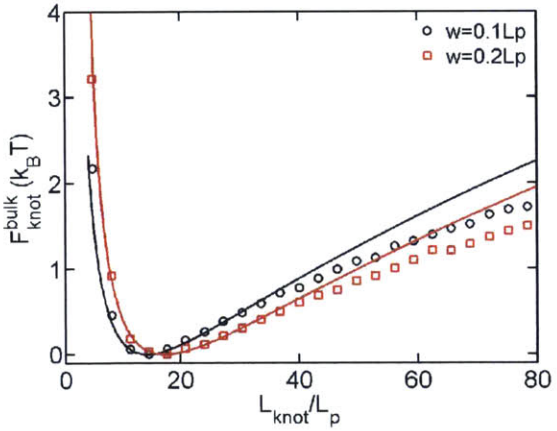

Fig. F.7: *Potential of mean force as a function of knot size for semiflexible chains in bulk. The symbols and lines correspond to simulation results and theoretical predictions.*

Appendix G

G.1 Simulation Methods

We have used a Brownian dynamics technique, parameterized to dsDNA by Vologodskii[170, 191] and others[192]. We briefly recap these simulation techniques below, noting the purposeful similarity to Ref. [170].

The molecule is discretized into n beads. These beads have positions $\mathbf{r}_1$, $\mathbf{r}_2, ...,$ $\mathbf{r}_n$. The n-1 bead-bead bond vectors are $\mathbf{b}_i = \mathbf{r}_{i+1} - \mathbf{r}_i$. The chains are connected by stiff Hookean springs, with a total stretching energy of:

$$E^s = \frac{h}{2} \sum_{i=1}^{n} (b_i - b_o) \tag{G.1}$$

where $h = 200 k_b T / b_o^2$ is the spring constant and $b_o = 0.2 l_p = 10$ nm is the equilibrium spring length.

The semiflexible nature of dsDNA is enforced by penalizing the bending angle θ_i between bonds b_i and b_{i+1}. The form of the bending energy is given as:

$$E^b = \frac{g}{2} \sum_{i=1}^{n-2} \theta_i^2 \tag{G.2}$$

where $g = 4.81 k_B T$ is chosen to set $l_p = 50$ when $b_o = 10$ nm[218].

A screened Debye-Huckle potential is enforced by quantizing the charge along the chain at the bead positions and calculating their interactions based on their relative positions, $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$ by:

$$E^{\mathrm{e}} = \frac{\nu b_o^2}{D} \sum_{i,j}^{n} \frac{\exp\left(-\kappa r_{ij}\right)}{r_{ij}} \tag{G.3}$$

where $\kappa^{-1} =$ is the Debye length, D is the dielectric constant of water, and $\nu = 2.43\ e/\mathrm{nm}$ is the effective linear charge density of DNA at an ionic strength of 10 mM.

An additional close range excluded volume is added for the purpose of limiting self-crossings:

$$E^{\mathrm{ev}} = -\sum_{i,j}^{n} \mu r_{ij} \text{ if } r_{ij} < 2 \text{ nm} \tag{G.4}$$

For any of these energies, E^j, the resulting force exerted on a bead is $\mathbf{F}_i^j = -\frac{\partial E^j}{\mathbf{r}_i}$. These calculations are relatively straightforward with the exception of the calculation of bending energies, which are derived in Ref. [219].

In a subset of simulations, an external tensile force, F^T, is applied to the terminal beads:

$$\mathbf{F}_1^{\mathrm{T}} = -F^T \hat{\mathbf{x}} \tag{G.5}$$

$$\mathbf{F}_{i=2,\ldots,n-1}^{\mathrm{T}} = 0\hat{\mathbf{x}} \tag{G.6}$$

$$\mathbf{F}_n^{\mathrm{T}} = F^T \hat{\mathbf{x}} \tag{G.7}$$

In the majority of simulations, the chain is extended by a planar elongational flow described by

$$\mathbf{u}\left(\mathbf{r}_i\right) = \dot{\varepsilon}\left(\hat{\mathbf{x}} - \hat{\mathbf{y}}\right) \cdot \mathbf{r}_i \tag{G.8}$$

The resulting drag force on bead i is given as:

$$\mathbf{F}_i^{\mathrm{d}} = \zeta_b \left[\mathbf{u}\left(\mathbf{r}_i\right) - \frac{d\mathbf{r}_i}{dt}\right] \tag{G.9}$$

where ζ_b is the drag coefficient of a single bead.

The Brownian dynamics technique takes advantage of the negligibility of inertia at the scale of discretization to set $\sum \mathbf{F}_i^j \simeq 0$. For simulation purposes, length is made dimensionless by the bond length, $\tilde{r} = r/b_o$, force by k_BT/b_o, $\tilde{F} = F/(k_BT/b_o)$, and time by the bead diffusion time, $\tilde{t} = t/(b_o^2\zeta_b/k_BT)$. This yields the Langevin equation:

$$\frac{d\tilde{\mathbf{r}}_i}{d\tilde{t}} = \tilde{\mathbf{F}}_i^{\mathrm{s}} + \tilde{\mathbf{F}}_i^{\mathrm{b}} + \tilde{\mathbf{F}}_i^{\mathrm{e}} + \tilde{\mathbf{F}}_i^{\mathrm{ev}} + \tilde{\mathbf{F}}_i^{\mathrm{B}} + \tilde{\mathbf{u}}\left(\mathbf{r}_i\right) \tag{G.10}$$

where $\tilde{\mathbf{F}}_i^{\mathrm{B}}$ is the Brownian force on particle i. Brownian forces must be chosen to satisfy the Fluctuation-Dissipation theorem, i.e. $\langle \mathbf{F}_i^{\mathrm{B}}(t) \rangle = 0$ and $\langle \mathbf{F}_i^{\mathrm{B}}(t) \mathbf{F}_j^{\mathrm{B}}(t') \rangle = 2k_BT\zeta_b(t - t')$.

The Langevin equation is solved in time with forward Euler integration and a midstepping

algorithm[220]. This scheme results in Brownian forces of the form:

$$\tilde{\mathbf{F}}_i^B = \sqrt{\frac{24}{d\tilde{t}}}\mathbf{U} \tag{G.11}$$

where U is 3 dimensional random variable with each component sampled from a uniform distribution between -1/2 and 1/2 and $d\tilde{t}$ is the size of the time step in the integration scheme.

For all simulations, $n = 300$ corresponding to a contour length of $L_c = 2.99$ μm. For the previous parameterization, there are 5 bonds per persistence length. We also introduce the drag on a persistence length of segments, $\zeta_p = 5\zeta_b$ in order to provide a physical scale for drag to nondimensionalize results (as opposed to ζ_b, *discretization* drag scale). The longest relaxation time of an unknotted chain, $\lambda = 128\left[\zeta_p l_p^2(k_B T)\right]$, was determined by fitting the long time decay of the squared end-to-end distance of an initially stretched chain to a single exponential.

As a knot is not rigorously defined on an open polymer, we developed an algorithm to calculate bounds on a knot on an open polymer. Our method is similar to those developed by other researchers for this same problem [153, 195, 196]. An artificial self closing procedure was used whereby we assume the chain may be closed without generating any additional self crossings. This is a decent approximation for a knot in a polymer under a strong flow or tension. An invariant of the knot, the Alexander polynomial [186], is calculated for this artificially closed loop. For a given knot, this polynomial will not change, and since the initial conditions are specified, it is known what the prescribed value for this polynomial should be. We then successively remove self-crossings from the left and right of the polymer chain. When the Alexander polynomial deviates upon removal of a crossing from its prescribed value, we take that crossing to be a boundary of the knot. This procedure yields both a left and right bound of a given knot. While this procedure is simplistic, its suitability is confirmed by the snapshots in Fig. S7.1 of the manuscript.

Initial conditions for knots were introduced via harmonic [1] or (p, q) torus parameterizations. For a given set of simulation parameters, knots were relaxed at the simulation conditions for at least 1000 $\tilde{t}$. Their positions were held near the chain center by periodic reptation moves whereby segments from one end of the chain were cut and added to the other end of the chain. This procedure generated an equilibrated knot for a given set of simulation parameters. To generate simulation runs, a unique random seed was specified for the Brownian forces. The equilibrated knot was then held at the desired initial condition (via reptation moves) for at least another 1000 $\tilde{t}$ until the knot position was within 1/50 l_p of the desired initial position. After this time, the knot position was recorded as the initial position, reptation moves were halted, and the knot trajectory was allowed to evolve freely.

G.2 Simplification of the Diffusion-Convection Equation

In the manuscript, we present the diffusion-convection equation for the probability distribution function, Ψ, for knot positions as:

$$\frac{\partial \Psi}{\partial t} = \nabla \cdot (D\nabla\Psi) - \nabla \cdot (\mathbf{u}\Psi) \tag{G.12}$$

where D is the diffusivity of the knot.

We simplify the equation to

$$\frac{\partial \Psi}{\partial t} = D\nabla^2\Psi - \dot{\varepsilon}\left(x\frac{\partial \Psi}{\partial x} - \Psi\right) \tag{G.13}$$

based on the following assumptions:

1. The chain coordinates are co-linear with the stretching axis of the flow. This should be reasonable for the moderate to strong flows shown in this manuscript, Wi $\geq$ 1.6.

2. The simplification of the diffusive term in the diffusion-convection requires $\frac{\partial D}{\partial x} \approx 0$. The diffusivity of a knot can depend on tension in a chain[195, 196], and an approximately quadratic distribution of tension develops along a chain extended by an elongational flow. We have measured the diffusivity of the 3_1 knot over a range of tensions felt in the center of the chain, shown in Fig. SG.1. For these tensions, the diffusivity of the 3_1 knot is roughly constant.

3. Sub-diffusive "breathing" modes[196] are ignored. There is evidence for these in the short-time data in Fig. SG.1, but the transport of knots off the chain occurs at length scales when the MSDs have become linear.

4. End effects are ignored.

G.3 Isotension Results

A measure of knot diffusivity is required to collapse the data in Fig. 2 of the manuscript. For a knot of a given topology on a chain extended by an applied tension, the magnitude of the applied tension can affect the diffusivity of the knot[195, 196]. For a polymer extended by a planar elongational flow, an approximately quadratic distribution of tension results with an apex near the center. Furthermore, this peak tension is itself a function of Weissenberg number. We ran a series of simulations where the diffusion coefficients of a knot were calculated at a constant tension.

As the diffusion-dominated regime is confined to a small fraction of the chain near the center, we estimate the approximate tension, F^{T*}, felt by knots in this regime to be the tension at the center of a completely linearized chain at the flow conditions:

$$F^{T*} = \sum_{i=n/2}^{n} \dot{\varepsilon}\zeta_b \mathbf{b}_i \cdot \hat{\mathbf{x}} \approx \dot{\varepsilon}\zeta_b b_o \sum_{i=1}^{n/2} i = \frac{1}{8}\dot{\varepsilon}\zeta_b b_o (n+2)n \tag{G.14}$$

The approximate tensions used for the flow strengths considered in the manuscript are shown in Table G.1. Tensions above the range shown are not feasibly investigated by the simulation model we employed as they lead to non-physical crossings of strands and untying of knots.

The mean squared displacement of the 3_1 knot along a chain extended by tensions is shown in Fig. S1. The calculated diffusivities are shown in the inset. The diffusivity for the 3_1 knot did not vary over the range of simulation flow strengths, so a characteristic diffusivity was taken at $F^T = 50k_bT/l_p$ when collapsing all the data in Fig. 2 of the manuscript. In Table G.2, the calculated diffusivities are shown for the 3_1 knot (same as Fig. SG.1 inset) as well as the 4_1, 5_1, 5_2, 7_1, and 10_{28} knots.

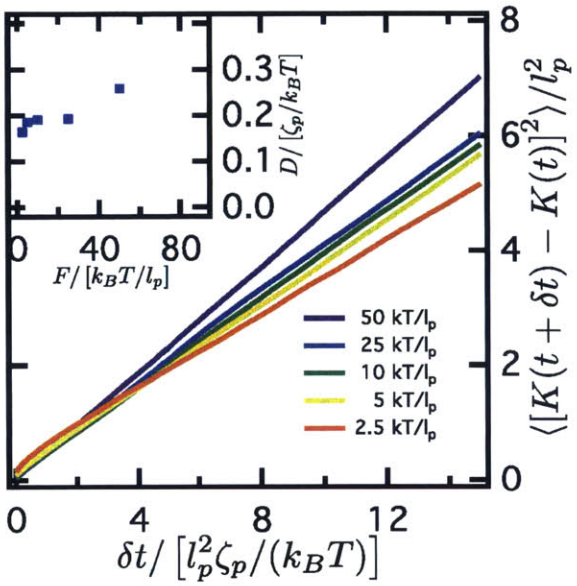

Fig. G.1: *Mean-squared displacements of the* 3_1 *knot versus lagtime for varying applied tension. (Inset) Diffusivities of the* 3_1 *knot calculated by fitting the knot mean-squared displacements in the linear regime.*

Table G.1: *Approximate tension for various flow strengths*

Strain rate, $\dot{\varepsilon} \times \left[l_p^2\zeta_p/k_BT\right]$	Wi	Approximate Tension, $F^{T*}/(k_BT/l_p)$
0.0125	1.6	5.6625
0.0395	5.0	17.9064
0.1250	16	56.6250
0.1875	24	84.9375

The length of the diffusion-dominated regime, l, is shown for various knots at the strain rates used in the manuscript in Table G.3. For even the small chains used in this manuscript ($L_c = 59.8l_p$), the length of the diffusion-dominated regime is a small fraction of the total chain.

Table G.2: *Knot metrics in tensioned DNA*

Knot	Tension, $F^T/(k_BT/l_p)$	Diffusivity, $D/(k_BT/\zeta_p)$
3_1	2.5	0.163
3_1	5.0	0.186
3_1	10	0.191
3_1	25	0.193
3_1	50	0.259
4_1	50	0.085
5_1	50	0.097
5_2	50	0.064
7_1	50	0.067
10_{28}	50	0.042

Table G.3: *Length of diffusion-dominated regime, l/l_p, for selected knots at various flow strengths*

Knot \ Wi	1.6	5.0	16	24
3_1	4.55	2.56	1.44	1.17
4_1	2.61	1.47	0.82	0.67
5_1	2.79	1.57	0.88	0.72
5_2	2.26	1.27	0.71	0.58
7_1	2.31	1.30	0.73	0.60
10_{28}	1.84	1.04	0.58	0.48

Bibliography

[1] Koseleff, P.-V.; Pecker, D. 2012, arXiv:1203.4376. arXiv.org e-Print archive. `http://arxiv.org/pdf/1203.4376.pdf` (accessed July 28, 2014).

[2] Watson, J. D.; Crick, F. H. *Nature* **1953**, *171*, 737–738.

[3] Group, T. I. S. M. W. *Nature* **2001**, *409*, 928–933.

[4] The Chimpanzee Sequencing and Analysis Consortium, *Nature* **2005**, *437*, 69–87.

[5] Lindblad Toh, K. et al. *Nature* **2005**, *438*, 803–819.

[6] Ihmels, H.; Otto, D. *Supermolecular dye chemistry*; Springer, 2005; pp 161–204.

[7] Perkins, T.; Smith, D.; Larson, R.; Chu, S. *Science* **1995**, *268*, 83.

[8] Bustamante, C.; Smith, S. B.; Liphardt, J.; Smith, D. *Curr. Opin. Struct. Biol.* **2000**, *10*, 279–285.

[9] Shaqfeh, E. *J. Non-Newtonian Fluid Mech.* **2005**, *130*, 1–28.

[10] Tegenfeldt, J. O.; Prinz, C.; Cao, H.; Huang, R. L.; Austin, R. H.; Chou, S. Y.; Cox, E. C.; Sturm, J. C. *Anal. Bioanal. Chem.* **2004**, *378*, 1678–1692.

[11] Smith, S.; Finzi, L.; Bustamante, C. *Science* **1992**, *258*, 1122–1126.

[12] Smith, S. B.; Cui, Y.; Bustamante, C. *Science* **1996**, *271*, 795–799.

[13] Wang, M. D.; Yin, H.; Landick, R.; Gelles, J.; Block, S. M. *Biophys. J.* **1997**, *72*, 1335.

[14] Marko, J. F.; Siggia, E. D. *Macromolecules* **1995**, *28*, 8759–8770.

[15] Zhang, X.; Chen, H.; Fu, H.; Doyle, P. S.; Yan, J. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 8103–8108.

[16] Randall, G. C.; Doyle, P. S. *Phys. Rev. Lett.* **2004**, *93*, 058102.

[17] Randall, G. C.; Schultz, K. M.; Doyle, P. S. *Lab Chip* **2006**, *6*, 516–525.

[18] Tang, J.; Doyle, P. S. *Appl. Phys. Lett.* **2007**, *90*, 224103.

[19] Balducci, A. G.; Tang, J.; Doyle, P. S. *Macromolecules* **2008**, *41*, 9914–9918.

[20] Tang, J.; Trahan, D.; Doyle, P. *Macromolecules* **2010**, *43*, 3081–3089.

[21] Jo, K.; Dhingra, D. M.; Odijk, T.; de Pablo, J. J.; Graham, M. D.; Runnheim, R.; Forrest, D.; Schwartz, D. C. *Proc. Natl. Acad. Sci.* **2007**, *104*, 2673–2678.

[22] Reisner, W.; Morton, K. J.; Riehn, R.; Wang, Y. M.; Yu, Z.; Rosen, M.; Sturm, J. C.; Chou, S. Y.; Frey, E.; Austin, R. H. *Phys. Rev. Lett.* **2005**, *94*, 196101.

[23] Reisner, W.; Pedersen, J. N.; Austin, R. H. *Rep. Prog. Phys.* **2012**, *75*, 106601.

[24] Bonthuis, D. J.; Meyer, C.; Stein, D.; Dekker, C. *Phys. Rev. Lett.* **2008**, *101*, 108303.

[25] Graham, M. D. *Annu. Rev. Fluid Mech.* **2011**, *43*, 273–298.

[26] Balducci, A.; Mao, P.; Han, J.; Doyle, P. *Macromolecules* **2006**, *43*, 6273–6281.

[27] Tang, J.; Levy, S. L.; Trahan, D. W.; Jones, J. J.; Craighead, H. G.; Doyle, P. S. *Macromolecules* **2010**, *43*, 7368–7377.

[28] Kim, Y.; Kim, K. S.; Kounovsky, K. L.; Chang, R.; Jung, G. Y.; Jo, K.; Schwartz, D. C. *Lab Chip* **2011**, *11*, 1721–1729.

[29] Yeh, J.-W.; Taloni, A.; Chen, Y.-L.; Chou, C.-F. *Nano Lett.* **2012**, *12*, 1597–1602.

[30] Turner, J. C.; Van de Griend, P. *History and science of knots*; World Scientific, 1996.

[31] Edwards, S. *Proc. Phys. Soc.* **1967**, *91*, 513.

[32] de Gennes, P.-G. *J. Chem. Phys.* **1971**, *55*, 572.

[33] Doi, M.; Edwards, S. F. *The theory of polymer dynamics*; Oxford University Press, 1988; Vol. 73.

[34] Ashley, C. W. *The Ashley book of knots*; Doubleday Books, 1944; Vol. 13.

[35] Raymer, D.; Smith, D. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 16432.

[36] Wang, J. C. *Q. Rev. Biophys.* **1998**, *31*, 107–144.

[37] Wang, J. C. *Annu. Rev. Biochem.* **1996**, *65*, 635–692.

[38] Fortune, J. M.; Osheroff, N. *Prog. Nucleic Acid Res. Mol. Biol.* **2000**, *64*, 221–253.

[39] Kim, Y. S.; Kundukad, B.; Allahverdi, A.; Nordensköld, L.; Doyle, P. S.; van der Maarel, J. R. *Soft Matter* **2013**, *9*, 1656–1663.

[40] Meluzzi, D.; Smith, D. E.; Arya, G. *Annu. Rev. Biophys.* **2010**, *39*, 349–366.

[41] Arsuaga, J.; Vázquez, M.; Trigueros, S.; Sumners, D. W.; Roca, J. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5373–5377.

[42] Arsuaga, J.; Vazquez, M.; McGuirk, P.; Trigueros, S.; Sumners, D. W.; Roca, J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 9165–9169.

[43] Taylor, W. *Nature* **2000**, *406*, 916–919.

[44] Lua, R.; Grosberg, A. *PLoS Comput. Biol.* **2006**, *2*, e45.

[45] Virnau, P.; Mirny, L. A.; Kardar, M. *PLoS Comput. Biol.* **2006**, *2*, e122.

[46] Kardar, M. *Eur. Phys. J. B* **2008**, *64*, 519–523.

[47] Mallam, A. L.; Rogers, J. M.; Jackson, S. E. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 8189–8194.

[48] Micheletti, C.; Di Stefano, M.; Orland, H. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 2052–2057.

[49] Sumners, D.; Whittington, S. *J. Phys. A* **1988**, *21*, 1689.

[50] Kantor, Y. *Pramana* **2005**, *64*, 1011–1017.

[51] Quake, S. R. *Phys. Rev. Lett.* **1994**, *73*, 3317.

[52] Lai, P.-Y. *Phys. Rev. E* **2002**, *66*, 021805.

[53] Virnau, P.; Kantor, Y.; Kardar, M. *J. Am. Chem. Soc.* **2005**, *127*, 15102–15106.

[54] Rabin, Y.; Grosberg, A.; Tanaka, T. *Europhys. Lett.* **1995**, *32*, 505.

[55] Lee, N.; Abrams, C.; Johner, A.; Obukhov, S. *Macromolecules* **2004**, *37*, 651–661.

[56] De Gennes, P. G. *Macromolecules* **1984**, *17*, 703–704.

[57] Lee, N.; Abrams, C.; Johner, A.; Obukhov, S. *Phys. Rev. Lett.* **2003**, *90*, 225504.

[58] Yoshinaga, N. *Phys. Rev. E* **2008**, *77*, 061805.

[59] des Cloizeaux, J. *J. Phys., Lett.* **1981**, *42*, 433–436.

[60] Deutsch, J. *Phys. Rev. E* **1999**, *59*, R2539.

[61] Grosberg, A. Y.; Rabin, Y. *Phys. Rev. Lett.* **2007**, *99*, 217801.

[62] Matthews, R.; Louis, A. A.; Yeomans, J. M. *Phys. Rev. Lett.* **2009**, *102*, 088101.

[63] Marenduzzo, D.; Micheletti, C.; Orlandini, E.; Sumners, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 20081–20086.

[64] Huang, L.; Makarov, D. E. *J. Chem. Phys.* **2008**, *129*, 121107.

[65] Szymczak, P. *Eur. Phys. J.: Spec. Top.* **2014**, *223*, 1805–1812.

[66] Rosa, A.; Di Ventra, M.; Micheletti, C. *Phys. Rev. Lett.* **2012**, *109*, 118301.

[67] Dietrich-Buchecker, C. O.; Sauvage, J.-P. *Angew. Chem., Int. Ed. Engl.* **1989**, *28*, 189–192.

[68] Ayme, J.-F.; Beves, J. E.; Leigh, D. A.; McBurney, R. T.; Rissanen, K.; Schultz, D. *Nat. Chem.* **2012**, *4*, 15–20.

[69] Arai, Y.; Yasuda, R.; Akashi, K.; Harada, Y.; Miyata, H.; Kinosita, K.; Itoh, H. *Nature* **1999**, *399*, 446–448.

[70] Bao, X.; Lee, H.; Quake, S. *Phys. Rev. Lett.* **2003**, *91*, 265506.

[71] Metzler, R.; Reisner, W.; Riehn, R.; Austin, R.; Tegenfeldt, J.; Sokolov, I. *Europhys. Lett.* **2006**, *76*, 696.

[72] Viovy, J.-L.; Miomandre, F.; Miquel, M.-C.; Caron, F.; Sor, F. *Electrophoresis* **1992**, *13*, 1–6.

[73] Tang, J.; Du, N.; Doyle, P. S. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 16153–16158.

[74] Zhou, C.; Reisner, W. W.; Staunton, R. J.; Ashan, A.; Austin, R. H.; Riehn, R. *Phys. Rev. Lett.* **2011**, *106*, 248103.

[75] Reifenberger, J. G.; Dorfman, K. D.; Cao, H. *Analyst* **2015**,

[76] Dekker, C. *Nat. Nanotech.* **2007**, *2*, 209–215.

[77] Clarke, J.; Wu, H.-C.; Jayasinghe, L.; Patel, A.; Reid, S.; Bayley, H. *Nat. Nanotechnol.* **2009**, *4*, 265–270.

[78] Plesa, C.; Verschueren, D.; Ruitenberg, J. W.; Witteveen, M. J.; Jonsson, M. P.; Grosberg, A. Y.; Rabin, Y.; Dekker, C. *Biophys. J.* **2015**, *108*, 166a.

[79] Adams, C. *The knot book: an elementary introduction to the mathematical theory of knots*; AMS Bookstore, 2004.

[80] Kawauchi, A. *A survey of knot theory*; Birkhäuser, 2012.

[81] Cozzarelli, N. R.; Wang, J. C. *DNA topology and its biological effects*; Cold Spring Harbor Laboratory Press, 1990.

[82] Tubiana, L.; Orlandini, E.; Micheletti, C. *Prog. Theor. Phys. Supp* **2011**, *191*, 192.

[83] Van Rensburg, E. J.; Sumners, D.; Wasserman, E.; Whittington, S. *J. Phys. A* **1992**, *25*, 6557.

[84] Millett, K.; Dobay, A.; Stasiak, A. *Macromolecules* **2005**, *38*, 601–606.

[85] Barber, C. B.; Dobkin, D. P.; Huhdanpaa, H. *ACM Trans. Math. Software* **1996**, *22*, 469–483.

[86] Renner, C. B.; Du, N.; Doyle, P. S. *Biomicrofluidics* **2014**, *8*, –.

[87] Bousse, L.; Cohen, C.; Nikiforov, T.; Chow, A.; Kopf-Sill, A. R.; Dubrow, R.; Parce, J. W. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 155–181.

[88] West, J.; Becker, M.; Tombrink, S.; Manz, A. *Anal. Chem.* **2008**, *80*, 4403–4419.

[89] Auroux, P.-A.; Koc, Y.; deMello, A.; Manz, A.; Day, P. J. R. *Lab Chip* **2004**, *4*, 534–546.

[90] Easley, C. J.; Karlinsey, J. M.; Bienvenue, J. M.; Legendre, L. A.; Roper, M. G.; Feldman, S. H.; Hughes, M. A.; Hewlett, E. L.; Merkel, T. J.; Ferrance, J. P.; Landers, J. P. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 19272–19277.

[91] Landers, J. P. *Anal. Chem.* **2003**, *75*, 2919–2927.

[92] O'Connor, L.; Glynn, B. *Expert Rev. Med. Devices* **2010**, *7*, 529–539.

[93] Viovy, J.-L. *Rev. Mod. Phys.* **2000**, *72*, 813.

[94] Mitnik, L.; Heller, C.; Prost, J.; Viovy, J. *Science* **1995**, *267*, 219–222.

[95] Magnúsdóttir, S.; Isambert, H.; Heller, C.; Viovy, J.-L. *Biopolymers* **1999**, *49*, 385–401.

[96] Isambert, H.; Ajdari, A.; Viovy, J.; Prost, J. *Phys. Rev. E* **1997**, *56*, 5688.

[97] Isambert, H.; Ajdari, A.; Viovy, J.-L.; Prost, J. *Phys. Rev. Lett.* **1997**, *78*, 971.

[98] Bruinsma, R. F.; Riehn, R. *ChemPhysChem* **2009**, *10*, 2871–2875.

[99] Kojima, M.; Kubo, K.; Yoshikawa, K. *J. Chem. Phys.* **2006**, *124*, 024902.

[100] Jones, J. J.; van der Maarel, J. R.; Doyle, P. S. *Nano Lett.* **2011**, *11*, 5047–5053.

[101] Zhang, C.; Shao, P. G.; van Kan, J. A.; van der Maarel, J. R. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 16651–16656.

[102] Schwinefus, J. J.; Morris, M. D. *Macromolecules* **1999**, *32*, 3678–3684.

[103] Shi, X.; Hammond, R. W.; Morris, M. D. *Anal. Chem.* **1995**, *67*, 1132–1138.

[104] Sunada, W. M.; Blanch, H. W. *Electrophoresis* **1998**, *19*, 3128–3136.

[105] Shau-Chun, W.; Morris, M. D. Mechanistic Studies of DNA Separations by Capillary Electrophoresis in Ultradilute Polymer Solutions: Effects of Polymer Rigidity. Proceedings of IUPAC International Congress on Analytical Sciences 2001. Tokyo, 2001; pp i173–i176.

[106] Dorfman, K. D. *Rev. Mod. Phys.* **2010**, *82*, 2903.

[107] Rubinstein, M.; Colby, R. H. *Polymer physics*; OUP Oxford, 2003.

[108] Jones, G. E. *J. Membr. Biol.* **1974**, *16*, 297–312.

[109] Jan, K.-M.; Chien, S. *J. Gen. Physiol.* **1973**, *61*, 638–654.

[110] Tolmachov, O. E. *Med. Hypotheses* **2012**, *78*, 632–635.

[111] Diakité, M. L. Y.; Champ, J.; Descroix, S.; Malaquin, L.; Amblard, F.; Viovy, J.-L. *Lab Chip* **2012**, *12*, 4738–4747.

[112] Renner, C. B.; Doyle, P. S. *Soft Matter* **2015**, *11*, 3105–3114.

[113] Michaeli, Y.; Ebenstein, Y. *Nat. Biotechnol.* **2012**, *30*, 762–763.

[114] Marciel, A. B.; Schroeder, C. M. *J. Polym. Sci., Part B: Polym. Phys.* **2013**, *51*, 556–566.

[115] Mai, D. J.; Brockman, C.; Schroeder, C. M. *Soft Matter* **2012**, *8*, 10560–10572.

[116] Juang, Y.-J.; Wang, S.; Hu, X.; Lee, L. *Phys. Rev. Lett.* **2004**, *93*, 268105.

[117] Gerashchenko, S.; Steinberg, V. *Phys. Rev. E* **2008**, *78*, 040801.

[118] Randall, G. C.; Doyle, P. S. *Macromolecules* **2006**, *39*, 7734–7745.

[119] Ou, J.; Cho, J.; Olson, D. W.; Dorfman, K. D. *Phys. Rev. E* **2009**, *79*, 061904.

[120] Chan, E. Y.; Goncalves, N. M.; Haeusler, R. A.; Hatch, A. J.; Larson, J. W.; Maletta, A. M.; Yantz, G. R.; Carstea, E. D.; Fuchs, M.; Wong, G. G.; Gullans, S. R.; Gilmanshin, R. *Genome Res.* **2004**, *14*, 1137–1146.

[121] Kim, J. M.; Doyle, P. S. *Lab Chip* **2007**, *7*, 213–225.

[122] Dylla-Spears, R.; Townsend, J. E.; Jen-Jacobson, L.; Sohn, L. L.; Muller, S. J. *Lab Chip* **2010**, *10*, 1543–1549.

[123] Balducci, A.; Hsieh, C.; Doyle, P. *Phys. Rev. Lett.* **2007**, *99*, 238102.

[124] Xu, W.; Muller, S. J. *Lab Chip* **2012**, *12*, 647–651.

[125] Perkins, T. T.; Smith, D. E.; Chu, S. *Science* **1997**, *276*, 2016–2021.

[126] Smith, D. E.; Chu, S. *Science* **1998**, *281*, 1335–1340.

[127] Schroeder, C. M.; Babcock, H. P.; Shaqfeh, E. S.; Chu, S. *Science* **2003**, *301*, 1515–1519.

[128] de Gennes, P. G. *Science* **1997**, *276*, 1999–2000.

[129] Larson, R.; Hu, H.; Smith, D.; Chu, S. *J. Rheol.* **1999**, *43*, 267.

[130] Randall, G.; Doyle, P. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 10813.

[131] Balducci, A.; Doyle, P. *Macromolecules* **2008**, *41*, 5485–5492.

[132] Micheletti, C.; Marenduzzo, D.; Orlandini, E. *Phys. Rep.* **2011**, *504*, 1–73.

[133] Micheletti, C.; Orlandini, E. *Macromolecules* **2012**, *45*, 2113–2121.

[134] Micheletti, C.; Orlandini, E. *Soft Matter* **2012**, *8*, 10959–10968.

[135] Dai, L.; van der Maarel, J. R.; Doyle, P. S. *ACS Macro Lett.* **2012**, *1*, 732–736.

[136] Kundukad, B.; Yan, J.; Doyle, P. S. *Soft Matter* **2014**, *10*, 9721–9728.

[137] Hsieh, C.; Balducci, A.; Doyle, P. *Macromolecules* **2007**, *40*, 5196–5205.

[138] Sing, C. E.; Alexander-Katz, A. *Macromolecules* **2010**, *43*, 3532–3541.

[139] Sing, C. E.; Alexander-Katz, A. *J. Chem. Phys.* **2011**, *135*, 014902.

[140] Cifre, J. G. H.; de la Torre, J. G. *J. Chem. Phys.* **2001**, *115*, 9578–9584.

[141] Pierański, P.; Przybył, S.; Stasiak, A. *Eur. Phys. J. E* **2001**, *6*, 123–128.

[142] Schroeder, C.; Shaqfeh, E.; Chu, S. *Macromolecules* **2004**, *37*, 9242–9256.

[143] Di Stefano, M.; Tubiana, L.; Di Ventra, M.; Micheletti, C. *Soft Matter* **2014**, *10*, 6491–6498.

[144] Dai, L.; Renner, C. B.; Doyle, P. S. *Macromolecules* **2014**, *47*, 6135–6140.

[145] Frank-Kamenetskii, M.; Lukashin, A.; Vologodskii, A. *Nature* **1975**, *258*, 398 – 402.

[146] Taylor, W. R.; Lin, K. *Nature* **2003**, *421*, 25.

[147] Tubiana, L.; Rosa, A.; Fragiacomo, F.; Micheletti, C. *Macromolecules* **2013**, *46*, 3669–3678.

[148] Marenduzzo, D.; Orlandini, E.; Stasiak, A.; Sumners, D. W.; Tubiana, L.; Micheletti, C. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 22269–22274.

[149] Saitta, A. M.; Soper, P. D.; Wasserman, E.; Klein, M. L. *Nature* **1999**, *399*, 46–48.

[150] Kivotides, D.; Wilkin, S.; Theofanous, T. *Phys. Rev. E* **2009**, *80*, 041808.

[151] Saitta, A. M.; Klein, M. L. *J. Chem. Phys.* **2002**, *116*, 5333–5336.

[152] Coluzza, I.; van Oostrum, P. D.; Capone, B.; Reimhult, E.; Dellago, C. *Phys. Rev. Lett.* **2013**, *110*, 075501.

[153] Katritch, V.; Olson, W. K.; Vologodskii, A.; Dubochet, J.; Stasiak, A. *Phys. Rev. E* **2000**, *61*, 5545–5549.

[154] Orlandini, E.; Tesi, M.; Whittington, S. *J. Phys. A* **2005**, *38*, L795.

[155] Mobius, W.; Frey, E.; Gerland, U. *Nano Lett.* **2008**, *8*, 4518–4522.

[156] Matthews, R.; Louis, A. A.; Likos, C. N. *ACS Macro Lett.* **2012**, *1*, 1352–1356.

[157] Poier, P.; Likos, C. N.; Matthews, R. *Macromolecules* **2014**, *47*, 3394–3400.

[158] Guitter, E.; Orlandini, E. *J. Phys. A* **1999**, *32*, 1359.

[159] Marcone, B.; Orlandini, E.; Stella, A.; Zonta, F. *J. Phys. A* **2005**, *38*, L15.

[160] Nakajima, C. H.; Sakaue, T. *Soft Matter* **2013**, *9*, 3140–3146.

[161] Ercolini, E.; Valle, F.; Adamcik, J.; Witz, G.; Metzler, R.; De Los Rios, P.; Roca, J.; Dietler, G. *Phys. Rev. Lett.* **2007**, *98*, 058102.

[162] Metzler, R.; Hanke, A.; Dommersnes, P. G.; Kantor, Y.; Kardar, M. *Phys. Rev. Lett.* **2002**, *88*, 188101.

[163] Shaw, S. Y.; Wang, J. C. *Science* **1993**, *260*, 533–536.

[164] Odijk, T. *Macromolecules* **1983**, *16*, 1340–1344.

[165] Yang, Y.; Burkhardt, T. W.; Gompper, G. *Phys. Rev. E* **2007**, *76*, 011804.

[166] Dai, L.; Jones, J. J.; van der Maarel, J. R.; Doyle, P. S. *Soft Matter* **2012**, *8*, 2972–2982.

[167] Grassberger, P. *Phys. Rev. E* **1997**, *56*, 3682.

[168] Dai, L.; van der Maarel, J.; Doyle, P. S. *Macromolecules* **2014**, *47*, 2445–2450.

[169] Zheng, X.; Vologodskii, A. *Phys. Rev. E* **2010**, *81*, 041806.

[170] Vologodskii, A. *Biophys. J.* **2006**, *90*, 1594–1597.

[171] Marcone, B.; Orlandini, E.; Stella, A.; Zonta, F. *Phys. Rev. E* **2007**, *75*, 041105.

[172] Zhou, Q.; Larson, R. G. *Macromolecules* **2006**, *39*, 6737–6743.

[173] Quake, S. R.; Babcock, H.; Chu, S. *Nature* **1997**, *388*, 151–154.

[174] Günther, K.; Mertig, M.; Seidel, R. *Nucleic Acids Res.* **2010**, *38*, 6526–6532.

[175] Hsieh, C.-C.; Balducci, A.; Doyle, P. S. *Nano Lett.* **2008**, *8*, 1683–1688.

[176] Rybenkov, V. V.; Cozzarelli, N. R.; Vologodskii, A. V. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5307–5311.

[177] Dai, L.; Renner, C. B.; Doyle, P. S. *Macromolecules* **2015**, *48*, 2812–2818.

[178] Farago, O.; Kantor, Y.; Kardar, M. *Europhys. Lett.* **2002**, *60*, 53.

[179] Wang, Y.; Tree, D. R.; Dorfman, K. D. *Macromolecules* **2011**, *44*, 6594–6604.

[180] Gupta, D.; Sheats, J.; Muralidhar, A.; Miller, J. J.; Huang, D. E.; Mahshid, S.; Dorfman, K. D.; Reisner, W. *J. Chem. Phys.* **2014**, *140*, 214901.

[181] Lam, E. T.; Hastie, A.; Lin, C.; Ehrlich, D.; Das, S. K.; Austin, M. D.; Deshpande, P.; Cao, H.; Nagarajan, N.; Xiao, M.; Kwok, P.-Y. *Nat. Biotechnol.* **2012**, *30*, 771–776.

[182] Levy, S. L.; Craighead, H. G. *Chem. Soc. Rev.* **2010**, *39*, 1133–1152.

[183] Persson, F.; Tegenfeldt, J. O. *Chem. Soc. Rev.* **2010**, *39*, 985–999.

[184] Reifenberger, J.; Dorfman, K.; Cao, H. *Bull. Am. Phys. Soc.* **2014**, *59*.

[185] Jun, S.; Thirumalai, D.; Ha, B.-Y. *Phys. Rev. Lett.* **2008**, *101*, 138101.

[186] Frank-Kamenetskiĭ, M.; Vologodskiĭ, A. *Phys.-Usp.* **1981**, *24*, 679.

[187] Renner, C. B.; Doyle, P. S. *ACS Macro Lett.* **2014**, *3*, 963–967.

[188] Staczek, P.; Higgins, N. P. *Mol. Microbiol.* **1998**, *29*, 1435–1448.

[189] Virnau, P.; Kardar, M.; Kantor, Y. *Chaos* **2005**, *15*, 41103–41103.

[190] Hsieh, C.; Doyle, P. *Korea-Australia Rheology Journal* **2008**, *20*, 127–142.

[191] Jian, H.; Vologodskii, A. V.; Schlick, T. *J. Comput. Phys.* **1997**, *136*, 168–179.

[192] Klenin, K.; Merlitz, H.; Langowski, J. *Biophys. J.* **1998**, *74*, 780–788.

[193] Stigter, D. *J. Colloid Interface Sci.* **1975**, *53*, 296–306.

[194] Stigter, D. *Biopolymers* **1977**, *16*, 1435–1448.

[195] Huang, L.; Makarov, D. E. *J. Phys. Chem. A* **2007**, *111*, 10338–10344.

[196] Matthews, R.; Louis, A.; Yeomans, J. *Europhys. Lett.* **2010**, *89*, 20001.

[197] Belmonte, A. In *Physical and Numerical Models in Knot Theory*; Calvo, J. A., Ed.; Series on Knots and Everything; World Scientific, 2005; Vol. 36.

[198] Borujeni, E. E.; Zydney, A. L. *J. Membr. Sci.* **2014**, *450*, 189–196.

[199] Klotz, A. R.; Duong, L.; Mamaev, M.; de Haan, H. W.; Chen, J. Z. Y.; Reisner, W. W. *Macromolecules* **2015**, *48*, 5028–5033.

[200] Hsieh, C.; Balducci, A.; Doyle, P. *Nano Lett.* **2008**, *8*, 1683–1688.

[201] Jones, J. J.; van der Maarel, J. R. C.; Doyle, P. S. *Nano Lett.* **2011**, *11*, 5047–5053.

[202] Jones, J. J.; van der Maarel, J. R.; Doyle, P. S. *Phys. Rev. Lett.* **2013**, *110*, 068101.

[203] Burde, G.; Zieschang, H. *Knots*; Walter de gruyter, 2003; Vol. 5.

[204] Bogle, M.; Hearst, J.; Jones, V.; Stoilov, L. *J. knot theory ramif.* **1994**, *3*, 121–140.

[205] Lamm, C. Zylinder-Knoten und symmetrische Vereinigungen. Ph.D. thesis, University of Bonn, 1999.

[206] Rechnitzer, A. Minimal knots on cubic lattices.

[207] Van Rensburg, E. J.; Whittington, S. *J. Phys. A* **1991**, *24*, 5553.

[208] Tang, J. Single Molecule DNA Dynamics in Micro- and Nano-fluidic Devices. Ph.D. thesis, Massachusetts Institute of Technology, 2010.

[209] Senti, F. R.; Hellman, N. N.; Ludwig, N.; Babcock, G.; Tobin, R.; Glass, C.; Lamberts, B. *J. Polym. Sci.* **1955**, *17*, 527–546.

[210] Mitnik, L.; Salomé, L.; Viovy, J. L.; Heller, C. *J. Chromatogr. A* **1995**, *710*, 309–321.

[211] Rief, M.; Fernandez, J. M.; Gaub, H. E. *Phys. Rev. Lett.* **1998**, *81*, 4764–4767.

[212] Armstrong, J.; Wenby, R.; Meiselman, H.; Fisher, T. *Biophys. J.* **2004**, *87*, 4259–4270.

[213] Sorci, G. A.; Reed, W. F. *Langmuir* **2002**, *18*, 353–364.

[214] Striegel, A.; Yau, W. W.; Kirkland, J. J.; Bly, D. D. *Modern Size-Exclusion Liquid Chromatography: Practice of Gel Permeation and Gel Filtration Chromatography*; John Wiley & Sons, 2009.

[215] Heller, C. *Electrophoresis* **1998**, *19*, 1691–1698.

[216] Gray, D. G. *Faraday Discuss. Chem. Soc.* **1985**, *79*, 257–264.

[217] Bockris, J. O.; Reddy, A. K. *Modern Electrochemistry*; Springer, 2000.

[218] Vologodskii, A. V.; Frank-Kamenetskii, M. D. *Methods Enzymol.* **1992**, *211*, 467–480.

[219] Allison, S. A. *Macromolecules* **1986**, *19*, 118–124.

[220] Grassia, P.; Hinch, E.; Nitsche, L. *J. Fluid Mech.* **1995**, *282*, 373–403.